Optics and Lasers in Engineering 106 (2018) 119–131

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

journal homepage: www.elsevier.com/locate/optlaseng

High-speed 3D shape measurement with structured light methods: A review

Song Zhang

School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA

ARTICLE INFO

Keywords:
High speed
3D Shape measurement
3D Imaging
Structured light
High accuracy

ABSTRACT

High-speed 3D shape measurement (or imaging) has seen tremendous growths over the past decades, especially the past few years due to the improved speed of computing devices and reduced costs of hardware components. 3D shape measurement technologies have started penetrating more into our daily lives than ever before with the recent release of iPhone X that has an built-in 3D sensor for Face ID, along with prior commercial success of inexpensive commercial sensors (e.g., Microsoft Kinect). This paper overviews the primary state-of-the-art 3D shape measurement techniques based on structured light methods, especially those that could achieve high measurement speed and accuracy. The fundamental principles behind those technologies will be elucidated, experimental results will be presented to demonstrate capabilities and/or limitations for those popular techniques, and finally present our perspectives on those remaining challenges to be conquered to make advanced 3D shape measurement techniques ubiquitous.

1. Introduction

With recent advancements on personal computers, mobile devices, and cloud computing, high-speed and high-accuracy 3D shape measurement (or imaging) techniques have been increasingly sought by scientists in fields such as biomedical engineering and computer science, by engineers from various industries including the manufacturing and entertainment, and even by ordinary people with different technical backgrounds. The commercial success of consumer level real-time 3D imaging technologies including Microsoft Kinect, Intel RealSense, and recently Apple iPhone X propels the application developments and simultaneously drives the needs for better 3D imaging technologies.

High-speed 3D imaging technologies can be classified into two major categories: the passive and the active methods. The passive techniques use no active illumination for 3D reconstruction with stereo vision [1,2] being one of the most popular methods. The stereo-vision system captures images from at least two different perspectives, and analyzes the images to find corresponding points from those images for 3D coordinate calculation based on triangulation. The stereo vision method is very simple since only cameras are used, and can also be very fast, as fast as the camera can capture images. Hinging on detecting the corresponding pairs from different images, the measurement accuracy of this method varies depending upon the object to be measured, and could be very low if an object does not present rich surface texture. Furthermore, it is difficult for such a technique to achieve camera pixel spatial resolution due to the use of various image correlation methods for stereo correspondence determination.

The active methods, in contrast, actively illuminate the object to facilitate 3D reconstruction. As one of the extensively adopted active methods, the time-of-flight (TOF) technique uses an active emitter to modulate the light in time domain, an optical sensor collects the light scattered back by the object, recovers depth information by calculating the time delay from the signal leaves the device and the signal returns to the device [3]. Unlike the stereo-vision method, the TOF method does not require triangulation for 3D reconstruction, and thus the entire system can be very compact, making it applicable for mobile applications. However, because light travels very quickly, the achievable depth resolution is typically not high for short range measurement. Kinect II employs the TOF technique for real-time 3D imaging, and successfully finds its applications in human computer interaction where neither accuracy or spatial resolution requirement is high.

The structured light technique belongs to one of the active methods and utilizes a projection device to actively project structured patterns. The structured light system is similar to a stereo system with the difference of replacing one camera with a projector. The projected structured patterns carry encoded information to resolve the fundamentally difficult correspondence problem of the stereo vision technique. Numerous codification methods have been developed with some being discussed by Salvi et al. [4]. Due to the flexibility and versatility of structured light methods, 3D shape measurement using structured light methods has been a vibrant field with increased interest in development and employment. In particular, high-speed and high-accuracy 3D shape measurement techniques become more and more important with new applications found almost every day. Therefore, this paper will primarily

E-mail address: szhang15@purdue.edu

https://doi.org/10.1016/j.optlaseng.2018.02.017
Received 16 December 2017; Received in revised form 21 February 2018; Accepted 23 February 2018

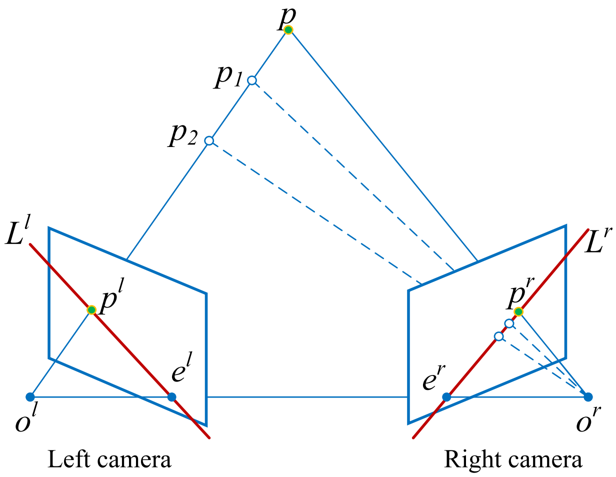

Fig. 1. Epipolar geometry for a standard stereo vision system.

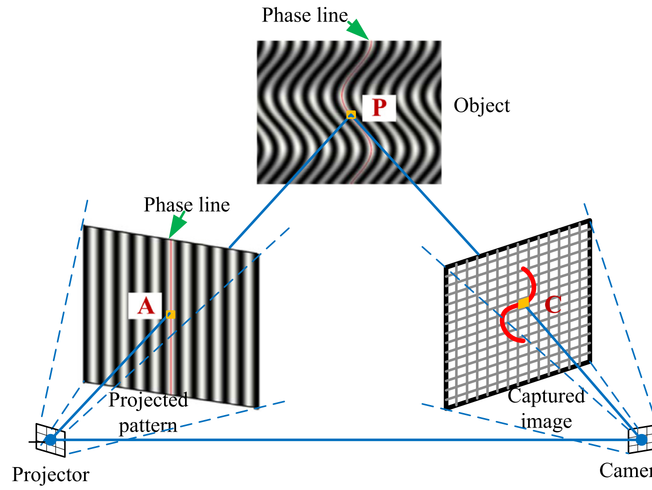

Fig. 2. Schematic diagram of a typical structured light system.

focus on structured light methods that could achieve high measurement speed and high measurement accuracy. Specifically, this paper elucidates the principles of various coding methods, discusses their advantages or shortcomings, and presents some experimental data obtained by structured light methods.

It should be noted that, due to our limited knowledge and the page constraints, this paper, by no means, intends to elaborate all state-of-the-art 3D shape measurement/imaging technologies, albeit we endeavor to cover as many as existing techniques as possible. The readers are encouraged to refer to some other review papers on 3D shape measurements such as [5–10].

The paper is organized as follows: Section 2 presents the basics of structured light technique including system calibration. Section 3 introduces relevant principles of structure encoding methods along with some experimental results to demonstrate their performances; Section 4 discusses our perspectives on the challenges in this field; and Section 5 summarizes this paper.

2. Basics of structured light techniques

This section briefly explains the basic principles behind structured light techniques that use triangulation for 3D reconstruction, the epipolar geometry that could simplify structured pattern design strategies, and then the calibration methods that estimate physical properties of the structured light system.

2.1. Basics of epipolar geometry

Structured light techniques originated from the conventional stereo vision method that recovers 3D information by imitating human perception system. The stereo-vision system captures two images from different perspectives, as illustrated in Fig. 1. For a given point p in a 3D space, pl and p^r are the imaging points on two 2D imaging planes. If the corresponding pair (i.e., p^l and p^r) can be found by analyzing those two images, (x, y, z) coordinates of point p can be calculated using triangulation assuming the optical parameters (e.g., focal length for camera lens, principal point) and geometric parameters (e.g., transformation from one camera to the other) are known. The parameters required for 3D reconstructions can be calibrated, which will be detailed in Section 2.3. In computer vision, *epipolar geometry* is developed to increase the robustness and simplify the correspondence determination [11,12]. Epipolar geometry essentially constrains the stereo searching by using the geometric constraints of the stereo vision system. The focal points of the lenses o^l and o^r and the object point p forms a plane called *epipolar plane*, and the intersection between the epipolar plane and a imaging plane is a line that is called *epipolar line.* L^l and L^r shown on Fig. 1 is the epipolar line on left and right camera image, respectively. Since the true corresponding point pairs must lie on the epipolar line, the original 2D corresponding searching problem becomes 1D, making it more efficient and robust. The intersection point between the line $\overline{o^l o^r}$ and the camera image plane is called epipole. e^l and e^r, shown in Fig. 1, is the epipole for the left and right camera, respectively. For a given point on the left camera p^l, the epipolar plane can be formed by combining the point p^l with two other known points, the focal point o^l and the epipole e^l, and thus the epipolar lines L^r and L^l can be mathematically calculated once the system is calibrated.

To further improve the correspondence searching speed, stereo images are rectified such that the corresponding point only occurs on the same row; and the process of rectifying stereo images is often referred as *image rectification.* Image rectification essentially translates and rotates the original images to align those epipolar lines (e.g., make L^l and L^r on the same line) using the stereo-vision system calibration data. Numerous global or semi-global stereo-matching algorithms [13–19] have developed to find the corresponding points using the epipolar geometry with some optimization strategies. The stereo-matching algorithm typically generates a disparity map that stores the pixel shift of a corresponding pairs from the left camera image to the right camera image. The disparity map is then used to reconstruct (x, y, z) coordinates for each point based on the calibrated system parameters. Since only two cameras are used, the stereo-vision technique has obvious advantages: the simplicity of hardware configuration and straightforward calibration for the system [20]. However, heavily relying on natural texture for correspondence establishments, the accuracy of stereo-vision techniques varies from one object to another; and the measurement accuracy is not high if an object has no obvious distinctive features.

2.2. Basics of structured light technique

The structured light technique fundamentally eliminates the stereo-vision problem by replacing one of the cameras of the stereo-vision system with a projector and actively projecting known feature points [4]. Fig. 2 shows the schematic diagram of a 3D shape measurement system based on one type of structured light technique. The projector shines structured patterns onto the object whose geometry distorts structured patterns. A camera captures the distorted structured images from another perspective. In such a system, the correspondence is established by analyzing the distortion of captured structured images with known features (e.g., phase line) projected by the projector. Once the system is calibrated and the correspondence is known, (x, y, z) coordinates can be reconstructed, using a method similar to that used by stereo vision techniques.