3D Profile Measurement by Color Pattern Projection and System Calibration

C. Sinlapeecheewa and K. Takamasu
Department of Precision Engineering, The University of Tokyo
7-3-1 Bunkyo-ku, Hongo, Tokyo, JAPAN
Phone/Fax: +81-3-5841-6472
E-mail: chanin@nano.pe.u-tokyo.ac.jp, takamasu@pe.u-tokyo.ac.jp

Abstract
In this paper, we present a new color structured lighting for 3D profile measurement by projecting a pattern of color stripes. The advantage of using color pattern is that it simplifies the difficult matching problem of using a multiple-line stripe pattern. The problem of finding the correct color stripe correspondence between light source and images is solved by accurately calibrating the system parameters. A technique for camera-projector calibration using calibration points that projected from projector is presented. The mean error of this calibration method is about 0.2 mm. A pattern of color stripes is projected onto the objects when taking images with cameras from various viewpoints, stripe pattern information are extracted from the acquired images and then used for finding the correct projected-observed stripe match. After matched stripes information was obtained, 3D profile is reconstructed by means of triangulation. The selection of color used to generate the color-stripe pattern is presented. An experiment using human hand as a model is also demonstrated.

Keywords: 3D profile measurement, Structured light, Color pattern projection, Projector-cameras calibration

1. Introduction

To reconstruct accurate shape from image is a challenging problem in computer vision. One of the most popular approaches in structured lighting is to use a single-line stripe [1-5]. The advantage of this approach is that it greatly simplifies the matching problem, whereas the drawback of this approach is that only one single stripe of 3D data points can be obtained with each image shot. To speed up the data acquisition process, multiple-line stripe patterns are used instead. Osawa et al. [3] used pattern of four gray levels with space-time scanning method. Various gray code patterns are projected onto the object in time series when taking images with a camera. The advantage of this approach is that it speeds up the acquisition process and still simplifies the matching problem. However, it still needs several image shots to generate time series code of lighting. In order to reduce the number of input image, color-stripe pattern is generated, and several cameras are used to take images from various viewpoints. Chen et al. [4] combined color-structured light and stereo vision by using an un-calibrated structured light source to project a pattern of color stripes. Two cameras were used to observe projected pattern, then the correct stereo correspondence of both cameras based on intra-scanline dynamic programming were carried out. Zhang et al. [5] used only one camera for capturing an image of projected stripe pattern, and then matched the observed edges in the image. The correspondence problem is solved using multi-pass dynamic programming algorithm.

In this paper, we developed a new technique using complete calibrated system with color structured light projection to measure 3D-object. This technique focuses on trying to reduce the number of input image in order to speed up the acquisition process. The correspondence problem is solved using a simple matching algorithm to match the projected pattern with two observed stripe patterns taken from two cameras. All of the system parameters, projector and cameras parameters, are completely calibrated. The camera model described in section 3 is based on directly using all the optical and geometric parameters of the camera [6-9]. The configuration of DLP-projector is similar to CCD-camera in the point that the former consists of optical lens and digital micromirror whereas the latter, the optical lens and image sensor. The projector model is then applied from the camera model. An ordinary calibration procedure is implemented base on the triangulation and parameters are estimated using least square method. In our experiments, the color multi-stripe pattern is projected onto the object when CCD cameras capture the reflected light. This method needs only one image taken from each camera. As a result, it is possible to measure moving object such as human body. However, the calibration of system parameters becomes very important and must be very precise, since many cameras and projector are used to match the points in 3D space.

The approach consists of five steps: 1) establishing the projector and camera models, a novel projector-cameras calibration procedure to support triangulation computations, 2) structured light projection and 2D image acquisition, 3) image processing for extracting 2D stripe information; pixel coordinate, color and brightness, 4) solving the matching problem among projected stripe pattern and observed stripe pattern, 5) 3D profile solution via triangulation. This paper is organized as follows. Section 2 presents an arrangement of the proposed system. In section 3, mathematical model and system calibration procedure are described. Section 4 gives an explanation of 3D profile measurement techniques and the use of color.

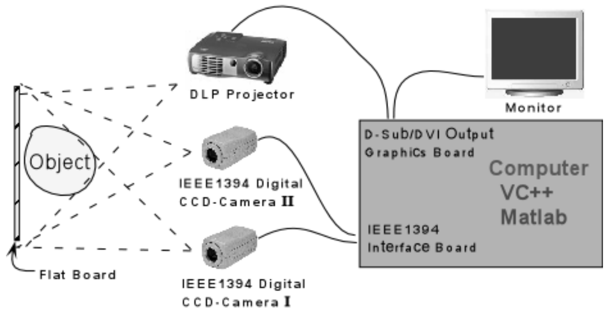

Fig. 1. System Arrangement

2. Color-Stripe Pattern Projection Method

2.1 System Arrangement

The arrangement of the 3D profile measurement using color-pattern projection method is shown in Fig. 1. Custom hardware solutions have been developed to do the rapid shape measurement. The system consists of a projector, two CCD-cameras, a computer and white flat panel. The projector is a *PLUS* V-1080, Digital Light Projector (DLP) designed for general presentation. It is connected to the computer in parallel to the monitor via a graphics adaptor card. The 1024*768 pixels spatial resolution of presentation-type projector is a convenient source of structured light. However, owing to its large aperture, limitation in this unusual application is the restricted depth of field and geometric lens aberrations. The cameras are Sony DFW-X700, provide a resolution of 1024*768 pixels. The interface to the computer is done with a high data transfer rate of IEEE1394 interface card, *Zenkuman* PFW-41. The 866 MHz Pentium PC generates the projection patterns, acquires the images, and carries out the 3D reconstruction procedure. The white flat panel's surface should be very smooth and well light reflected. We used a thick plate of glass covered with a thin white sticker sheet.

2.2 How the Method Work

The color stripes pattern generated by computer is projected onto the object. Stripe lights projected on the object surface are distorted by the shape of object. Two of CCD-cameras, placed at different viewpoints, capture the image of distorted color stripes. The images are saved in the computer's harddisk for further image processing. According to the triangulation, we can extract the position of point on the object if position of projection image and its correspondence stripe images, captured by both cameras, are known. We then developed a procedure for matching up projection stripes and distorted stripes. The inputs of this procedure are images taken from each camera and system parameters. Since we were dealing with points in 3D space, it is important to accurately calibrate the parameters. White flat panel, fixed to the coordinate measuring machine (CMM), is a measured object for system parameters calibration. System parameters can be categorized into external parameter; position and orientation of device and internal parameter; focal length and distortion coefficients of lens [6-9].

3. Mathematical Model and System Calibrations

In order to deduce the object's position and orientation from an image, we need details of the projector and camera's position and orientation in space relative to some reference coordinate system, called the *world coordinate system*. Furthermore, we must also have some geometrical model of the projector and the cameras arrangement. Also some processes of finding the various parameters presented in the model were needed. The latter process is called *projector-camera calibration*.

3.1 Camera model

A camera model is defined as the mathematical relationship between the 3D coordinates of a point in the scene space and its corresponding coordinates on the image plane. The adopted camera model [6-9] is shown in Fig. 3. Since we are interested in attributing to each parameter a precise physical meaning, the relation should be defined by directly using all the internal and external parameters of the camera.

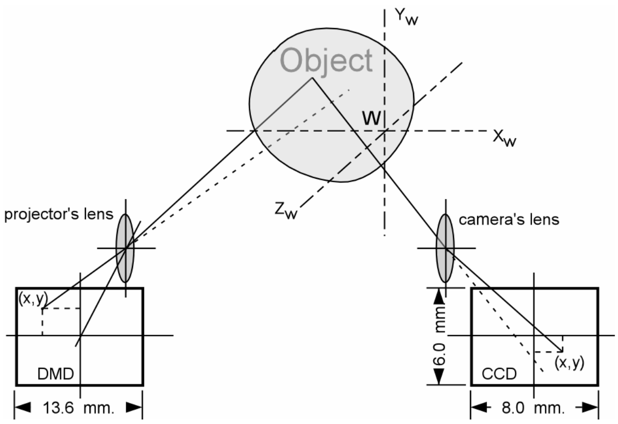

Fig. 2. Camera and projector coordinates

Three reference frames are defined in the camera model shown in Fig. 2;

- *World reference frame:* rigidly attached to the scene; used for specifying the world coordinates of any point $[x_w\ y_w\ z_w]$ of the 3D scene;
- *Camera reference frame:* rigidly attached to the camera; the optical axis is coincided of the z_c axis, while x_c and y_c are respectively parallel to the horizontal and vertical axes of the image plane (which is assumed to be orthogonal to the optical axis). The origin is an optical point of the lens. The camera coordinates are specified as $[x_c\ y_c\ z_c]^T$ The intersection between optical axis and image plane is called principal point;
- *Image reference frame:* defined on the image plane; the origin is the principal point and defined to be at

the center of the image sensor, the image coordinates $[n_{cx}\ n_{cy}]^T$ are expressed in pixels.

The relationships between the world coordinate and image coordinates are described in the following steps:

(1) *Conversion from image coordinates to camera coordinates*

$$x_d = (n_{cx} - \frac{H_p}{2}) \cdot \frac{L}{H_p} \tag{1}$$

$$y_d = (\frac{V_p}{2} - n_{cy}) \cdot \frac{W}{V_p} \tag{2}$$

where *Hp* and *Vp* are horizontal and vertical number of pixels of the image sensor (CCD), while *L* and *W* are the horizontal and vertical size of the image sensor, respectively.

(2) *Lens distortion*

In order to obtain more accurate model, both radial and tangential lens distortion components should be considered. With radial distortion, image coordinates are radial shifted from the principal point, while the tangential distortion accounts for the component that is perpendicular to the radial direction. We only consider radial distortion since the tangential distortion is often negligible with respect to the radial one. The radial distortion is modeled by the power series that expresses the undistorted image coordinate $[x_u\ y_u]$ as a function of the distorted ones $[x_d\ y_d]$.

$$\begin{aligned} x_u &= x_d \cdot (1 + k_3 r_d^2 + k_5 r_d^4 + \cdots) \\ y_u &= y_d \cdot (1 + k_3 r_d^2 + k_5 r_d^4 + \cdots) \\ r_d^2 &= x_d^2 + y_d^2, \end{aligned} \tag{3}$$

where r_d is the distance from principal point; center of CCD. The first two terms of the series (k_3, k_5) are usually sufficient for an accurate parameterization of the radial distortion.

(3) *Rotating the camera coordinates to be paralleled to the world coordinates*

$$\begin{bmatrix} x_u^r \\ y_u^r \\ z_u^r \end{bmatrix} = R * \begin{bmatrix} x_u \\ y_u \\ f \end{bmatrix}, \tag{4}$$

$$R = \begin{bmatrix} \cos(c) & -\sin(c) & 0 \\ \sin(c) & \cos(c) & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(a) & -\sin(a) \\ 0 & \sin(a) & \cos(a) \end{bmatrix} * \begin{bmatrix} \cos(b) & 0 & \sin(b) \\ 0 & 1 & 0 \\ -\sin(b) & 0 & \cos(b) \end{bmatrix},$$

where *R* is a rotation matrix which specifies the rigid displacement between world reference frame and camera frame, while f is the focal length of the optical lens.

(4) *Perspective projection from the point on the image plane*

Perspective projection vector is the vector from the optical center of the lens to the distorted image position on the image plane.

$$\bar{C}_P = \begin{bmatrix} x_u^r \\ y_u^r \\ z_u^r \end{bmatrix} - \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \tag{5}$$

The parameters for defining the camera properties of the *i*-th camera specified by the following nine elements vector;

$$C_i = [a_i \quad b_i \quad c_i \quad x_{ci} \quad y_{ci} \quad z_{ci} \quad k_{3ci} \quad k_{5ci} \quad f_{ci}]^T, \tag{6}$$

Typically, in the perspective projection, the ratio of actual size of image sensor and focal length could be defined as one specific value. Since the size of an image sensor (*L* and *W*) is assumed a priori known from the camera specification, only the focal length was included in the set of camera parameters.

3.2 Projector Model

Projector model is similar to the previous camera model. Projector projects the image by reflecting light from the Digital Micromirror Device (DMD) that passed through the lens, whereas that camera captures the light passed through the lens and fell down onto the image sensor. The parameters set of the projector are specified by vector of nine elements:

$$P = [a \quad b \quad c \quad x_p \quad y_p \quad z_p \quad k_{3p} \quad k_{5p} \quad f_p]^T, \tag{7}$$

In our system, we used a projector to project the color stripes pattern and two cameras to capture image simultaneously. The number of system parameter is 27.

3.3 Parameters Calibration

System calibration is the process of determining the internal geometrical-optical characteristics and the 3D position-orientation of the cameras and projector relative to the chosen world coordinate system. Typically, to calibrate any system parameter, it needs some known values to be compared with. Here, we used flat panel fixed to the moving bridge of CMM, which is well known as a very precision measuring machine. Then accurate position of flat panel could be obtained. However, number of system parameters is increased, since normal vector of flat panel is unknown. The world coordinate system is chosen to be a projected position of center of pixel array (512, 384) point on the flat panel. Position of flat panel is set to be zero here. The orientation of world coordinate is paralleled to the CMM's coordinates.

The calibration is performed as follows. A known pixel coordinate point is projected onto the flat panel. Cameras monitor projected point then pixel coordinate of observed point on camera image is extracted by such an image processing. Then now, we have known positions of white flat panel, points on projector image plane and their correspondence points on camera image plane.

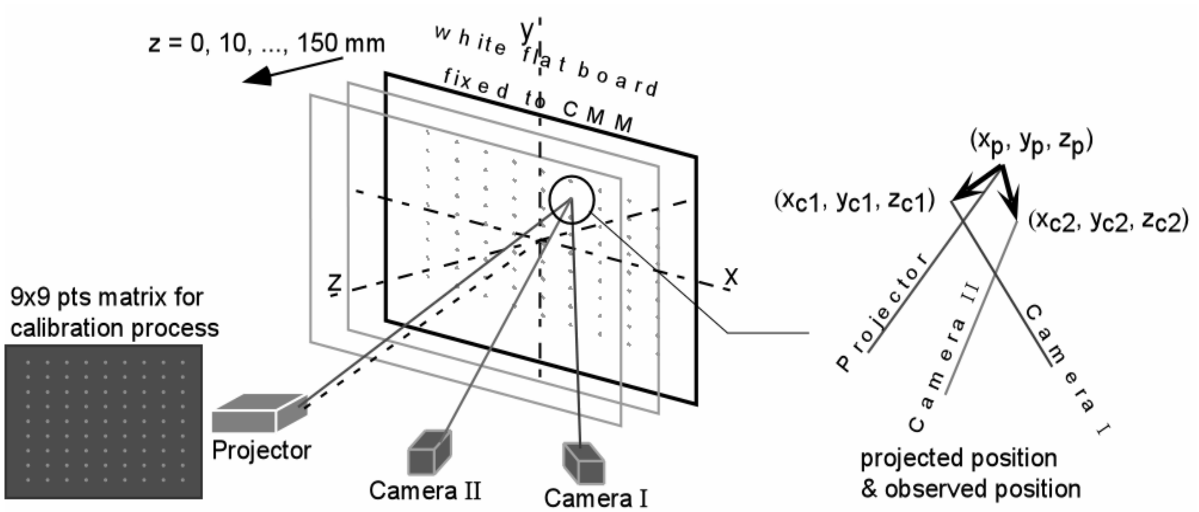

Fig. 3. Calibration model

Information of each point could be written as,

$$pt = [n_{px} \quad n_{py} \quad Z \quad n_{cx1} \quad n_{cy1} \quad n_{cx2} \quad n_{cy2}]^T, \qquad (8)$$

Calibration		#1	#2	#3
Projector Parameters	Xp	-4.83	-6.16	-5.28
	Yp	-280.95	-279.07	-279.43
	Zp	1,045.42	1,043.70	1,045.95
	ψ	0.0071	0.0067	0.0061
	θ	-0.0035	-0.0057	-0.0035
	ϕ	-0.0163	-0.0181	-0.0175
	fp	25.499	25.470	25.512
	kp3	6.36E-04	6.02E-04	6.21E-04
	kp5	-2.75E-05	-2.37E-05	-2.96E-05
Camera I Parameters	Xc1	799.98	801.02	799.78
	Yc1	-10.11	-10.07	-10.20
	Zc1	903.17	904.54	903.46
	ψ1	-0.0372	-0.0383	-0.0379
	θ1	0.7451	0.7437	0.7448
	ϕ1	-0.0158	-0.0175	-0.0171
	fc1	15.808	15.866	15.817
	kc31	1.31E-03	1.36E-03	1.30E-03
	kc51	-1.27E-05	-1.60E-05	-1.27E-05
Camera II Parameters	Xc2	414.68	414.06	414.33
	Yc2	-5.35	-4.86	-4.93
	Zc2	904.25	903.75	903.94
	ψ2	-0.0205	-0.0215	-0.0213
	θ2	0.4479	0.4466	0.4478
	ϕ2	0.0138	0.0121	0.0126
	fc2	12.007	12.033	12.007
	kc32	1.61E-02	1.57E-02	1.77E-02
	kc52	5.47E-03	6.26E-03	5.94E-03
Flat Panel Parameters	ψ	-7.36E-05	-7.48E-05	-6.31E-05
	θ	1.59E-06	1.51E-06	1.43E-06
Average E1				
Average E2				

Fig. 4. Calibration results

Then *projected point* (X_p, Y_p, Z_p) could be calculated from system parameters and pixel coordinate of projected point (n_{px}, n_{py}). Also with system parameters and pixel coordinate of observed point (n_{cx}, n_{cy}), *monitored point* (X_c, Y_c, Z_c) could be calculated. Actually, *projected point* and *monitored point* are the same point on the flat panel but derived from different sources. The difference, error, of these two points could be written as Eq. (9).

$$E_i = \sqrt{(X_{pi} - X_{ci})^2 + (Y_{pi} - Y_{ci})^2 + (Z_{pi} - Z_{ci})^2}, i = 1, 2, ... \qquad (9)$$

where i is camera number. By minimizing these errors, E_1 and E_2, an accurate estimation of the system parameters could be obtained. In order to precisely estimate the system parameters, large number of calibration points should be acquired. In our experiment, we projected the 9*9 points matrix pattern on flat panel that placed at 16 various positions then 1,296 calibration points are obtained. Non-linear least square optimization technique is used to estimate the parameters for the projector-cameras model. The calibration results are shown in Fig. 4.

4. 3D Profile Measurements

4.1 Using of Color

Owing to the performance of CCD sensor and variation of reflecting light by surface and ambient light, the different between projected color and captured color is occurred which causes the limitation of usable color. For one shot scanning that only one stripes-pattern is used, several stripes are generated with limited number of color. Therefore, matching of color-stripes between captured pattern and projected pattern becomes necessary. Generally, the correctness of matching is bad as the increase of stripe:color ratio. In order to select the suitable colors for generating a pattern, many of stripe-patterns that varied by color are projected on the flat panel. The colors of stripe-pattern are varied by changing of three color-components; R, G and B, as step of 0, 30, 55, 80, 105, 130, 155, 180, 205, 230 and 255. With the information of projected-stripe color extracted from the images, we selected colors that are obviously distinct from one another.

In the RGB colorspace, each color is represented by a three number 'triple'. The components of this triple specify, respectively, the amount of red, the amount of green, and the amount of blue in the color. The HSV colorspace works somewhat differently. It is considered to be more intuitive to use, closer to how an artist actually mixes colors. In the HSV colorspace, each color is again determined by a three-component 'triple'. The first component, *Hue*, describes the basic color. The second and third component, *Saturation* and *Value*, describes purity of color and intensity respectively. HSV colorspace is then suitable to use for specifying the colors since shade of colors is depended on only H component. The selected colors with their range of H values are shown in fig. 5.

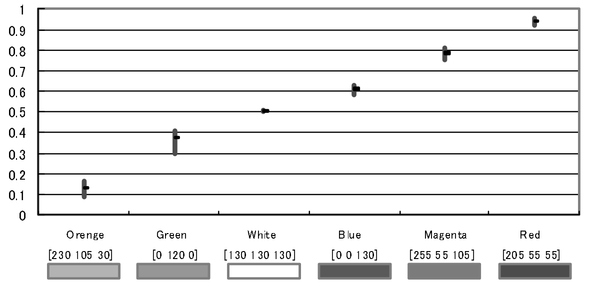

Fig. 5. Range of hue of captured stripe colors

According to Fig. 5, *Hue* of each color is distinctly different. For example, Hue of orange color is between 0.09 and 0.18; green color is between 0.3 and 0.42 and so on.

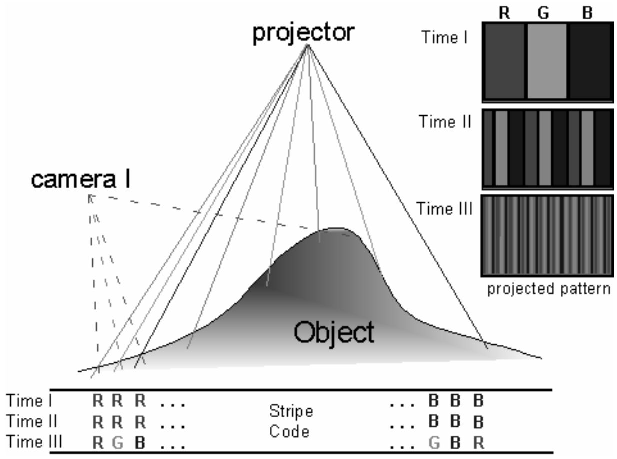

Fig. 6. Principal of space time scanning technique

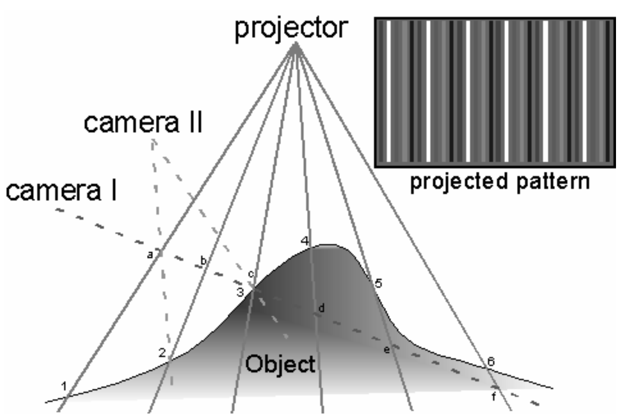

Fig. 8. Principal of one shot scanning

4.2 Space Time Scanning Technique

The principle of the space time scanning technique is described in Fig. 6. The color of stripes pattern are changed in time series as in pattern 1, 2 and 3.

Each stripe has individual stripe-code; the left most stripe has RRR code, the next stripe has RRG code and so on. The captured stripe and corresponding projected stripe then could be matched. The 3D points are calculated by means of triangulation. By this projection method of m patterns, the maximum number of stripe is N^m. Therefore; 6-colored pattern with 3 projected stripes-patterns can generate the pattern of 216 stripes. To reduce a number of patterns, new technique called *one shot scanning* is developed.

4.3 One Shot Scanning Technique

Rather than taking several images from one camera, we could instead take one image from each of several cameras put at the various viewpoints. This technique is called *one shot scanning*.

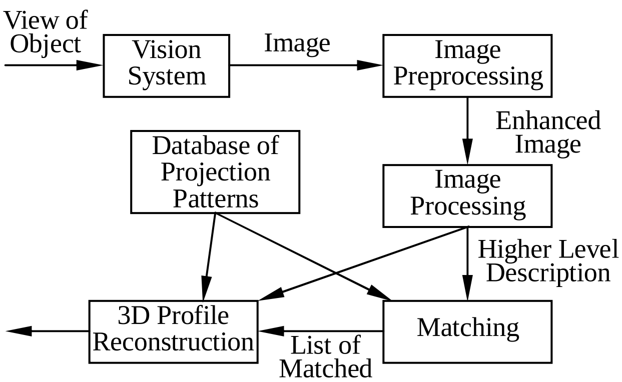

Fig. 7. Schematic diagram of 3D profile measurement

The technique works as follows. The color stripes pattern is projected onto the object. More than two of CCD-cameras simultaneously capture the image of projected pattern that distorted by object's shape. The captured images are managed by such an image processing for preparing a higher-level description of image such as pixel coordinate and color of stripe. The projected stripe-pattern is then matched to the captured stripe-pattern. Finally, having a matched, the 3D profile can then be reconstructed by means of triangulation. Clearly, it needs a very accurate parameter estimation, since numerous points on the images are to be precisely matched. Bad parameters would be caused an incorrect match. The principal of the *one shot scanning* technique is described in Fig. 8.

The measure space is divided into 66 sub-spaces; stripes with various shade. The Colors of each set of 11 stripes are 1-white, 2-red, 3-green, 4-blue, 5-magenta and 6-orange. The color stripes structure is ordered as 1, 2, 3, 4, 5, 6, 1, 2, 3, ... from left to right. The problem is how to correctly match this pattern.

The stripes projection of such one color is shown in Fig. 8. In order to simplify the explanation, only 6 stripes projection is considered. Projected points are point *1, 2, 3, 4, 5* and *6*. Now, consider point *3*, Camera I could observe this point as point *a, b, c, d, e* or *f*. By checking with another view point camera, camera II, point *3-c* could be matched. However, there is a case that camera II could observed more than one point, as can be seen in Fig. 8, both points *a* and *c* could be observed. Point *3-a* could be incorrectly matched. It needs a contribution to minimize this error. Obviously, the matching process is the most difficult and complex in 3D profile measurement that since it could cause an error as seen in Fig. 9.

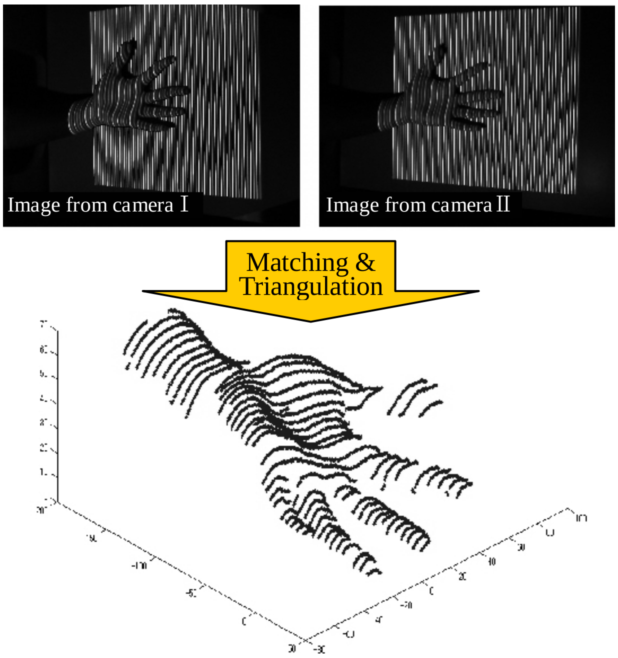

Fig. 9. 3D profile measurement result

5. Conclusion

In this work, we have successfully developed and implemented a new method for 3D profile measurement by using color stripes pattern projection called *one shot scanning*. The advantage of *one shot scanning* that is the use of only one image taken from each camera; yields the possibility of profile measurement for moving object such as human body. However, the moving speed of the object that could be reconstructed is limited by the shutter speed of the cameras. The more shutter speed, the more moving speed of object is allowed. However, when the shutter speed is fast, the brightness of image is low. The suitable shutter speed then has to be adjusted. We also developed a procedure for completely projector-camera parameters calibration using points pattern projected from projector as the calibration points. For the image processing, it still needs to be improved in the point of matching process to minimize the computing time and obtain more correctly matching result. We also hope to implement a realtime capture (possibly offline processing) system by repeat the process at the image caption rate.

References

[1] H. A.M. Daanen, G. Jeroen van de Water, Whole Body Scanners, Displays, Vol. 19, pp. 111-120, 1998

[2] G. Sansoni, M. Carocci, R. Rodella, Calibration and Performance Evaluation of a 3-D Imaging Sensor Based on the Projection of Structured Light, IEEE Trans. Instrumentation and measurement, Vol. 49, No. 3, June 2000

[3] S. Osawa, R. Furutani, K. Takamasu, S. Osono, H. Asano, 3-D Shape Measurement by Self-referenced Pattern Projection Method, Measurement, Vol. 26, 1999, pp. 157-166

[4] C.S. Chen, Y.P. Hung, C.C. Chiang, J.L. Wu, Range Data Acquisition Using Color Structured Lighting and Stereo Vision, Image and Vision Computing, Vol. 15, 1997, pp. 445-456

[5] L. Zhang, B. Curless, S. M. Seitz, Rapid Shape Acquisition Using Color Structured Light and Multi-pass Dynamic Programing, IEEE Proceedings on 3D Data Processing Visualization and Transmission, 2002

[6] R.Y. Tsai, A Versatile Camera Calibration technique for high-accuracy 3D Machine Vision Metrology Using Off-the-shelf TV Cameras and lenses, IEEE Journal on Robotics and Automation, 1987, pp. 323-344

[7] H. Zhuang, A Self-calibration Approach to Extrinsic Parameter Estimation of Stereo Cameras, Robotics and Autonomous System, Vol. 15, 1995, pp. 189-197,

[8] F. Pedersini, A. Sarti, S. Tubaro, Accurate and Simple Geometric Calibration of Multi-camera Systems, Signal Processing, Vol. 77, 1999, pp. 309-334

[9] J. Salvi, X. Armangue, J. Battle, A Comparative Review of Camera Calibrating Methods with Accuracy Evaluation, Pattern Recognition, Vol. 35, pp. 1617-1635, 2002