

# Resolution adjustable 3D scanner based on using stereo cameras

Tzung-Han Lin\*

\*Graduate Institute of Color and Illumination Technology,  
National Taiwan University of Science and Technology, Taiwan.

E-mail: thl@mail.ntust.edu.tw Tel: +886-2-27303717

**Abstract**—This paper addresses a stereo-based 3D scanner system, which is able to acquire various resolution range data. The system consists of stereo cameras and one slit laser. In each stereo image pair, we cast one laser stripe on the surface of object, and analyze their disparities for determining their depth values. Utilizing a super-sampling filter, the sub-pixel features are generated for enhancing the native resolution of CCD component. In this system, we use one slit laser for sweeping the surface of objects and generating correspondences under the epipolar constrain. Since the correspondences are generated by the positions of the cast stripes, their resolution is controllable.

## I. INTRODUCTION

3D scanner now is a kind of consumer-level products for various applications. It was widely used in reverse engineering, industrial inspection, heritage protection, *et al.* In these fields, non-contact measurement methods are often used for efficiently acquiring range data. In non-contact measurements, there are two main categories called time-of-flight and triangulation methods [1]. The triangulation method usually casts light stripes on the surface of objects, and then a receiver acquires images for disparity analysis. It is much easy being carried out by using cost effective components. However, it is still very sensitive to the material of objects.

In 2000, Levoy, *et al.* had accomplished the digital Michelangelo project [2]. It is one famous project for 3D reconstruction for statues. Their objective is to preserve the capital creation which was accomplished at 16 century. They used high resolution laser scanners to acquire the shapes of David statue. Their spatial resolution was 0.25 mm, and finally two billion polygons models were created. The 3D shape is sometime used not only for data perseveration but also for shape modification. It also can be reused for the shape design and shape analysis.

Most commercial 3D scanners are designed for portability and user can readily gather 3D data from arbitrarily directions. Borghese, *et al.* [3] proposed a low cost portable 3D scanner system. Their prototype consists of a laser pointer and a pair of video camera. By recognizing the cast spots, 3D points are acquired. Rocchini, *et al.* [4] used a video projector and a camera to fabricate a scanner device, and successfully reconstructed cultural heritage artifacts. In their work, the stripped color light is projected on artifacts, and the stripped patterns are analyzed for determining the surfaces of artifacts.

Their solution is one kind of non-contact measurement methods and cost effective.

In some 3D scanner, the stereo camera is used for reducing blind spots and improving the data stability. Lv and Zhang build a 3D scanner system by using binocular vision [5]. In their method, a slit laser is used for generating corresponding features. However, the acquired pixels are discrete from Sobel edge operator. And the 3D points will be discrete. To avoid this phenomenon, our design has an improvement by using sub-pixel operation and epipolar constrains under a non-parallel camera configuration. Besides, the resolution depends on how fast the motion of this slit laser is.

Calibration is usually the most important part in the development of high precision 3D scanners. Vilaca, *et al.* [6] proposed a 3D scanner based on stereo vision and laser triangulation. They used a transparent checkerboard for simultaneously calibrating the extrinsic parameters of both cameras. And the laser plane is constrained for exactly laying on the checkerboard. As a consequence, a lookup table of cast laser points is generated. Applying an additional translation on the scanner, the range data will be obtained.

Obviously, the integration of the scanned patched is always necessary to accomplish a complete model. Since the coordinates of scanned patches are not consistent in different scanning procedures, the coordinates' transformations among them are critical. Besl and McKay proposed a closed form solution named ICP (iterative closest points) for solving registration problems [7]. The basic assumption is that two

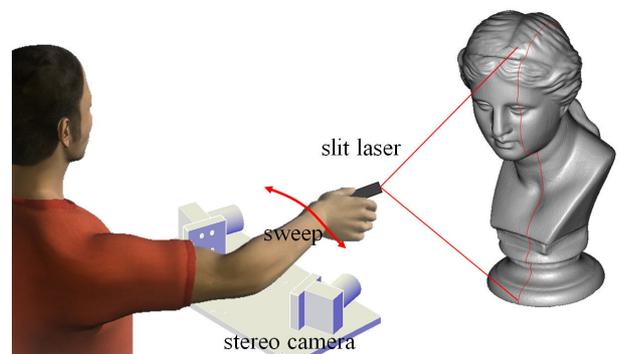


Fig. 1 Schematic of our system. The user sweeps the laser beam, and the corresponding features are collected for 3D reconstruction

objects are rigid bodies and they have an overlap region. With ICP method, all scanned patches can readily be integrated.

In this paper, we developed a 3D scanner device based on stereo camera and one slit laser. In our prototype, its size is compact for small sculptures. We focus on developing a high flexibility 3D scanner whose resolution is adjustable. The operation has high flexibility and it allows users to readily sweep a slit laser under different speeds and various tilt angles. Fig. 1 illustrates the schematic of our system. The stereo camera aims at the object and the user sweeps a slit laser for casting laser lines on this object. All of the captured images from this stereo camera are synchronized. And then, they are transmitted to one computer for 3D reconstruction.

## II. PROPOSED SYSTEM

### A. Fabrication of 3D scanner

Traditionally, 3D laser scanner consists of at least one charged coupled device (CCD) and one slit laser. The tilt angle between the laser and CCD is usually set to be a constant. An example is shown in the left of Fig 2. The position between the CCD and the slit laser is fixed, and both of them are driven in a specific motion, which is usually a translation or a rotation, for acquiring 3D range data. By increasing the tilt angle, it will induce a better spatial resolution. However, it will induce more self-occluded conditions as well. The cast stripe will be observed in the CCD, and their depths depend on how far the pixels shift in horizontal direction. It can be stored as a lookup table as well.

In order to improve the flexibility, the slit laser in our design is allowed to have arbitrary motions. And, both CCDs are used for generating stereo images shown in the right of Fig. 2. Once the slit laser casts a stripe on the object, two CCDs will take a stereo pair for 3D reconstruction. Since the slit laser is allowed having arbitrary motions, the synchronization of these two CCDs is extremely required. In practice, we select IEEE1394 interface for synchronization. The difference of timestamp between two CCDs is less than 0.125 ms. Some stereo pairs which are not synchronized will be rejected to avoid outliers in 3D calculation. With regard of the geometrical relation between two CCDs, it depends on the work zone. In our prototype, the convergence angle of the stereo camera is 30 degree. The photon sensor size of CCD is 4.8 mm by 2.8 mm with one 8 mm lens.

### B. Calibration of stereo cameras

The calibration of stereo cameras follows Zhang's method [8]. Initially, both of CCDs are individually calibrated to obtain the intrinsic parameters and distortion factors. Then, their extrinsic parameters are determined from the same features. In calibration procedure, the 3D projection can be written as equation (1), where  $\mathbf{P}$  is the projection matrix consisting of an intrinsic parameter  $\mathbf{K}$  and an extrinsic parameter  $[\mathbf{R}|\mathbf{t}]$ . For convenience, the rotation matrix  $\mathbf{R}$  is decomposed into three column vectors as in equation (2). Utilizing the homography of at least three planes in Euclidean coordinate, first two columns of the homography of one plane,

says  $\mathbf{h}_1$  and  $\mathbf{h}_2$ , are rewritten as in equation (3). The intrinsic parameter  $\mathbf{K}$  will be solved by Cholesky factorization of equation (4). Finally, the extrinsic parameters are obtained.

$$\mathbf{x} = \mathbf{P}\mathbf{X} = \mathbf{K}[\mathbf{R}|\mathbf{t}]\mathbf{X} \quad (1)$$

$$\mathbf{x} = \mathbf{K}[\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3 \quad \mathbf{t}]\mathbf{X} \quad (2)$$

$$\begin{cases} \mathbf{h}_1^T \omega \mathbf{h}_2 = 0 \\ \mathbf{h}_1^T \omega \mathbf{h}_1 = \mathbf{h}_2^T \omega \mathbf{h}_2 \end{cases} \quad (3)$$

$$\omega = (\mathbf{K}\mathbf{K}^T)^{-1} \quad (4)$$

### C. Feature extraction

Feature extraction is the main procedure for generating correspondences. Since the slit laser casts one bright stripe on the surface on the object, every sub-pixel of this stripe is considered as one feature. Here, we use the Gaussian filter to determine the brightest pixel in each row of one image instead of the operation of conventional edge detection. To obtain a stable value, the brightest pixel of each row is considered as the centroid of the intensity distribution of a specific width. In practice, the slit laser in our system is not necessary to be a perfect flat plane. In other words, the emission of a laser line is not necessary to be exactly straight. However, the only limitation is that the laser line casts only one brightest feature in one row of the image. Nevertheless, a straight slit laser is recommended.

### D. Correspondences under epipolar constrain

The epipolar geometry in the stereo cameras is determined after calibration. To determine the fundamental matrix between these two CCD, we directly calculate it from the known intrinsic and extrinsic parameters following Faugeras's work [9]. We assume  $[\mathbf{R}|\mathbf{t}]$  and  $[\mathbf{R}'|\mathbf{t}']$  is the extrinsic parameter of left and right CCDs, respectively. The translation and rotation between two CCDs will be  $\mathbf{t}^*$  and  $\mathbf{R}^*$ , as shown in equation (5). Then, the fundamental matrix is rewritten as the combination of both intrinsic parameter and the essential matrix as in equation (6). This procedure can be

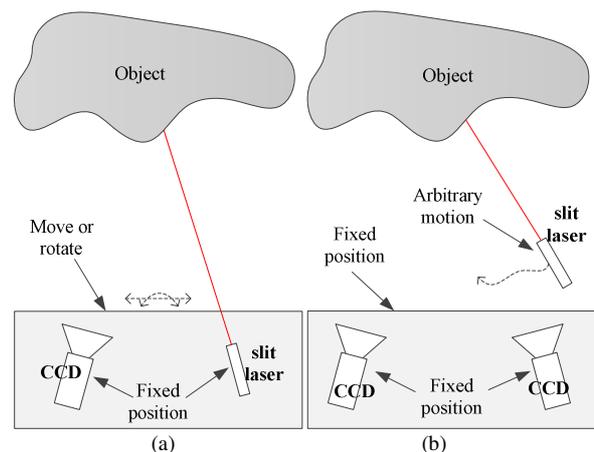


Fig. 2 Schematics of traditional 3D scanner compared to our design. The traditional 3D scanner is shown in (a). Our design has high flexibility shown in (b).

determined by selecting at least 8 correspondences and solved by RANSAC algorithm as well.

$$\begin{bmatrix} \mathbf{R}^* & \mathbf{t}^* \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R}' & \mathbf{t}' \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{bmatrix}^{-1} \quad (5)$$

$$\mathbf{F} = \mathbf{K}'^{-T} [\mathbf{t}^*]_{\times} \mathbf{R}^* \mathbf{K}^{-1} \quad (6)$$

Fig. 3 shows how we interpolate a correspondence by the epipolar line. In a stereo pair, the brightest features of each row of both views are determined and stored as sub-pixel values. All brightest features in an image are sorted according to their  $y$ -components. For example in Fig. 3,  $\mathbf{x}_i$  is one feature in left view, and its corresponding point  $\mathbf{x}'_i$  is the interpolation of its epipolar line  $\mathbf{l}'_i$  and the line formed by  $\mathbf{x}'_k$  and  $\mathbf{x}'_{k+1}$ . These two points  $\mathbf{x}'_k$  and  $\mathbf{x}'_{k+1}$  are the brightest features in right view. To have a good estimation, the distance from either  $\mathbf{x}'_k$  or  $\mathbf{x}'_{k+1}$  to the epipolar line must be less than a specific value. The value we recommend is 1.414 pixels. In this example, one correspondence  $(\mathbf{x}_i, \mathbf{x}'_i)$  is generated. By the similar procedure, the interpolated features in left view can be found to be corresponding points of the features of right view as well.

#### E. Closed form solution for 3D points

The closed form solution for 3D estimation is based on the assumption: the cross product of the desired and estimated 2D points is a zero vector. This is basically a direct linear solution. We assume the 3D point  $\mathbf{X}$  is projected as  $\mathbf{x}$  and  $\mathbf{x}'$  on the left and right images, respectively, as shown in equation (1). And, the projection matrix  $\mathbf{P}$  can be rewritten as three row vectors shown in equation (7).

Based on the direct triangulation method, the cross products of desired and estimated 2D points in two-view images are induced as equation (8). That means one correspondence will determine one 3D point when two projective matrixes are known. Finally, the solution will be the format like equation (9), where  $\mathbf{x}$  and  $\mathbf{x}'$  are the 2D points in different CCDs and defined in equations (10) and (11). So, it can be solved by singular value decomposition (SVD) method. Finally, all of the range points are converted into triangular meshes via Delaunay triangulation. Some polygons having irregular sizes or extreme sharp shapes will be

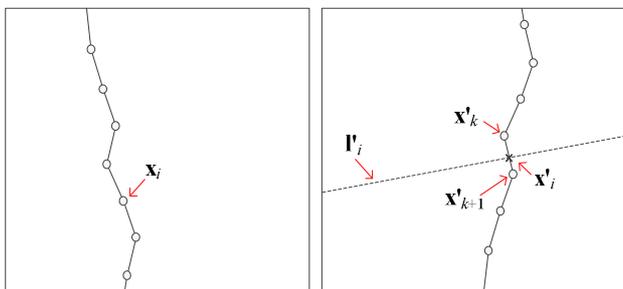


Fig. 3 Schematics of determining one correspondence from a stereo pair. The cast stripes of the object in different views is shown in left and right figure, respectively. The vertex  $\mathbf{x}'_i$ , which is the corresponding point of  $\mathbf{x}_i$ , is the intersection of the epipolar line and two successive vertices.

removed.

$$\mathbf{P} = [\mathbf{p}_1^T; \mathbf{p}_2^T; \mathbf{p}_3^T] \quad (7)$$

$$\mathbf{x} \times [\mathbf{P}\mathbf{X}] = 0 \quad (8)$$

$$\begin{bmatrix} x\mathbf{p}_3^T - \mathbf{p}_1^T \\ y\mathbf{p}_3^T - \mathbf{p}_2^T \\ x'\mathbf{p}_3^T - \mathbf{p}_1^T \\ y'\mathbf{p}_3^T - \mathbf{p}_1^T \end{bmatrix} \mathbf{X} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad (9)$$

$$\mathbf{x} = [x \quad y \quad 1]^T \quad (10)$$

$$\mathbf{x}' = [x' \quad y' \quad 1]^T \quad (11)$$

#### F. How the resolution adjustable

In our design, the feature of adjustable resolution comes out. This is because the correspondences are arbitrarily generated in various densities. In some applications, the complexity of surface is not equal. For example in a portrait sculpture, the shape on the hair is much complex than that on the cheek. Sometime, the complex region needs huge amount of sampling points for representing its detail.

However, the resolution in vertical and horizontal directions has a little difference. In the horizontal direction, the sampling features depend on how fast the horizontal motion of the slit laser is. Even though one can use a very slow motion of the slit laser for acquiring more sampling points, the horizontal resolution has its native limitation, says the grey bits and pixel numbers; In the vertical direction, the resolution depends on not only the pixel numbers but also the slope of epipolar line. As a consequence, our method adeptly induces different resolution for various kinds of surface complexities.

### III. EXPERIMENT RESULT

We fabricated a 3D scanner consisting of one slit laser and two CCDs. Basically, 3D points are obtained by the triangulation method. Our method has high flexibility for adapting various complexities of 3D objects. Fig. 4 shows a snapshot when the user is scanning a sculpture. The prototype



Fig. 4 Snapshot of scanning a sculpture.

which is connected to one computer by IEEE1394 interfaces is shown in the right-bottom corner of Fig. 4. The user, on the right-top of figure, is sweeping a slit laser and generating a stripe on the surface of this sculpture. Since our calculation for 3D reconstruction is real time, the monitor is displaying the acquired 3D points simultaneously.

In operation, the relative position between the object and the scanner is fixed till finishing current range data. During a scan procedure, a trig is sent to start receiving images for analysis, and then, another trig is sent again to stop collecting images. To generate more 3D points, we set CCDs to be able to take portrait images. That means the pixel resolution of the CCD is 480 by 640. So, our method can generate at most 1280 3D points in one image pair. The frame rate of our design is 60 fps. When one image pair is transmitted to the computer, the brightest feature in each row is determined immediately. After finding the brightest feature, a software threshold for rejecting outliers may be needed. Besides, the band-pass filter for a specific spectrum plays the same role as well. Since the parameter is calibrated before starting using the scanner, its epipolar geometry will assist in finding an accuracy correspondence. Currently, the computer has ability to solve equations such as equation (9) in a very short period. This performance conducts that around 76 thousand 3D points could be acquired within one second.

With regard of the error analysis, our re-projection error in the calibration procedure is 0.25 pixel for a 0.3M pixel image. However, it depends on how far the object is. Fig. 5 shows the result of scanning a distant plate under different resolutions. In the left and middle, the operator sweeps the slit laser once with a fast and a slow motion, respectively. In the right of Fig. 5, the laser is swept back and forth for several times to generate a lot of 3D points. In other words, the resolution increases from left to right figures. Nevertheless, the error distributions are almost the same in different resolutions. With comparison to scanning a close plate, the error distribution is totally different as shown in Fig. 6. Although the depth resolution in the close region is better than those in a distant region, the average re-projection error is not good. This is because we had converted CCDs into the pinhole camera model, and the calibration parameter is optimized in a specific distance, which is so called work zone.

The other phenomenon in Fig. 7 shows compromised re-projection errors. This 3D sculpture is scanned by different resolutions. In Fig. 7, many creases come out in the high

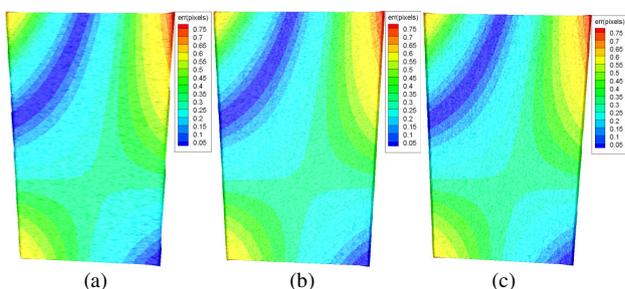


Fig. 5 Scanning a distant plate with different resolutions. (a) and (b) figures indicate the results of one-pass fast and one-pass slow scans. (c) is a multi-pass slow scan.

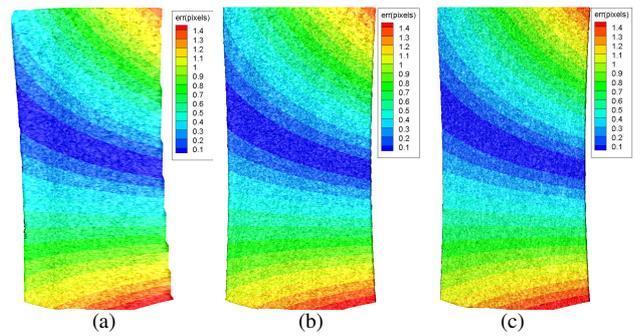


Fig. 6 Scanning a close plate with different resolutions. (a) and (b) figures indicate the results of one-pass fast and one-pass slow scans. (c) is a multi-pass slow scan.

resolution scan. This is because there are too many sampling data in a small region. The other reason is that the CCD induces noise when converting photons into electronic signals in a high contract image. Nevertheless, it is a benefit for reducing the measurement uncertainty. In practice, we resample 3D points that are in a specific volume to obtain a stable 3D estimation. And, another treatment on improving the surface roughness is to use Gaussian smooth operator.

Due to the operation flexibility, the user can readily increase the resolution for spcific regions. As mentioned in previous section, the slit laser is not necessary to be a perfect line. It can be a short line or a curve to occupy a brightest pixel in one row. Even using one laser point is also available in our system. This behavior induces a more flexiabile resolution distribution but it is more time consuming.

With regard of the limitation of our system, we do not treat non-lambertian surfaces. This is because we use an active laser light for generating the cast stripes. If the laser light has obvious reflection or translucancy on the surface of objects, there will be many outliers on the undesired poistions.

The analysis of re-projection error will assist user avoiding poor quality regions. Some physical behaviors of the surface may affect 3D acquisition results. For example, surface reflectivity is a critical factor for representing the quality of cast stripes. Some dark regions having high light absorption rate are usually difficult to be scanned. In Fig. 8, the scan of composite color materials shows that the 3D points on the black region are rejected, although the overall error is small. In these scan conditions, an additional process for hole-filling is usually necessary. It sometimes needs extra works via 3D

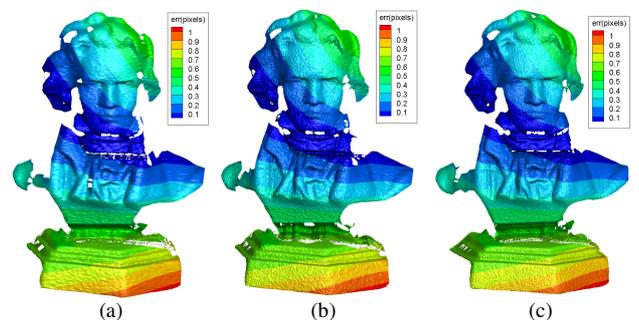


Fig. 7 Scanning a sculpture with different resolutions. The resolution increase from (a) to (c).

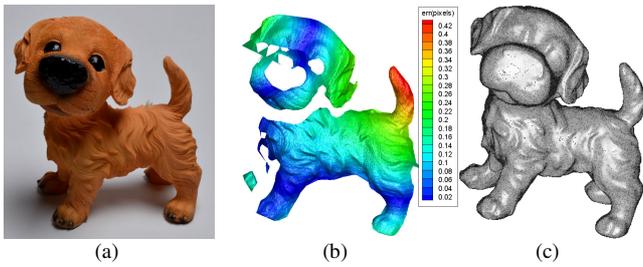


Fig. 8 Scan result of a dog model consisting of composite materials. (a) is the original object, (b) is the range data and (c) is a final result. The reconstructed 3D model in (c) consists of 10 scan patches.

manipulation software. After hole filling, a defect appears on the nose and eyes of the dog model. All scan patches are registered interactively by ICP method. And then, these patches are integrated together by boolean operation. The estimated error on the overlap region of two patches is 0.1 mm.

To estimate the ability of the proposed system, we have tested several sculptures having different sizes. In Fig. 9, three sculptures are scanned. In the left column of Fig. 9, the scan result compared to the original model has lost details on hair, mouth and ear. Since the model is small enough, the surface details are not easy being recognized by our design. However, it is not a native limitation for this kind of design. The scale can be changed for adapting different size objects. Utilizing a small baseline of stereo camera, it improves the spatial resolution but get a small work zone. In the middle and right columns of Fig. 9, the scan results of two large sculptures show good surface details. We believe that adjustable



Fig. 9 The scan results after integration. (a), (b) and (c) shows the original models having 7cm, 24cm and 35 cm heights, respectively. (d), (e) and (f) are the 3D reconstruction results by our proposed system.

resolution is helpful for scanning 3D objects.

#### IV. CONCLUSION

In this paper, we carry out a 3D scanner system for range data acquisition. We use two CCDs and one moving slit laser to fabricate a high operation-flexibility 3D scanner. This enables users to duplicate 3D shapes as multi-resolution data. Besides, the measurement uncertainty can be improved by casting more stripes on the surface of objects. We have provided a complete analysis to find the correspondences and to reconstruct 3D points. We believe that adjustable resolution will be a key feature for more comprehensive scanning applications.

#### ACKNOWLEDGMENT

This work was supported in part by the National Science Council of Taiwan under Grant NSC-102-2221-E-011-094-MY2.

#### REFERENCES

- [1] J. Davis and X. Chen, "A laser range scanner designed for minimum calibration complexity," *IEEE International Conference on 3D Digital Imaging and Modeling 2001*, pp. 91–98, 2001.
- [2] M. Levoy, K. Pulli, B. Curless, S. Rusinkiewicz, D. Koller, L. Pereira, *et al.*, "The digital Michelangelo project: 3D scanning of large statues," *ACM SIGGRAPH 2000*, pp. 131–144, Jul. 2000.
- [3] N. A. Borghese, G. Ferrigno, G. Baroni, A. Pedotti, S. Ferrari and R. Savare, "Autoscan: A Flexible and Portable 3D Scanner," *IEEE Computer Graphics and Applications*, vol. 18, pp. 38–41, 1998.
- [4] C. Rocchini, P. Cignoni, C. Montani, P. Pingi, and R. Scopigno, "A low cost 3D scanner based on structured light," *Eurographics 2001*, vol. 20(3), pp. 299–308, 2001.
- [5] Z. Lv and Z. Zhang, "Build 3D scanner system based on binocular stereo vision," *Journal of Computers*, vol. 7(2), pp. 399–404, 2012.
- [6] J. L. Vilaca, J. C. Fonseca, and A. M. Pinho, "Non-contact 3D acquisition system based on stereo vision and laser triangulation," *Machine Vision and Applications*, vol. 21(3), pp. 341–350, Mar. 2009.
- [7] P. Besl and N. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14(2), pp. 239–256, 1992.
- [8] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22(11), pp. 1330–1334, 2000.
- [9] O. D. Faugeras and Olivier D. Faugeras, "The fundamental matrix: theory, algorithm, and stability analysis," *International Journal of Computer Vision*, vol. 17, pp. 43–75, 1996.