Optics and Lasers in Engineering 75 (2015) 88–94

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

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

Robust and efficient decoding scheme for line structured light

Yunfei Long [a], Shuaijun Wang [a], Wei Wu [b], Kai Liu [a,*]

[a] School of Electrical Engineering and Information, Sichuan University, China
[b] College of Electronics and Information Engineering, Sichuan University, China

ARTICLE INFO

Article history:
Received 1 December 2014
Received in revised form
20 May 2015
Accepted 25 June 2015
Available online 25 July 2015

Keywords:
3-D surface measurement
Line structured light
Peak detection

ABSTRACT

Accurately and robustly extracting the peak positions of line stripes has been the focus of 3-D reconstruction by using line structured light. However, typical peak detection algorithms need to employ neighborhood techniques, and, thus, their accuracies are susceptible to line stripe ambiguities. In this paper, we present a pixel-wise decoding scheme for line-scanning patterns. The proposed method can achieve: (1) robustly decoding without neighboring operations and (2) efficient implementation without searching maximum intensities or computing derivatives. Compared with typical spatial peak-detection algorithms, our experimental results demonstrate that, with holding the same scanning accuracy, the proposed method achieves much higher computational efficiency and more robustly decoding.

1. Introduction

Structured light illumination (SLI) is widely used for 3-D reconstruction [1]. Among various SLI techniques, line structured light (including laser scanning) [2] is one of the most applied methods for its accuracy and simplicity. Typically, the line structured light technique projects a sequence of line stripes, which move across the scanned objects, and uses a camera to record the illuminated scene which is distorted by the shape of the objects. Then, by processing the captured images, a decoding algorithm retrieves camera–projector correspondences. Finally, based on the camera–projector correspondences, 3-D coordinates can be computed through triangulation [1].

Determining the peak position of line stripes in captured images is a crucial procedure for decoding line structured light and significantly affects measurement accuracy. Typical peak-detection methods include Gaussian approximation (GA) [3], Centre of mass (CM) [3], Linear approximation (LA) [3], Blais and Rioux detector (BR) [3], and Steger's method [4,5]. A comprehensive introduction of the peak detectors is given by Naidu and Fisher in [3]. Most line structured light systems perform peak detection in the spatial domain. In other words, they treat each column or row of an image as independent signals and locate the peak position within each column or row. Peak detectors including GA [3], CM [3], LA [3], and BR [3] assume a single peak in each image column or row from which a maximum intensity and its neighboring intensities are searched in order to estimate a sub-pixel peak position. While Steger's method [4,5], by utilizing the differential properties of curvilinear structures, is able to detect multiple peaks appearing in an image column or row. Over the years, various approaches have been proposed to improve the robustness of peak detection under different measurement conditions, such as different optical properties of surfaces [6], variable ambient luminance [7] or high noise levels [6–8].

This paper solves another problem associated with line peak detection, namely, line stripe ambiguities. For example, in the case of vertical scanning setup, the line stripe moves vertically across recorded images, and the line stripe ambiguities appear when there is more than one line peak in certain columns of a captured image. Typically, there are two types of line stripe ambiguities. The first type, caused by the simultaneous projection of multiple line stripes, can be straightforwardly solved by projecting additional structured light patterns, *e.g.* Gray code patterns [8,9]. The second type, as shown in Fig. 1, is caused by the geometry structure of scanned objects as well as the placement of scanning devices, and arises under even single line stripe projection. In the presence of the second type of ambiguities, typical spatial peak detectors like GA [3], CM [3], LA [3], and BR [3] are not applicable because they can only handle a single peak in each column. Steger's method [4,5] can address such ambiguities by successfully extracting multiple peak positions; however, it is time-consuming since the derivatives of large numbers of line-scanning images have to be computed. Thus, how to solve the ambiguities with high efficiency is still a problem. In this paper, we focus on dealing with the second type of line stripe ambiguities. Hence, if not specified, the phrase "line stripe ambiguities" indicates the second type.

In correspondence to spatial decoding, temporal decoding represents another category of decoding strategies, which are

* Corresponding author.
E-mail address: kailiu@scu.edu.cn (K. Liu).

http://dx.doi.org/10.1016/j.optlaseng.2015.06.008
0143-8166/

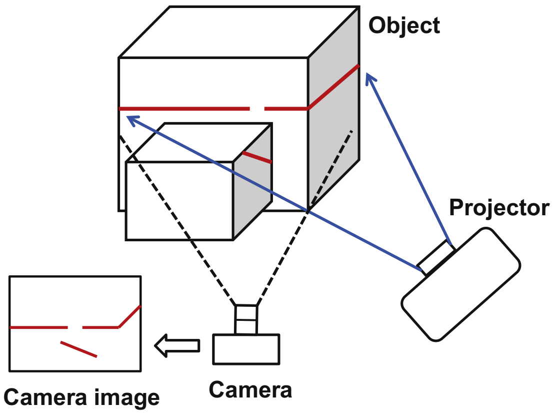

Fig. 1. Line stripe ambiguities.

typically adopted in some SLI techniques, *e.g.* phase measuring profilometry (PMP) [10], triangular-pattern phase shifting method (TPP) [11] and Gray coding scheme [12]. The temporal decoding is a pixel-wise operation, *i.e.* considering the intensities of each pixel over a temporal sequence of images as independent signals. From the temporal sequence of a camera pixel, sub-pixel projector coordinates corresponding to that camera pixel can be computed. Some laser scanning systems operating in temporal domain have also been reported [13–15]. As the line stripe moves across the scene, they find the time when peak intensity is seen at each pixel to construct correspondences. Curless and Levoy [16] further exploit the advantages of temporal peak detection to solve some problems associated with laser scanning, including reflectance discontinuity and shape discontinuity. Bouguet and Perona [17] build a simple 3-D scanning system which casts a moving shadow on objects and adopts both spatial and temporal strategies to locate the shadow. However, the benefits of temporal decoding in dealing with line stripe ambiguities have not been explored.

In this paper, in order to solve the issue of ambiguities for line stripe scanning with high efficiency, we propose a temporal centre of mass (TCM) algorithm to non-ambiguously extract line stripe. TCM takes advantage of the fact that although a spatial signal (an image column or row) may have more than one peak, the temporal signal at a camera pixel only has a single peak, which can be non-ambiguously and straightforwardly located with high efficiency. Note that, although the strategy of temporal decoding has been adopted in laser scanning to address measurement distortions caused by reflectance discontinuity and shape discontinuity, to our best knowledge, there are no publications reporting applying temporal peak detection to solve ambiguities in line structured light. Compared with typical spatial peak-detection algorithms, our experimental results demonstrate that, by holding the same scanning accuracy, the proposed method achieves much higher computational efficiency and more robustly decoding.

2. Method

Fig. 2 shows the schematic diagram of our 3-D measurement procedures. For our optical 3D scanning system consisting of a projector and a camera, the line structured light patterns are expressed as

$$I_n^p(x^p, y^p) = M^p \delta(n - y^p), \tag{1}$$

where (x^p, y^p) is the coordinates of patterns in the projector; the term I_n^p is the intensity at (x^p, y^p); $\delta(\cdot)$ is the unit-impulse function which satisfies $\delta(x) = 1$ for $x=0$ and $\delta(x) = 0$ for $x \neq 0$; constant M^p is the modulation of the projected line stripes, and equals to 255 for an 8-bit depth projector; and the integer number n, ranging from 0 to $H^p - 1$ (H^p is the vertical resolution of the projector), represents line shift index.

Apart from line patterns described above, two additional patterns, *i.e.* a pure black pattern

$$I_B^p(x^p, y^p) = 0 \tag{2}$$

and a pure white pattern

$$I_W^p(x^p, y^p) = M^p, \tag{3}$$

are needed to normalize captured images.

After scanning an object with the patterns described in Eqs. (1), (2) and (3), respectively, we obtain images as follows.

For the images patterned by Eq. (1), we use Gaussian distribution [18] to model them as

$$I_n^c(x^c, y^c) = M^c(x^c, y^c) \exp\left[-\frac{(n - y^p)^2}{2\sigma_w^2}\right] + A^c(x^c, y^c), \tag{4}$$

where (x^c, y^c) is the coordinates in the camera, the term I_n^c is the intensity at (x^c, y^c), the term M^c represents the modulation at (x^c, y^c), the term σ_w is the standard deviation width [18] of the line stripe, and A^c represents the intensity of ambient light at (x^c, y^c).

For the images patterned by Eqs. (2) and (3), they are directly modeled as

$$I_B^c(x^c, y^c) = A^c(x^c, y^c) \tag{5}$$

and

$$I_W^c(x^c, y^c) = M^c(x^c, y^c) + A^c(x^c, y^c), \tag{6}$$

respectively. In order to eliminate the effects of surface reflectivity and ambient light, we firstly normalize Eq. (4) by

$$F_n(x^c, y^c) = \frac{I_n^c(x^c, y^c) - I_B^c(x^c, y^c)}{I_W^c(x^c, y^c) - I_B^c(x^c, y^c)}, \tag{7}$$

where F_n is within the range of $[0, 1]$.

In order to theoretically compare the proposed TCM with conventional CM, here, we briefly introduce the procedures of the conventional CM [7]. For each normalized image F_n, CM extracts the peak position in column-wise, and the peak position in column x^c can be calculated as [7]

$$C_n(x^c) = \frac{\sum_{Y_n^{max}(x^c) - W/2}^{Y_n^{max}(x^c) + W/2} F_n(x^c, y^c) y^c}{\sum_{Y_n^{max}(x^c) - W/2}^{Y_n^{max}(x^c) + W/2} F_n(x^c, y^c)}, \tag{8}$$

where $C_n(x^c)$ represents the Y coordinate of a peak position extracted from column x^c, the term $Y_n^{max}(x^c)$ is the maximum intensity found in column x^c, and W denotes the window size around $Y_n^{max}(x^c)$ and determines the number of pixels involving CM computation. From C_n, which is the vertical camera coordinates corresponding to projector vertical coordinate n, the camera–projector mapping can be established. However, when line stripe ambiguities occur, there will be at least two local intensity maxima in certain columns, so $Y_n^{max}(x^c)$, obtained by searching the maximum intensity in column x^c, does not result in correct correspondences.

Differing from the standard CM which spatially detects peaks in each captured image, the proposed TCM algorithm determines peak intensity along time axis for each camera pixel. According to Eq. (4), the intensities of a temporal sequence at (x^c, y^c), *i.e.* $I_n^c(x^c, y^c)$ for $n = 0, 1, \ldots, H^p - 1$, should be Gaussian-shaped. Thus the index n of a detected peak is the projector coordinate y^p corresponding to (x^c, y^c). In order to avoid errors caused by random noise, in peak extraction, we only use the I_n^c with significant values for TCM computation. Thus, we determine the