3D Object Modeling by Structured Light and Stereo Vision

UĞUR ÖZENÇ[*], OĞUZHAN TAŞTAN[**], M. KEMAL GÜLLÜ[*]
[*]Department of Electronics and Telecomm. Eng.
Kocaeli University
Umuttepe Kampusu, 41380, Kocaeli
TURKEY
[**]Department of Computer Eng.
Middle East Technical University
Inonu Bulvari, 06800, Ankara
TURKEY

kemalg@kocaeli.edu.tr   http://akademikpersonel.kocaeli.edu.tr/kemalg/

Abstract: - In this paper, we demonstrate a 3D object modeling system utilizing a setup which consists of two CMOS color cameras and a DLP projector by making use of structured light and stereo vision. The calibration of the system is carried out using a calibration pattern. Images are taken with stereo camera pair by projecting structured light onto the object, and a 3D point cloud is reconstructed by using both epipolar constraint of stereo vision and gray-code constraint of structured light. Additionally, separate 3D point clouds are extracted performing ray-plan intersection between camera-projector pairs. Finally, obtained three 3D point clouds are superposed using iterative closest point algorithm. Experimental results show that the proposed 3D modeling system is able to cope with occlusions arose from a single camera-projector setup, and objects models are reconstructed successfully.

Key-Words: - 3D modeling, Coded light, Stereo vision, Occlusion regions

1 Introduction

3D object modeling is the process of building a digital mathematical representation from a 3D real world using different systems such as laser or computer vision based specialized systems. 3D modeling systems are used in industrial part design, prototyping, quality control, entertainment, reverse engineering, medical applications, cultural heritage applications etc.

Computer vision based 3D modeling systems are generally use stereo vision or structured/coded light [1-3]. Stereo vision based 3D modeling is a passive technique and can recover the structure of the object by matching features detected in multiple images of the same object [1]. This technique is computationally intensive and the depth data could be noisy. Additionally, modeling performance depends on the surface of the object and ambient light. Another is an active technique that form a structured or coded light strips using a light source such as projectors, and construct a 3D model based on deformations of the stripes in the captured image [2-3]. A color-coded structured light based is method is proposed in [4]. In this method, the correspondence problem is solved using multi-pass dynamic programming algorithm. Authors also evaluated the performance of single stripe image and multiple images comprised of time shifted stripes, and they concluded that reconstructed 3D models using multiple images include more detail and less noise than the obtained by single images. In [3], a method that combines color structured light and stereo vision principle is proposed. The aim of the stereo vision in this work is to eliminate the correspondence problem between the color stripes projected by the light source and the color stripes observed in the images. The dense range map is generated with only one pair of stereo images, but they are not used both images. To overcome correspondence problem, another method is to use coded structured light. In [2], authors combined gray-coded and time shifted structured light patterns with the stereo vision methodology for this purpose. Stereo vision is exploited to find 3D points utilizing epipolar constraints.

This work is intended to fill empty occlusion regions originated from the angle between the light source and the camera due to the obstruction. We use stereo vision with coded structured light source and three point clouds are separately extracted

ISBN: 978-1-61804-329-0

taking epipolar constraints and coding scheme of the projected light into account. Obtained point clouds are combined in order to obtain final point cloud then. The rest of the paper is organized as follows: Section 2 introduces stages of the 3D model reconstruction, experimental results and discussions are provided in Section 3, and Section 4 concludes the paper.

2 3D Model Reconstruction

The setup used for 3D model reconstruction is comprised of two cameras, a projector and a computer. The structure of the 3D reconstruction system is given in Fig.1.

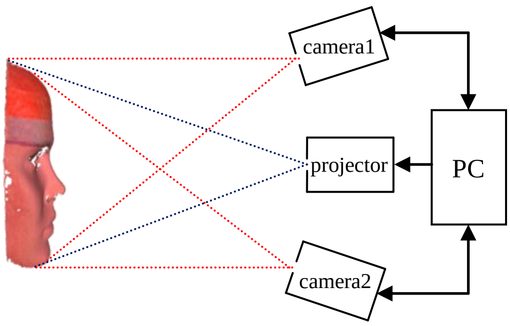
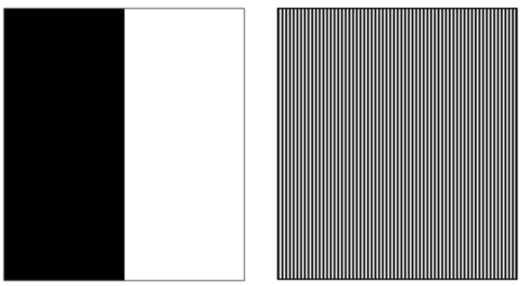


Figure 1. 3D model reconstruction system

In this setup, the projector has 1200x800 pixel resolution with 500 ANSI lumens of brightness, the cameras are colored and have 2040x2040 pixel resolution.

2.1 Gray-coded structured light

The gray-coded structured light is utilized to solve the correspondence among the points in the stereo pairs. The object is illuminated by a set of n encoded black and white patterns that correspond to binary code. Width of the patterns are progressively halved for each capture. Therefore, number of edges falling onto the object are increased sequentially. Fig. 2 illustrates first and last binary patterns projected onto the object for 7-bit code in this work. This coding scheme allows to distinguish 128 stripes. Higher bit depths can be used related to accuracy of the line extraction process.



(a) (b)

Figure 2. First and last patterns for gray-coding

2.2 Data Acquisition

In this work a set of 7 vertical patterns are used for 7-bit gray-code. Transition regions from black to white or white to black in acquired images do not present an abrupt change of gray levels depending on surface characteristics of the object and light conditions of the environment. So as to detect the edges effectively, negative images created by inverting the light intensities of the gray-coded patterns are projected onto the object and captured images are then subtracted from positive images. This operation provides improved detection performance and consequently gives better decoding performance.

2.3 System Calibration

The goal of this step is to obtain the intrinsic and extrinsic parameters for the cameras and projector Note that the projector is considered as an inverse camera. Focal length, coordinates of the principal points, radial and tangential distortions, positions and orientations of the cameras and projector are considered in this stage. In particular, intrinsic and extrinsic parameters of each camera and projector are obtained by multiple images of a reference checkerboard, and then computed by an optimization procedure given in [5].

2.4 3D Model Reconstruction

Given a pair of projection points $\mathbf{P_L}$ and $\mathbf{P_R}$, the corresponding 3D point (observation) $\mathbf{P}$ in Fig. 3 can be calculated by triangulation based on the extrinsic and intrinsic parameters in a stereo vision based 3D model extraction [6]. In Fig 3. $\mathbf{O_L}$ and $\mathbf{O_R}$ corresponds to optical centers of the cameras, $\mathbf{E_L}$ and $\mathbf{E_R}$ are the epipoles.

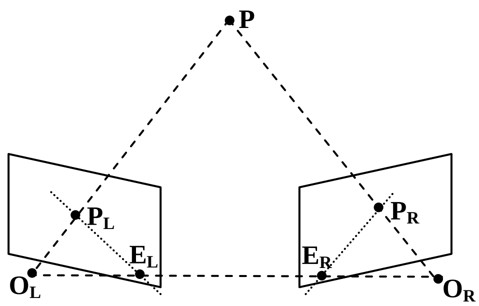


Figure 3. Ray-to-ray intersection and epipolar geometry

A 3D point on the line between points $\mathbf{P_L}$ and $\mathbf{P}$ corresponds to a point on the line between $\mathbf{E_R}$ and $\mathbf{P_R}$. If the projection points $\mathbf{P_L}$ and $\mathbf{P_R}$, and the distance between camera optical centers $\mathbf{O_L}$ and $\mathbf{O_R}$ are known, 3D point $\mathbf{P}$ in real world can be computed using triangulation. Similarly, a line in

ISBN: 978-1-61804-329-0

structured light in projector plane and a point P in real world form a light plane. Intersecting camera ray with the light plane gives the depth of the point P from ray-to-plane projection.

Flowchart of the proposed 3D model reconstruction system is given in Fig. 4.

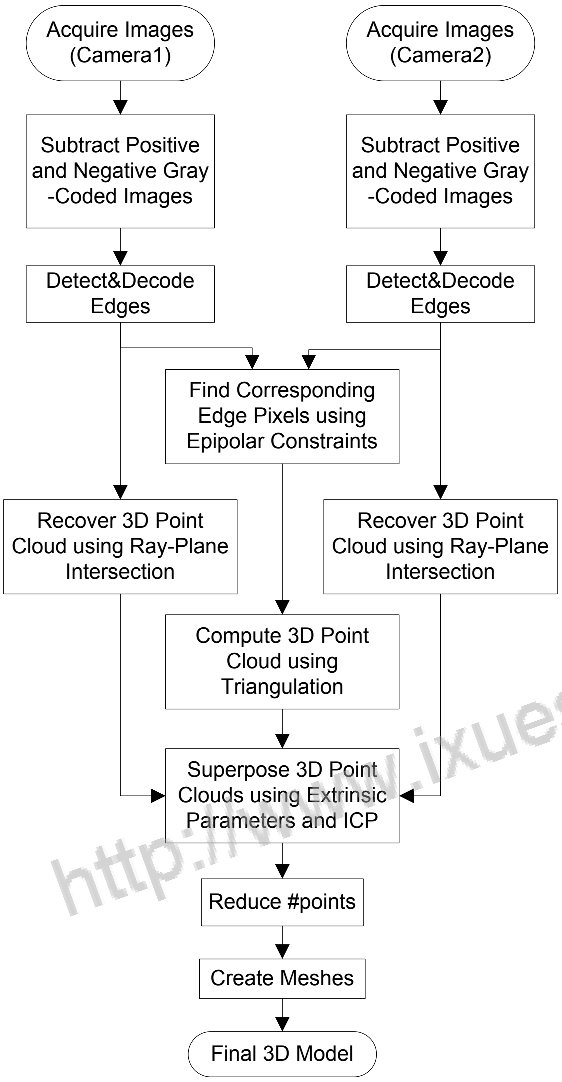
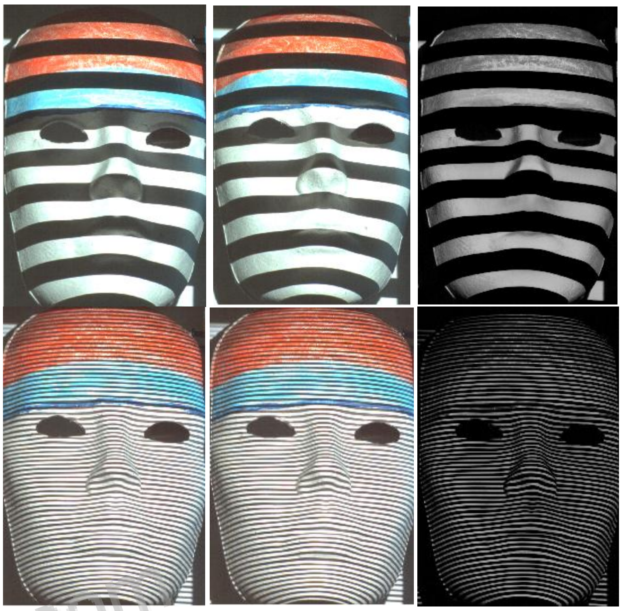


Figure 4. Flowchart of the 3D model reconstruction system.

After system calibration, gray-coded patterns are projected onto object consecutively and image sequences are acquired for both cameras. Subtraction is performed between positively and negatively gray-coded images as given in Fig. 5. Afterwards, edges are extracted by filtering with 10x10 Prewitt kernel followed by thresholding. The left eye regions cropped from both edge images are shown in Fig. 6. A given point (red star) in camera1 edge image is searched in each edge pixel in the camera1 image taking epipolar constraint into account, and a point cloud is generated. Afterwards, 3D point clouds for camera1-projector and camera2-projector pairs are computed separately using ray-to-plane projection and triangulation. The point clouds are superposed using iterative closest point algorithm [7] to obtain final 3D point cloud. Finally, the point cloud is converted to triangular meshes and then 3D model is reconstructed.



(a) (b) (c)

Figure 5. a) Cropped gray-coded images for camera corresponding to 4th (upper) and 7th (bottom) bit levels, b) negative images, c) images obtained after subtraction

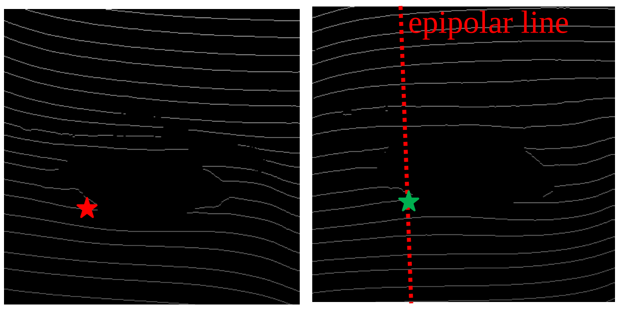


Figure 6. Cropped left eye region from both edge images and the corresponding pixels found using epipolar geometry (red and green stars).

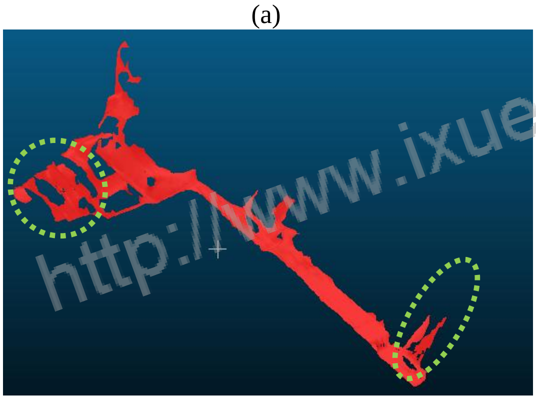
3 Experimental Results

In this section we evaluated visual performance of the 3D reconstruction system with two kind of materials: a face model which include smooth surface and an industrial machine part that contains abrupt surface changes. Fig. 7 shows an image of the industrial machine part and 3D models extracted by gray-coded structured light technique to represent the effect of the occlusion in 3D model reconstruction and how this effect can be handled utilizing multiple camera instead of a single camera.

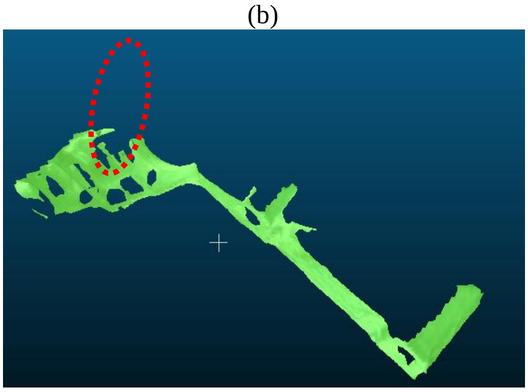
ISBN: 978-1-61804-329-0

Note that, missed model parts are marked with dotted ellipses in Fig. 7b and c. Superposing these two models gives better shape compared to the models obtained by single camera as given in Fig. 7d. For the face model, stereo vision based triangulation produce the 3D model given in Fig. 8a. Fig. 8b shows the 3D models extracted by gray-coded structured light technique for both cameras. Final 3D model reconstructed superposing 3D models given in Fig 8a and b is shown in Fig. 8c from different viewpoints. It is seen from Fig. 8c that the occluded areas are filled after merging into a complete 3D model.

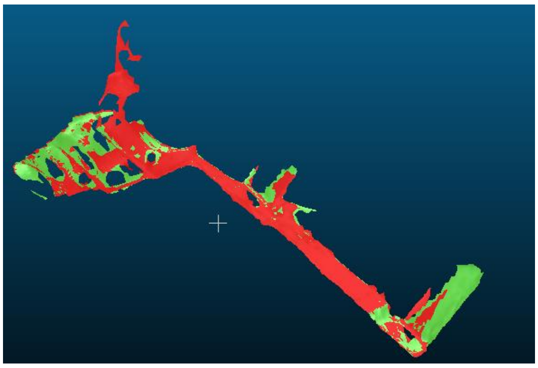
(a)



(b)

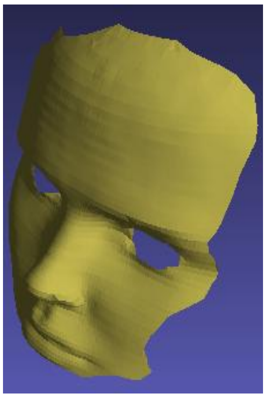


(c)

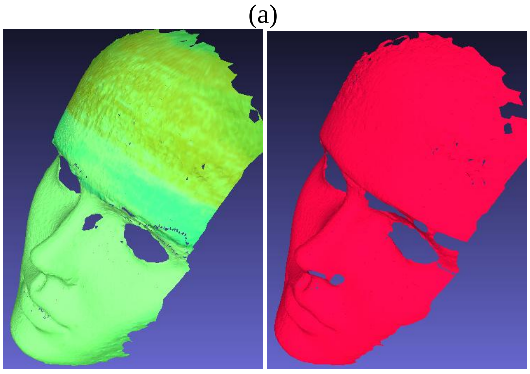


(d)

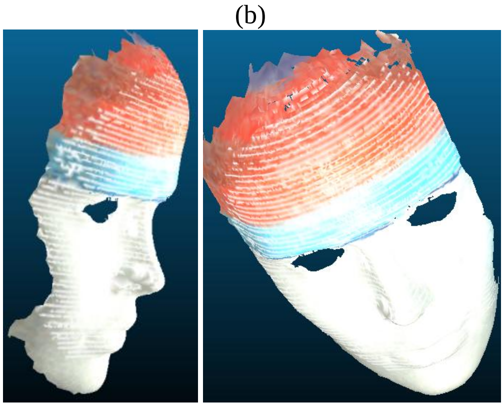
Figure 7. a) An image of the industrial machine part, b-c) 3D models extracted by gray-coded structured light technique, d) superposed model



(a)



(b)



(c)

Figure 7. 3D model(s) produced by: a) stereo vision based triangulation, b) gray-coded structured light technique, and c) final model

4 Conclusion

In this work, a 3D scanning system utilizing a setup which has two CMOS cameras and a DLP projector making use of structured light and stereo vision has been demonstrated. Three point clouds, one of them has been extracted by the use of stereovision and epipolar constraint and additional two of them has been extracted via gray coded structured light technique. These three point clouds have been superposed using iterative closest point algorithm. Then these point clouds have been merged into a complete point cloud, uniformly sampled, converted to triangular meshes and the final 3D model is reconstructed. It can be seen from the experimental results part that the proposed 3D modeling system is able to cope with occlusions arose from a single camera-projector setup and stereo correspondence and objects models are reconstructed successfully.

References:

[1] S. Se, O. Jasiobedzki, Stereo-vision based 3D modeling for unmanned ground vehicles, *International Journal ofIntelligent Control and Systems*, Vol.13, No.1, 2008, pp. 46-57.

[2] M. Pashaei, S.M. Mousavi, Implementation of a low cost structured light scanner, *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, Vol. XL-5/W2, 2013, pp.477-482.

[3] C.S. Chen, Y.P. Hung, C.C. Chiang, J.L. Wu, Range data acquisition using color structured lighting and stereo vision, *Image Vision Computing*, Vol. 15, No. 6, 1997, pp. 445-456.

[4] L. Zhang, B. Curles, S.M. Seitz, Rapid shape acquisition using color structured light and multi-pass dynamic programming, *1st IEEE International Symposium on 3D Data Processing, Visualization, and Transmission*, 2002, pp. 705-708.

[5] Z. Zhang, A Flexible New Technique for Camera Calibration, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 11, 2000, pp. 1330–1334.

[6] R. Hardley, A. Zisserman, Multiple View Geometry in Computer Vision, *Cambridge University Press (Second Edition)*, 2004.

[7] Z. Zhang, Iterative Point Matching for Registration of Free-Form Curves and Surfaces, *International Journal of Computer Vision*, Vol. 13, No. 2, 1994, pp. 119-152.

ISBN: 978-1-61804-329-0