Interactive Shape Acquisition using Marker Attached Laser Projector

Ryo Furukawa
Faculty of Information Sciences,
Hiroshima City University, Japan
ryo-f@cs.hiroshima-cu.ac.jp

Hiroshi Kawasaki
Institute of Industrial Science,
The University of Tokyo, Japan
h-kawa@sak.iis.u-tokyo.ac.jp

Abstract

In this paper, a 3D shape acquisition system consisting of a laser projector and a video camera is described. A user projects a stripe of light to the measured 3D scene by hand. The projected light and LED markers attached to the laser projector are captured with the video camera. By estimating the 3D pose of the laser projector from the 2D locations of the markers, the distance to the surface lit by the laser can be calculated. The use of a laser projector that projects a stripe of light enables correction of shape estimation errors from known 3D surfaces. With this system, the user can measure the 3D shapes of objects in real time, thereby confirming the shape acquisition of each part of the surface.

1. Introduction

As CG technology has become common, the specs needed for 3D scene measurement systems have varied. All kinds of measurement methods have been studied, from air borne land-form measurement systems to desktop range finders, with each having specific pros and cons.

Using the currently available laser range finders, long range, very precise and dense depth images can be acquired [1], but the costs of these specially designed devices are very high. Structured light based stereo or laser stereo systems are for middle or short range use, with even more precise and dense measurements[2], but they also require special light projecting devices, making their total costs very expensive.

Stereo vision or motion stereo systems can be much less expensive since they basically use common digital still cameras or video cameras. Some of them are capable of reconstructing very large scale scenes, including architectural features like buildings or landscapes [11]. The disadvantage of these systems is that most of them are feature point based, making it difficult to get dense shape data without some kind of sophisticated interpolation mechanism.

Since universal device interfaces like IEEE 1394 or USBs have become common, connecting still or video cameras to computers has gotten much easier, suggesting that vision-based, personal or desktop shape acquisition systems can be developed. Currently, there are several research studies which describe low cost, easily available 3D measurement systems [3] [7] [4] [14] [5] [13].

This paper describes a vision based, low cost shape recovery system. The devices needed for this system are a video camera and a laser projector mounted with LED markers for pose detection. The cost of the system is low because the laser projectors and LEDs can be purchased inexpensively. Since this system doesn't require a special platform for the objects being measured, it is very portable and can be used to measure small portions of large objects, for example, reliefs on relics. The laser projector in our system emits a laser plane rather than a laser beam. This makes it easier to obtain dense depth images, and also makes it possible to correct errors of 3D position estimation using information on known surface positions. The measurement process is done in real time as the scene is captured by video camera. Therefore, users can confirm which part of the depth image lacks data while measuring the scene.

Also described in this paper are the results of error analyses on our device. The analyses imply that attaching additional LED marker can improve the estimated mean squared errors of the final depth estimations. We could get some experimental results which supports the results of our analyses.

2. Related Works

Recently, a number of low cost and easily available 3D measurement systems have been proposed. In order to be low cost, most of these systems avoid to using a mechanical system. Without using a mechanical system, it is necessary to estimate the position of each sensor in order to employ triangulation to acquire 3D information. So far, various types of positioning methods have been proposed.

Bouguet and Perona[3] are using shadows to calibrate the camera and light source positions. They cast shadows on the object, and estimate the 3D values using these calibrated

parameters. Fisher et al.[7] have modified this idea.

Woo and Jung[4] have proposed another solution. They put a cubic frame in the field of view of the camera, and put the object inside this frame. They then emit a line beam to the object, and detect the bright point on the cube to estimate the pose of the beam plane and the 3D information.

Takatsuka et al.[14] have adopted a more active method. They put LED markers on the sensor itself, and capture it with a single camera to estimate the sensor position. They use the laser pointer as a sensor, and have proposed a clever method for efficient 3D data calculation using the detected LED markers.

Davis and Chen [5] have applied the stereo vision technique, using mirrors to archive an efficient stereo algorithm using a single camera.

Since we have attached LED markers on a laser projector and capture the markers with a video camera to estimate position of the projector, our method is close to Takatsuka's work [14]. However, we used a laser projector which emits a fan-shaped sheet of light spreading within a plane instead of a beam of light. Therefore, the calibration and estimation method is closely related to formerly mentioned research [3][7][4].

We have also recognized the importance of the interactive scanning system stated by Rusinkiewicz et al.[13], and therefore one of our main research purposes is to realize a real-time system.

3. Shape measurement using a marker attached laser projector

3.1. System configuration

Our system consists of a video camera and a laser projecting device. The laser projecting device emits a plane of light, rather than a beam. While measuring, the user hold the laser projector by hand and projects the laser onto the measured object. The sheet of laser produces a stripe on the object, and the user moves the laser so that the stripe scans the entire object surface. The projected stripe is observed by the video camera. The shape of the target object is obtained by analyzing the video sequence. In this paper, the laser projecting device is referred to as "a wand." Figure 1 shows the wand and the concept of the 3D measurement.

The location and pose of the laser pointer are also obtained from visual information. To avoid instability of complex image processing, four LED markers are attached to the laser pointer as shown in figure 1 (a). Four of the markers are located so that they form a square and so that the square is placed on the same plane as the sheet of light as much as possible. The other two markers are located with certain distances from the square.

These markers define the local coordinates in 3D space.

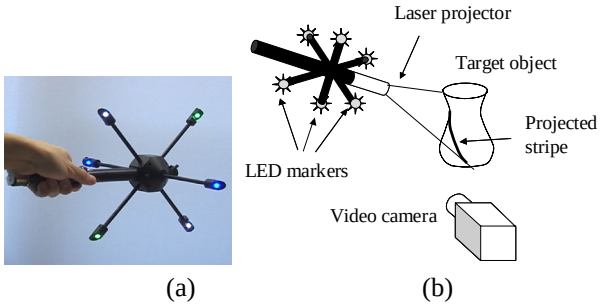


Figure 1. System configuration: (a) shows the wand and (b) shows the concept the of 3D measurement.

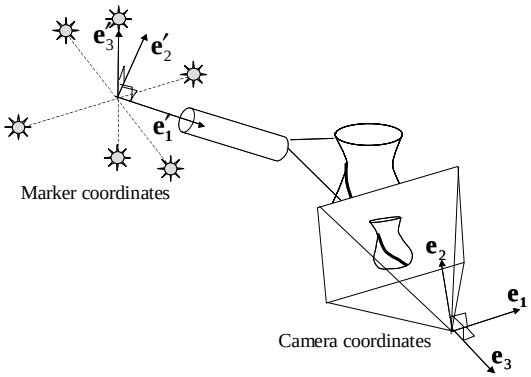


Figure 2. Coordinate systems

The origin of the "marker coordinates" is placed at the center of the square formed by the four markers. We define the marker coordinate system by three unit vectors, $\mathbf{e}'_1$,$\mathbf{e}'_2$ and $\mathbf{e}'_3$. The first and the second coordinates, whose directions are expressed by $\mathbf{e}'_1$ and $\mathbf{e}'_2$ are taken so that they are parallel to the edges of the square formed by markers. The first coordinate is parallel to the laser projector. The direction of the third coordinate $\mathbf{e}'_3$ is orthogonal to $\mathbf{e}'_1$ and $\mathbf{e}'_2$. We also define camera coordinates defined by $\mathbf{e}_1$,$\mathbf{e}_2$ and $\mathbf{e}_3$, where $\mathbf{e}_1$ and $\mathbf{e}_2$ are respectively parallel to horizontal and vertical directions of image plane, and $\mathbf{e}_3$ directed towards the backward direction of the camera. Figure 2 shows the definitions of the coordinate systems. Here, we name 2 planes. The first one is the plane spanned by $\mathbf{e}'_1$ and $\mathbf{e}'_2$, which we call the "marker plane." The second one is the plane where the laser light is projected, which we call the "laser plane." If the markers are attached precisely, the marker plane and the laser plane overlap each other.

When the shape of the target object is scanned, the markers of the laser pointer are also captured into the frames. From the 2D locations of the markers in the video frame, the

parameters of the transformation between the camera coordinates and marker coordinates are estimated. The transformation consists of a 3D rotation matrix $\mathbf{R}$ and translation vector $\mathbf{t}$. Parameters of the transformation are estimated using the methods described in the following session.

3.2. Estimation of marker plane

From the captured image frame, the location the LEDs are obtained using simple thresholding and calculating the center of gravity of the connected pixels. From the detected LEDs, four LEDs on the marker plane are selected. To make the selection easy, the four LEDs on the marker plane have different colors from the other two.

After detecting the images of the four LED markers, a description of the lines on the image which go through the markers is calculated. Then, we select two pairs of lines which correspond to the opposite sides of the square that the markers form. For each pair of lines, simple equations are solved to acquire the coordinate of the crossing point of the lines. The crossing point is called a vanishing point. If the lines are parallel on the image, the VP is defined as the point at infinity at the direction of the parallel lines.

There are two vanishing points in image which correspond to two unit vectors of the marker coordinate system, one for the direction of $\mathbf{e}'_1$ and the other for the direction of $\mathbf{e}'_2$. Let the 2D coordinates of these vanishing points be described as (g_{0n}, g_{1n}) in pixels, where the origin of the coordinates is placed at the optical center of the image.

Then the 3D directions of those vanishing points ($\mathbf{r}_1$ and $\mathbf{r}_2$) are

$$\mathbf{r}_n = (g_{0n}, g_{1n}, -f)^{\mathbf{t}} \tag{1}$$

where f denotes focal length in pixel and $\circ^{\mathbf{t}}$ denotes transpose of vectors and matrices. We define $\hat{\mathbf{e}'_i} = \mathbf{r}_i/\|\mathbf{r}_i\|, (i = 1, 2)$. The unit vectors $\hat{\mathbf{e}'_i}$ are used as estimation of $\mathbf{e}'_i$. Estimation $\hat{\mathbf{e}'_3}$ is defined by taking a cross product of the two unit vectors as $\hat{\mathbf{e}'_3} = \hat{\mathbf{e}'_1} \times \hat{\mathbf{e}'_2}$. Figure 3 shows relationships between vanishing points and estimations $\hat{\mathbf{e}'_i}, (i = 1, 2)$.

Using three unit vectors, estimation $\hat{\mathbf{R}}$ of rotation matrix $\mathbf{R}$ can be described as:

$$\hat{\mathbf{R}} = \begin{pmatrix} \hat{\mathbf{e}_1} & \hat{\mathbf{e}_2} & \hat{\mathbf{e}_3} \end{pmatrix} \tag{2}$$

Let 2D coordinates of images of first and second markers be $(p_{11}, p_{12}), (p_{21}, p_{22})$. 3D locations in camera coordinates of the markers are

$$\mathbf{m}_1 = (p_{11}u, p_{12}u, -fu)^{\mathbf{t}}, \mathbf{m}_2 = (p_{21}v, p_{22}v, -fv)^{\mathbf{t}}. \tag{3}$$

where u, v are unknown variables. Let the 3D distance between those markers, which is known value, be D_m. Then

$$\|\mathbf{m}_1 - \mathbf{m}_2\| = D_m. \tag{4}$$

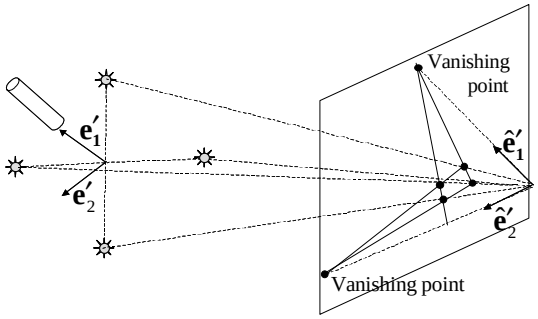


Figure 3. Estimating of direction of marker coordinates

The equation of the marker plane is expressed as $\mathbf{e}'^{\mathbf{t}}_{31}\mathbf{x} + d = 0$, where $\mathbf{x}$ is a 3D point expressed in camera coordinates. Since $\mathbf{m}_1$ and $\mathbf{m}_2$ are on the marker plane,

$$\mathbf{e}'^{\mathbf{t}}_{31}\mathbf{m}_i + d = 0 \ (i = 1, 2). \tag{5}$$

From equations (3),(4) and (5), d, u, v can be solved. Then we obtain $\hat{\mathbf{t}}$, as an estimation of $\mathbf{t}$, by calculating the center of gravity of 4 locations of the markers.

As mentioned above, the position and pose of the laser pointer are obtained as a transformation between the camera coordinates and the marker coordinates. To improve the estimation of the transformation, we optimize it with nonlinear optimization. All the detected LEDs including the ones which are not on the marker plane are used for the optimization. The real locations of the LEDs in the marker coordinates can be measured in advance. By transforming the measured locations with the estimated transformation, the estimated locations of markers' images can be obtained. The estimation of the transformation is refined by minimizing the sum of the squared distances between the actual LED positions on the images and the estimated marker positions, which is described as

$$f(\hat{\mathbf{R}}, \hat{\mathbf{t}}) = \sum_i \{proj(\hat{\mathbf{R}}\mathbf{m}'_i + \hat{\mathbf{t}}) - \mathbf{p}_i\}, \tag{6}$$

$$proj((x_1, x_2, x_3)^{\mathbf{t}}) = ((fx_1)/x_3, (fx_2)/x_3)^{\mathbf{t}} \tag{7}$$

where $\mathbf{p}_i$ is a location of image of the ith marker, $\mathbf{m}'_i$ a location of the ith detected marker in marker coordinates. $proj(\cdot)$ is an operation of projection from 3D position in camera coordinates into image coordinates. The rotation matrix $\mathbf{R}$ is expressed by roll, pitch and yow. The translation vector $\mathbf{t}$ is expressed by 3 components (t_1, t_2, t_3). Thus, the optimization is processed for 6 parameters. We use simplex descending algorithm as the optimization algorithm because of ease of implementation.

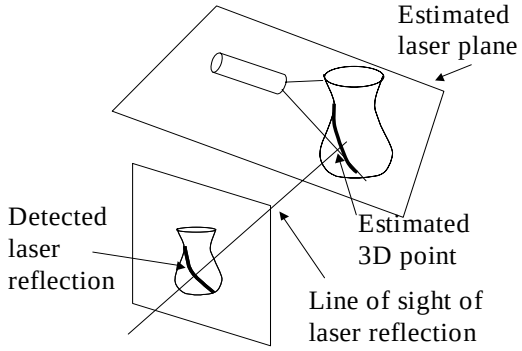


Figure 4. Estimating locations of surface points

3.3. Shape acquisition

Using the estimated parameters of the marker plane, which is approximately the same as the laser plane, we can obtain the location of the surface points by triangulation. First, the stripe of light projected onto the target object is extracted from the image. The pixels on the stripe are picked up by thresholding and skeletonizing. The line going through the origin of the camera coordinates and the surface point (line of sight) can be calculated for each pixel using intrinsic parameters, and the 3D location of the surface can be determined by taking the intersection of the line of sight and the estimated laser plane. Figure 4 shows how to estimate the surface locations by triangulation.

All calculations are employed in real-time and the estimated depth values are stored in a depth image; this depth image has the same width and height as captured video frames and thus, its data size is $16bit \times width \times height$. The system display the estimated depth values as an image to the user in real-time and it helps the user to recognize whether there is an estimated data or not for each pixel. Therefore, the user can interactively and efficiently scan the object.

Since this system scan the same pixel repeatedly, all the estimated depth values for each pixel are averaged to increase the data accuracy. The system also calculates the standard deviation for each pixel in real-time (after outliers are rejected) and shows the user this value as an evaluation of the reliability of the measured depth value to assist the scanning.

After a scanning has finished, x, y values for all the pixels are calculated by using the depth values and the intrinsic camera parameters as the final results.

3.4. Calibration of the laser plane in the marker coordinates

The LED markers are mounted on the laser pointer so that the markers plane and the laser plane overlaps as precisely as possible. Since the mounting is done manually, there are always errors in the placement of the LEDs. To avoid incorrectness in the measurement of the 3D surfaces, we calibrate the relationship between the laser plane and the markers.

The relationship can be defined as the parameters of the laser plane expressed in marker coordinates. If the markers are mounted correctly, the equation of the laser plane expressed in marker coordinates (x_1, x_2, x_3) is $x_3 = 0$.

For preparation of the calibration, the following measurement must be conducted in advance. First, a rectangular parallelepiped, or a box-shaped object, of a known size is captured. From the image, the location and pose of the object are estimated using methods of camera calibration. From the pose, the equations of the planes enveloping the box are estimated.

We project the laser light onto the surface of the target object, and capture the lit surface and the markers. From the markers, the transformation between the marker coordinates and camera coordinates can be estimated as described in the previous section. From the transformation and the laser plane expressed in marker coordinates, the laser plane in the camera coordinates can be calculated.

Next, the intersecting lines of the laser plane and the surface of the box shaped object are calculated. Then, the images of the lines can be determined. These lines are the locations where the light stripes are "supposed" to be projected if the given laser plane is correct.

The true location of laser reflection can be detected from the image. If there are errors in the given parameters of the laser plane, the true locations and calculated locations of the laser reflections don't match as described in Figure 5.

By calculating the sum of directions between the pixels of the real stripe and the presumed location of stripe, we can evaluate how those stripes match. The calibration is done by minimizing this value with non linear optimization for parameters of the laser plane. We then can get the laser plane parameters in the marker coordinates such that the estimated and detected stripes in the image match well. For the implementation for this paper, we applied the simplex descending algorithm [12] for the non-linear optimization using parameters for plane $x_3 = 0$ as initial parameters.

3.5. Correction of the laser pointer pose estimation using known 3D points

When the 3D locations of some part of the measured surface are known, there are constraints between the position of the laser plane and the location of the image of the light stripe projected on the known surface. If some part of the projected stripe of laser is on the region of surface whose depth values are known as shown in figure 6, the known depth value and the depth value calculated from the laser plane estimation should match. We use these constraints to

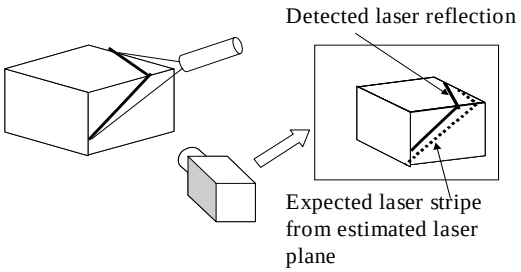


Figure 5. Calibration of laser plane

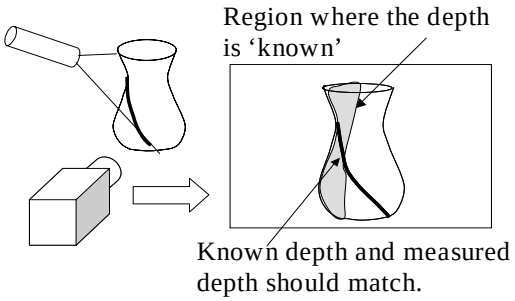


Figure 6. Constraints from known surface points

correct errors in the laser plane estimation due to image processing errors, asymmetry in the shape of LEDs themselves, errors in the intrinsic parameters of the cameras, etc.

We use differences between the known depth and measured depth to evaluate how the constraint is fulfilled. As already described, the pose of the light plane is expressed as a transformation between the camera and the marker coordinates, which is estimated by minimizing the error function (6). By using the sum of the differences between the known and measured depth and the function (6) instead of using only the function (6), we can correct the estimation of the laser plane.

In a practical 3D scene measurement, there is often no a priori information about surface of the scene. In this case, "known points" should be decided by certain criteria. In our work, the depth image pixels satisfying the following conditions are treated as "known":

- Depth of the pixel is measured more than N times.
- Standard deviation of measured depth values of the pixel is less than σ_d.

For the experiments described later, we defined $N = 5$ and $\sigma_d = 0.01$

4. Experiments and evaluations

4.1. Examples of measurements

To show the performance of the system described in this paper, we demonstrate two examples of the shape acquisition. The subject of the first example is a light stand shown in Figure 7(a). We measured the light stand with an image for the texture. Figures 7(b)-(e) show the textured 3D model rendered with different viewpoints. Figure 7(f) shows the acquired polygons around the top of the lightstand.

For the second example, we measured shape of a plastic toy from 11 different directions and registered the data sets. We used algorithms proposed by Miyazaki et al. [10] for the procedure of generating the polygon model. Figure 8 shows (a) a photograph of the subject, (b) the acquired 3D model with texture, (c) the acquired polygons rendered as wireframes, and (d) the registerd models with textures.

and (b) is the registered point set. For the first and the second example, each of the scanning from one direction took around 30 seconds.

4.2. Error analyses

To evaluate the errors of our shape acquisition system, we conducted error analyses. Let $\mathbf{y} = (y_1\, y_2\, \cdots\, y_n)^t$ be an observation vector, whose components are observed coordinates of the marker images, and x be the the objective value (the distance from the camera to a surface point lit by the laser sheet). From an observed vector $\mathbf{y}$, a pose estimation of the wand $\mathbf{p} = (p_1\, p_2\, \cdots\, p_m)^t$, is estimated first, then x is estimated from $\mathbf{p}$. Here, $\mathbf{p}$ denotes 6 parameters of $\mathbf{R}$ (roll, pitch and yow) and $\mathbf{t}$ (a 3D vector). We can define a mapping of projection from $\mathbf{p}$ to $\mathbf{y}$, and a mapping of distance calculation from $\mathbf{p}$ to x. The mappings are expressed as

$$x = f_c(\mathbf{p}) \quad (8)$$

$$\mathbf{y} = \mathbf{g}_c(\mathbf{p}) = (g_{c,1}(\mathbf{p})\, g_{c,2}(\mathbf{p})\, \cdots, g_{c,n}(\mathbf{p}))^t. \quad (9)$$

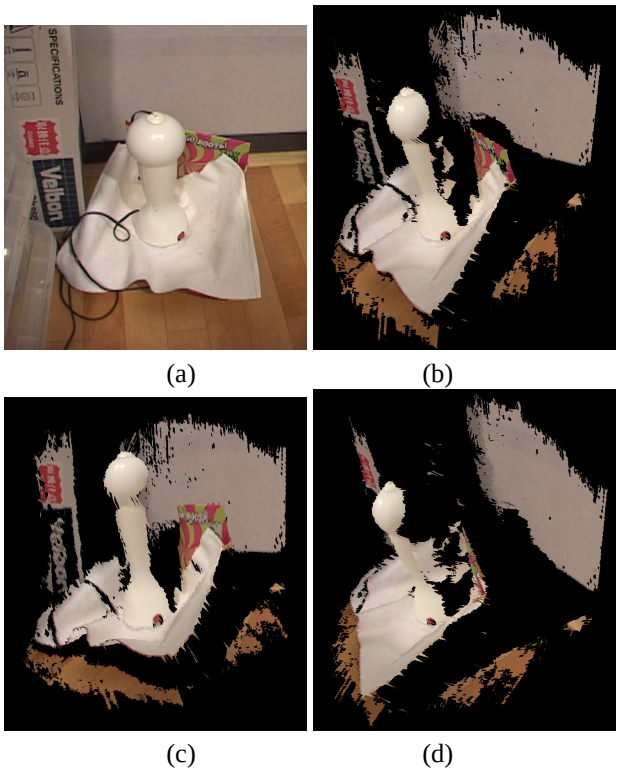
where c denotes the condition of measurement, such as the pose and location of the wand and the distance between the wand and the objective surface.

Let $\delta\mathbf{y} = (\delta y_1, \cdots, \delta y_m)$ be the observation error vector added to $\mathbf{y}$, $\delta\mathbf{p} = (\delta p_1, \cdots, \delta p_n)^t$ be the the estimation error vector for $\mathbf{p}$, and δx be the estimation error x. Using an abridged notation

$$\partial_{\mathbf{p}} = \left(\frac{\partial}{\partial p_1}\, \frac{\partial}{\partial p_2}\, \cdots\, \frac{\partial}{\partial p_m} \right)^t,$$

the relationships between the errors above are

$$\delta x = (\partial_{\mathbf{p}} f_c)^t\, \delta\mathbf{p} \quad (10)$$



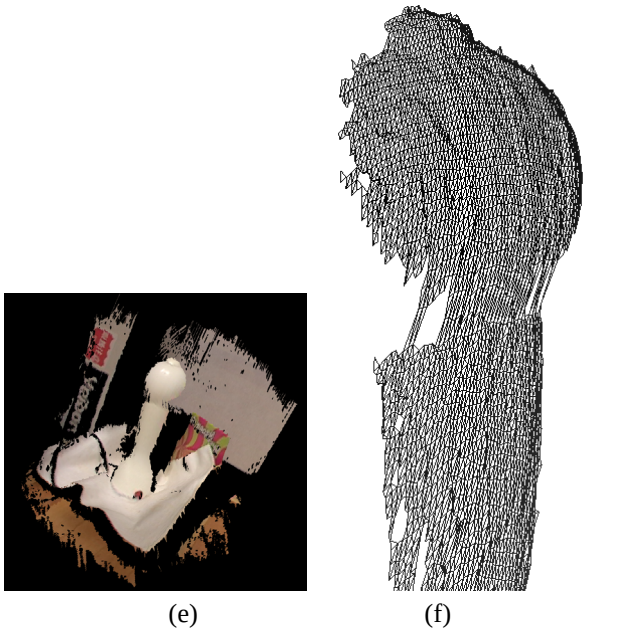


Figure 7. Results of shape acquisition from a light stand: (a) the measured object,(b)-(e) the acquired 3D model rendered from different views, and (f) the acquired polygons around the top of the light-stand.

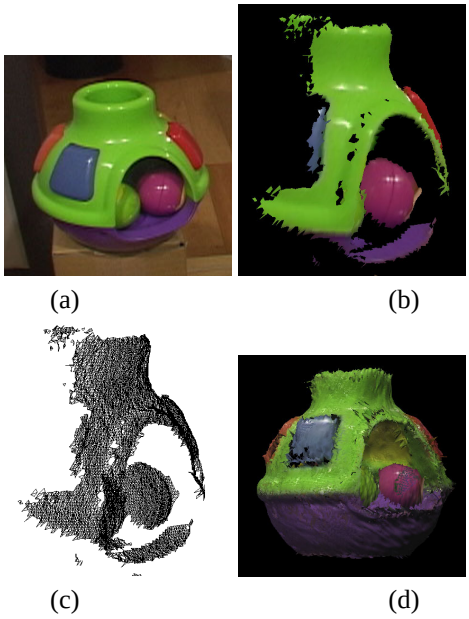


Figure 8. Results of shape acquisition from a plastic toy: (a) the measured object, (b) the acquired 3D model with texture, (c) the acquired polygons rendered as wireframes, (d) the registerd models with textures.

$$\delta \mathbf{y} \quad = \quad \partial_{\mathbf{p}}(\mathbf{g}_c^t)\,\delta \mathbf{p} = \begin{bmatrix} (\partial_{\mathbf{p}} g_{c,1})^t \delta \mathbf{p} \\ \vdots \\ (\partial_{\mathbf{p}} g_{c,n})^t \delta \mathbf{p} \end{bmatrix} \qquad (11)$$

.

To analyze the estimated measurement errors relative to the observation errors, we applied statistical approaches which are normally used in the field of computer vision[6],[8],[9], [15]. Assuming that the probability distributions of $\delta \mathbf{y}$ are isotropic and independent normal distributions with the same variances, the most likelihood estimation of $\mathbf{p}$ is obtained by minimizing $\|\mathbf{y} - \mathbf{g}_c(\mathbf{p})\|$, where $\|\cdot\|$ means a square norm. In this case, the ratio of the standard deviation of δx on measurement condition c and the common standard deviation of components of $\delta \mathbf{y}$ can be estimated by

$$\epsilon(c) = \|(\partial_{\mathbf{p}} f_c)^t [\{\partial_{\mathbf{p}}(\mathbf{g}_c^t)\}^t \{\partial_{\mathbf{p}}(\mathbf{g}_c^t)\}]^{-1} \{\partial_{\mathbf{p}}(\mathbf{g}_c^t)\}^t \|. \qquad (12)$$

Although only four of the six LEDs are used for estimating the initial pose of the wand, all the detected LEDs are used for the refinement process by non-linear optimization. To examine the effects of using redundant LEDs, we analyzed two different models, the one with 4 LEDs (4-LED) and the one with 5 LEDs (5-LED). The reason why we analyzed the model with 5 LEDs instead of the one with 6

LEDs is that, in real cases, one of the LEDs is usually hard to detect because of occlusion of the wand itself.

We calculated $\epsilon(c)$ for various conditions c in case of both 4-LED and 5-LED models, which are estimations of relative measurement errors. To validate the estimations, we also evaluated the relative errors by simulations. The simulations are conducted with assumptions of isotropic and independent normal distributions of $\delta\mathbf{y}$ and the errors of the objective value x are caused only by the observation errors of $\mathbf{y}$. For each condition c, we repeated the measurement process 100 times with samples of $\delta\mathbf{y}$ generated by the distributions, and the mean square of estimation errors of x is calculated, which we call $MeanSqrErrorRate(c)$.

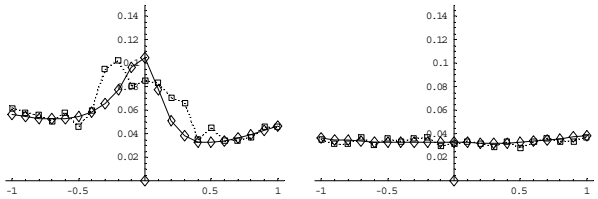
$\epsilon(c)$ and $MeanSqrErrorRate(c)$ plotted for various conditions c of 4-LED and 5-LED are shown in figures 9-10 with (a):4-LED and (B):5-LED. Markers represent the values of $\epsilon(c)$ (blank diamonds), $MeanSqrErrorRate(c)$ (blank squares). Here, a measurement condition c includes the rotation matrix $\mathbf{R}$, the translation vector $\mathbf{t}$, and the distance between the wand and the measured surface point. The rotation $\mathbf{R}$ is expressed by he parameters of $roll_c$, $pitch_c$ and yow_c. The translation $\mathbf{t}$ is expressed by (x_{c1}, x_{c2}, x_{c3}). The subscript c means that the value is a part of measurement condition c. The initial condition of the experiment was $(x_{c1}, x_{c2}, x_{c3}) = (0, 0, -0.8)$, $roll_c = pitch_c = yow_c = 0$ and the distance from the wand to the measured surface was 0.3. For the experiment of figure 9, x_{c1} was changed over the domain of $-1 \leq x_{c1} \leq 1$. For the experiment of figure 10, $pitch_c$ was changed over the domain of $-\frac{17\pi}{36} \leq pitch_c \leq \frac{17\pi}{36}$.

For all the results, values of $\epsilon(c)$ were good estimations of $MeanSqrErrorRate(c)$ for both 4-LED and 5-LED. Comparing the cases of 4-LED and 5-LED, the errors of 5-LED were less than half of those of 4-LED. We could also see that the values of $\epsilon(c)$ for the 4-LED model have a local peak at $x_{c1} = 0$ or $pitch_c = 0$, whereas values for 5-LED don't. The same tendencies were observed for the simulated errors. This implies the effectiveness of using an extra LED for observations.

We also conducted an experiment of real 3D measurement with various measurement conditions c. On measuring, we captured 240 frame images (8 sec with video camera) with the wand being wiggled at each condition. Then we estimated 3D location of the points lit by the laser for each frame and calculated the standard deviation and maximum error values for evaluation.

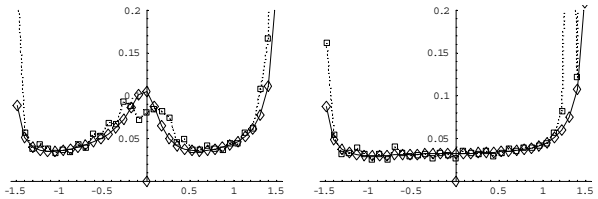
The initial position and pose of the wand was $x_{c1} = x_{c2} = 0$, $x_{c3} = -0.8$, and $roll_c = pitch_c = yow_c = 0$. About the rotational variation, we changed the yow_c of the pose of the wand from $-\pi$ to π with intervals of $\pi/12$. About the translational variations, we moved the wand from $x = -0.5$ to 0.2 with intervals of 0.1.

The experiments for both models of 4-LED and 5-LED



(a):4-LED (b):5-LED

Figure 9. Error estimations and simulated errors plotted for $-1 \leq x_{c1} \leq 1$; blank diamonds indicate $\epsilon(c)$; blank squares indicate $MeanSqrErrorRate(c)$.



(a):4-LED (b):5-LED

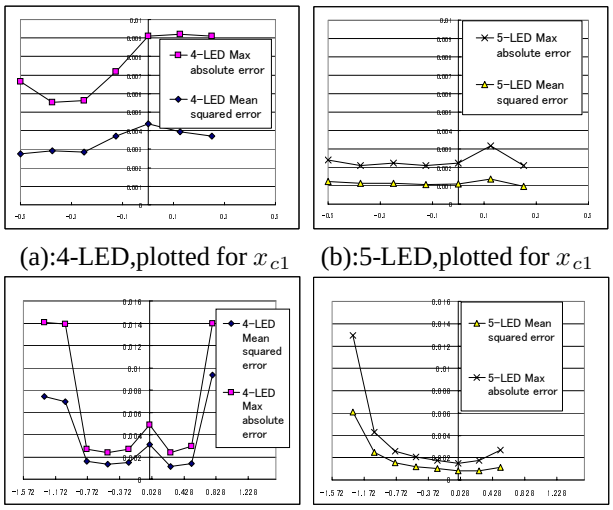
Figure 10. Error estimations and simulated errors plotted for $-\frac{17\pi}{36} \leq pitch_c \leq \frac{17\pi}{36}$.

are done based on the same observed data. One captured sequence of images is processed with both using four of the LEDs and using five.

Figure 11 shows the results of the experiments. There were some unmeasured data because of physical constraints of the wand, the measured surface and the camera (e.g. occlusion by the measured surface, or the LEDs getting out of the frame). We can see the similarities in tendencies between the values obtained from the analytical estimations shown in figures 9 and 10 and the values obtained from the actual measurement shown in figure 11. For example, the errors of the real measurements of 4-LED has a local peak at $x_{c1} = 0$ or $pitch_c = 0$, whereas 5-LED doesn't. Another similarity is that the errors of 5-LED model were less than half of those of 4-LED model.

5. Conclusion

In this paper, we proposed a low cost 3D shape acquisition system. The devices required for this system are a video camera, a laser projector, and LEDs. The LEDs are attached to the laser projector and are captured by video camera when scanning to estimate the pose and position of the laser projector. To scan objects efficiently, we use a laser projector which emits a fan-shaped laser plane rather than a



(a):4-LED,plotted for x_{c1} (b):5-LED,plotted for x_{c1}

(c):4-LED,plotted for $pitch_c$ (d):5-LED,plotted for $pitch_c$

Figure 11. Errors of real measurements, (a),(b): plotted for $-0.5 \leq x_c \leq 0.2$, (c): plotted for $-\frac{5\pi}{12} \leq pitch_c \leq \frac{3\pi}{12}$, and (d): plotted for $-\frac{5\pi}{5} \leq pitch_c \leq \frac{2\pi}{12}$.

laser beam. The advantages of using a laser plane projector are not only we can easily scan large area by sweeping the surface with the stripe of the laser reflection, but we can also reduce the scanning errors using information of known 3D points. In addition, we also proposed an efficient method to calibrate the relationship between the laser plane and the LED markers. By using acquired relationship, scanning errors are also successfully reduced.

For the experiments to evaluate the effectiveness of our 3D measurement system, we measured objects and created textured 3D models of them. We also conducted error analyses for the system, whose results implied the effectiveness of using additional LED images to refine pose estimation of the laser projector. Experimental results of real measurement showed the same tendencies to our error analyses, which implied the correctness of the analyses.

In the future, we plan to implement some post-processing algorithms to refine the captured data.

References

[1] http://www.cyra.com/.

[2] http://www.minoltausa.com/vivid/.

[3] J. Y. Bouguet and P. Perona. 3D photography on your desk. In *International Conference on Computer Vision*, pages 129–149, 1998.

[4] Chang Woo Chu, Sungjoo Hwang, and Soon Ki Jung. Calibration-free approach to 3D reconstruction using light stripe projections on a cube frame. In *Third International Conference on 3D Digital Imaging and Modeling*, pages 13–19, 2001.

[5] James Davis and Xing Chen. A laser range scanner designed for minimum calibration complexity. In *Third International Conference on 3D Digital Imaging and Modeling*, pages 91–98, 2001.

[6] Martin A. Fischler and Robert C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm ACM*, 24(6):381–395, 1981.

[7] R. B. Fisher, A. P. Ashbrook, C. Robertson, and N. Werghi. A low-cost range finder using a visually located, structured light source. In *Second International Conference on 3D Digital Imaging and Modeling*, pages 24–33, 1999.

[8] K. Kanatani. *Statistical Optimization for Geometric Computation: Theory and Practice*. Elsevier Science, Amsterdam, The Netherlands, 1996.

[9] Bogdan Matei and Peter Meer. Optimal rigid motion estimation and performance evaluation with bootstrap. In *Computer Vision and Pattern Recognition99*, volume 1, pages 339–345, 1999.

[10] Daisuke Miyazaki, Takeshi Ooishi, Taku Nishikawa, Ryusuke Sagawa, Ko Nishino, Takashi Tomomatsu, Yutaka Takase, and Katsushi Ikeuchi. The great buddha project: Modeling cultural heritage through observation. In *6th International Conference on Virtual Systems and Multimedia*, pages 138–145, 2001.

[11] M. Pollefeys, R. Koch, M. Vergauwen, and L. Van Gool. Hand-held acquisition of 3D models with a video camera. In *Second International Conference on 3D Digital Imaging and Modeling*, pages 14–23, 1999.

[12] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannnery. *Numerical Recipe in C: the art of scientific computing, 2nd edition*. Cambridge Univ. Press, 1994.

[13] Szymon Rusinkiewicz. Real-time 3D model acquisition. *ACM SIGGRAPH*, pages 438–446, 2002.

[14] Masahiro Takatsuka, Geoff A.W. West, Svetha Venkatesh, and Terry M. Caelli. Low-cost interactive active monocular range finder. In *Computer Vision and Pattern Recognition*, volume 1, pages 444–449, 1999.

[15] Tong Zhang and Carlo Tomasi. Fast, robust, and consistent camera motion estimation. In *Computer Vision and Pattern Recognition99*, volume 1, pages 164–170, 1999.