Chapter 2

SHAPE RECTIFICATION OF 3D DATA OBTAINED BY A MOVING RANGE SENSOR BY USING IMAGE SEQUENCES

Atsuhiko Banno and Katsushi Ikeuchi

Abstract   For a large object, scanning from the air is one of the most efficient methods of obtaining 3D data. We have been developing a novel 3D measurement system, the Flying Laser Range Sensor (FLRS), in which a range sensor is suspended beneath a balloon. The obtained data, however, have some distortion due to movement during the scanning process. Then we propose a novel method to rectify the shape data obtained by a moving range sensor. The method rectifies them by using image sequences. We are conducting the Digital Bayon Project, in which our algorithm is actually applied for range data processing and the results show the effectiveness of our methods. Our proposed method is applicable not only to our FLRS, but also to a general moving range sensor.

1.   Introduction

We have been conducting some projects to model large scale cultural heritage objects such as great Buddhas, historical buildings and suburban landscapes [21, 16]. Basically, to scan these large objects, a laser range finder is usually used with a tripod positioned on stable locations. In the case of scanning a large scale object, however, it often occurs that some part of the object is not visible from the laser range finder on the ground. In spite of such a difficulty, we have scanned large objects from scaffolds temporally constructed nearby the object. However, this scaffold method requires costly, tedious construction time. In addition, it may be impossible to scan some parts of the object due to the limitation of available space for scaffold-building.

We are now conducting a project [15] to model the Bayon Temple [33] in Cambodia; the temple's size is about 150×150 square meters with over 40 meters in height. Scanning such a huge scale object from several scaffolds is unrealistic. To overcome this problem, several methods have been proposed.

For example, aerial 3D measurements can be obtained by using a laser range sensor installed on a helicopter platform[31]. High frequency vibration of the platform, however, should be considered to ensure that we obtain highly accurate results. To avoid irrevocable destruction, the use of heavy equipment such as a crane should be eschewed when scanning a cultural heritage object.

Figure 2.1.   The FLRS and the Bayon Temple

Based upon the above considerations, we proposed a novel 3D measurement system, a Flying Laser Range Sensor (FLRS)[14]. This system digitizes large scale objects from the air while suspended from the underside of a balloon platform (Fig.2.1). Our balloon platform is certainly free from high frequency vibration such as that of a helicopter engine. The obtained range data are, however, distorted because the laser range sensor itself is moving during the scanning processes (Fig.2.2).

In this study, we propose a method to rectify 3D range data obtained by a moving laser range sensor. Not only can this method be used in the case of our FLRS, it is also applicable to a general moving range sensor.

In this method based on "Structure from Motion", we use distorted range data obtained by a moving range sensor and image sequences obtained by a video camera mounted on the FLRS. The motion of the FLRS is roughly estimated only by the obtained images. And then the more refined parameters are estimated based on an optimization imposing some constraints, which include information derived from the distorted range data itself. Finally, using the refined camera motion parameters, the distorted range data are rectified.
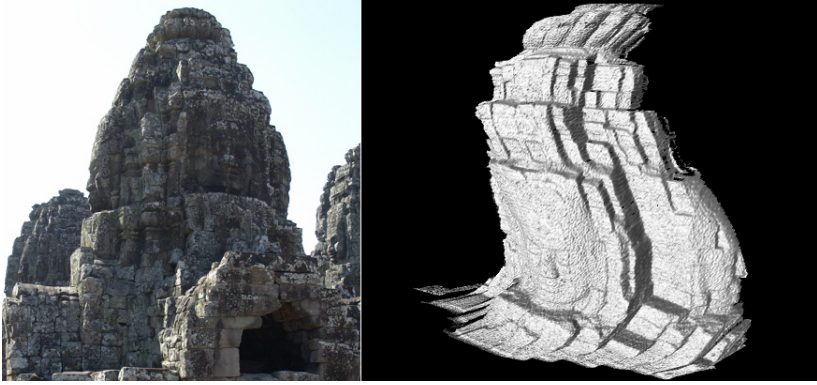
Figure 2.2. An sample snap shot and the distorted range data obtained by the FLRS.

This method is not limited to the case of our FLRS but also applicable to a general moving range sensor that has smooth motion. In this study, we do not utilize physical sensor such as gyros, INS and GPS for estimation of self position and pose.

2. Full Perspective Factorization

Estimations of the shape of an object or of camera motion by using images are called "Shape from Motion " or "Structure from Motion ", and are main research fields in computer vision.

The factorization method proposed in [32] is one of the most effective algorithms for simultaneously recovering the shape of an object and the motion of the camera from an image sequence. Then the factorization was extended to several perspective approximations and applications [8, 23, 7, 25, 12, 11].

[25] also presented perspective refinement by using the solution under the para-perspective factorization as the initial value. In [12] a factorization method with a perspective camera model was proposed. Using the weak-perspective projection model, they iteratively estimated the shape and the camera motion under the perspective model.

2.1 Weak-Perspective Factorization

Given a sequence of F images, in which we have tracked P interest points over all frames, each interest point p corresponds to a single point $\vec{S_p}$ on the object. In image coordinates, the trajectories of each interest point are denoted as $\{(u_{fp}, v_{fp})|f = 1, ..., F, p = 1, ..., P, \ 2F \geq P\}$.

Using the horizontal coordinates u_{fp}, we can define an $F \times P$ matrix U. Each column of the matrix contains the horizontal coordinates of a single point

in the frame order, while each row contains the horizontal coordinates for a single frame. Similarly, we can define an $F \times P$ matrix V from the vertical coordinates v_{fp}.

The combined matrix of $2F \times P$ becomes the measurement matrix as follow.

$$W = \left(\frac{U}{V} \right) \tag{2.1}$$

Each frame f is taken at camera position $\vec{T_f}$ in the world coordinates. The camera pose is described by the orthonormal unit vectors $\vec{i_f}$, $\vec{j_f}$ and $\vec{k_f}$. The vectors $\vec{i_f}$ and $\vec{j_f}$ correspond to the x and y axes of the camera coordinates, while the vector $\vec{k_f}$ corresponds to the z axis along the direction perpendicular to the image plane (Fig.2.3).
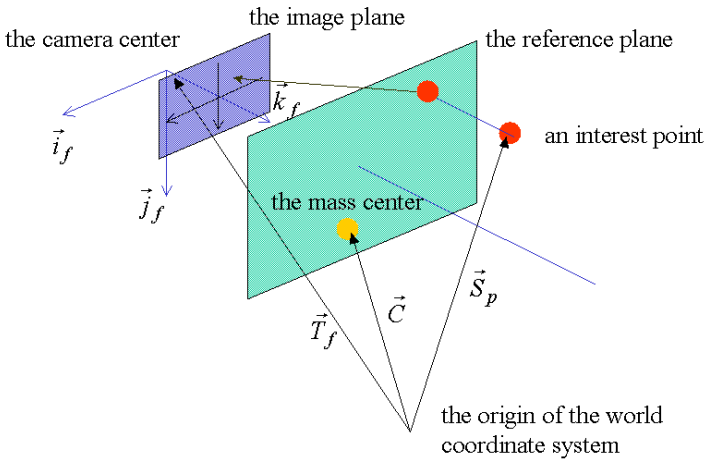

Figure 2.3. The Coordinate System: $\vec{T_f}$ denotes the position of the camera at time of frame f. The camera pose is determined by three unit basis vectors.

Under the weak-perspective camera model, a single point in the world coordinates $\vec{S_p}$ is projected onto the image plane f as (u_{fp}, v_{fp}).

$$u_{fp} = \frac{f}{z_f} \vec{i_f}^t \cdot (\vec{S_p} - \vec{T_f}) \tag{2.2}$$

$$v_{fp} = \frac{f}{z_f} \vec{j_f}^t \cdot (\vec{S_p} - \vec{T_f}) \tag{2.3}$$

$$where \; z_f = \vec{k_f}^t \cdot (\vec{C} - \vec{T_f}) \tag{2.4}$$

The vector $\vec{C}$ is the center of mass of all interesting points. Without loss of generality, the origin of the world coordinates can be placed at the centroid, that is $\vec{C} = 0$. Then this means that $z_f = -\vec{k_f} \cdot \vec{T_f}$ to simplify the expansion

of the following formulation. To summarize,

$$\begin{cases} u_{fp} = \vec{m_f}^t \cdot \vec{S_p} + \mathbf{x_f} \\ v_{fp} = \vec{n_f}^t \cdot \vec{S_p} + \mathbf{y_f} \end{cases} \tag{2.5}$$

$$where \quad \vec{m_f} = \frac{f}{z_f}\vec{i_f}, \quad \mathbf{x_f} = -\frac{f}{z_f}\vec{i_f}^t \cdot \vec{t_f}$$

$$\vec{n_f} = \frac{f}{z_f}\vec{j_f}, \quad \mathbf{y_f} = -\frac{f}{z_f}\vec{j_f}^t \cdot \vec{t_f}$$

Using that the center of all interest points is the origin,

$$\sum_{p=1}^{P} u_{fp} = \sum_{p=1}^{P} \vec{m_f}^t \cdot \vec{s_p} + \sum_{p=1}^{P} \mathbf{x_f} = P\mathbf{x_f} \tag{2.6}$$

similarly,

$$\sum_{p=1}^{P} v_{fp} = P\mathbf{y_f} \tag{2.7}$$

We obtain the registered measurement matrix $\tilde{W}$, after translation $\tilde{W} = W - (\mathbf{x_1}\ \mathbf{x_2}\ \ldots\ \mathbf{x_F}\ \mathbf{y_1}\ \ldots\ \mathbf{y_F})^t$
$(1, \ldots 1)$ as a product of two matrices M and S.

$$\tilde{W} = M \cdot S \tag{2.8}$$

$$where \quad M : 2F \times 3 Matrix \quad S : 3 \times P Matrix$$

The rows of the matrix M represent the orientation of the camera coordinates axes throughout the sequence, while the columns of the matrix S represent the coordinates of the interest points in the world coordinates. Both matrices are at most rank 3. Therefore, by using the Singular Value Decomposition (SVD), we can find the best approximation to $\tilde{W}$.

2.2   Extension to Full-Perspective Factorization

The above formulation is under the weak perspective projection model, which is a linear approximation of the perspective model. Next, using an iterative framework, we obtain approximate solutions under the non-linear, full-perspective projection model.

Under the perspective projection model, the projective equations between the object point $\vec{S_p}$ in 3D world and the image coordinate (u_{fp}, v_{fp}) are written as

$$u_{fp} = f\frac{\vec{i_f}^t \cdot (\vec{S_p} - \vec{T_f})}{\vec{k_f}^t \cdot (\vec{S_p} - \vec{T_f})} \tag{2.9}$$

$$v_{fp} = f\frac{\vec{j_f}^t \cdot (\vec{S_p} - \vec{T_f})}{\vec{k_f}^t \cdot (\vec{S_p} - \vec{T_f})} \tag{2.10}$$

Replacing $z_f = -\vec{k}_f^{\,t} \cdot \vec{T}_f$, we obtain the following equations.

$$(\lambda_{fp} + 1)u_{fp} = \frac{f}{z_f}\vec{i}_f^{\,t} \cdot (\vec{S}_p - \vec{T}_f) \tag{2.11}$$

$$(\lambda_{fp} + 1)v_{fp} = \frac{f}{z_f}\vec{j}_f^{\,t} \cdot (\vec{S}_p - \vec{T}_f) \tag{2.12}$$

$$\lambda_{fp} = \frac{\vec{k}_f^{\,t} \cdot \vec{S}_p}{z_f} \tag{2.13}$$

Note that the right hand sides of Eq.2.11 and Eq.2.12 are the same form under the weak-perspective model (see Eq.2 and 3). This means, multiplying a image coordinate (u_{fp}, v_{fp}) by a real number λ_{fp} maps the coordinate in the full-perspective model space into the coordinate in the weak-perspective model space. Solving for the value of λ_{fp} iteratively, we can obtain motion parameters and coordinates of interest points under the full perspective projection model in the framework of weak-perspective factorization.

The entire algorithm of the perspective factorization is as follows:

Input: An image sequence of F frames tracking P interest points.

Output: The 3D positions of P interest points $\vec{S}_p$. The camera position $\vec{T}_f$ and poses $\vec{i}_f$, $\vec{j}_f$, $\vec{k}_f$ at each frame f.

  1 Given $\lambda_{fp} = 0$

  2 Supposing the Equations 2.11 and 2.12, solve for $\vec{S}_p, \vec{T}_f, \vec{i}_f, \vec{j}_f, \vec{k}_f$ and z_f through the weak-perspective factorization .

  3 Calculate λ_{fp} by Equation 2.13.

  4 Substitute λ_{fp} into step (2) and repeat the above procedure.

Until: λ_{fp}'s are close to ones at the previous iteration.

2.3   Tracking

As input, we need P interest points at each frame of a sequence, which are tracked identified points in the 3D world. There are several methods to derive interest points from images [22, 29]. Among them, we adopt *Harris operator* [13] and *SIFT key* [18] for derivation of interest points. SIFT key is robust against scale, rotation and affine transformation changes. The main reason why we adopt the method is its stability of points derivation and usefulness of the key, which has 128 dimensional elements and can be used for the identification for each point.

3.   Refinement

Without noise in the input, the factorization method leads to the excellent solution. As a result, the rectified 3D shape through the estimated camera parameters is valid. Real images, however, contain a bit of noise. Therefore, it is not sufficient to rectify range data obtained by the FLRS only through the factorization. For the sake of a more refined estimation of motion parameters, we impose three constraints: for tracking, movement, and range data. The refined camera motion can be found through the minimization of a global functional. To minimize the function, the solution by the full-perspective factorization is utilized as the initial value to avoid local minimums.

3.1   Tracking Constraint

As the most fundamental constraint, any interest point $\vec{S}_p$ must be projected at the coordinates (u_{fp}, v_{fp}) on each image plane. This constraint is well known as Bundle Adjustment [5]. When the structure, motion and shape have been roughly obtained, this technique is utilized to refine them through the image sequence. In our case, the constraint conducts the following function:

$$F_A = \sum_{f=1}^{F} \sum_{p=1}^{P} \left(\left(u_{fp} - f \frac{\vec{i}_f^{\,t} \cdot (\vec{S}_p - \vec{T}_f)}{\vec{k}_f^{\,t} \cdot (\vec{S}_p - \vec{T}_f)} \right)^2 \right.$$
$$\left. + \left(v_{fp} - f \frac{\vec{j}_f^{\,t} \cdot (\vec{S}_p - \vec{T}_f)}{\vec{k}_f^{\,t} \cdot (\vec{S}_p - \vec{T}_f)} \right)^2 \right) \tag{2.14}$$

The minimization of F_A leads to the correct tracking of fixed interest points by a moving camera. However, we can see that the presence of parameters we are trying to estimate in the denominator makes this equation a difficult one. We have to seek the optimal solution via some non-linear minimization techniques. Then, suppose that instead, we consider the following function:

$$F'_A = \sum_{f=1}^{F} \sum_{p=1}^{P} \left(\left(\vec{k}_f^{\,t} \cdot (\vec{S}_p - \vec{T}_f) u_{fp} - f \cdot \vec{i}_f^{\,t} \cdot (\vec{S}_p - \vec{T}_f) \right)^2 \right.$$
$$\left. + \left(\vec{k}_f^{\,t} \cdot (\vec{S}_p - \vec{T}_f) v_{fp} - f \cdot \vec{j}_f^{\,t} \cdot (\vec{S}_p - \vec{T}_f) \right)^2 \right) \tag{2.15}$$

The term $\vec{k}_f^{\,t} \cdot (\vec{S}_p - \vec{T}_f)$ is the depth, the distance between the optical center of camera f and a plane, which is parallel to the image plane and include the point $\vec{S}_p$. The cost function F_A is the summation of squared distances on the image plane while the cost function F'_A is estimated on the plane of the point $\vec{S}_p$.

3.2   Smoothness Constraint

One of the most significant reasons for adopting a balloon platform is to be free from the high frequency that occurs with a helicopter platform [14]. A balloon platform is only under the influence of low frequency: the balloon of our FLRS is held with some wires swayed only by wind. This means that the movement of the balloon is expected to be smooth. Certainly, the movement of the balloon is free from rapid acceleration, rapid deceleration, or acute change in course. Taking this fact into account, we consider the following function:

$$F_B = \int \left(w_1 \left(\frac{\partial^2 \vec{T}_f}{\partial t^2} \right)^2 + w_2 \left(\frac{\partial^2 \mathbf{q}_f}{\partial t^2} \right)^2 \right) dt \qquad (2.16)$$

Here, $\vec{T}_f$ denotes the position of the camera, t is time, w_1, w_2 are weighted coefficients, and $\mathbf{q}_f$ is a unit quaternion that represents the camera pose. The first term of the above integrand represents smoothness with respect to the camera's translation while the second represents smoothness with respect to the camera's rotation. When the motion of the camera is smooth, the function F_B becomes a small value.

We implement in practice the following discrete form:

$$F_B' = \sum_{f=1}^{F} \left(w_1 \left(\frac{\partial^2 \vec{T}_f}{\partial t^2} \right)^2 + w_2 \left(\frac{\partial^2 \mathbf{q}_f}{\partial t^2} \right)^2 \right) \qquad (2.17)$$

3.3   Range Data Constraint

Taking a broad view of range data obtained by the FLRS, the data are distorted by the swing of the sensor. We can find, however, that these data contain instantaneous precise information locally; that information is utilized for refinement of the camera motion.

The FLRS re-radiates laser beams in raster scan order. This means that we can instantly obtain the time when each pixel in the range image is scanned because the camera and the range sensor are calibrated. If the video camera is synchronized with the range sensor, we can find the frame among the sequence when the pixel is scanned. With the video camera calibrated with the range sensor, we can also obtain the image coordinate of each interest point in the 3D world with respect to the instantaneous local coordinate.

Considering this constraint, we can compensate the camera motion.

When the range sensor scans interest point $\vec{S}_p$, we can conduct the third constraint to be minimized as follows:

$$F_C = \sum_{p=1}^{P} \left\| \mathbf{x}_{fp} - R^t (\vec{S}_p - \vec{T}_{fp}) \right\|^2 \qquad (2.18)$$

Here, the index fp denotes the frame number when the range sensor scans interest point $\vec{S}_p$. It is very significant to note that $\mathbf{x}_{fp}$ is the 3D coordinate val-

ues not described in the sensor-oriented coordinate system but in the camera-oriented one, which is rewritten based on the range data and camera-sensor calibration. In practice, we find sub-frame fp by using a linear interpolating technique for the motion of interest points between frames. The main purpose of the above constraint is to adjust the absolute scale.

As $\mathbf{x}_{fp} = (x_{fp}, y_{fp}, z_{fp})$, the above function can be rewritten as the stronger constraint:

$$
\begin{aligned}
F'_C &= \sum_{p=1}^{P} \Big(\big(x_{fp} - \vec{i}_{fp}^{\ t} \cdot (\vec{S_p} - \vec{T_{fp}}) \big)^2 \\
&+ \big(y_{fp} - \vec{j}_{fp}^{\ t} \cdot (\vec{S_p} - \vec{T_{fp}}) \big)^2 \\
&+ \big(z_{fp} - \vec{k}_{fp}^{\ t} \cdot (\vec{S_p} - \vec{T_{fp}}) \big)^2 \Big)
\end{aligned} \tag{2.19}
$$

3.4 The Global Cost Function

Based on the above considerations, we can understand that the next cost function should be minimized. Consequently, the weighted sum

$$
F = w_A F'_A + w_B F'_B + w_C F'_C \tag{2.20}
$$

leads to a global function. The coefficients w_A, w_B and w_C are determined experimentally, and we will discuss them later.

To minimize this function, we employ Fletcher-Reeves method or Polak-Ribiere method [26, 17, 30], which are types of the conjugate gradient method (in the next section, we explain the conjugate gradient method briefly). Then, we use the golden section search to determine the magnitude of gradient directions. For optimization, Levenberg-Marquardt method [19] is generally employed to minimize a functional value. Levenberg-Marquardt method is very effective in estimating function's parameters, especially in fitting a certain function. However in our function, minimizing the value of F'_B is not a parameter fitting problem. All we have to do is to simply decrease F'_B. Therefore we adopt the conjugate gradient method.

4. FLRS

FLRS(Flying Laser Range Sensor) has been developed to measure large objects from the air by using a balloon without constructing any scaffolds (Fig. 2.4).

We have two types of FLRSs. Each FLRS is composed of a scanner unit, a controller and a personal computer (PC). These three units are suspended beneath a balloon.

The scanner unit includes a laser range finder, especially designed to be suspended from a balloon. Figure 2.5 shows the interior of the scanner unit. It consists of a spot laser radar unit and two mirrors. We chose the LARA25200

Figure 2.4.   The FLRS (25m sensor)

and LARA53500 supplied by Zoller+Fröhlich GmbH[2] as laser radar units
because of their high sampling rate. Each laser radar unit is mounted on each
FLRS scanner unit. Two systems equipped with Lara25200 and LARA53500
 are respectively referred to as "25m sensor" and "50m sensor". The specifica-
tions of two units are shown in Table 2.1.

Table 2.1.   The specifications of the 25m (LARA25200) and 50m (LARA53500) Sensors

	25m Sensor	50m Sensor
Ambiguity interval	25.2 m	53.5 m
Minimum range	1.0 m	1.0 m
Resolution	1.0 mm	1.0 mm
Sampling rate	$\leq$ 625,000 pix/s	$\leq$ 500,000 pix/s
Linearity error	$\leq$ 3 mm	$\leq$ 5mm
Range noise at 10m	$\geq$ 1.0 mm	$\geq$ 1.5mm
Range noise at 25m	$\geq$ 1.8 mm	$\geq$ 2.7mm
Laser output power	23 mW	32mW
Laser wavelength	780nm	780nm

Both sensors have the similar mirror configurations. There are two mirrors
inside each unit to give a direction to the laser beam. One is a polygon mir-
ror with 4 reflection surfaces, which determines the azimuth of the beam. In
normal use, the polygon mirror, which rotates rapidly, controls the horizontal
direction of the laser beam. Another is a plane mirror (swing mirror) which de-

termines the elevation of the beam. The plane mirror swings slowly to controls the vertical direction of the laser beam.
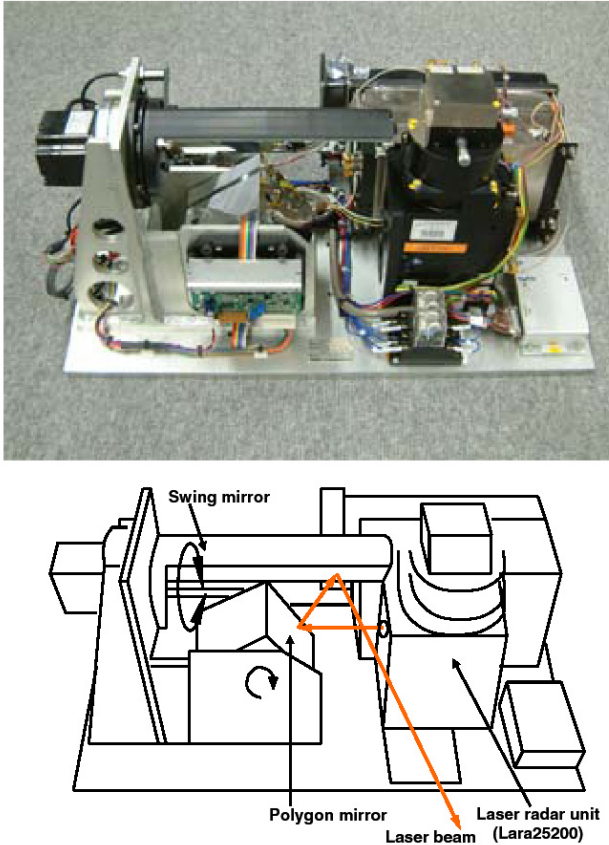

Figure 2.5.   The interior of scanner unit (25m sensor)

The lase beam emitted from the LARA is hit on a surface of the polygon mirror at first. Then the polygon mirror reflects the laser beam onto the plane mirror. The plane mirror also reflects the beam outside the unit(lower of Fig.2.5).

The combination of two mirror demonstrate the specifications as in Table 2.

5.   Experiments

We have been conducting the "Digital Bayon Project", in which the geometric and photometric information of the Bayon Temple is preserved in digital form. With respect to the acquisition of the geometric data, large parts of

Table 2.2.   The specifications of the 25m sensor and 50m sensor

	25m Sensor	50m Sensor
Angle Resolution		
Horizontal	0.05 deg	0.05 deg
Vertical	0.02 deg	0.02 deg
Horizontal field	$\leq$ 90 deg	$\leq$ 90 deg
Vertical field	$\leq$ 30 deg	$\leq$ 30 deg
Scanning period/image	$\leq$ 15 sec	$\leq$ 1 sec

the temple visible from the ground are scanned by range sensors placed on the ground. On the other hand, some parts invisible from the ground, for example, roofs and tops of towers, are scanned by our FLRS system.

The left side of Fig.2.6 shows a photo of the scanned area. On the right side of Fig.2.6, the dense fine model is the correct shape obtained by the Cyrax-2500 [1] fixed on the ground.

Figure 2.6.   A scene for this experiment. Left - a photo of an object; Right - 3D model obtained by the Cyrax-2500 fixed on the ground.

There are data missing in the model. To fill in the missing pieces of the model obtained by the sensor on the ground, we utilize our FLRS effectively. Figure 2.7 shows a sample image of the sequence obtained by the video camera. In this experimental data set, it takes one second for a range image: thirty pictures are saved in the meantime.

The result is shown in Fig.2.8. The upper shape in Fig.2.8 is the original one obtained from the FLRS. We can see that the shape is widely deformed. In the middle of Fig.2.8, the rectified shape by full-perspective factorization is shown. With respect to motion parameters, the ambiguity in scale is removed manually. At a glance, the factorization seems to rectify the shape properly. In

Figure 2.7. A sample shot of the image sequence

detail, however, the distortion in S shape is still left. Especially, the shape of the entrance is skewed. On the other hand, the lower shape is rectified correctly by our method. It is clear that the distortion in S shape is removed and the shape of the entrance is correctly recovered into a rectangle.

To evaluate the accuracy of our shape rectification algorithm, we compare the rectified shape with other data, which are obtained by a range finder, the Cyrax-2500, positioned on the ground. Aligning two data sets by using the conventional ICP algorithm [3] [6], we analyze the overlapping area.

Figure 2.9 indicates the point-to-point distances in the ICP algorithm. The region where the distances between them are less than 6.0 cm is colored light gray. The area where the distances are farther than 6.0 cm is displayed in dark gray. The upper figure shows the comparison between the correct shape and the original distorted one obtained by the FLRS. The middle one shows the rectified shape by the full-perspective factorization without ambiguity in scale. The lower shows the rectified shape by our method.

At a glance, the light gray region is clearly expanded by our rectification algorithm. Some parts of the rectified shape are colored dark gray because of the lack of corresponding points. Taking account of the fact that the correct shape of the parts invisible from the ground could not be measured, the proposed method could rectify the 3D shape correctly.

Table 2.3 shows a quantitative evaluation for our method. This table indicates the ratios of match region and the average distances between the Cyrax's model and the above three models. These numbers show that our method increases the match region and bring the distorted model by the FLRS to the correct one. We can see that our method was able to rectify the FLRS data properly.

Table 2.3.   The evaluation of the rectified models. (a)The original distorted model. (b)The rectified model by the full-perspective factorization removing the scale ambiguity manually. (c)The rectified model by our method.

	(a)	(b)	(c)
match region (%)	37.2	49.8	62.7
error (average) [cm]	20.46	10.55	2.11

Figure 2.10 shows several samples of the method.

6.   Conclusions

In this chapter, we have described FLRS system and a proposed method to rectify 3D range data obtained by a moving laser range sensor.

We described how an outstanding measurement system FLRS was built to scan large objects from the air. This system allowed us to measure the large cultural heritage objects by using a balloon. To rectify the distorted shapes obtained from the FLRS, we proposed a rectification method based on the "Structure from Motion" techniques by using image sequences.

We utilized distorted range data obtained by a moving range sensor and image sequences obtained by a video camera mounted on the FLRS. First, the motion of the FLRS was estimated through full perspective factorization only by the obtained image sequences. Then the more refined parameters were estimated based on an optimization imposing three constraints: the tracking, smoothness and range data constraints. Finally, refined camera motion parameters rectified the distorted range data.

This method has shown proper performance and practical utilities.

Our method can be generally applied to a framework in which a range sensor moves during the scanning process, and is not limited to our FLRS because we impose only the smooth movement constraint.

Acknowledgments

References

[1] www.leica-geosystems.com.

[2] www.zf-lase.com.

[3] P. J. Besl and N. D. McKay. A method for registration of 3-D shapes. *IEEE Trans. on PAMI*, 14:239–256, 1992.

[4] G. Blais and M. D. Levine. Registering multiview range data to crate 3D computer objects. *IEEE Trans. on PAMI*, 17(8):820–824, 1995.

[5] D. Brown. The bundle adjustment – progress and prospect. In *XIII Congress of the ISPRS*, Helsinki, 1976.

[6] Y. Chen and G. Medioni. Object modeling by registration of multiple range images. *Image and Vision Computing*, 10(3):145–155, 1992.

[7] S. Christy and R. Horaud. Euclidean shape and motion from multiple perspective views by affine iterations. *IEEE Trans. on PAMI*, 18(11):1098–1104, 1996.

[8] J. Costeira and T. Kanade. A multi-body factorization method for motion analysis. In *Proc. of ICCV1995*, pages 1071–1076, 1995.

[9] J. H. Friedman, J. L. Bentley, and R. A. Finkel. An algorithm for finding best-matches in logarithmic time. *ACM Trans. on Mathematical Software*, 3(3):209–226, 1977.

[10] P. Gill, W. Murray, and M. Wright. *Practical Optimization*. Academic Press, London, 1981.

[11] A. Gruber and Y. Weiss. Multibody factorization with uncertainty and missing data using the em algorithm. In *Proc. of CVPR2004*, volume 1, pages 707–714, 2004.

[12] M. Han and T. Kanade. Perspective factorization methods for euclidean reconstruction. Technical Report :CMU–RI–TR–99–22, Robotics Institute, Carnegie Mellon University, 1999.

[13] C. Harris and M. Stephens. A combined corner and edge detector. In *Proc. of Alvey Vision Conference*, pages 147–152, 1988.

[14] Y. Hirota, T. Masuda, R. Kurazume, K. Ogawara, K. Hasegawa, and K. Ikeuchi. Designing a laser range finder which is suspended beneath a balloon. In *Proc. of ACCV2004*, volume 2, pages 658–663, 2004.

[15] K. Ikeuchi, K. Hasegawa, A. Nakazawa, J. Takamatsu, T. Oishi, and T. Masuda. Bayon digital archival project. In *Proc. of VSMM2004*, pages 334–343, 2004.

[16] K. Ikeuchi, A. Nakazawa, K. Hasegawa, and T. Ohishi. The great buddha project: Modeling cultural heritage for VR systems through observation. In *Proc. of ISMAR2003*, 2003.

[17] D. A. Jacobs. *The State of the Art in Numerical Analysis*. Academic Press, London, 1977.

[18] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2):91–110, 2004.

[19] D. W. Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*, 11:431–441, 1963.

[20] T. Masuda, Y. Hirota, K. Nishino, and K. Ikeuchi. Simultaneous determination of registration and deformation parameters among 3D range images. In *Proc. of 3DIM2005*, pages 369–376, 2005.

[21] D. Miyazaki, T. Oishi, T. Nishikawa, R. Sagawa, K. Nishino, T. Tomomatsu, Y. Yakase, and K. Ikeuchi. The great buddha project: Modelling cultural heritage through observation. In *Proc. of VSMM2000*, pages 138–145, 2000.

[22] H. P. Moravec. Towards automatic visual obstacle avoidance. In *Proc. 5th International Joint Conference on Artificial Intelligence*, page 584, 1977.

[23] T. Morita and T. Kanade. A sequential factorization method for recovering shape and motion from image streams. *IEEE Trans. on PAMI*, 19(8):858–867, 1997.

[24] P. Neugebauer. Geometrical cloning of 3D objects via simultaneous registration of multiple range images. In *Proc. of the International Conference on Shape Modeling and Application*, pages 130–139, 1997.

[25] C. Poelmann and T. Kanade. A paraperspective factorization method for shape and motion recovery. *IEEE Trans. on PAMI*, 19(3):206–218, 1997.

[26] E. Polak. *Computational Methods in Optimization*. Academic Press, New York, 1971.

[27] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in C*. Cambridge University Press, 1988.

[28] S. Rusinkiewicz and M. Levoy. Efficient variant of the ICP algorithm. In *Proc. of 3DIM2001*, pages 145–152, 2001.

[29] S. M. Smith and M. Brady. SUSAN - a new approach to low level image processing. *IJCV*, 23(1):45–78, 1997.

[30] J. Stoer and R.Bulirsh. *Introduction to Numerical Analysis*. Springer-Verlag, New York, 1980.

[31] S. Thrun, M. Diel, and D. Haehnel. Scan alignment and 3-D surface modeling with a helicopter platform. In *Proc. of the 4th International Conference on Field and Service Robotics*, 2003.

[32] C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: a factorization method. *IJCV*, 9(2):137–154, 1992.

[33] J. Visnovcova, L. Zhang, and A. Gruen. Generating a 3D model of a bayon tower using non-metric imagery. In *Proc. of the International Workshop Recreating the Past –Visualization and Animation of Cultural Heritage*, 2001.

[34] E. Walter and L. Prontazo. *Identification of Parametric Models from Experimental Data*. Springer, 1997.

[35] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *IJCV*, 13:119–152, 1994.
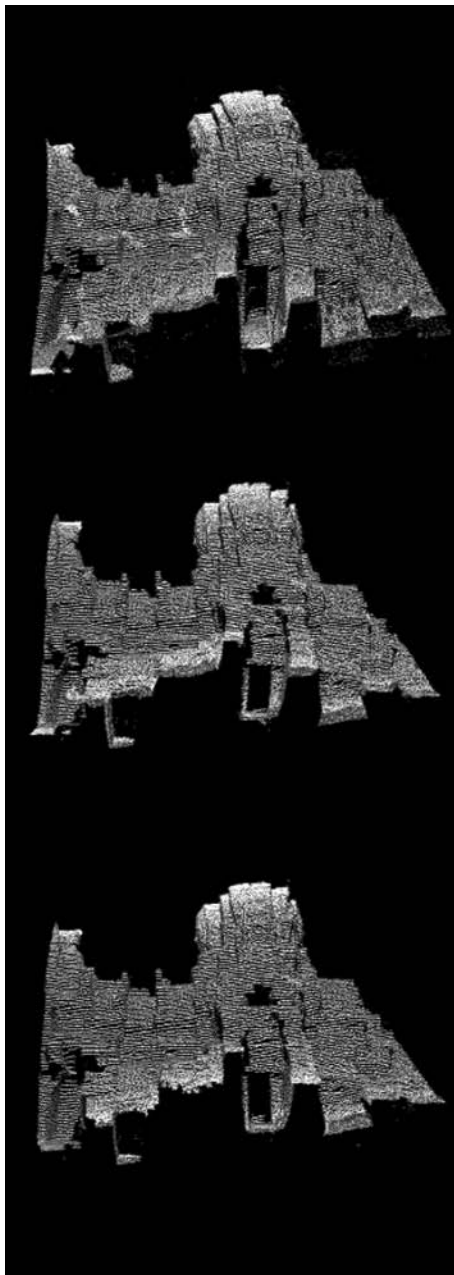
Figure 2.8.   The upper figure shows the original distorted shape obtained by the FLRS. The middle one shows the rectified shape by the full-perspective factorization without ambiguity in scale. The lower shows the rectified shape by our method.
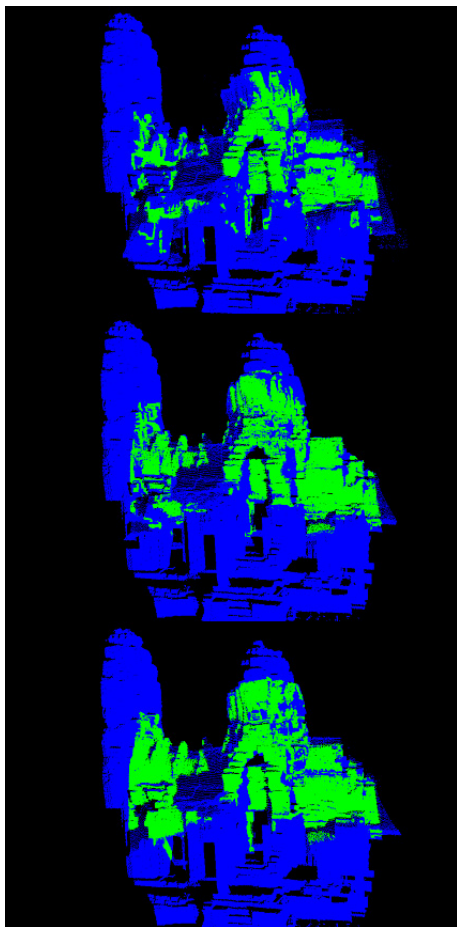
Figure 2.9. The upper figure shows the comparison between the correct shape and the original distorted one obtained by the FLRS. The light gray region indicates where the distance of two shapes is less than 6.0 cm. The middle one shows the rectified shape by the full-perspective factorization without ambiguity in scale. The lower shows the rectified shape by our method.
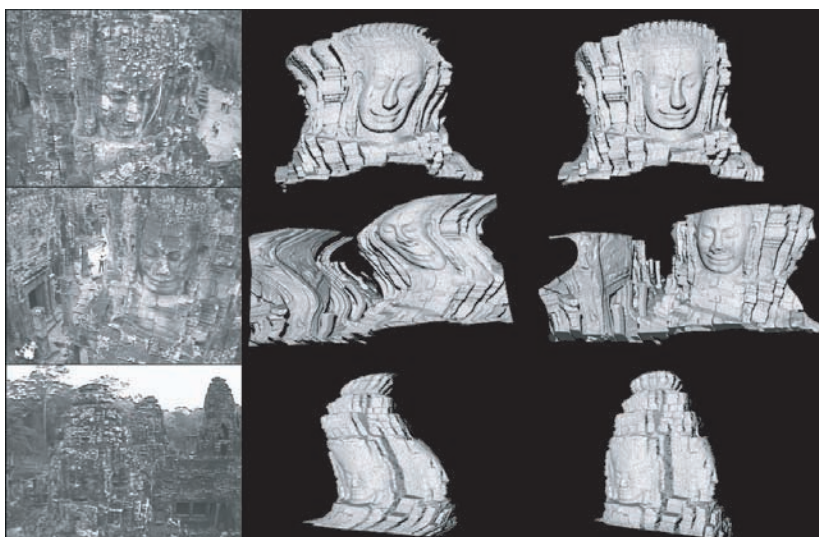
Figure 2.10.   Some sample photos by FLRS (left), the original distorted data sets (center) and the rectified sets (right)