Height Estimation of Target Objects Based on Structured Light

Wei Liu and Yongsheng Zhao

Department of Modern Education Technology, Ludong University, Yantai, China

Correspondence should be addressed to Wei Liu; ldulw@sina.com

Received 20 June 2018; Accepted 25 September 2018; Published 1 November 2018

Guest Editor: Shengping Zhang

Copyright © 2018 Wei Liu and Yongsheng Zhao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The height estimation of the target object is an important research direction in the field of computer vision. The three-dimensional reconstruction of structured light has the characteristics of high precision, noncontact, and simple structure and is widely used in military simulation and cultural heritage protection. In this paper, the height of the target object is estimated by using the word structure light. According to the height dictionary, the height under the offset is estimated by the movement of the structured light to the object. In addition, by effectively preprocessing the captured structured light images, such as expansion, seeking skeleton, and other operations, the flexibility of estimating the height of different objects by structured light is increased, and the height of the target object can be estimated more accurately.

1. Introduction

In recent years, with the development of science and technology, three-dimensional reconstruction technology as an important part of machine vision has attracted more and more attention, especially in industrial product design and cultural heritage protection. However, based on the three-dimensional reconstruction of the surface of the structured light, it is possible to reconstruct the surface of the object by laser scanning without touching the object, which can greatly protect the original culture from damage in the cultural heritage. This can make a great contribution to the excavation of ancient excellent culture and the spread of Chinese civilization. Therefore, the three-dimensional reconstruction based on structured light has important practical significance for the protection of cultural heritage and the design of industrial products [1–3]. At present, in the three-dimensional reconstruction of the main use of word-structured optical scanning method and three-dimensional reconstruction of structured light technology as a noncontact active measurement technology, with low cost, high precision, vision, real-time, anti-interference ability, and so on, these characteristics will inevitably make the next few years of this reconstruction will have a better development prospects [4, 5].

3D surface reconstruction is to rebuild the actual shape of the real life of the object, which has become an important topic in computer vision. And researchers from all over the world have made considerable achievements in this regard. The structure of three-dimensional reconstruction system for structured light mainly includes cameras and lasers; you can use the ordinary camera to complete the task of detection, but because of the different structure of the light, the experimental results will be affected. According to the laser projection of different ways can be divided into point, line, and multiline structure of light. Point structure light for the laser projector projection of a beam of light, measured the surface of the measured object a point; the camera can only get this photo of the three-dimensional coordinates of the information; the amount of information is too small; a word line structure light projector projects a light plane; the intersection of the light plane and the measurement object can draw a cross section information; the algorithm is easy to use; multirow structured light projects multiple light planes; the surface of the object forms multiple laser lines; pictures can give us multiple cross-section information, which is large amount of information; however, it is necessary to increase the matching of light bars, which greatly improves the difficulty and complexity of the algorithm and is still in the stage of experimental research [6–8].

At home and abroad for the 3D surface reconstruction conducted in-depth study, Horn [9] proposed the concept of SFS, which is a widely concerned three-dimensional shape
reconstruction of the important ideas. The main content of this idea is to reconstruct the three-dimensional shape of the surface of the object by identifying and analyzing the shape information of the direction of the light, the brightness, the surface shape of the object, and the grayscale variation of the reflection model. Ikeuchi and Horn [10] are used to solving the three-dimensional reconstruction by using the illuminance equation and the smoothing criterion as the constraint of reconstruction. So, the problem of 3D reconstruction is transformed into the minimization problem of solving function. In this case, Horn proposed another smoothing standard. The main content is a smooth surface, where the surface obeys the integrable constraint, because the algorithm is seeking to directly recover complex surface unit normal vectors, so that reconstruction cannot get the absolute height of the surface. Fuqiang Zhou [11] used to use the same way to achieve the cross-laser plane calibration. In the experiment, four edge feature points of the space disk are obtained and the radius of the disk is calculated by fitting the feature points. The absolute error of the radius is 0.059 mm. Harbin Institute of Technology Dongbin Zhao [12] and other scholars put forward a new monocular image restoration object surface height and gradient algorithm is an iterative calculation of the composite image, obtaining accurate surface height, and they also validate the feasibility of the algorithm for actual solder joint images. Ruiling Liu [13], for high light and shadow, put forward a four-light source vector selection algorithm; she compares the normal vector of different pixels recovery with the mirror reflection direction and chooses the nearest normal vector to restore the shape of the required vector, which avoids the error caused by the threshold elimination of high light and shadow in the traditional algorithm, and remove the high light and shadow constraints on the algorithm to expand the scope of application of the algorithm.

2. Based on the Gradient of Moving Objects Detection

In the process of 3D reconstruction of structured light, in order to reconstruct the 3D structure of the 2D image taken by the camera, the camera parameter must be calibrated and the geometric model of camera imaging should be built; that is, the camera’s internal and external parameters should be measured. Then, the correspondence between the image and the spatial point is constructed; that is, the laser plane equation is calibrated. This paper mainly used Zhenyou Zhang camera calibration method [14].

2.1. Camera Parameter Calibration. The camera model is very similar to the model used by Heikkila and Silven of the University of Oulu in Finland. We especially recommend their CVPR’97 paper: the function of the four-step camera calibration program with an implicit image correction [15].

In the camera model, the parameters are as follows:
- Focal length: stored in pixels in 2×1 vector \( f_c \).
- Main points: the coordinates of the primary point are stored in the 2×1 vector \( cc \).
- Skew factor: define the skew factor for the angle between the \( x \) and \( y \) axes in the scalar \( \alpha_x \).
- Distortion: the image distortion factor (radial and tangential distortion) is stored in the 5×1 vector \( k_c \).

Let \( p \) be the spatial point of the coordinate vector \( XX_c = [X_c; Y_c; Z_c] \) in the reference frame of the camera. And then the projection is performed on the image plane based on the intrinsic parameter \( (f_c, cc, \alpha_x, k_c) \).

Let \( x_n \) be normalized (pinhole) image projection:

\[
x_n = \begin{bmatrix} X_c \\ Y_c \\ Zc \\ Zc \end{bmatrix}
\]

Let \( r^2 = x^2 + y^2 \); after the lens is distorted, the new normalized point coordinates \( x_d \) are defined:

\[
x_d = \begin{bmatrix} x_d (1) \\ x_d (2) \end{bmatrix} = \left( 1 + k_c (1) x^2 + k_c (2) r^4 + k_c (5) r^6 \right) x_n + d_x
\]

where \( d_x \) is the tangential distortion vector:

\[
d_x = \begin{bmatrix} 2k_c (3) xy + k_c (4) \left( r^2 + 2x^2 \right) \\ k_c (3) \left( r^2 + 2x^2 \right) + 2k_c (4) xy \end{bmatrix}
\]

Therefore, \( k_c (5) \) contains the radial, tangential distortion coefficient [16]. It is worth noting that this distortion model was first introduced by Brown in 1966, called the “Plumb Bob” model (radial polynomial + “thin prism”). The tangential distortion is due to the incorrect alignment of the “eccentric” in the composite lens or other manufacturing defects of the lens assembly.

Once the distortion is applied, the final pixel coordinates of the P on the projection plane are \( x_{pix d} = [x_p; y_p]: \)

\[
x_p = f_c (1) (x_d (1) + \alpha_x x_d (2)) + cc (1)
y_p = f_c (2) x_d (2) + cc (2)
\]

Thus, the pixel coordinate vector \( x_{pix d} \) and the normalized (distorted) coordinate vector \( x_d \) are related to each other by a linear equation [17]:

\[
\begin{bmatrix} x_p \\ y_p \\ 1 \end{bmatrix} = KK \begin{bmatrix} x_d (1) \\ x_d (2) \\ 1 \end{bmatrix} \tag{5}
\]

where \( KK \) is called the camera matrix and is defined as follows:

\[
KK = \begin{bmatrix} f_c (1) & alpha_f c(1) & cc (1) \\ 0 & f_c (2) & cc (2) \\ 0 & 0 & 1 \end{bmatrix} \tag{6}
\]
2.2. External Condition Variable Setting. In reconstructing the surface height of the object, we are using triangles similar to the surface reconstruction of the object. In the case of similar judgments, the degree of use of the triangles is the same. But when the triangles are similar, we use the same degree of the angle of the two similar triangles [18]. Thus, in the world coordinate system with the intersection of the center of view and the center of the object, there exists a proportional relationship between the tangent values of the corners in the right two angles, so that the degree of the angle $\alpha$ and the distance $L$ between the optical center and the object need to be known. These two variables can change the position of the camera structure by artificial changing. Therefore, when setting up the camera and the structure of the light we need to measure the angle $\alpha$ and length $L$. And these two are invariants; that is to say in the whole process of shooting we must ensure that the two variables remain unchanged or reconstructed objects will be distorted. So, in the process of taking pictures we must ensure that the external condition variable setting.

3. Basic Principle of Three-Dimensional Reconstruction of Structured Light

In order to obtain the three-dimensional information of the object in the structured light measurement, the basic idea is to use the geometric information in the structured light image to help provide the geometric information in the scene [19]. According to the geometric relationship inside the camera, we can determine the structure of light and the geometric relationship between objects, thus rebuilding the surface of the object.

3.1. The Correspondence between Pixels and a World Coordinate Point. As shown in Figure 1, the angle between the structural smooth and the optical axis of the camera is $\alpha$, and the origin $O_w$ of the world coordinate system $O_p - X_wY_wZ_w$ is located at the intersection of the camera’s optical axis and the structured light plane. The $X_w$-axis and the $Y_w$-axis are parallel to the camera coordinate systems $X_c$ and $Y_c$, respectively, and $Z_w$ and $Z_c$ coincide but are opposite [20]. The distance between $O_w$ and $O_c$ is $l$. Thus, the world coordinate system and the camera coordinate system have the following relationship:

$$
X_c = X_w
$$

$$
Y_c = -Y_w
$$

$$
Z_c = l - Z_w
$$

$A'$ is the image of $A$ in the world coordinate system; the line of sight $OA'$ is

$$
\frac{X_w}{X} = \left(\frac{Y_w}{y}\right) = \frac{(l - Z_w)}{f}
$$

(8)

In the world coordinate system, the plane equation of structured light is

$$
X_w = Z_w \tan \alpha
$$

(9)

Thus, it can be learned that the correspondence between pixel points and world coordinate points is

$$
X_w = \frac{(f_x (u - u_0) l \tan \alpha)}{(f_x (u - u_0) + \tan \alpha)}
$$

where $\alpha$ is the angle between the camera and the laser pen. The solutions of (9) are

$$
X_w = \frac{(x \tan \alpha)}{(x + f \tan \alpha)}
$$

$$
Y_w = \frac{(-y \tan \alpha)}{(x + f \tan \alpha)}
$$

$$
Z_w = \frac{x_l}{(x + f \tan \alpha)}
$$

(10)

Because $O_p - uv$ is the Cartesian coordinate system defined on the digital image [21], $(u, v)$ is the coordinates of the pixels, and $u$ and $v$ represent the number of rows and columns of pixels in the image array, respectively. Establish the coordinate system $O_1 - x'y'$ expressing in physical units parallel to the $u$-axis and the $v$-axis, the origin is the camera optical axis and image. The plane is usually located in the center of the image, but in reality there will be a small offset; $O_p - x'y'$'s coordinates are recorded as $(u_0, v_0)$. The physical dimensions of each pixel in the $x$-axis and $y$-axis directions are $S_x$ and $S_y$; the coordinates of any one of the two coordinate systems are represented by a uniform coordinate and a matrix, with the following relationship:

$$
\begin{bmatrix}
1 & 0 & u_0 \\
0 & 1 & v_0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
X \\
y \\
1
\end{bmatrix} = \begin{bmatrix}
\frac{1}{S_x} & 0 & 0 \\
0 & \frac{1}{S_y} & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
u \\
v \\
1
\end{bmatrix}
$$

(11)

The inverse relationship is

$$
\begin{bmatrix}
u \\
v \\
1
\end{bmatrix} = \begin{bmatrix}
S_x & 0 & -u_0S_x \\
0 & S_y & -v_0S_y \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
X \\
y \\
1
\end{bmatrix}
$$

(12)

FIGURE 1: Dimensional reconstruction of structured light.
Figure 2: Schematic diagram of experimental laser photography.

\[
Y_w = \frac{(-f_x (v - v_0) l \tan \alpha)}{(f_x (u - u_0) + \tan \alpha)}
\]

\[
Z_w = \frac{(f_x (u - u_0) l)}{(f_x (u - u_0) + \tan \alpha)}
\]

(13)

3.2. Surface Height Calculation Principle. As shown in Figure 2, the corresponding relationship between the pixel and the point in the world coordinate system is shown in (13). In the experiment, we simplified the shooting method. The laser angle and the vertical direction remained unchanged at 30°, so it was easy to calculate [22, 23]. When the shooting platform has no objects, the laser light directly to the platform will not be offset, but when the object is placed on the platform, the laser light to the surface of the object will occur after a certain shift. As shown in Figure 1, \( r_0 \) is the reference laser line, and \( r \) is the laser line that is offset after adding the object. Since the angle is 30°, \( \tan \alpha = h/d = \sqrt{3}/3 \), so the relationship between the horizontal offset \( d \) of the laser and the height \( h_0 \) of the point \( A \) of the object is known [24]. It can also be seen from Figure 1 that if the distance \( L \) between the light and the object and \( \alpha \) changes, the reconstruction will change.

4. Denoising after Loading the Mask

Due to shooting methods and other reasons, there is a certain amount of noise in the loaded laser mask. Here to solve the two main noises, other light source interference and laser line breakage, the main method is to filter the connected domain to remove other light source interference, through the expansion of the skeleton to avoid laser line breakage.

4.1. Filter the Connected Domain to Remove Other Light Sources. Filtering the connected domain is to keep the connected domain in the image and remove those nonconnected pixels. Here is the use of bwareaopen function; this function is also called delete the minimum area function; you can set the minimum size of the connected domain, which has the default value of 8. In the experiment, this value is set to 2 in this paper. In the design of the function, after loading the laser mask, the image will be converted to a black and white image; the image matrix is shown as 0-1 matrix. However, due to other light sources, there are some interference points in the image (as shown in Figure 3).

These interference points do not exist in the form of communication, but in the form of pixels scattered in the image, so you can filter the connection domain and remove these interference points, and this operation will have a sharp effect on the laser itself. That is, around the laser line “burr” will be deleted, which will make the reconstruction results more smooth. The effect after screening is shown in Figure 4.

4.2. Expand the Skeleton to Obtain the Laser Line. Laser mask image screening connection domain processing will be loaded; the laser line itself will be interference. The biggest problem is that the laser line is broken. In view of this situation, first of all, we have carried out the expansion operation and first broken the laser line through the expansion of the connection; the effect is shown in Figure 5.

After the laser line is inflated; the laser light becomes thicker. Obviously, we cannot use this inflated image directly to a high degree of reconstruction. All we have to do is to get a
thin continuous laser line, and we cannot change the shape of the original laser line. So we took the skeleton operation. This operation will be the same as the original laser line shape of the laser line, and this laser line is a single row of pixels of the laser line, which is in line with our requirements. The effect is shown in Figure 6.

5. Perform a High Degree of Summation and Interpolation

The main content of this part is taking the main process after rebuilding a single laser height: superposition and interpolation. The superposition is mainly a comprehensive display of each reconstructed laser height. Interpolation is the linear interpolation of the resulting discrete data, making it appear continuously and smoothly.

5.1. Height Superimposed and Evenly Displayed. This paper is designed to reconstruct the height of the surface of the object by using a single word structure, but a laser can only reproduce the height of the laser line (Figure 7).

Therefore, if the we use word structure of light on the surface of the three-dimensional reconstruction, there are two ways. One is to take the image into video and then a frame of a video of the laser line in a high degree of reconstruction, which will get a relatively smooth surface of the object, but this method is more difficult to shoot, the data being many. The second is that the isometric image is highly reconstructed and then interpolated. This method is relatively simple. No matter what method is used, the final reconstruction is a section of the height matrix. Therefore, to sum up the height of each reconstruction, each height matrix in a world coordinate system is displayed (shown in Figure 8).

Because we need to ensure the laser line and the location of the mandrel and their angle in the shooting of the image, we can only be moving objects when we shoot an object. Only in this way can we ensure the same angle between the laser line plane and the camera object, in order to accurately rebuild the height of the object, that is, to ensure that the angle between XcOw and OcOw. Therefore, when shooting a number of laser lights to rebuild the height we can only move the object to shoot, but the image will be in the same position if we shoot laser line. Then, the reconstruction of the laser height will be superimposed. Therefore, it is necessary for man-made reconstruction of the laser height according to the distance when the object is moving evenly distributed.
such that each height line is shown to be scattered (shown in Figure 8).

5.2. Interpolate to Reconstruct a Smooth Surface. As the design uses a word structure of light on the surface of the three-dimensional reconstruction of the object, the word structure of light can only rebuild a laser line under a height. After the above superposition, we will get a lot of high degrees of reconstruction, but these are not continuous but a height line. In order to rebuild these lines into the surface, there are two ideas: one is to take a lot of height lines for superposition; the other is to take a limited height line for superposition and then interpolation. These two methods can get a smooth surface reconstruction of the object, but the former method of workload is too large; here we use the second method, the height of the superposition of the interpolation operation, and the superposition of the use of the griddata function and of the discrete height of the linear interpolation to get a smooth surface of the object and the effect is shown in Figure 9.

6. Experimental Results and Analysis

This chapter is mainly to reconstruct the experimental results according to the physical comparison and analyze the advantages and disadvantages of the reconstruction of the experimental results.

6.1. Comparison of Physical and Reconstruction Results. In order to better test the continuity of reconstruction of a high degree, this paper is selected as a hemisphere, because the hemisphere in the rise or fall is continuous, so this can better reflect the effect of reconstruction. And in order to reduce the reflection of interference that the laser irradiation on the surface of the object caused, we then select the rough diffuse reflector to take pictures. As can be seen from Figure 10, the hemisphere's tennis is exactly in line with our basic requirements, and the rough surface of the tennis is just a diffuse material.

The three-dimensional reconstruction of structured light is based on the degree of deviation of the laser line and then
multiplied by the height of the offset to reconstruct the height of the object. But in the process of reading because the reconstruction of the height is too small, basically we do not see the surface of the reconstruction of the object.

Therefore, this article will rebuild the height in accordance with a certain proportion of the amplification. But for the comparison of the actual object and the reconstruction height, it can be seen that the height of the reconstruction is higher than the actual height. The results are shown in Figures 11 and 12.

6.2. Analysis of Other Groups of Results. From the comparison to the hemisphere reconstruction results and the actual object, the reconstruction results have been reconstructed out of the hemisphere, but for the reconstruction of the hemisphere there is a certain error. For example, the hemisphere is not very standard, and there is an error in the reconstructed hemisphere surface. First of all, from Figure 8 we can see that before the interpolation of their height of the degree of bending the laser line hit the object on the degree of bending more consistently. After the interpolation, we can see that his image has better reconstructed the surface of the object, and the interpolation is relatively smooth.

The design of this paper, from the beginning has been the use of the hemisphere for debugging and a series of operations; when the program is completed introducing a number of other objects, the compatibility of the program was tested. The first is to introduce a rectangular model (shown in Figure 13)

The laser data taken in the program according to the experimental data taken in Figure 13 is shown in Figure 14; as a result of shooting, the experimental data is the deviation of the laser line in the reconstruction of the experimental results that are skewed. The actual height is shown in Figure 15.

Figures 13, 14, and 15 show the rejoined results of a rectangle introduced into the program.

6.3. Height Comparison. According to the actual height of the object measurement results and reconstruction results to do a comparison, as shown in Table 1. From the table we can see, in the reconstruction of the height of the object, the accuracy is very high.
7. Summary

In this paper, we deeply study the using of word structure of light on the object surface reconstruction. Given an image for denoising, we can minimize the impact of other lights to the photographic picture by increasing the compatibility of the given photos. To get each of the height lines of the sum and interpolation operations, we then get a smooth three-dimensional reconstruction of the surface of the object. The latter part of the research process will focus on three-dimensional high-precision, high-speed, and real-time reconstruction for further study.

Data Availability

The datasets used in the experiment are from previously reported studies and datasets, which have been cited.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The study was supported by a Project of Shandong Province Higher Educational Science and Technology Program (no. J14LN64).

References


