Entire Model Acquisition System using Handheld 3D Digitizer

Hiroshi Kawasaki
Saitama University, Saitama, Japan
kawasaki@mm.ics.saitama-u.ac.jp

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

Abstract

In this paper, a real-time, handheld 3D model acquisition system consisting of a laser projector, a video camera and a turntable is described. The user projects a stripe of light at the 3D object by hand while rotating the object on a turntable. The projected light and LED markers attached to the laser projector and turntable are captured by the video camera. By estimating the 3D orientation of the laser projector and the turntable angle from the 2D locations of the markers, the 3D location of the surface lit by the laser can be calculated. In addition, post-processing algorithms for refining the estimated 3D data have been proposed. The algorithm not only improves the accuracy of the 3D measurement, but also achieves to decrease the number of LEDs for 3D data estimation; therefore, it significantly improves the user's convenience in scanning the object. With this system, users can measure an entire 3D object in real-time.

1 Introduction

As CG technology becomes more common, a large amount of 3D model data is naturally required for various purposes such as filmmaking, digital archiving and so on. So far, 3D models are usually created manually using 3D software, making their acquisition a laborious and time-consuming task. As systems that would reduce the cost and time for 3D model creation, 3D digitizers have attracted a great deal of interest.

While demands for 3D digitizers are rapidly multiplying, commonly used 3D scanners (i.e., laser range scanners) are usually large, heavy and expensive[3, 9]. Certainly, there exist some simple and inexpensive 3D acquisition systems using vision-based techniques such as stereo vision. However, these techniques are still being researched and are difficult to apply for practical uses. At the present time, portable and inexpensive 3D scanners are the most viable solution to many usersf needs.

Ordinary 3D scanners can only retrieve a 3D depth image from a single viewpoint. Thus far, several methods have been proposed to acquire the entire shape of the object. One straightforward solution is to align and integrate multiple geometries by software. Another solution is the use of the turntable. However, both techniques have limitations. In the former case, the algorithms for alignment still often require manual intervention and those for merging sometimes yield unstable results. In the latter case, estimating a precise rotation axis and angle is not easy without the added cost of mechanical sensors.

Based on these facts, we propose a real-time shape acquisition method which can capture the entire shape of the object interactively using the handheld 3D digitizer described in [5]. To achieve the entire-shape acquisition, we adopt a turntable in our system and propose an original method to accurately estimate the turntable's rotation angle using a vision-based technique. The proposed digitizer is of handheld size; this allows users to take the scanner anywhere, including outdoors.

Also described in this paper is a post-processing algorithm for reducing errors of 3D reconstruction. The error reduction is based on consistencies of 3D points whose locations are estimated multiple times. The parameters for 3D estimation are iteratively corrected to achieve better consistencies. Also, the error correction algorithm is useful for calibrating parameters of the 3D digitizer.

This paper is organized as follows. In Section 2, we explain the basic idea of 3D model acquisition with regard to related works, and in Section 3, describe our system configuration and its characteristics. Section 4 presents the post-processing algorithm, while Section 5 presents the results of experiments using our 3D digitizer. Section 6 contains our conclusions regarding this method.

2 Related Works

2.1 Low-cost 3D Scanning Device

So far, a great deal of research has been conducted on 3D model acquisition. Our purpose was to make a portable and inexpensive 3D scanner that could retrieve the entire model of the object in real-time. Therefore, we mainly survey simple and low-cost 3D measurement systems. In order to be simple and low-cost, most of these avoid mechanization. Rather, these systems need to estimate the positions of sensors in order to employ triangulation in acquiring 3D information. So far, various types of positioning methods have been proposed. Bouguet and Perona[1] used shadows to calibrate the camera and light source positions. They cast shadows on the object, and estimated the 3D values using these calibrated parameters. Fisher et.al.[4] modified this idea. Woo and Jung[2] proposed another solution. They put a cubic frame in the field of view of the camera, and placed the object inside this frame. Then, they emitted a line beam to the object, and detected the bright point on the frame to estimate the position of the beam plane and the 3D information. Takatsuka et.al.[8, 6] adopted a more active method. They placed LED markers on the sensor itself, and captured it with a single camera to estimate the sensor position. They used a laser pointer as a sensor, and proposed a simple 3D estimation method using the detected LED markers. We extended Takatsuka's method, using a line laser projector instead of a beam of light[5].

2.2 Entire Shape Acquisition System

Usually, a 3D modeling of the entire shape of an object is accomplished by performing the following three steps:

1. Acquiring range images (Scanning).
2. Aligning the range images acquired from different viewpoints (Aligning).
3. Integrating the aligned range images into a mesh (Merging).

Since it is impossible to detect holes before alignment, the first two steps of the 3D model acquisition process are usually performed iteratively to retrieve the entire model without holes. If, therefore, the 3D scan and alignment stage could be done in real-time, it would significantly reduce the total time and cost of acquiring an entire 3D model. Rusinkiewicz et.al. [7] proposed a real-time 3D model acquisition system based on such an idea. However, their proposed system adopts a structured light-based 3D acquisition technique at the scanning stage; therefore, system configuration is complex and it would be limited to a fixed facility.

Our system is based on the same concepts, and the system works in real-time. However, unlike the Rusinkiewiczs' system, we propose a handheld scanner that is simple and lightweight enough to carry anywhere the user likes.

Figure 1: A Handheld 3D scanning device.

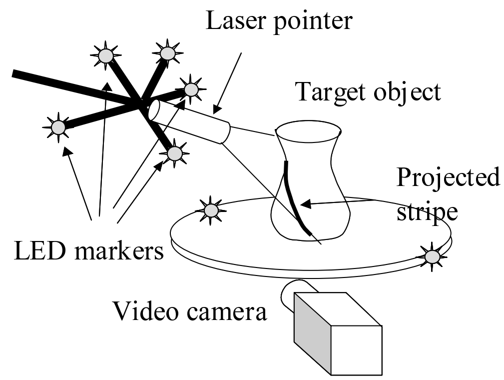
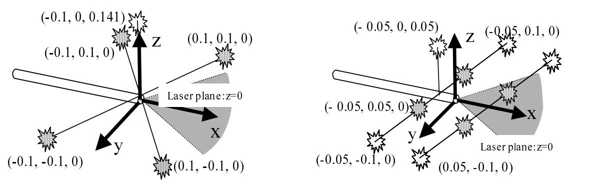


Figure 2: An entire 3D shape measurement system.

3 Entire Shape Acquisition System using Turntable

3.1 System Configuration

Our shape measurement system consists of a video camera, a PC, the 3D digitizer and a turntable (Fig.1 and 2). While measuring, the user holds the 3D digitizer by hand and projects the laser onto the measured object. At the same time, the user rotates the object using the turntable. The sheet of laser light produces a stripe on the object, and the user moves the 3D digitizer and turntable so that the stripe scans the entire object surface. The projected stripe and the LED lights are observed by the video camera. The video sequence is transmitted from the video camera to a PC via a USB cable or IEEE1394 link. The shape of the target object is recovered by analyzing the video sequence. The camera is fixed and the object lies on the rotating turntable to expose its entire shape to the camera.



(a)Original configuration (b)New configuration

Figure 3: A 3D digitizer.

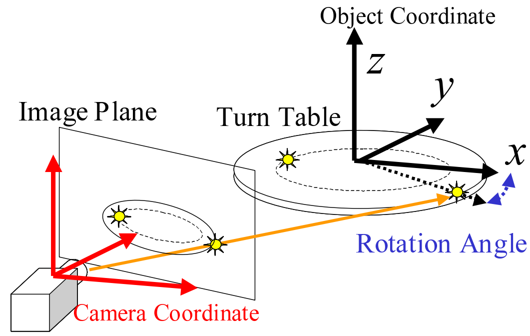


Figure 4: A 3D measurement equipment.

3.2 Real-Time Handheld 3D Digitizer

3.2.1 3D Scanning Device and Shape Recovery Algorithm

We briefly explain our 3D scanning device described in [5]. 3D shapes are recovered by a triangulation method. Since the 3D digitizer and the video camera form the baseline for the triangulation, the position and angle of the 3D digitizer are important to determine. We avoided mechanical devices or sensors in order to make the system simple and handy, instead adopting computer vision techniques (LEDs) to estimate these parameters. To make the estimation accurate, robust and efficient, LEDs attached to the 3D digitizer as shown in Fig.3-(a) are used. These LEDs define the local coordinates in 3D space. From the 2D locations of the LED lights in the video frame, the position and the angle of the 3D digitizer can be estimated. Using these parameters, we can acquire 3D shapes by simply applying a triangulation method. Details are described in [5].

3.2.2 Additional LEDs

Theoretically only 3 LEDs are required to estimate a laser plane. However, if there are 4 LEDs configured as a square shape, efficient initial scanner orientation estimations can be achieved[5]. Inversely, the 4 LEDs must be captured in the same frame as the object for efficient 3D scanning; this makes the actual scanning process difficult because the size of the scanner must be large in order to achieve these accuracies.

In this paper, we extend the 3D digitizer to solve this problem. For initial position estimation, we configured 9 LEDs, and only used 4 of them, which are positioned to form a small square (Fig.3-(b)). The other 5 LEDs are used to improve the accuracy as redundant information and are not necessarily captured in the same frame as the object. Thanks to this algorithm, we can scan the object more freely than before.

3.3 Turntable with LEDs

The original 3D digitizer can only capture the 3D data from a single direction. Therefore, to capture the entire model of the object in real-time, we propose a simple solution that utilizes a turntable with LEDs.

To scan the object, we place the object on the turntable and scan it freely by rotating the turntable and projecting a line laser onto the object. In each frame, we estimate the rotation angle using the LEDs attached to the turntable, and then directly transform the original estimated 3D data in the camera coordinate system to the object's coordinate system. Accumulating these transformed 3D points in the object coordinates enables us to retrieve an entire model of the object.

Estimation of the rotation angle of the turntable can be achieved by vision-based techniques; therefore, we do not need extra special devices for the system. Thus, the system preserves the original advantages of the proposed digitizer, such as the requirement for only a single camera, the handheld size of the digitizer, and real-time processing.

3.3.1 Estimation of Rotation Angle

To transform 3D data from the camera coordinate system to the object coordinate system, our system requires the rotation axis and the rotation angle of the turntable. To estimate the rotation angle, a precise positioning of the rotation table must be performed. We adopt a simple method for this purpose. We put a positioning marker on the rotation table and capture it by gradually rotating it. Then we apply optimization methods to calculate both external parameters and rotation axis simultaneously; this calculation usually converges stably. By using these parameters, we can calculate the transformation matrix **A** which converts the camera coordinate system to the object coordinate system. We define the object coordinate system as Fig.4, and make [the] z axis of the object coordinate system [] coincide with the rotation axis. Using the matrix **A**, we can estimate the rotation angle as follows:

1. Convert the camera center coordinate **o** from the camera coordinate system to the object coordinate system as **Ao**.
2. Convert LED coordinate **u** from the camera coordinate system to the object coordinate system as **Au**. (Here, $\mathbf{u} = (u, v, f)$ where LED position on image plane is (u, v))
3. Emit lines from the camera center to the LED positions on the image plane and calculate the cross section of the line and the turntable($z = 0$ at the object coordinate system). The value can be calculated as

$$\mathbf{p} - (\mathbf{q} - \mathbf{p}) \cdot \frac{\mathbf{e_z} \cdot \mathbf{p}}{\mathbf{e_z} \cdot (\mathbf{q} - \mathbf{p})} \quad (1)$$

where $\mathbf{p} = \mathbf{Au}$, $\mathbf{q} = \mathbf{Ao} and \mathbf{e_z} = (\mathbf{0}, \mathbf{0}, \mathbf{1})$.

4. The rotation angle can be calculated as the angle from the x axis of the object coordinate system.
5. Calculate the radius as the distance from the origin to the LED position to distinguish individual LEDs.

Since the LEDs are sometimes occluded by the object, we put several LEDs on the turntable, each with a different radius. Therefore, at any time, we can detect at least one LED attached to the turntable and identify it by its radius.

3.3.2 Voxel Representation

With our proposed method, 3D point data are non-uniformly estimated in 3D space; therefore, we adopted a voxel-based representation for our data structure. Once 3D points were estimated by our 3D digitizer and converted to the object coordinate system by using rotation angle(Fig.5), they were quantized to the voxel grid. They then voted for the voxel(Fig.6).

After the completion of the scanning, we produced a 3D point at the center of any voxel that has been voted more than the threshold count. Our target object is not usually larger than 20cm and our target precision is less than 1mm; therefore, we set the voxel resolution as 256x256x256 for a 20cm cube.

4 Post-Process to Increase Accuracy

We implement an offline geometry refinement algorithm as post-process to improve the accuracy of the 3D data estimated by our 3D digitizer. There are several reasons for bad 3D estimation, a significant one of which is the mis-estimation of the laser planes.

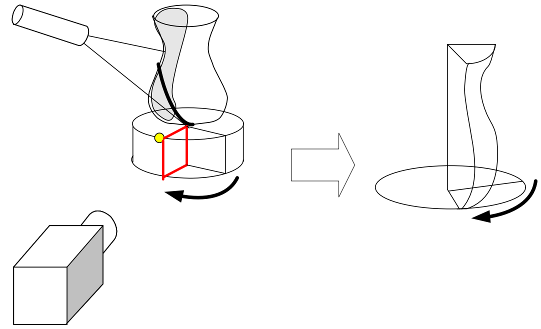


Figure 5: Rotation of the acquired 3D data.

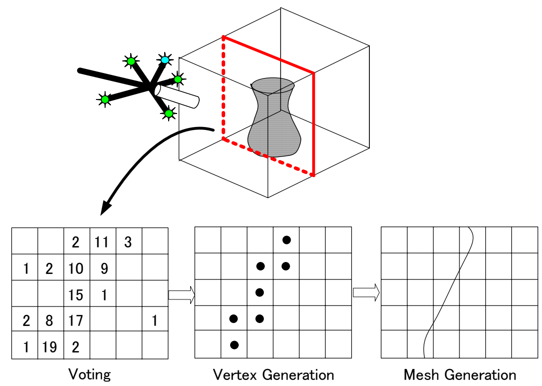


Figure 6: Voting system for voxel representation.

With regard to the mis-estimation of the laser planes, two main reasons can be considered.

1. Image processing limitation.
2. Differences between the laser planes and the LED marker planes.

In the former case, for example, when a part of the LED areas cannot be observed due to occlusion or pixel noise, incorrect laser plane parameters are estimated. In the latter case, manufacturing limitations cause the mis-alignment between the laser plane and LED marker planes.

The latter case is sometimes more critical than the former case, because all of the 3D estimations are then affected by the errors of the alignment of the laser plane. In particular, when the plane of the laser sheet is nearly parallel to the view direction, error values drastically increase.

Therefore, in this paper, we try to improve the accuracy of the 3D data by improving the accuracy of the laser plane by our post-process method. To solve the former case, we use a multiple linear regression analysis (Sec.4.1) to solve the latter case, we propose a g*laser plane calibration*h described in Sec.4.2.

4.1 Multiple Linear Regression Analysis

Assume that F is the set of all frames and f is an arbitrary frame of F($f \in F$). Then we define the estimated plane parameters for frame f as $p(f)$, and the set of pixels that are detected as the laser area in the frame f as $R(f)$. The parameters $p(f)$ include information on the position of the 3D digitizer, the 3D digitizer coordinate system described later, and the position of the laser plane relative to the 3D digitizer coordinate system. D is the set of depth estimations for all the pixels. Each element of D is a tuple (x,y,d), where (x,y) is the pixel location and d is the depth estimation at the pixel. $D(R(f))$ means a set of extracted elements of D whose pixel locations are limited to the area $R(f)$.

Since the described algorithm is a post-process, the depth values in D are calculated using observations of all the frames. Therefore, the depth values from D tend to be more accurate than the depth estimations obtained from $p(f)$, which is obtained from the observation of a single frame f. In this paper, we correct $p(f)$ so that $p(f)$ becomes more consistent with D

To implement this idea, we can apply a multiple linear regression analysis to fit a plane to the points in $D(R(f))$, since elements of $D(R(f))$ are points on a laser plane.

The 3D digitizer coordinate system is defined as the following:

3D digitizer coordinate system The 3D Digitizer coordinate system consists of an x-axis, a y-axis, and a z-axis whose origin is at the center of the LEDs mounted on the 3D digitizer. The x-y plane coincides with the LED plane and the x-axis direction is the same as the laser projecting direction. Thus, if the 3D digitizer is created without misalignment, the laser plane is $z = 0$ (see also Fig.7).

The actual process is as follows. Multiple linear regression analysis is applied to the estimated points in $D(R(f))$ expressed in the 3D digitizer coordinate system. In the 3D digitizer coordinate system, if the orientation estimation is correct and there is no mis-alignment of the LEDs, the z value of all points will be 0, and all regression coefficients will also be calculated as 0. Fitting a plane to the points in $D(R(f))$ using the multiple linear regression analysis, we can evaluate the error of the plane estimation. If the fitted plane deviates too much from the initial estimation of the laser plane for a certain frame, it is likely that the position and angle estimations of the device include large errors. In that case, the data of the frame is removed from the post-process iteration so as not to affect the final 3D depth estimation.

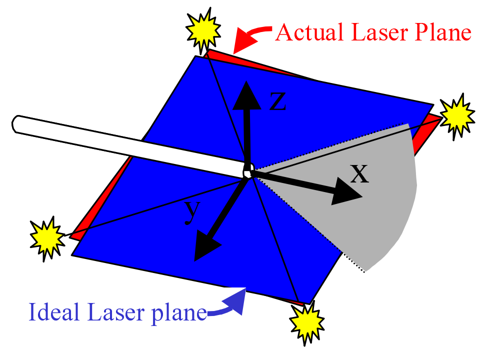


Figure 7: Error of the 3D digitizer.

We consider that an initial orientation estimation $\acute{p}(f)$ is also important for an estimation of $p(f)$; therefore, we add extra points which are on the $\acute{p}(f)$ plane to $D(R(f))$ to reflect it. The number of the extra points is calculated as follows:

1. Calculate the weight function w_p, which is 1.0 when 6LEDs are observed and linearly decrease to 0.0 (only 4 LEDs are observed).
2. Calculate the number of the extra points as floor($w_p \times 4$) (Maximum number is 4).

By conducting the process iteratively, precise estimations of both the 3D digitizer's orientation and the 3D shape can be simultaneously achieved.

4.2 Laser Plane Calibration

The deviation of the actual laser plane of the 3D scanner and the ideal laser plane can be acquired as the mean distance between the laser plane estimated from the LEDs and the laser plane obtained from the laser reflections; we call this process a g*laser plane calibration*h. Once the laser plane calibration parameters are obtained, they can be used in initial 3D estimations for measurements performed later. Since this deviation affects all the 3D estimations, accuracies of all the 3D scanning processes are expected to improve.

The actual laser plane calibration parameters can be calculated as the mean deviation of the laser planes estimated from the laser reflections and the plane $z = 0$ when the post-process has converged. Note that there are two main reasons that can be considered for the deviation as we have already described in this section; the deviations vary from frame to frame. Therefore, we adopt median values of the laser plane coefficients of all frames as the laser plane calibration results.

4.3 Algorithm

The whole process of the post-process for the refinement of the laser planes is as follows:

1. Calculate the depth of each frame f using estimated plane parameters (only the LEDs are used for the estimation at this stage) and integrate them using Bayes' theorem to create the initial depth image D.
2. Conduct the following processing to each frame f.
   (a) Extract elements of D at area $R(f)$, which is $D(R(f))$, and convert the tuples into 3D points C expressed in the camera coordinate system.
   (b) Transform points from C (the camera coordinate system) into the 3D digitizer coordinate system using parameters $p(f)$. The set of the resulting points is depicted as C'.
   (c) Add extra points to C', based on $\acute{p}(f)$ as described in Section 4.1
   (d) Apply multiple linear regression analysis to C' to fit the points into a plane. The model formula of the analysis is $z = ax + by + c$. The laser plane parameters in $p(f)$ are replaced by the new plane.
   (e) Calculate the residual $e(f)$ between 3-Dimensional points in C' and the new plane of $p(f)$.
3. Re-calculate the depth using updated plane parameters in $p(f)$ for each frame f and finally update D. Also, calculate E as $\sum_{f \in F} e(f)^2$
4. Repeat **2** and **3** until total residual E converges to minimum value.

5 Experiments and Evaluation

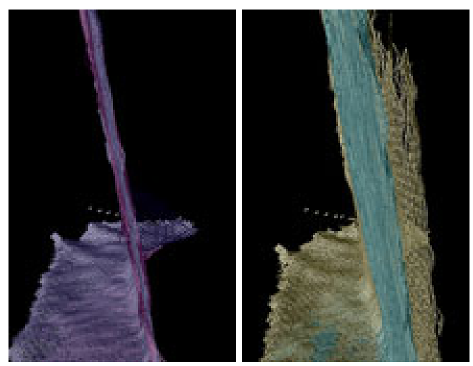
To demonstrate the effectiveness of our proposed method, we conducted several experiments. The first experiment was for the evaluation of the post-processing algorithm of the plane refinement. Then we conducted another experiment to evaluate the post-processing algorithm of the g*laser plane calibration*h. Here, we show several entire 3D models acquired by our 3D digitizer with turntable for demonstration followed by an evaluation.

5.1 Evaluation of Laser Plane Refinement

We first scanned a box for evaluation. The size of the box was 10cm×10cm×20cm and all planes were flat. To conduct the experiment under the same conditions, we captured the whole scanning scene by video camera and used the same scene as the input. The results are shown in Fig.8. Fig.8-(a) is the result with our proposed post processing algorithm and Fig.8-(b) and (c) are the results without post-processing. When we compare Fig.8-(a) and (b), we can see that planes of the box of Fig.8-(a) are smoother than those of (b). In addition, our proposed method preserves the sharp edges of the object while a simple smoothing algorithm such as median filter rounds off the edges as shown in Fig.8-(c). We can also see the improvements of our method in Table 1.

5.2 Result of Laser Plane Calibration

Further, to evaluate the effectiveness of the g*laser plane calibration*h, scanning results of the same object projecting the laser from the left and right side with and without the laser plane calibration are shown in Fig.9-(a) and (b). In Fig.9-(a), we can see that the two measured planes are almost the same; however, in Fig.9-(b), the two planes are apart.

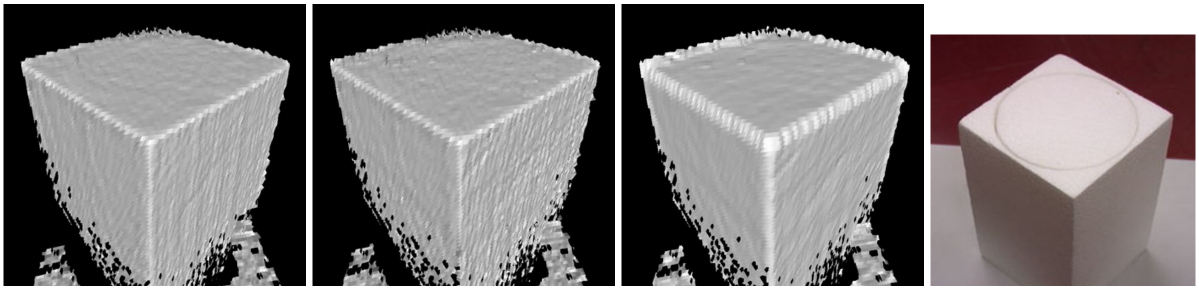


(a)With calibration (b)Without calibration

Figure 9: Laser calibration result.

5.3 Results of Entire Model Acquisition

To evaluate the accuracy of the estimated rotation angle and the voxel voting method, we scanned a simple-shaped object with several certain rotation angles. Table 2 and Fig.10 and 11 show the results. We can see that all scanned objects



(a) Proposed Method (b) Original Method (c) (b) + Median Filter. (d) Actual box

Figure 8: Result of scanning calibration box.

Table 1: Evaluation results

	Proposed Method	Original Method	Original Method + Median Filter.
Standard deviation (mm)	0.36	0.49	0.65
Angle difference (deg.)	0.61	0.89	093

are correctly converted to the object coordinate and the entire model is successfully created. Table 2 also tells us that angle differences between 90 degrees and 134 degrees are larger then those of other degrees; we consider that quantization errors of LED caused this tendency.

Table 2: evaluation of rotation angle

Data index	estimated rotation angle	difference
00	90.45	-0.84
01	133.94	+0.80
02	177.96	-0.39
03	225.40	+0.39
04	314.39	-0.29
05	4.68	+0.17
06	52.65	+0.55

Finally, several objects were scanned by our 3D scanner as shown in Fig.12 and 13. We can see that entire shapes of the objects are successfully captured by our 3D scanner.

6 Conclusion

In this paper, we have proposed a real-time handheld 3D digitizer that can scan the entire shape of an object. The devices required for this system are a video camera, a laser projector, a turntable and LEDs. The LEDs are attached to the laser projector and are captured by video camera when scanning to estimate the angle and position of the laser projector. To scan the entire shape of an object efficiently, we used the LED attached to a turntable. The advantages of using the LED-attached turntable are not only that we can

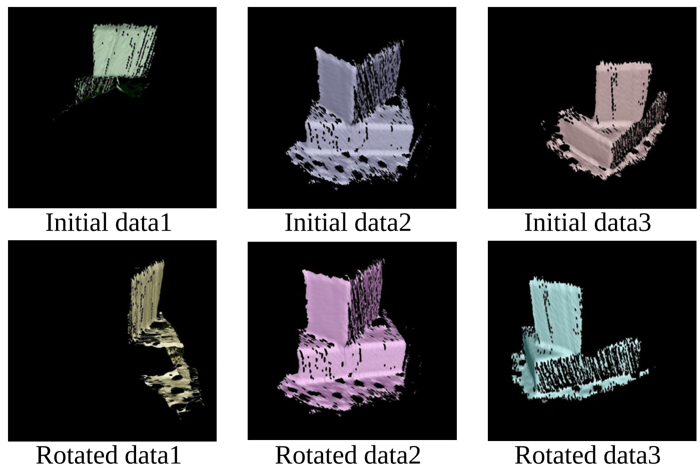


Figure 10: Scanned 3D data

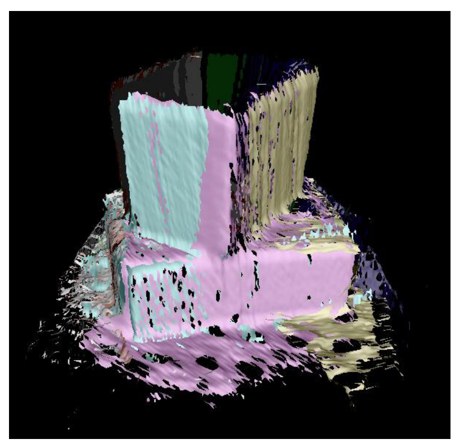


Figure 11: Integrated 3D data.

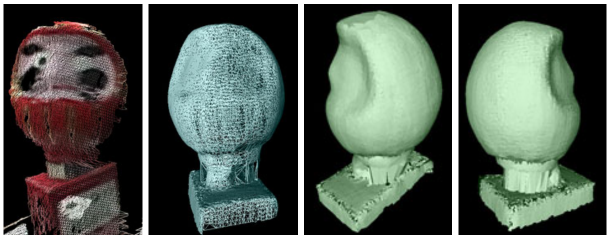


Figure 12: Wooden toy.

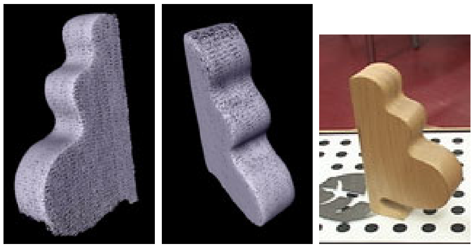


Figure 13: Bookshelf

easily scan the entire body of the object, but also that it preserves the original advantages of the 3D digitizer, such as its handheld size and its capability of working interactively in real-time.

We have also proposed a post-processing technique that can significantly improve the accuracy of the acquired 3D data. The post-process mainly improves the mis-estimation of the laser plane by multiple linear regression analysis.

To evaluate the effectiveness of our 3D measurement system, we measured objects and created 3D models of them. We also conducted several experiments to test the effectiveness of our proposed methods. Experimental results of real measurement showed the apparent strength of our proposed method.

References

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

[2] C. W. Chu, S. Hwang, and S. K. 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.

[3] Cyrax. Lica inc. http://www.cyra.com/.

[4] 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.

[5] R. Furukawa and H. Kawasaki. Interactive shape acquisition using marker attached laser projector. In *International Conference on 3D Digital Imaging and Modeling2003*, pages 491–498, 2003.

[6] T. Masahiro, G. A. West, S. Venkatesh, and T. M. Caelli. Low-cost interactive monocular range finder. *Machine Vision and Applications*, 14(3):139–144, 7 2003.

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

[8] M. Takatsuka, G. A. West, S. Venkatesh, and T. M. Caelli. Low-cost interactive active monocular range finder. In *CVPR99*, volume 1, pages 444–449, 1999.

[9] Vivid. Minolta inc. http://www.minoltausa.com/vivid/.