Research Article

3D Reconstruction of Traditional Handicrafts Based on Binocular Vision

Yi Qin[1] and Zhipeng Xu [2]

[1]*Hebei Academy of Fine Arts, Shijiazhuang 050700, China*
[2]*Human Settlements and Design, Shijiazhuang Institute of Technology, Shijiazhuang 050000, China*

Correspondence should be addressed to Zhipeng Xu; xianyasanren@163.com

As an indispensable and essential material part of China's traditional culture, traditional handicrafts have played a more important role in modern life. It is not only the crystallization of the wisdom of the Chinese people of all nationalities but also an excellent platform to display China's long-standing culture. However, the protection and reconstruction of traditional handicrafts are relatively lagging. To realize the reconstruction of traditional handicrafts, this paper proposes an improved AD-Census stereo matching algorithm of a binocular measurement system for object positioning and three-dimensional reconstruction. Methods based on the principle of binocular vision measurement include firstly, histogram equalization, adaptive threshold Canny edge extraction, and expansion are used for image preprocessing; Secondly, the calibration method is used to complete camera calibration, remove camera distortion through stereo correction, and improve the AD-Census stereo matching algorithm based on gradient division of weak texture and edge area. Finally, the parallax map generated by the improved algorithm is used to realize three-dimensional reconstruction. The experimental results show that this method can control the error standard deviation within 0.5 mm, can realize the reconstruction of traditional handicrafts, have high accuracy, and can be widely used in practice.

1. Introduction

Traditional handicrafts are arts and crafts with artistic styles made by human handicrafts. It is the embodiment of the living state of the nation and the integration of life and technology. More and more attention has been paid to the protection and inheritance of traditional handicrafts [1]. With the maturity of virtual reality technology, protectors pay more attention to scientific and technological means such as digitization and informatization to realize the digital protection and inheritance of traditional handicrafts [2]. Virtual reality technology is to build a virtual scene in the real space with the help of comprehensive information processing technologies such as computer technology, simulation technology, and artificial intelligence technology, combined with virtual wearable devices, to connect virtual and reality to increase the sensory experience of the audience [3]. The addition of virtual reality technology will provide a new way to protect and disseminate traditional handicrafts.

Machine vision can be divided into monocular vision imaging and binocular vision imaging according to the number of image sensors used. Monocular vision and binocular vision are widely used in different fields. Literature [4] used the feature point matching method based on random number classification to study the grasping operation of monocular vision and mobile manipulators. Literature [5] uses monocular vision ranging to assist robots in painting. Literature [6] applies monocular vision to a low-cost indoor robot to realize robot azimuth estimation. The most extensive application of binocular vision is measurement. Literature [7] realizes the noncontact measurement of the free-form surface by binocular vision. Literature [8] uses machine learning and binocular vision to measure the size of assembly parts and realizes and guides the robot to complete the intelligent equipment of spaceborne equipment.

The classical method of 3D reconstruction in paper 12 is the motion restoration structure SFM (structure from motion) algorithm. Literature [9] restores the 3D point

cloud of the target through the same name points of two images and the spatial reference relationship of camera points. Since then, the three-dimensional reconstruction method based on SFM has attracted extensive attention, and relevant scholars have researched it. Literature [10] proposes an incremental SFM (incremental structure from motion) reconstruction method. With the continuous addition of images, this method needs to estimate the camera pose and optimize the adjustment parameters iteratively. Therefore, the algorithm's time complexity is high, and the cumulative error is significant. To solve this problem, literature [11] improves and optimizes the incremental SFM algorithm and uses SIFT algorithm to describe and match the feature points of the image. At the same time, a random sampling and consensus (RANSAC) algorithm is introduced to eliminate the wrong matching, which makes the solution of the essential matrix between images more accurate. Literature [12] analyzes the time complexity of each link of the incremental SFM algorithm and adopts different methods to optimize the SFM algorithm for nodes such as feature matching, beam adjustment, and reconstruction. Literature [13] improves the stability of initialization and the parallel computing ability of the algorithm by constructing the relationship graph hierarchically, reorganizing it in the form of a binary tree, and reconstructing it layer by layer. Literature [14] summarizes the previous experience and makes corresponding improvements to the robustness, accuracy, and integrity of the incremental SFM algorithm. However, no matter how the above methods are improved, they are always limited by the incremental SFM algorithm defects, which have high time complexity and low computational efficiency.

In this paper, an improved AD-Census stereo matching algorithm based on a binocular measurement system is proposed for 3D reconstruction, which solves the problem of false matching in the measured scene due to the change of background and illumination. Firstly, in the preprocessing stage, histogram equalization is used to enhance the image contrast. The Canny adaptive threshold is used for edge extraction, and then, the image boundary is enhanced by edge expansion. Secondly, camera calibration and stereo correction are used to remove distortion. Finally, the improved AD-Census stereo matching algorithm generates a parallax map, which improves the matching accuracy of weak texture areas and realizes the accurate three-dimensional reconstruction of traditional handicrafts.

2. State of the Art

Binocular stereo vision is to find the pixel parallax between the points with the same name according to two images taken from different angles and simulate the human eye to restore the depth information of the target point. Figure 1 shows the measurement model. The imaging plane is parallel and coplanar in the standard stereo camera model. In Figure 1, O_l and O_r are the optical centers of the camera, and the K axis is parallel to the optical axis and perpendicular to the imaging plane IOJ. f is the focal length and h is the baseline. U_l and U_r are the target point's mapping point on the two cameras' imaging plane, and the image coordinate
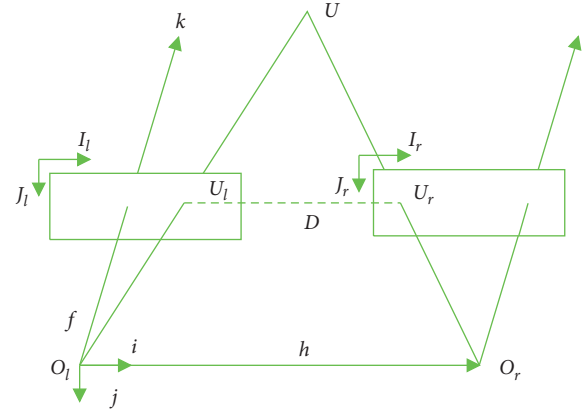

Figure 1: Binocular vision measurement model.

$U_l = (i_l, j_l), U_r = (i_r, j_r)$. In the standard stereo model, the optical center is in the same horizontal line, that is, $j_l = j_r$, so the parallax $d = i_l - i_r$.

The target point, two cameras' optical centers, and imaging points form two similar triangles. The depth information can be obtained according to the triangle similarity theorem, as shown in

$$\frac{h}{k_c} = \frac{h - (i_l - i_r)}{k_c - f} \Rightarrow k_c = \frac{f \cdot h}{i_l - i_r} = \frac{f \cdot h}{d}. \tag{1}$$

The further known three-dimensional coordinates are expressed as

$$\begin{cases} i_c = \dfrac{i_l \cdot h}{d} = \dfrac{i_l \cdot k_c}{f}, \\ \\ j_c = \dfrac{j_l \cdot h}{d} = \dfrac{j_l \cdot k_c}{f}. \end{cases} \tag{2}$$

3. Methodology

Accurate parallax calculation is the key to the accurate three-dimensional reconstruction of the binocular vision measurement system. The complete binocular vision measurement system mainly includes five steps: image acquisition and preprocessing, camera calibration, stereo correction, stereo matching, and 3D point cloud reconstruction [15].

3.1. Image Acquisition and Preprocessing. The image acquisition device used in the experiment is two high frame rate cameras with an adjustable baseline. The camera can output color and end compressed high-definition images and has the functions of synchronously capturing images and recording videos. The parameters of the left and right cameras are completely consistent. The software environment is VS2019 under the Windows 10 system, including OpenCV library, and PCL library, as well as MATLAB software.

The interference factors such as uneven illumination exposure and background noise in the pictures taken by the

actual camera will affect the final reconstruction results. In the experimental preprocessing stage, the effects of image enhancement algorithms such as histogram equalization, gray logarithm transformation, and bilateral filtering are compared. Gray histogram equalization makes the color distribution more uniform by widening the original histogram. After processing the original image, the local contrast is enhanced without affecting the overall difference [16]. Gray logarithm transform stretches the low brightness area through gray value mapping. At the same time, the region with higher brightness is compressed to enhance the dark details of the image, but the overall brightness is increased too much. Bilateral filtering considers the spatial proximity and gray similarity of image pixels at the same time, so as to remove noise and maintain the edge. The brightness is improved, and the details are more prominent. However, due to the smoothing of a large number of pixels in the image, it is unfavorable to the subsequent matching of points with the same name. After a comprehensive comparison, the histogram equalization algorithm is selected for image preprocessing. For the preprocessed image, the Canny adaptive threshold is further used for edge extraction, and then expansion operation is used to fill the holes in the edge area, which is convenient for accurate matching at the boundary.

3.2. Camera Calibration.

Camera calibration is to obtain the internal and external parameters of the camera. Calculating the conversion relationship of target points from a two-dimensional camera plane to three-dimensional space [17], the matrices $\mathbf{R}$ and $\mathbf{N}$ are established through rotation and translation, and the external parameter matrix is obtained to determine the relative position of the camera and the target. The pinhole imaging model can approximately represent the perspective projection, and the perspective projection matrix $\mathbf{W}_2$ is obtained according to the principle of triangular similarity. The image coordinates to pixel coordinates are normalized based on the camera imaging plane coordinate system to obtain the conversion matrix $\mathbf{W}_1\mathbf{W}_1$ as shown in

$$
s\begin{bmatrix} p \\ q \\ 1 \end{bmatrix} = \begin{bmatrix} \dfrac{1}{dI} & 0 & p_0 \\ 0 & \dfrac{1}{dJ} & q_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & N \\ 0^N & 1 \end{bmatrix} \begin{bmatrix} I_M \\ J_M \\ K_M \\ 1 \end{bmatrix} \tag{3}
$$

$$
= W_1 W_2 \begin{bmatrix} R & N \\ 0^N & 1 \end{bmatrix} \begin{bmatrix} I_M \\ J_M \\ K_M \\ 1 \end{bmatrix} = W \begin{bmatrix} R & N \\ 0^N & 1 \end{bmatrix} I_M.
$$

In the experiment, MATLAB is used to calibrate the camera. The chess boat images in each scene are captured from different angles, and the corner grid is extracted. Calculate the calibration parameters of forty groups of chessboard images [18]. Experience shows that the error below is the available accuracy, and most of the errors this time are concentrated around, so the expected effect has been achieved.

3.3. Stereo Correction.

Stereo correction uses distortion parameters to eliminate lens distortion and compensate for the fisheye effect around the image boundary. The two images with noncoplanar line arrangement are corrected to coplanar line arrangement by epipolar constraint, and the search space of corresponding matching points is reduced from two-dimensional to one-dimensional straight line.

3.4. Improvement of Stereo Matching Algorithm.

Stereo matching refers to comparing the similarity between the central pixel of the reference image and the pixel to be matched in another image within the parallax search range. Select the point with the lowest cost and the highest similarity to the same name point to calculate the parallax. Cost aggregation gathers the initial matching costs of adjacent pixels in the window area for addition or mean calculation, which is used for similarity matching to improve the reliability of matching costs.

The traditional AD-Census algorithm proposed in the literature [19] uses the window construction of adaptive cross-domain for cost aggregation. Set the color and distance as the constraints of the arm length extension and consider the color difference of adjacent pixels to form a cross arm for calculation. In order to improve the matching accuracy of local stereo matching in weak texture areas, the AD-Census stereo matching algorithm is improved. In the cost aggregation stage, the weak texture and edge area are divided first based on the gradient, and then the cross-domain is aggregated to calculate the parallax and generate the parallax map.

Setting the gradient threshold gives priority to delimiting whether the pixel is in the weak texture area or the edge area. Not only the color and distance constraints but also the gradient difference of adjacent pixels are considered. A looser color and distance threshold is set in the weak texture area, and the restriction is strengthened by reducing the color and distance threshold in the edge area to ensure that the arm length extends only in the areas with similar colors. The arm length constraint condition of the improved weak texture region is shown in

$$
A(u) < \beta_{\max} \begin{cases} D_a(u_x, u) < \beta_1, \\ D_d(u_x, u) < L_{\max}, \\ D_c(u_x, u) < \tau_{\max}, \text{and } D_c(u_x, u_x + (1,0)) < \tau_{\max}, \\ D_c(u_x, u) < \tau_2, \text{if } L_2 < D_d(u_x, u) < L_{\max}. \end{cases} \tag{4}
$$

The maximum color and distance threshold for the edge region need to be reduced appropriately. Avoid the excessive extension of the arm length in the edge area, increase the wrong matching, and the other conditions remain unchanged as shown in

$$A(u) > \beta_{\max}: L'_{\max} = \frac{1}{2}L_{\max}, L'_2 = \frac{1}{2}L_2, \tau'_{\max} = \frac{3}{4}\tau_{\max}, \qquad (5)$$

where $\beta_{\max}$ is the gradient threshold for dividing weak texture and edge region, taken as 120. Below $\beta_{\max}$ is the weak texture, and vice versa is the edge area. β_1 is the threshold value of the gradient difference between the central pixel and the pixel on the arm, taken as 40. D_g (p_i, p), $D_c(p_i, p)$, and $D_d(p_i, p)$ are, respectively, the gradient difference, color difference, and spatial distance between pixels p and p_i. $\tau_{\max}$ and τ_2 are two different color thresholds, 20 and 10, respectively. $L_{\max}$ and L_2 are two different distance thresholds, corresponding to 34 and 17, respectively.

The improved arm length constraint extends pixel by pixel with pixel u as the center. Determine the left, right, upper, and lower arm lengths, which are $\{b_u^-, b_u^+, q_u^+, q_u^-\}$, respectively. When any of the above conditions are not met, stop the extension of the arm length. Construct a cross-region composed of a horizontal line segment $B(u)$ and a vertical line segment $Q(u)$ as the aggregate local support skeleton. The construction process is shown in Figure 2.

3.5. 3D Point Cloud Reconstruction.

Point cloud registration unifies the pose of the point cloud data from each perspective, and finally, the overlapping areas of each part can be completely coincident. The commonly used algorithm in point cloud registration is the iterative nearest point algorithm. Because the adjacent point clouds are partially coincident, the final iteration fails, and an accurate model cannot be obtained. Therefore, this paper first obtains the best change matrix according to the feature points and extracts the feature points, such as NARF (normalized radial feature) algorithm, scale-invariant feature transform (SIFT), FPFH algorithm, and Harris algorithm. Then improve the iterative nearest point algorithm to modify the model.

3.5.1. Point Cloud Rough Registration.

This paper aims at the point cloud in the overlapping area of two point clouds, which has the same distribution relationship with the neighborhood. Therefore, this paper combines FPFH and RANSAC to eliminate the wrong matching points [20] and then carries out the rough registration of point clouds, so as to obtain a better initial model and improve the basis for fine registration. After the sample rotates every 60°, the point cloud data can be collected from multiple perspectives, and rough registration can be carried out based on frame 1 [21].

Suppose the target point cloud of the reconstructed object is $N(n_1, n_2, \ldots, n_t)$ and the source point cloud to be registered is $S(s_1, s_2, \ldots, s_t)$. Firstly, a binary tree ($k-d$ tree) is constructed to search for the nearest neighbor of the structured point cloud. Calculate the FPFH value of each point cloud in point cloud cluster N and source point cloud S and use open MP multithreading technology to speed up FPFH calculation. So far, point cloud pairs with similar characteristics can be found. However, there may be errors in point cloud pairs with similar characteristics. The

RANSAC algorithm eliminates the wrong point pairs. The rotation matrix **R** and translation matrix **N** that can coincide between the two frame point clouds are calculated. To avoid falling into the local optimum, the error function Huber, as shown in equation (6), is used to judge the current model's performance.

$$B(u_x) = \begin{cases} \frac{1}{2}u_x^2, & u_x \leqslant w_x, \\ \\ \frac{1}{2}w_x(2u_x - w_x), & u_x > w_x, \end{cases} \qquad (6)$$

where w_x is the preset setting value. The change matrix is the optimal solution when the error function is the smallest.

The specific process of rough registration is shown in Figure 3.

3.5.2. Precise Registration of Iterative Closest Points Based on Weight.

During rough registration, the point cloud of frame 0 is S_0 and the point cloud of frame i is S_i. When S_0 and S_1 are roughly registered, S'_1 is obtained. An error exists in point cloud registration of S_1 and S'_1. However, coarse registration of point cloud S'_1 and S_2 will transfer and accumulate errors. Therefore, when registering to S_n in turn, the error will become larger and larger. Thus, it is necessary to reduce the error accumulated in the registration process of two adjacent frames in the delicate registration stage. The iterative closest point algorithm (ICP) is the most widely used 3D reconstruction algorithm at present. This paper designs an iterative nearest point algorithm based on weight. For two adjacent frames, the point cloud has a good pose after entering the rough registration stage. At this time, the ICP algorithm [22] can be used for the fine registration of point cloud data. The weight is used to reduce the global impact of local error, and the threshold of the normal vector between point clouds is set to remove invalid point pairs. Finally, ICP iterative calculation based on weight is carried out.

The specific steps are as follows:

(1) Assuming S_x is the point cloud to be registered and N_y is the target point cloud, the point cloud is constructed into a k-d tree, and the corresponding point pair is found by a fast nearest neighbor search.

(2) Calculate the corresponding normal vector according to the corresponding point pair. Assuming that the normal vector of one of the point pairs is t_{S_x} and t_{N_x}, calculate the angle relationship value between the pair of normal vectors and compare it with the threshold t_r. If it is less than the threshold, the point pair will be eliminated.

$$\theta = \frac{t_{S_x} \cdot t_{N_x}}{t_{S_x} \cdot t_{N_x}}. \qquad (7)$$

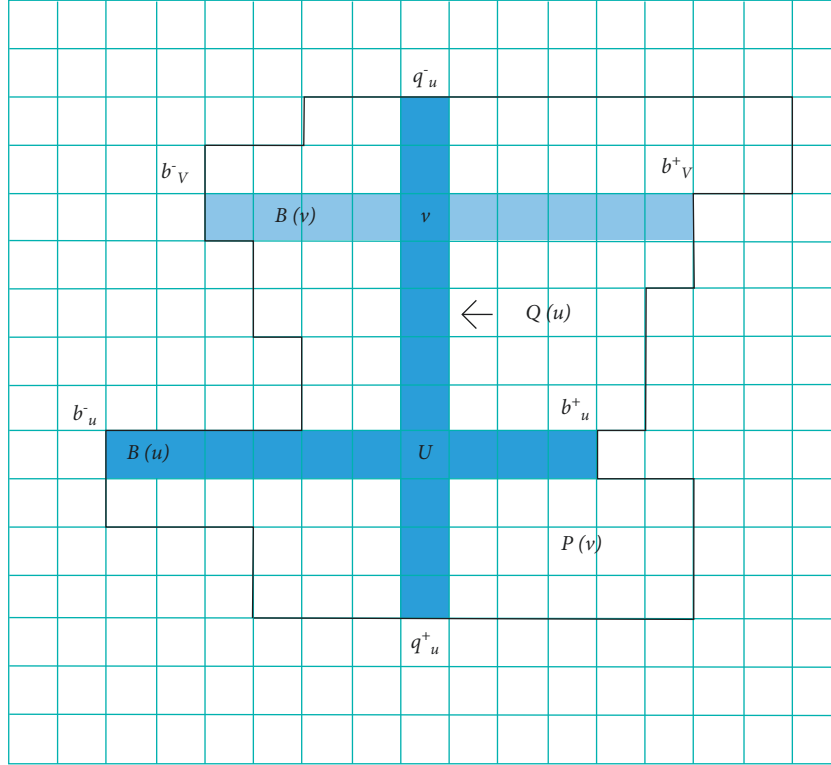(3) Calculate the weight according to the point pair, and the weight m_{xy} is calculated as follows:

FIGURE 2: Window construction of cost aggregation cross-domain.

$$d_{xy} = S_x - N_y^2,$$

$$\sigma = \frac{2}{z} \sum_{y=1}^{z} d_{xy}, \qquad (8)$$

$$m_{xy} = b(x, y)\exp\left(-\frac{d_{xy}}{\sigma^2}\right),$$

where d_{xy} is the distance between corresponding points. Since 60° is selected each time, the over-lapping part is $\sin 60°$ of the previous adjacent data frame. Therefore, the function $b(x, y)$ represents whether the selected point belongs to the benchmark.

$$b(x, y) = \begin{cases} 1, & N_y \in S_x, \\ \frac{1}{2}, & N_y \notin S_x. \end{cases} \qquad (9)$$

According to the weight value of the point pair, combined with the least square method, the corresponding spatial transformation parameters $\mathbf{R}$ and $\mathbf{N}$ are obtained.

$$E(\mathbf{R}, \mathbf{N}) = \min \sum_{x=1}^{z} b(S_x, N_x) N_x - (S_x \mathbf{R} + \mathbf{N})^2. \qquad (10)$$

(4) The obtained spatial transformation parameters $\mathbf{R}$ and $\mathbf{N}$ are applied to the source point cloud to obtain a new point cloud set $\mathbf{V}$.

$$d = \frac{1}{T}\left(\sum_{x=1}^{T} \|\mathbf{V} - \mathbf{N}\|\right). \qquad (11)$$

(5) Judge whether the Euclidean distance between the target point cloud and the new point cloud is less than the threshold r. If less, reiterate until convergence.

The specific flow of the algorithm is shown in Figure 4.

4. Result Analysis and Discussion

The hardware of the test system includes a six-axis industrial robot, two CCD cameras, a camera mounting bracket, and a light-emitting board. The three-dimensional reconstruction experiments were carried out through three groups of handicrafts of different sizes and shapes. The three groups of traditional handicrafts were embroidery ornaments, egg carving, and wood carving square boxes, which are shown in Figure 5.

4.1. Camera Calibration. The single size of the black-and-white chessboard calibration board is 3 mm × Take 15 units of grids in 3 mm and direction, respectively, and complete the camera calibration with MATLAB software. In this study, the left and right cameras collected 20 calibration plate pictures with different angles and positions for monocular and binocular calibration. According to the single target setting, the focal length of the camera, the
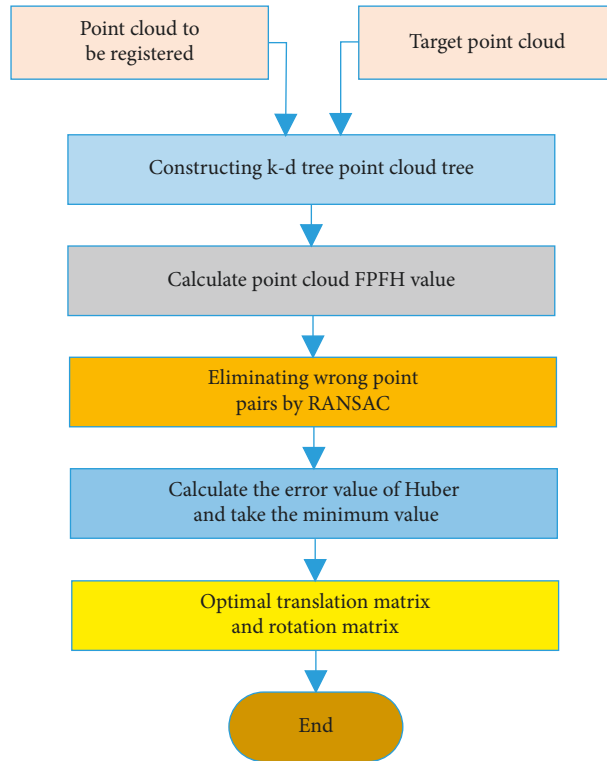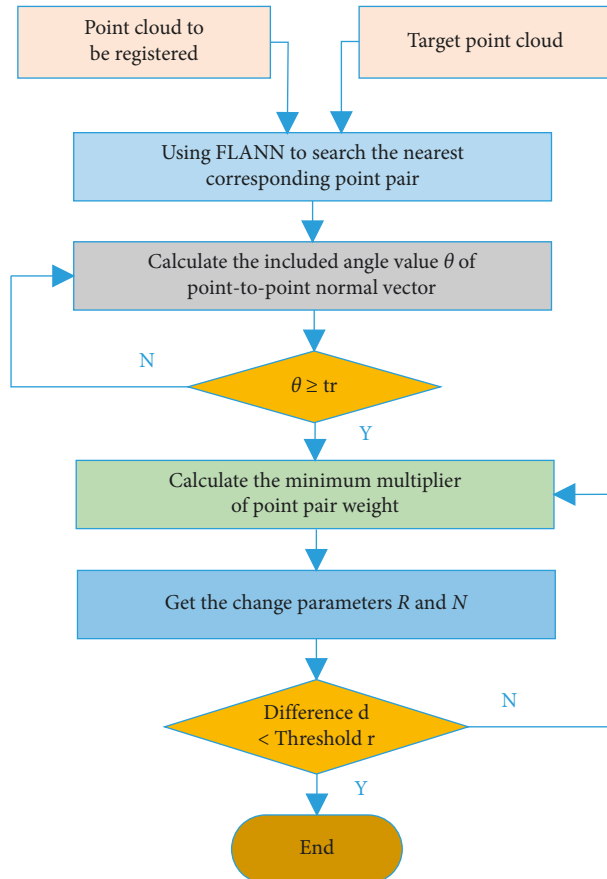
Figure 3: Rough registration process.

Figure 4: Fine registration process.

FIGURE 5: Traditional handicrafts.

TABLE 1: Calibration results of left and right cameras.

Parameter	Left camera	Right camera
Focal length/mm	11.542, 11.532	11.487, 11.575
Main point location/pixel	1227.351, 945.427	1227.356, 945.427

position of the main point (Table 1), and the internal parameters of the left and right cameras can be calculated, respectively, as follows: $\begin{bmatrix} 11844.741 & 0 & 1227.326 \\ 0 & 11837.387 & 945.524 \\ 0 & 0 & 1 \end{bmatrix}$ (left); $\begin{bmatrix} 11982.962 & 0 & 1227.322 \\ 0 & 11977.242 & 9.45425 \\ 0 & 0 & 1 \end{bmatrix}$ (right).

The rotation matrix $\mathbf{R}_c = \begin{bmatrix} 0.754 & 0.006 & -0.657 \\ -0.009 & 1.000 & -0.002 \\ 0.653 & 0.008 & 0.754 \end{bmatrix}$ of the left and right cameras is obtained according to the double target positioning, and the translation vector is $\mathbf{N}_c = \begin{bmatrix} -243.841 & -0.074 & 73.995 \end{bmatrix}^N$. The average focal length $f = 2.2 \times (f_{i1} + f_{j^1} + f_{ir} + f_{jr}) \times 0.25$, where the focal length components of the left camera in the i and j directions are f_{i1} and f_{j^1}, and the focal length components of the right camera in the i and j directions are f_{ir} and f_{jr}.

4.2. Point Circle Fitting and Matching Results. Three contour lines determine all vertices of the woodcut square box. Therefore, the point circle fitting method determines the feature points. The schematic diagram of the three-point fitting circle is shown in Figure 6.

In Figure 6, point 1, point 2, and point 3, respectively, represent the intersection of three adjacent contour lines of the woodcut square box. The center V is obtained by fitting the three points, that is, the feature point. Similarly, the above method is adopted when the vertex of the woodcut square box is greater than 2 contour lines. However, there will be errors in the multipoint fitting curve in the calculation process, and there will also be errors in the center of the final multipoint fitting circle. Therefore, in the experiment, a feature point of the handicrafts photographed at a specific position is known and compared with the point circle fitting method results. The coordinate values of point 1, point 2, point 3, and the fitting circle center are shown in
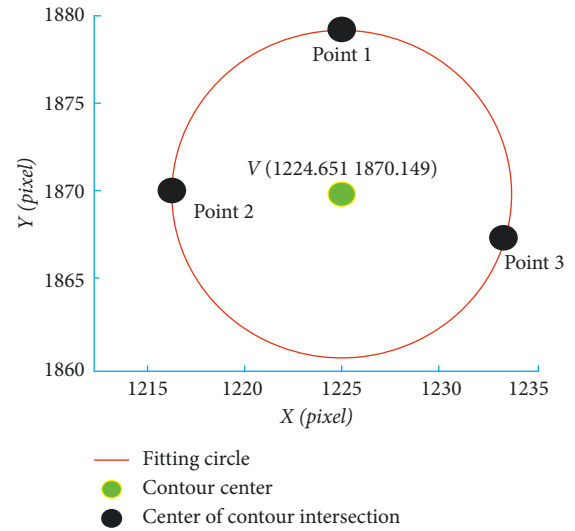

FIGURE 6: Point circle fitting results.

TABLE 2: Contour intersection and fitting center error.

Point	Coordinate value/pixel
1	1224.045, 1878.632
2	1216.127, 1870.144
3	1232.752, 1867.521
Fitting circle center	1224.651, 1870.149
Actual feature points	1220.000, 1875.000
Error value	4.652, −4.858

Table 2. It can be seen from Table 2 that the error of point circle fitting is within the allowable range.

Because the feature points of egg carving are not obvious, it is impossible to carry out effective stereo matching of feature points, so only two groups of handicrafts with obvious vertices are matched with the minimum distance of the common vertical line of different planes. The matching algorithm can accurately find the feature points and realize the matching of feature points.

FIGURE 7: Shooting results at two locations. (a) Position A. (b) Position B.
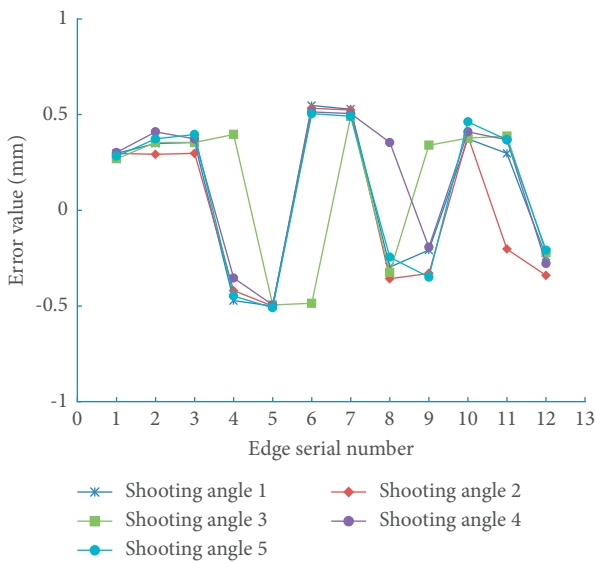

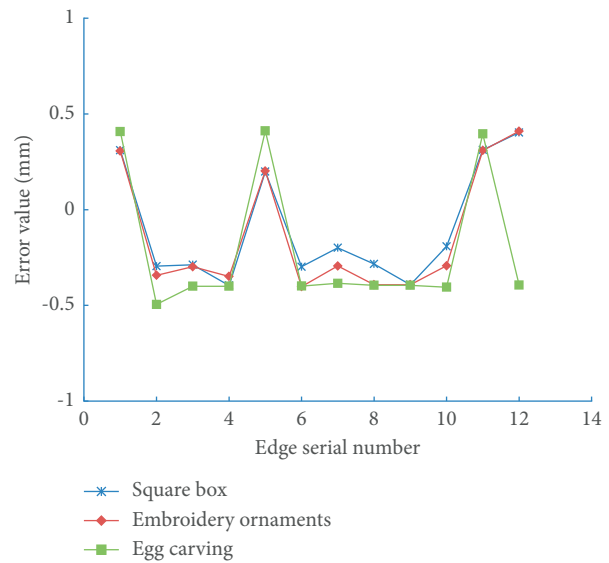FIGURE 8: Error analysis of reconstruction result from 5 different shooting angles.

FIGURE 10: Error of reconstruction result.

FIGURE 9: 3D reconstructed model of woodcut square box.

4.3. 3D Model Restoration and Error Analysis. The woodcut square box at two angles photographed by a binocular camera is shown in Figure 7, which is marked with overlapping feature points at two positions. The coordinate values of point 7 and point 8 in the coordinate system are converted to A in the coordinate system by using the secondary photographing and reconstruction method of overlapping feature points. Then, the spatial coordinates of all the remaining feature points of the woodcut square box are converted to coordinate system A to calculate the three-dimensional size of the woodcut square box. In order to verify the effectiveness of this method, the three-dimensional reconstruction of five groups of images under different shooting angles is carried out, and the size of 12 sides of the woodblock square box is calculated. Figure 8 shows the error curve between the size calculated by this method and the actual size.

For all crafts with the maximum outline size larger than the camera field of view and the number of surfaces exceeding 4, the same calculation method is used to obtain the full size of the crafts. The 3D reconstructed model of the woodcut square box is shown in Figure 9.

Figure 10 shows the measurement error distribution curve of each contour edge of different handicrafts under multiple groups of experiments. It can be seen that the error of the three-dimensional reconstruction results of handicrafts using the research method in this paper is controlled within ±0.5 mm, in which the egg carving error fluctuates the most, and the corresponding error value is also the largest.

This is because the egg carving has rounded corners, which will cause reconstruction deviation.

5. Conclusion

Taking traditional handicrafts as the research object, this paper studies the three-dimensional reconstruction of handicrafts. An improved AD-Census stereo matching algorithm of the binocular measurement system is proposed for object positioning and 3D reconstruction. The contour fitting algorithm is used to extract the edge contour of handicrafts. According to the intersection of multiple contour lines, the point circle fitting method determines the feature points. The feature point matching of handicrafts is realized using the minimum distance stereo matching algorithm based on the common vertical line of different planes. Aiming at the multisurface and large size of handicrafts, a secondary shooting reconstruction algorithm with overlapping feature points is introduced to realize the complete three-dimensional reconstruction of handicrafts. Finally, the three-dimensional reconstruction experiment is carried out through three groups of handicrafts with different sizes and shapes, which can restore the three-dimensional size and shape of the handicrafts as a whole. The error standard deviation of the main outline dimensions of each handicraft is within ±0.5 mm, reaching a high accuracy. However, when there are rounded corners on the edge of handicrafts, the fitting results of feature points will be biased, and the algorithm in this paper may not be able to accurately 3D reconstruction, which is also the direction of further optimization.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

References

[1] K.-K. Fan and T.-T. Feng, "Discussion on sustainable development strategies of the traditional handicraft industry based on su-style furniture in the ming dynasty," *Sustainability*, vol. 11, no. 7, p. 2008, 2019.

[2] X. Jing, F. Tan, and M. Zhang, "Digital application of intangible cultural heritage from the perspective of cultural ecology," *Journal of Smart Tourism*, vol. 1, no. 1, pp. 41–52, 2021.

[3] J. Abich, J. Parker, J. S. Murphy, and M. Eudy, "A review of the evidence for training effectiveness with virtual reality technology," *Virtual Reality*, vol. 25, no. 4, pp. 919–933, 2021.

[4] T. Watanabe, K. Yamazaki, and Y. Yokokohji, "Survey of robotic manipulation studies intending practical applications in real environments-object recognition, soft robot hand, and challenge program and benchmarking," *Advanced Robotics*, vol. 31, no. 19-20, pp. 1114–1132, 2017.

[5] Q. Zhao, X. Li, J. Lu, and J. Yi, "Monocular vision-based parameter estimation for mobile robotic painting[J]," *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 10, pp. 3589–3599, 2018.

[6] D. H. Yi, T. J. Lee, and D. I. Cho, "A new localization system for indoor service robots in low luminance and slippery indoor environment using afocal optical flow sensor based sensor fusion," *Sensors*, vol. 18, no. 1, p. 171, 2018.

[7] X. Wen, J. Wang, G. Zhang, and L. Niu, "Three-dimensional morphology and size measurement of high-temperature metal components based on machine vision technology: a review," *Sensors*, vol. 21, no. 14, p. 4680, 2021.

[8] Y. Xu, X. Liu, X. Cao et al., "Artificial intelligence: a powerful paradigm for scientific research," *Innovation*, vol. 2, no. 4, Article ID 100179, 2021.

[9] X. Liu, W. Zheng, Y. Mou, Y. Li, and L. Yin, "Microscopic 3D reconstruction based on point cloud data generated using defocused images," *Measurement and Control*, vol. 54, no. 9-10, pp. 1309–1318, 2021.

[10] T. Taketomi, H. Uchiyama, and S. Ikeda, "Visual SLAM algorithms: a survey from 2010 to 2016," *IPSJ Transactions on Computer Vision and Applications*, vol. 9, no. 1, pp. 1–11, 2017.

[11] M. Cao, H. Gao, and W. Jia, "Stable image matching for 3D reconstruction in outdoor," *International Journal of Circuit Theory and Applications*, vol. 49, no. 7, pp. 2274–2289, 2021.

[12] F. Huang, H. Yang, X. Tan, S. Peng, J. Tao, and S. Peng, "Fast reconstruction of 3D point cloud model using visual SLAM on embedded UAV development platform," *Remote Sensing*, vol. 12, no. 20, p. 3308, 2020.

[13] Z. Wu, S. Pan, F. Chen, G Long, C Zhang, and P S Yu, "A comprehensive survey on graph neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 4–24, 2020.

[14] Y. You, J. Cao, and W. Zhou, "A survey of change detection methods based on remote sensing images for multi-source and multi-objective scenarios," *Remote Sensing*, vol. 12, no. 15, p. 2460, 2020.

[15] G. Huo, Z. Wu, J. Li, and S. Li, "Underwater target detection and 3D reconstruction system based on binocular vision," *Sensors*, vol. 18, no. 10, p. 3570, 2018.

[16] C. O. Ancuti, C. Ancuti, C. De Vleeschouwer, and P. Bekaert, "Color balance and fusion for underwater image enhancement," *IEEE Transactions on Image Processing*, vol. 27, no. 1, pp. 379–393, 2017.

[17] L. Wang, "Symbol recognition system based on 3D stereo vision," *Journal of Intelligent and Fuzzy Systems*, vol. 40, no. 4, pp. 5985–5994, 2021.

[18] D. Li, L. Xu, X.-s. Tang, S. Sun, X. Cai, and P. Zhang, "3D imaging of greenhouse plants with an inexpensive binocular stereo vision system," *Remote Sensing*, vol. 9, no. 5, p. 508, 2017.

[19] S. Zhang, M. Wu, Y. Wu, G. Wu, and F. Liu, "Fixed window aggregation AD-census algorithm for phase-based stereo

matching," *Applied Optics*, vol. 58, no. 32, pp. 8950–8958, 2019.

[20] H. Lei, G. Jiang, and L. Quan, "Fast descriptors and correspondence propagation for robust global point cloud registration," *IEEE Transactions on Image Processing*, vol. 26, no. 8, pp. 3614–3623, 2017.

[21] R. A. Kuçak, S. Erol, and B. Erol, "An experimental study of a new keypoint matching algorithm for automatic point cloud registration," *ISPRS International Journal of Geo-Information*, vol. 10, no. 4, p. 204, 2021.

[22] Y. Xian, J. Xiao, and Y. Wang, "A fast registration algorithm of rock point cloud based on spherical projection and feature extraction," *Frontiers of Computer Science*, vol. 13, no. 1, pp. 170–182, 2019.