Optics and Lasers in Engineering 110 (2018) 54–62

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

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

An improved spatiotemporal correlation method for high-accuracy random speckle 3D reconstruction

Qijian Tang, Chang Liu, Zewei Cai, Huihe Zhao, Xiaoli Liu*, Xiang Peng*

Key Laboratory of Optoelectronic Devices and Systems of Ministry of Education and Guangdong Province, College of Optoelectronic Engineering, Shenzhen University, Shenzhen 518060, China

ARTICLE INFO

Keywords:
3D shape measurement
Digital speckles
Spatiotemporal correlation
Subpixel interpolation

ABSTRACT

The single-shot random speckle pattern method can be used for dynamic scene 3D reconstruction. However, the homonymous points search process based on spatial correlation is sensitive to changes in factors such as illumination, perspective, and curvature, which results in low reconstruction accuracy. This paper proposes an improved random speckle 3D reconstruction method with enhanced reconstruction accuracy. To accurately search homonymous points, the proposed method employs a novel spatiotemporal correlation model simultaneously with a subpixel interpolation strategy. Analysis conducted of the reconstruction accuracy associated with correlation region size and the number of patterns to be projected based on the model indicate that using a few (e.g., three) speckle patterns with an appropriate correlation size produces highly accurate results. Experimental results further verify that the proposed method is suitable for high-accuracy dynamic 3D shape measurement.

1. Introduction

Structured illumination-based 3D measurement is an active research area for applications in industrial inspection, healthcare, cultural education, and entertainment [1–6]. Among the methods employed, random speckle pattern is frequently used to identify homonymous image points corresponding to the same object point by means of spatial and temporal correlation. The spatial correlation method requires only one speckle pattern to calculate the maximal correlation coefficient between binocular capture images, which is suitable for dynamic 3D shape measurement. However, the spatial correlation relying on grayscale information statistics is facile and is influenced by camera perspective changes as well as strong curved surfaces [7,8], resulting in low reconstruction accuracy. Consequently, the temporal correlation method, with a sequence of time-varying speckle patterns, has been proposed to achieve more accurate 3D reconstruction. In general, more than 20 speckle patterns are required [9]. As a result, the temporal correlation method is difficult to adapt to dynamic scene 3D reconstruction.

To overcome the contradiction between accuracy and efficiency in random speckle 3D reconstruction, the spatiotemporal correlation method, which combines the advantages of spatial and temporal approaches, was proposed [10]. The method reduces the statistical measurement error with less projected patterns, which provides the possibility for high precision and dynamic scene reconstruction. Große and Kowarschik [11] combined both methods to verify homologous points. In their approach, when a pair of homologous points is found by the temporal correlation of the gray value-sequences, spatial correlation is used to determine its acceptance or rejection. Zhang et al. [12] matched a pair of video streams to reconstruct 3D scenes with a disparity function. They assumed the temporal disparity variations between frames to be locally linear and realized good reconstruction of smooth dynamic scenes using the stereo method. Harendt et al. [13] proposed a weighted spatiotemporal correlation method to reconstruct static and moving objects, where the weights are used to adjust the spatial and temporal extent of the matched regions. However, the weights are first estimated based on disparity variations between adjacent frames, which results in expensive computations.

Most of the existing studies on spatiotemporal correlation focus on its applications, without corresponding accuracy analysis. In this paper, we propose a novel spatiotemporal correlation model, and also analyze the influences of correlation region size and images required on the reconstruction accuracy based on this model in detail. The results of test conducted in which the proposed method was used to reconstruct static and dynamic scenes indicate that it is suitable for high-accuracy 3D measurement of strong curvature surfaces and dynamic scenes.

2. Principle

2.1. Binocular stereo vision

For a common binocular measurement system, as shown in Fig. 1, O_l and O_r are the optical centers of the left and right cameras. The coordinate of point M is X_w in the world coordinate system, and is X_l in

* Corresponding author.
E-mail addresses: lxl@szu.edu.cn (X. Liu), xpeng@szu.edu.cn (X. Peng).

https://doi.org/10.1016/j.optlaseng.2018.05.007
Received 15 February 2018; Received in revised form 6 May 2018; Accepted 9 May 2018

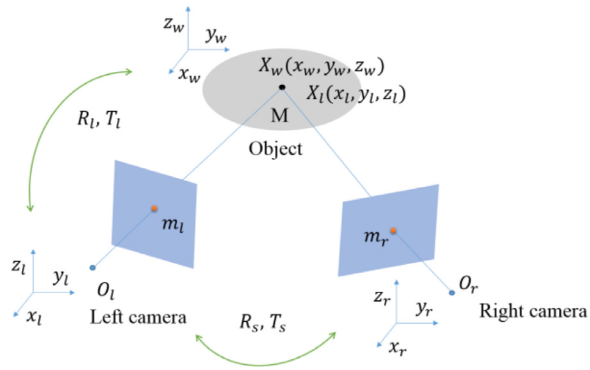

Fig. 1. Binocular stereo vision principle.

the left camera reference frame. Its image projections in the left and right image coordinate systems are denoted by m_l and m_r, respectively. $[R_s|T_s]$ is the position and orientation transformation matrix from the left camera to the right camera, and $[R_l|T_l]$ is the transformation from the world frame to the left camera. Generally, to simplify the model, the left camera reference frame is often chosen as the world frame, as a result, $R_l = I$, $T_l = 0$, where I is an identity matrix.

According to the principle of line of sight intersection, the coordinate of point M can be calculated by:

$$\begin{cases} s_l\hat{m}_l = K_l[I|0]X_l \\ s_r\hat{m}_r = K_r[R_s|T_s]X_l \\ K_l = \begin{bmatrix} f_{xl} & \alpha_l & u_{0l} \\ 0 & f_{yl} & v_{0l} \\ 0 & 0 & 1 \end{bmatrix} \\ K_r = \begin{bmatrix} f_{xr} & \alpha_r & u_{0r} \\ 0 & f_{yr} & v_{0r} \\ 0 & 0 & 1 \end{bmatrix} \end{cases} \tag{1}$$

where K_i (i denotes l and r) is the intrinsic camera parameter matrix; f_{xi} and f_{yi} are the equivalent focal length; (u_{0i}, v_{0i}) is the principal point coordinate; α_i is the skew coefficient defined by the angle between the u and v pixel axes; $\hat{m}_i$ is image coordinate after lens distortion correction based on Brown-Conrady model; s_i is the scalar factor. These parameters are easily acquired in the system calibration process.

In the process of 3D reconstruction, the fundamental technique is to find the corresponding image points in the left and right cameras belonging to the same object point. As shown in Fig. 2, a projector is often used to increase and improve the expressiveness of features. In the digital speckle correlation technology, the projector projects random speckle images onto the object, and two cameras capture two image sequences. Two correlation methods are always employed, spatial and temporal correlation methods. The former chooses the same size region in the same speckle illumination from two perspectives to calculate the correlation coefficient, whereas the temporal correlation method calculates the correlation coefficient using the gray value variations of each pixel.

2.2. Spatial correlation

Taking the point M in the left image in Fig. 3 as an example, the correlation region size is $(2w_m+1)\times(2w_m+1)$ pixels. The correlation objective is to find the homologous point in the right image for point M. The first step in this process is to shift the same size region onto the right image until the correlation weight achieves the global maximum.

To evaluate the correlation degree between the left and right regions, a correlation criterion has to be established beforehand. At present, although correlation coefficient functions have many definitions and deformations, Pearson's product-moment correlation is more robust and easier to handle [9,14–16]. For two one-dimensional signals, f_L and f_R, the correlation coefficient ω is calculated using Eq. (2):

$$\omega = \frac{cov(f_L, f_R)}{\sigma_L\sigma_R} = \frac{\sum_{k=1}^{NS}\left(f_{Lk}-\overline{f_{LS}}\right)\left(f_{Rk}-\overline{f_{RS}}\right)}{\sqrt{\sum_{k=1}^{NS}\left(f_{Lk}-\overline{f_{LS}}\right)^2}\sqrt{\sum_{k=1}^{NS}\left(f_{Rk}-\overline{f_{RS}}\right)^2}} \tag{2}$$

where NS is the size of the region f_{LS} and f_{RS} in the operation, and $\bar{f}$ is the mean value.

The equation has a value between +1 and −1, where 1 represents f_L and f_R having total positive linear correlation, 0 represents complete independence, and −1 represents total negative linear correlation. Although Eq. (2) is a one-dimensional correlation model, it is still applicable to two-dimensional image correlation. As shown in Fig. 4, a two-dimensional image region can be rearranged into a one-dimensional array.

Ideally, for the same object and system, $f_L = f_R$. However, as a result of camera perspective and curvature of object surface changes, $f_R = f_L + f_\delta$, where f_δ represents the difference. If f_L and f_R is sufficiently correlative, Eq. (2) can be rewritten as follows:

$$\omega = \frac{cov(f_L, f_R)}{\sigma_L\sigma_R} \approx \frac{{\sigma_L}^2}{\sqrt{{\sigma_L}^4 + {\sigma_L}^2{\sigma_\delta}^2}} \tag{3}$$

where σ represents the standard deviation and σ_δ is the standard deviation of the signal bias f_δ. When the perspective or curvature changes significantly, it results in notable σ_δ, and the coefficient ω is low. Further, a large NS will often increase the statistics uncertainty, and also causes a lower ω.

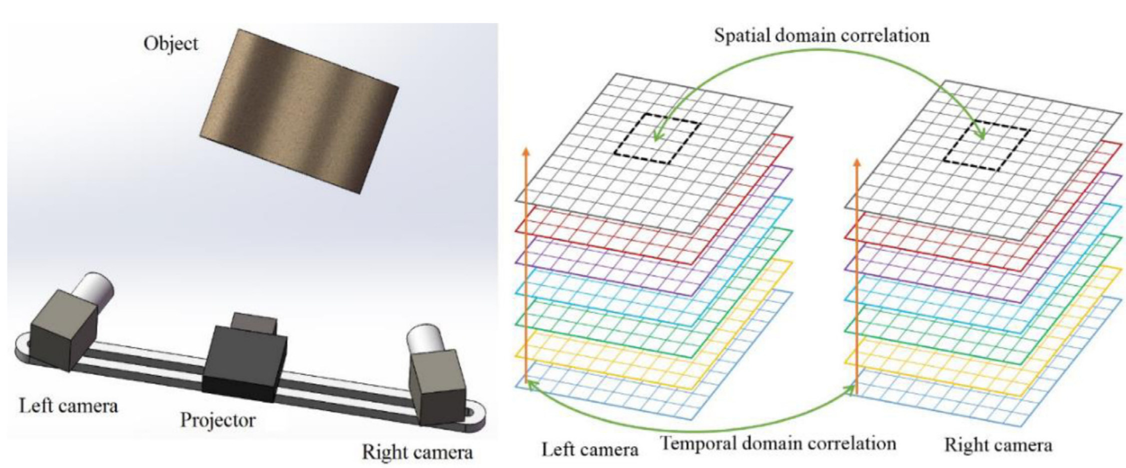

Fig. 2. Binocular measurement system based on random speckles.