Research Article

CAD-Aided 3D Reconstruction of Intelligent Manufacturing Image Based on Time Series

Liming Zhang, Lei Wang [ID], Xu du, and Fanbo Meng

College of Information Science and Electronic Technology, Jiamusi University, Jiamusi 154007, China

Correspondence should be addressed to Lei Wang; wangl@jmsu.edu.cn

To improve the three-dimensional (3D) reconstruction effect of intelligent manufacturing image and reduce the reconstruction time, a new CAD-aided 3D reconstruction of intelligent manufacturing image based on time series was proposed. Kinect sensor is used to collect depth image data and convert it into 3D point cloud coordinates. The collected point cloud data are divided into regions, and different point cloud denoising algorithms are used to filter and denoise the divided regions. With the help of CAD, FLANN matching algorithm is used to extract feature points of time-series images and complete image matching. Three-dimensional reconstruction of sparse point cloud and dense point cloud is carried out to complete 3D reconstruction of intelligent manufacturing images. The experimental results show that the image PSNR of this method is always above 52 dB, and the maximum reconstruction time is 4.9 s. The 3D reconstruction effect of intelligent manufacturing image is better, and it has higher practical application value.

1. Introduction

In recent years, in the field of information technology and industry, significant changes have taken place, such as large data, cloud computing, mobile Internet, 3D printing, and industrial robots, and these changes have brought a new round of revolution in global manufacturing, including intelligent manufacturing as a product of informatization and industrialization depth fusion, but also got great progress. It has been widely concerned and valued by the governments of various countries [1], such as the "Advanced Manufacturing National Strategy" of the United States, the "New Industry France Plan" of France, the "New Industry Creation Strategy" of Japan, and the "Industry 4.0" of Germany [2]. An important reason why developed countries attach importance to intelligent manufacturing is that the financial crisis in 2008 fully exposed the fragility of the virtual economy, so many countries have re-examined the important role of manufacturing in social development, taking intelligent manufacturing as a breakthrough to

improve the country's comprehensive competitiveness [3]. Intelligent manufacturing helps to improve the level of China's traditional manufacturing industry, achieve high-end technological innovation, relieve energy pressure, promote the transformation of new production modes, and accelerate the transformation from a manufacturing country to a manufacturing power. However, due to the continuous development of manufacturing technology, information technology, and network technology, the concept and connotation of intelligent manufacturing are also in constant change, enrichment, and improvement [4]. At this stage, the field has generally recognized the definition of intelligent manufacturing and believes that intelligent manufacturing integrates a variety of modern technologies, which can realize the monitoring and decision-making in the process of intelligent production and improve the intelligent degree and production efficiency of product production [5]. Therefore, in order to cope with the competitive market environment, ensure the clarity of intelligent manufacturing images, and enhance the competitiveness of enterprises.

Image 3D reconstruction has become the focus of current business and academic circles. So it is very important to study the 3D reconstruction method of intelligent manufacturing image.

There is not much existing image design 3D reconstruction methods in intelligent manufacturing, so we just need to migrate to other areas of image 3D reconstruction method, for example by Zhuang and Wan [6], who proposed a framework based on P2M single-image reconstruction of 3D model of the improved method, mainly because the traditional 3D reconstruction method cannot rebuild object's invisible part, it takes a long time to reconstruct, and it is also difficult to reconstruct objects without texture. Three-dimensional reconstruction is carried out around a single image. The main work is to replace the backbone of feature establishment, improve the feature extraction network by VGG-16 network, and propose an improved DenseNet network to obtain better two-dimensional feature points. To improve the reconstruction effects, Liu et al. [7] proposed a single-image 3D reconstruction method based on vanishing point optimization. First, the image is processed, the long lines in the image are extracted, the characteristics of the lines are analyzed, and the lines in different directions are grouped. Then, according to the linear distribution relationship of the lines in each direction, the linear model of the line parameters is obtained by the improved regression algorithm, the error lines are eliminated, and the extinction point is solved by the least square method. After the accurate extinction point is obtained, the internal and external parameters of the camera are obtained according to the properties of the extinction point. The minimum two-dimensional points of the image are obtained through interaction. The 3D reconstruction of the object is realized by calculating the 3D coordinates of the camera parameters and the geometric features of the object. Zhao et al. [8] proposed a 3D reconstruction method based on sea ice scene image classification. Aiming at the problem of large amount of computation and easy accumulation of errors in 3D reconstruction process, this method uses a classification network to screen suitable image sequence through training and uses 8-neighborhood filling algorithm to segment the reconstructed region of the obtained specific category sea ice scene image and finally carries out the 3D reconstruction of the segmented image. Yu et al. [9] take flange parts as an example to propose a reconstruction method of flange parts based on working drawing images. This method takes the scanned images of engineering drawings as the research object; carries out noise reduction, segmentation, and refinement processing; and then uses Harris corner recognition algorithm and Hough circle detection algorithm to extract and make statistics on their upper features. Back-propagation (BP) neural network was used to identify and extract the necessary parameters of parts reconstruction, and finally realized the three-dimensional reconstruction of flange model. Liu et al. [10] proposed a 3D reconstruction method of fuzzy image edge contour based on machine learning. Through the detection of image edge contour, the method of stereo matching is used to find the corresponding points in the detected image edge, so as to obtain 3D coordinates. According to the obtained 3D coordinates, the 3D point cloud of the object surface is obtained, and the 3D reconstruction of fuzzy image edge contour is realized based on machine learning.

But the aforementioned methods' image denoising processing effect is poor, the reconstruction effect is poorer, and the reconstruction takes longer, so this article proposes to solve the aforementioned problems existing in the method as the research target, to access the sensor as the tool of depth image data acquisition, combined with ordinary filter algorithm and bilateral filtering algorithm as the theoretical basis, to deal with the noise in point cloud data. Based on FLANN matching algorithm, feature points of time-series images were extracted, and 3D reconstruction of sparse point cloud and dense point cloud was carried out by CMVS algorithm and PMVS algorithm, respectively, so as to complete 3D reconstruction of intelligent manufacturing images.

2. CAD-Aided Intelligent Manufacturing Image 3D Reconstruction Method Design

2.1. Point Cloud Data Collection. As a 3D depth sensor, Kinect has the functions of real-time dynamic capture, voice recognition, microphone input, and intelligent interaction in addition to obtaining color images and depth images applied in this paper. The appearance is shown in the figure below. The external structure is simple and light, and the internal device is complex and powerful [11]. Kinect sensors emit infrared light using color and depth cameras to collect raw data, which are combined to generate 3D data [12, 13]. The imaging principle of Kinect camera is similar to that of pinhole imaging. In this paper, the pinhole model is used to introduce the camera imaging principle. Assume that P represents any object in space, the pinhole plane is the vertical plane where the camera view is located [14], and the imaging plane after P passes through the optical center is called the image plane. The horizontal ray passing through the optical center and perpendicular to the two planes is called the optical axis, and the intersection between the optical axis and the pinhole plane is called the main point. The distance between the image plane and the pinhole plane f is the focal length of the camera, the horizontal distance between P and the pinhole plane is Z, the real height of the object is X, the imaging height is x, and the imaging point is p [15]. According to the similar triangle principle:

$$\frac{x}{f} = -\frac{X}{Z}.$$
(1)

Considering the image coordinate system, axis u represents the rows of the image and axis v represents the columns of the image [16]. Then, the coordinate conversion relation of $p(u, v)$ and $P(x, y)$ is as follows:

$$\begin{cases} u = \dfrac{x}{k} + u_0, \\[3mm] v = \dfrac{y}{l} + v_0. \end{cases} \quad (2)$$

The relationship between camera coordinate system and image coordinate system can be expressed by the following formula:

$$\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \frac{1}{Z_c} \begin{bmatrix} f & -f\cot\theta & 0 & 0 \\ 0 & \dfrac{f}{\sin}\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}. \quad (3)$$

In the aforementioned formula, θ represents the included angle between two coordinate systems, and f represents the focal length of the camera. Combining the aforementioned two equations, the relationship between camera coordinate system and image pixel coordinate system can be expressed by the following formula:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \frac{1}{Z_c} \begin{bmatrix} \dfrac{f}{k} & -\dfrac{f}{k}\cot\theta & u_0 & 0 \\ 0 & \dfrac{f}{l\cdot\sin\theta} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix}$$

$$= \frac{1}{Z_c} A P_c. \quad (4)$$

In the aforementioned formula, (k, l, u, v, f, θ) represents camera parameters and A represents parameter matrix [17]. For the same pixel, the relationship between the world coordinate system and the camera coordinate system can be expressed by the following formula:

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} = \begin{bmatrix} R_{3\times3} & t_{3\times1} \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}. \quad (5)$$

In the aforementioned formula, $R_{3\times3} = (r_x r_y r_z)$ is the rotation matrix, $t_{3\times1} = (t_x t_y t_z)^T$ is the translation vector, and (R, t) is the external parameter matrix of the camera.

Combining the aforementioned two equations, the relationship between the image pixel coordinate system and the world coordinate system can be obtained, which can be described by the following formula:

$$s\tilde{m} = A(R, t)\tilde{M}. \quad (6)$$

In the aforementioned formula, $\tilde{m} = (u, v, 1)^T$ represents the coordinates of pixel points, and its corresponding world coordinate system is $\tilde{M} = (X_w, X_w, X_w, 1)^T$, and s represents the scale coefficient.

The camera coordinate system consists of optical center o, coordinate axes X_c, Y_c, and Z_c, and the optical axis is represented by axis Z_c. The u and v planes of the image coordinate system constitute the image plane and are perpendicular to axis Z_c. The origin is the intersection of the optical axis and the image plane. The projection of point $P(x, y, z)$ in space onto the image plane is $p(u, v)$. Firstly, the points in the world coordinate system are transformed into the camera coordinate system, and then the camera coordinates are transformed into the image physical coordinate system according to the similar triangle principle, and finally the coordinates $p(u,v)$ in the image pixel coordinate system are obtained.

The original data collected by the Kinect sensor cannot be used directly but need to be converted into data that can be understood and directly applied. The raw output data generated by the sensor device are usually mapped to the grayscale space in the form of a depth image. The image takes the origin of depth image, and the unit is pixel. Each pixel represents the depth distance from Kinect. The function R Depth Gen.Convert Projective To Real World in Open Source library Open NI is used for coordinate conversion. The following relationship exists between the world coordinate system and the depth image coordinate system:

$$X = (u - 300) \cdot Z \cdot \frac{1}{f_{xz}},$$

$$Y = (v - 300) \cdot Z \cdot \frac{1}{f_{yz}}. \quad (7)$$

In the aforementioned formula, X, Y, and Z, respectively, represent the X, Y, and Z axes in the world coordinate system with Kinect device as the origin, and u and v, respectively, represent the horizontal and vertical coordinate axes of the depth image coordinate system. $f_{xz} f_{yz}$ is the focal length of the corresponding depth camera. The X and Y coordinates of the camera center are 320 and 240, respectively (image resolution 640×480). The coordinates after transformation are the real-world coordinates. Coordinate values of X, Y, and Z coordinate form 3D point cloud, and the depth image data are converted into 3D point cloud coordinates for subsequent processing.

The specific conversion process is as follows:

Since the origin of the world coordinates and the origin of the camera coincide, that is, there is no rotation and translation, the 3×3 rotation matrix R and 3×1 translation matrix t are calculated as follows:

$$R = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} t$$

$$= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \quad (8)$$

At this point, the coordinate origin of the camera coordinate system and the world coordinate system coincide, so the same object under the camera coordinate system and the world coordinate system has the same depth, namely $Z_c = Z_w$. Then there is

$$Z_c \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \begin{bmatrix} \dfrac{f}{dx} & 0 & u_0 \\ 0 & \dfrac{f}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}. \tag{9}$$

From the aforementioned transformation matrix formula, the coordinate transformation formula of depth image point conversion into 3D point cloud can be calculated:

$$\begin{cases} x_w = z_c \cdot (u - u_0) \cdot \dfrac{dx}{f}, \\[2mm] y_w = z_c \cdot (v - v_0) \cdot \dfrac{dy}{f}, \\[2mm] z_w = z_c. \end{cases} \tag{10}$$

2.2. Point Cloud Data Filtering and Denoising. To obtain better denoising results of point cloud model, the features of the model can be preserved while removing noise. This paper first divides the scanned point cloud data into regions, and then adopts targeted point cloud denoising algorithms for the divided regions [18]. The specific region division method is as follows: the local surface fitting method is used to calculate the differential geometric information of the point cloud, and the local feature weights are set according to the mean curvature of the model point cloud obtained by calculation. Then the local feature weight threshold is given, and then the neighborhood feature weight of the sampling point is judged with the set threshold. Thus, the region of the point cloud data of the model is divided, and the flat region with less feature information and the region with more feature information are obtained [19].

Regional division is carried out by calculating the local feature weights of points. Given that the average curvature at sampling point p_i is H_i, the local feature weights of sampling points in k-nearest neighbor neighborhood are defined as follows:

$$s(p_i, k) = \sqrt{\frac{1}{k} \sum_{i}^{k} \left(\left| H_{p_i} \right| - \overline{H} \right)} + \sqrt{\left(H_i - \overline{H} \right)}. \tag{11}$$

where $\overline{H}$ is the average curvature in the neighborhood of the sampling point [20]:

$$\overline{H} = \frac{1}{n} \sum_{i=1}^{n} H_i. \tag{12}$$

If the local feature weight $s(p_i, k)$ at sampling point p_i is less than the threshold value $s_{\max}$ of the set feature weight, then sampling point p_i is judged to be a point in the flat region with less feature information [21]. If the local feature weight value $s(p_i, k)$ at sampling point p_i is greater than the threshold value $s_{\max}$ of the set feature weight value, then sampling point p_i is judged to be the point in the area with more feature information.

For point cloud data with well-divided feature types, targeted filtering denoising algorithm is adopted [22–24]. The neighborhood distance average filtering algorithm and the adaptive bilateral filtering algorithm are used to denoise the two different feature areas.

2.2.1. Point Cloud Denoising in Flat Area. Owing to the characteristic information of the few curvature changes more gently on the surface of the flat area of the model, the low complexity of denoising methods, namely using sampling point k-nearest neighborhood interior point to the average distance between neighboring points of statistical filtering algorithm for discrete points and noise points in the area to remove noise effectively after receiving new model of point cloud data [25], as follows:

$$P' = \left\{ p_i \in P \mid (\mu_n - \alpha\sigma_n) \leq \overline{d}_i \leq (\mu_n + \alpha\sigma_n) \right\}, \tag{13}$$

where $\overline{d}_i$ is the average distance between the inner point and the nearest neighbor point in the neighborhood μ_n of the sampling point K, μ_n is the average distance of the whole flat region, and σ_n is the standard deviation. The calculation formula for each parameter is as follows:

$$\overline{d}_i = \frac{1}{k} \sum_{j=1}^{k} \left\| p_{ij} - p_j \right\|. \tag{14}$$

$$\mu_n = \frac{1}{n} \sum_{i=1}^{N} \overline{d}_i,$$
$$\sigma_n = \sqrt{\frac{1}{N} \sum_{i=1}^{n} \left(\overline{d}_i - \mu_n \right)^2}. \tag{15}$$

The algorithm steps are as follows:

(1) For each sampled data point p_i in the flat region with small curvature changes, all points in its k-nearest neighbor neighborhood are searched;

(2) Formulas (10) to (12) were used to calculate the average distance from k-nearest neighbor to nearest neighbor $\overline{d}_i$ of each sampling point p_i, and the average distance and mean square error of all sampling points in point cloud of flat region $\overline{d}_i$.

(3) The mean-variance theory is used to determine the noise points and discrete points to be removed [26], that is, formula (9) is used to judge. If the average distance $\overline{d}_i$ of sampling point p_i exceeds the given threshold, the point is removed.

2.2.2. Point Cloud Smoothness in Feature-Rich Regions. Compared with other denoising algorithms, bilateral filtering algorithm has more advantages in the noise removal process of 3D scattered point cloud. It was initially used in image processing to preserve the image contour. Owing to its ability to maintain high-frequency features, this paper chooses the point cloud bilateral filtering method to denoising the feature-rich region of point cloud containing noise [27]. The specific method is shown in formula (16):

$$p' = p + \beta \cdot \vec{n}, \tag{16}$$

where p' is the filtered point data, p is the initial point cloud data, $\vec{n}$ is the normal vector, and β is the bilateral filtering factor. Its calculation formula is shown as follows:

$$\beta = \frac{\sum_{i=1}^{k} z_1\left(\|p - p_i\|\right) z_2\left(\langle p - p_i, n\rangle\right)\langle p - p_i, n\rangle}{\sum_{i=1}^{k} z_1\left(\|p - p_i\|\right) z_2\left(\langle p - p_i, n\rangle\right)}, \tag{17}$$

where k represents the sampling points in the neighborhood nearest to the sampling point [28], z_1 and z_2 are Gaussian filtering functions in the spatial domain and frequency domain of bilateral filtering functions, respectively, and their specific forms are shown as follows:

$$z_1(x) = e^{y^2/2\sigma_1^2},$$
$$z_2(y) = e^{-y^2/2\sigma_2^2}. \tag{18}$$

In the aforementioned formula, σ_1 represents the effect factor of the distance between sampling point p_i and points in its neighborhood on point p_i. The larger the value of σ_1 is, the more neighborhood points will be selected. In this case, the denoising result will be better [29, 30], but the effect of point cloud feature preservation will be reduced. σ_2 is the action factor of logarithmic point p_i on the projection of the distance vector from sampling point p_i to the point in its neighborhood on the normal direction $\vec{n}$ of the point, which mainly plays a role in regulating the degree of feature retention of point cloud data. The larger the value of p_i is, the better the feature retention effect of the point cloud model will be. In general, σ_1 is the neighborhood radius of the point, and σ_2 is the standard deviation of the neighborhood point.

The specific algorithm steps are as follows:

(1) Search k-nearest neighbor points of each data point p_i in the region with drastic curvature changes in the point cloud model, that is, in the feature-rich region;

(2) For all its nearest neighbor points, the value of $x = \|p - p_i\|$ in parameter $z_1(x)$ and $y = \langle p - p_i, n\rangle$ in parameter $z_2(y)$ are calculated;

(3) Calculate $z_1(x)$ and $z_2(y)$ according to the calculation formula of 6 and 7 and the selected value of 8;

(4) The values of $z_1(x)$ and $z_2(y)$ calculated in Step (3) are substituted into the calculation formula of bilateral filter factor β to calculate bilateral filter factor β;

(5) Then, formula (14) is used to move the points in the feature-rich region normally to obtain the filtered new point cloud data;

(6) After all points are calculated in turn and new point cloud data are obtained through transformation, end.

2.2.3. D Image Reconstruction Based on Time Series. In the real world, the collection of a large number of data is related to time, and the data are time related. These kinds of data are called time series. Strictly speaking, time series refers to a group of numerical sequences formed by sequential observation values of the same phenomenon at different times. It is a limited sequence of real values recorded in the direction of the time axis, reflecting the characteristics of attribute values in the time order. If the data series is continuous, it is called continuous time series [31]. If the data sequence is discrete, it is called a discrete time series. In daily life, a large amount of time-series data will be generated in different fields, which can be referred to as "time-series data." By collecting, recording, and organizing these time-series data and using advanced data mining tools, we can find a lot of valuable knowledge from time series, which can be used to guide our work and life. Therefore, 3D reconstruction of intelligent manufacturing image is carried out in this paper with the help of CAD.

By analyzing the structure of CAD-shaped DXF file, the corresponding object information can be extracted. In intelligent manufacturing image object data, the starting point coordinates of a line are the corresponding following values after block codes 10, 20, and 30, and the end coordinates are the corresponding following values after block codes 11, 21, and 31. In the entity extraction process, the first judgment is whether the group code is 10. If so, it indicates that the subsequent group value is the entity coordinate. The extracted object is projected (converted) onto the image according to the transformation parameters, and the projected value is taken as the initial value of the least square template matching. As the objects in some images may occlude each other, it is necessary to eliminate the lines that occlude each other. On this basis, FLANN matching algorithm is used to extract the feature points of intelligent manufacturing time-series image and complete the image matching. AKAZE algorithm introduces FED, a fast display diffusion mathematical framework, to solve partial differential equations quickly, which improves the solving speed and quality.

The nonlinear diffusion filter describes the evolution of image brightness and controls the diffusion process using the scaling parameter as the divergence factor of thermal diffusion function. Usually, partial differential equations are used to solve the problem [32]. Owing to the nonlinear nature of differential equations, the scale space is constructed through the diffusion of image brightness. The classical nonlinear diffusion equation is as follows:

$$\frac{\partial L}{\partial t} = \text{div}\left(c\left(x, y, t\right) * \nabla L\right). \tag{19}$$

In the aforementioned formula, L is the image brightness matrix, and div and ∇ represent the solution operation of divergence and gradient, respectively. Owing to the introduction of conduction function c into the diffusion equation, the diffusion can be adaptive to the local structure characteristics of the image. The conduction function depends on the local difference structure of the image and can be scalar or constant. The time parameter t corresponds to the scale factor and is controlled by image gradient in the diffusion process [33].

Assuming that the number of layers and towers are represented by O and S, respectively, the scale space is constructed by performing time t_i of the diffusion function:

$$\sigma_i(o, s) = 2^{(o+s/S)}. \tag{20}$$

AKAZE feature detection is also achieved by looking for Hessian local maximum points normalized at different scales [34, 35]. In the same scale, response values of eight field points and 18 upper-- and lower-level points are identified as extreme points, namely feature points. Hessian matrix formula is as follows:

$$L_H^i = \left(\frac{\sigma_i}{2^{oi}}\right)^2 * \left(L_{xx}^i L_{yy}^i - \left(L_{xy}^i\right)^2\right). \tag{21}$$

where σ_i represents the scale parameter, o is the corresponding group, L_{xx} and L_{yy} are the transverse and longitudinal differential, and L_{xy} represents the second-order cross-differential.

Based on the successful extraction of image feature points, FLANN matching algorithm is used to complete the rough matching process [36]. The specific process is as follows:

(1) Find the feature point P_1 in the image R_1 and the feature point P_1 in the image R_2, and get the feature set (P_1, P_2), traverse the images of all feature points R_1, be made up of all minimum distance matching points for collection, calculate the minimum distance from the collection of min d, $T = \mu \times \min d$, setting the threshold value if $d < T$, p_2 as p_1 candidates will match point.

(2) Perform step (1) for all feature points in image R_2 to obtain FLANN matching point pair (p_2, p_3) from figure R_2 to figure R_1. If $p_3 = p_1$, the matching is successful; otherwise, the point is removed, and the cycle continues until all image matching is completed.

According to the geometric constraints between images, the point cloud model is reconstructed by triangulation. On the basis of the camera motion, the 3D spatial position of feature points can be estimated. However, the depth information of pixels cannot be obtained only from a single image, so the method of triangulation (triangulation) is adopted to estimate. The principle of trigonometry is to determine the distance of a point by observing the angle of the same point in two places.

Consider that there are two images I_1 and I_2, and their camera light centers are O_1 and O_2. Taking the image on the left as a reference, the transformation matrix of the image on the right can be determined as T. Assuming that feature point P_1 in figure I_1, and I_2 of the feature points P_1 form corresponding, regardless of other factors, under ideal conditions, O_1P_1 and O_2P_2 will intersect at one point in the space, which is two feature points of the map point position in the 3D scene but cannot be totally eliminated due to the effects of noise, the two straight lines generally cannot smooth rendezvous in the actual situation, so the least square method is used to find the point closest to the intersection of the two lines. According to the definition in polar geometry theory, if x_1 and x_2 are set as the normalized coordinates of the two feature points, the relationship between them is as follows:

$$s_1 x_1 = s_2 R x_2 + t. \tag{22}$$

In the aforementioned formula, s_1 and s_2 are the depths of the two feature points. Multiply the left and right sides of the aforementioned formula by $\hat{x}_1$ to obtain the following formula:

$$\begin{aligned} s_1 \hat{x}_1 x_1 &= 0 \\ &= s_2 \hat{x}_2 r x_2 + \hat{x}_1 t. \end{aligned} \tag{23}$$

According to the formula, we can solve for s_1 and s_2 in turn. In practice, the influence of noise cannot be completely eliminated, so solving R and t may not make the result of the aforementioned equation exactly 0. Therefore, in practice, it is more common to replace the zero solution with the least square solution.

The point cloud results obtained by SFM algorithm can be optimized by means of clustering multiview stereo (CMVS) algorithm for image clustering classification. CMVS can reduce the time of dense matching, and on this basis, dense point clouds can be obtained through patch-based multiview STEREO (PMVS) processing. There are three necessary conditions for clustering: the clustering should be as small as possible under the condition that every cluster can be reconstructed; maintain density while removing mismatches and redundancies; and the result of reconstruction should be relatively complete. The overall process of the algorithm is shown in Figure 1.

The purpose of SFM screening is to screen out some points and take out the average value of each field location of feature points. Select images with high resolution and remove images that do not meet the requirements. The standard image segmentation method was adopted to segment the screened images, and the SFM feature points that were not included were constructed to construct image clusters. The efficiency basis was set, and the maximum efficiency was regarded as the behavior of the added clusters. The process of clustering classification and increasing images is repeated to obtain clusters that meet the requirements.

PMVS is a dense matching based on facet model, which can reconstruct the cluster of CMVS independently. Patches are generated based on sequential images and feature points,
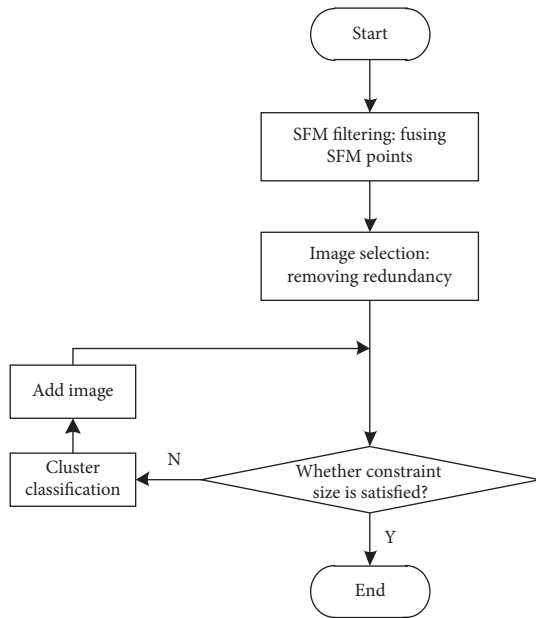
Figure 1: CMVS flow chart.

which are called seed points. The algorithm flow is shown in Figure 2.

It can be found that the algorithm consists of three parts: the first part is the generation of seed surface, the second part is the diffusion of surface, and the third part is the filtering of surface. The last two parts are iterated three times according to experience. After three iterations of surface diffusion and surface filtering, the number of outpoints in the diffusion reconstruction can be reduced, and the result is better after the third iteration. The normal section of a part of the surface of the target object represents a plane, and the seed plane refers to the sparse reconstruction of the spatial point cloud plane. First, corner detection is carried out on the sequence, and the obtained feature corner points are matched and 3D information is recovered according to triangulation measurement, and finally the face is generated. This process is called sparse reconstruction of 3D point cloud. For a given facet, it is impossible to ensure that all image blocks are included in the acquisition of facet, so the results of sparse reconstruction are extended to ensure that each image block has at least one facet. In this process, there will be wrong surfaces, which need to be filtered. The filtering is based on gray consistency and geometric consistency.

To sum up, this paper mainly uses Kinect sensor to collect depth image data and convert it into 3D point cloud coordinates. The collected point cloud data are divided into regions and the 3D point cloud data are filtered and denoised. With the help of CAD, FLANN matching algorithm was used to extract the feature points of time-series images, and then the images were matched. Combined with the image matching results, 3D reconstruction of sparse point cloud and dense point cloud was carried out to complete 3D reconstruction of intelligent manufacturing images.

3. Experimental Design and Result Analysis

3.1. Experimental Design. To verify the 3D reconstruction effect of CAD-aided intelligent manufacturing image based on time series designed in this paper, an experimental design is carried out. To ensure the reliability of the experimental results, this experiment needs to be carried out in the same environment, and the specific experimental environment parameters are shown in Table 1.

Multiple intelligent manufacturing images were collected, and the collected images were integrated to construct the experimental sample set. The experimental sample set was input into the computer, and the data in the sample set were processed using the reconstruction method based on P2M framework improvement, the reconstruction method based on vanishing point optimization and the reconstruction method based on time series. In this way, 3D reconstruction results of intelligent manufacturing images with different methods are obtained.

3.2. Analysis of Experimental Results. The PSNR of intelligent manufacturing image is one of the important indicators to verify the 3D reconstruction effect of intelligent manufacturing image. The higher the PSNR is, the better the 3D reconstruction effect of the image is. The specific comparison results are shown in Figure 3.

The analysis of the data in Figure 3 shows that with the increase of the amount of experimental data, the PSNR of images reconstructed by different methods changes in a fluctuating manner. Among them, the PSNR of images reconstructed by the improved P2M framework changes between 37.5 dB and 44.1 dB. Image reconstruction method based on vanishing point optimization PSNR between 33.9 dB and 40 dB changes, and reconstruction method based on improved P2M framework, based on the reconstruction of the vanishing point optimization method, the reconstruction method based on time-series images of the PSNR is always above 52 dB, reconstruction after using this method of intelligent manufacturing high image quality, intelligent manufacturing image 3D reconstruction effect is better.

The 3D reconstruction time of intelligent manufacturing image is one of the important indicators to verify the 3D reconstruction efficiency of intelligent manufacturing image. The lower the 3D reconstruction time of intelligent manufacturing image, the higher the 3D reconstruction efficiency of intelligent manufacturing image. The specific comparison results are shown in Figure 4.

By analyzing the data in Figure 4, it can be seen that with the increase of the amount of experimental data, the time of 3D reconstruction of intelligent manufacturing images by different methods presents an increasing trend. When the amount of experimental data is 800 MB, the reconstruction methods based on P2M frame improvement, vanishing point optimization, and
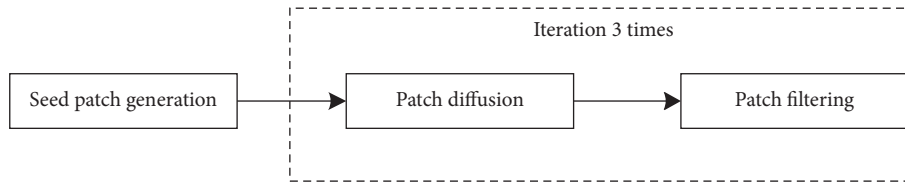
FIGURE 2: Flow chart of PMVS algorithm.

TABLE 1: Experimental environmental parameters.

Environmental parameters	Configuration	Parameter
	CPU	Intel(R)Core(TM)i5-9400
Hardware environment	Frequency	2.90 GHz
	RAM	16.0 GB
	Operating system	Windows 10
	Version	18,362.1082 pro
Software environment	Digits	64 bit
	Analog software language	APDL
	Simulation software	MATLAB 7.2
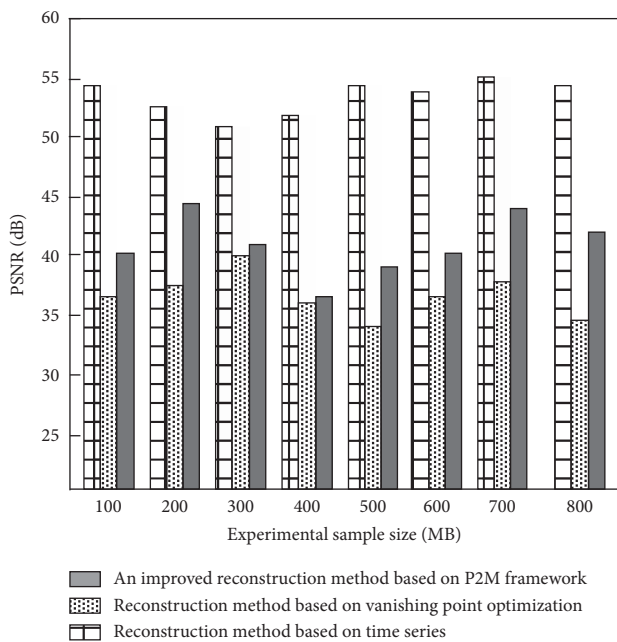
FIGURE 3: Comparison of PSNR of different methods.

FIGURE 4: Comparison of reconstruction time between different methods.

time series have reached the maximum time of 3D reconstruction of intelligent manufacturing image. Among them, the maximum time of 3D reconstruction of intelligent manufacturing image based on the improved reconstruction method of P2M frame is 9.3 s, vanishing point optimization is 7.7 s, and time series is 4.9 s. Compared with the reconstruction methods based on P2M frame improvement and vanishing point optimization, the reconstruction time of image 3D reconstruction based on time series always remains at a low level, indicating that this method has a very high efficiency of 3D reconstruction of intelligent manufacturing images.
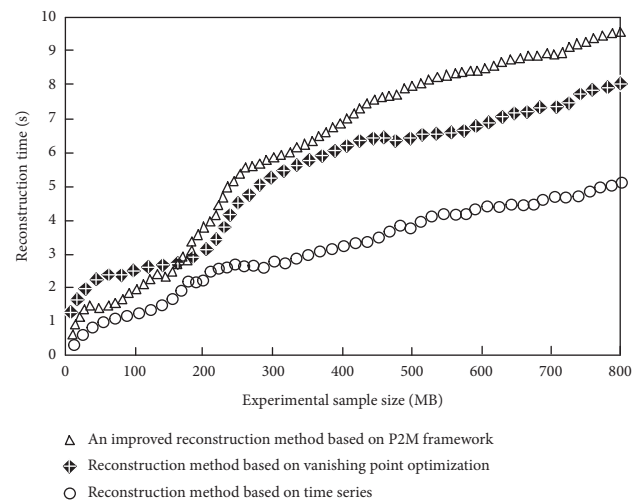
To further verify the 3D reconstruction effects of intelligent manufacturing images by different methods, two images are taken as examples to verify the 3D reconstruction of intelligent manufacturing images by different methods, and the results are shown in Figures 5 and 6.

By analyzing the 3D reconstruction results of CAD assisted intelligent manufacturing images in Figures 5 and 6, the reason is that this method uses different point cloud denoising algorithms to filter and denoise the divided area. With the help of CAD, FLANN matching algorithm is used to extract feature points of time-series images and complete image matching. Three-dimensional reconstruction of sparse point cloud and dense point cloud is respectively carried out to complete 3D reconstruction of intelligent manufacturing images.
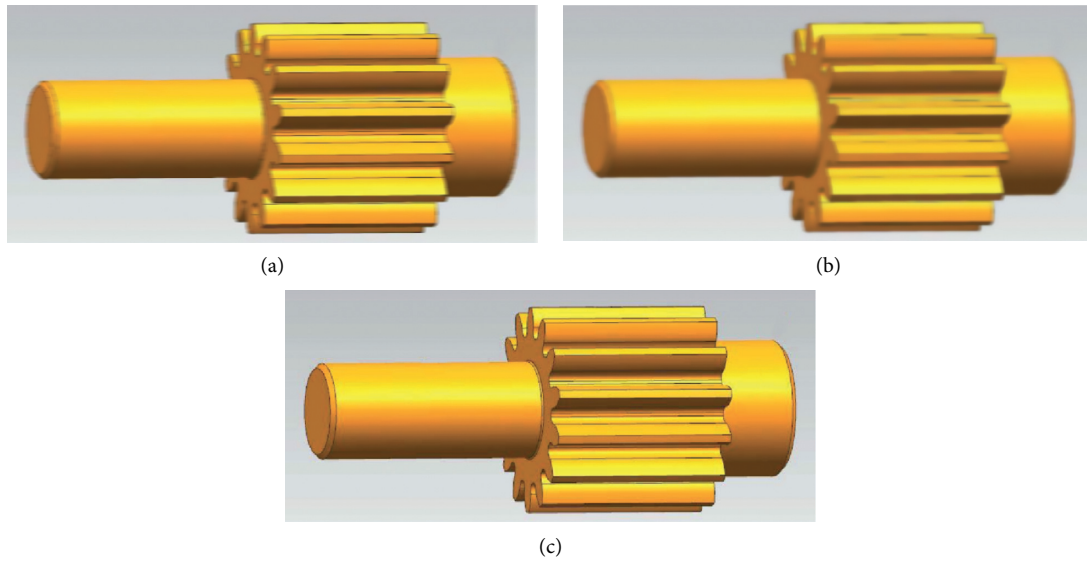
FIGURE 5: 3D reconstruction effect of image A: (a) improved reconstruction method based on P2M framework, (b) reconstruction method based on vanishing point optimization, and (c) reconstruction methods based on time series.
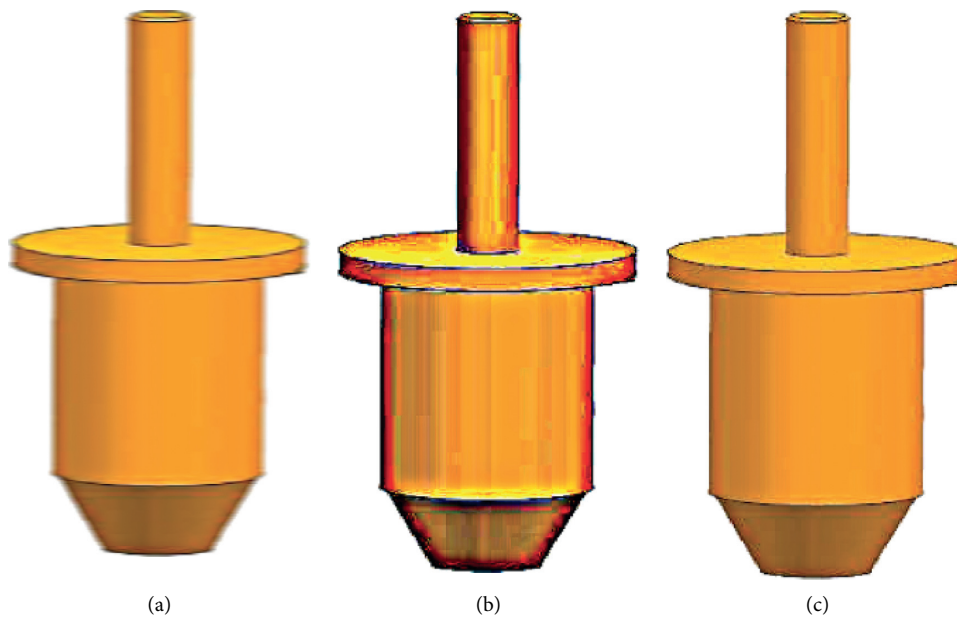

FIGURE 6: 3D reconstruction effect of image B: (a) improved reconstruction method based on P2M framework, (b) reconstruction method based on vanishing point optimization, and (c) reconstruction methods based on time series.

4. Conclusion

As the economic development situation enters the new normal, the manufacturing industry faces increasing external constraints such as resources and environment, and the manufacturing cost of enterprises keeps increasing. At this time, the extensive development path that mainly relies on input of factor resources is difficult to continue. Turn the current situation, structure, methods, promoting development, and it is urgent to form a new economic growth, shape the new international competition advantage focus on manufacturing, and in the process of manufacturing production, image clarity is to ensure that all production can proceed smoothly, the basis of which can accurately control the production progress and the problems existing in the production process, so to improving the quality of industrial production. Research of intelligent image 3D reconstruction method is of great significance, so in this paper, based on the time series of CAD auxiliary intelligent manufacturing image 3D reconstruction method, the experimental results show that the method of image peak signal-to-noise ratio is always above 52 dB, intelligent manufacture image 3 d reconstruction takes a maximum of 4.9 s, and intelligent manufacture image 3D reconstruction effect is better.

Data Availability

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

Acknowledgments

References

[1] A. K. Bashir, S. Mumtaz, V. G. Menon, and K. F. Tsang, "Guest editorial: cognitive analytics of social media for industrial manufacturing," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2899–2901, 2021.

[2] D. Zhu, Z. Xu, X. Xu, Q. Zhao, L. Qi, and G. Srivastava, "Cognitive analytics of social media services for edge resource pre-allocation in industrial manufacturing," *IEEE Transactions on Computational Social Systems*, vol. 8, no. 2, pp. 500–511, 2021.

[3] P. A. Bucur, P. Armbrust, and P. Hungerländer, "On the propagation of quality requirements for mechanical assemblies in industrial manufacturing," *Expert Systems with Applications*, vol. 174, no. 15, pp. 114608–114619, 2021.

[4] F. Ji, "Research on the construction of an intelligent manufacturing maturity model based on the three basic paradigms of "digitalization networking and intelligence"," *Modern Management*, vol. 11, no. 7, pp. 761–773, 2021.

[5] D.-D. Li, W.-M. Zhang, Y.-S. Li, F. Xue, and J. Fleischer, "Chatter identification of thin-walled parts for intelligent manufacturing based on multi-signal processing," *Advances in Manufacturing*, vol. 9, no. 1, pp. 22–33, 2021.

[6] Y. F. Zhuang and W. G. Wan, "3D reconstruction of single image based on p2m framework," *Electronic Measurement Technology*, vol. 43, no. 9, pp. 61–64, 2020.

[7] Z. Liu, J. Y. Wang, and L. J. Chen, "3D reconstruction of single image based on vanishing point optimization," *Journal of Zhejiang University of technology*, vol. 47, no. 2, pp. 180–185, 2019.

[8] C. H. Zhao, R. C. Zhao, and S. Feng, "A 3D reconstruction method based on sea ice scene image classification," *Journal of Engineering of Heilongjiang University*, vol. 11, no. 2, pp. 64–73, 2020.

[9] H. Yu, J. M. Yang, and X. G. Wang, "Research on flange 3D reconstruction method based on engineering drawing image," *Mechanical design and manufacturing*, no. 11, pp. 221–223, 2020.

[10] X. Liu, M. Y. Zeng, Y. C. Duan, and G. Fan, "Research on 3D reconstruction method of fuzzy image edge contour based on machine learning," *Electronic testing*, vol. 17, no. 1, pp. 39–41, 2019.

[11] E. Hannink, T. Shannon, H. Dawes, and K. Barker, "The validity of the kinect sensor for the measurement of sagittal spine curvature against the gold standard lateral spinal radiograph," *Physiotherapy*, vol. 107, no. 5, pp. e28–e29, 2020.

[12] Y. M. Lee, S. Lee, K. E. Uhm, G Kurillo, J. J Han, and J Lee, "Upper limb three-dimensional reachable workspace analysis using the kinect sensor in hemiplegic stroke patients: a cross-sectional observational study," *American Journal of Physical Medicine & Rehabilitation*, vol. 99, no. 5, pp. 397–403, 2020.

[13] Y. Tian, G. Wang, L. Li, T. Jin, F. Xi, and G. Yuan, "A universal self-adaption workspace mapping method for human-robot interaction using kinect sensor data," *IEEE Sensors Journal*, vol. 20, no. 99, pp. 1–10, 2020.

[14] C. L. Ramirez, C. Escobedo, J. Edwin, D. Miranda, D. Menotti, and G. Chávez, "A multimodal LIBRAS-UFOP Brazilian sign language dataset of minimal pairs using a microsoft Kinect sensor," *Expert Systems with Applications*, vol. 167, no. 1, pp. 114179–114183, 2020.

[15] T. Guzsvinecz, V. Szucs, and C. Sik-Lanyi, "Suitability of the kinect sensor and leap motion controller—a literature review," *Sensors*, vol. 19, no. 5, pp. 1–25, 2019.

[16] Y. Ma, D. Liu, and L. Cai, "Deep learning-based upper limb functional assessment using a single kinect v2 sensor," *Sensors*, vol. 20, no. 7, pp. 1903–1921, 2020.

[17] T. L. Banh and V. B. Bui, "First experiences with microsoft kinect V2 for 3D modelling of mechanical parts," *Applied Mechanics and Materials*, vol. 889, no. 1, pp. 329–336, 2019.

[18] J. B. Qu, Y. Wang, and Q. Zhao, "Application of DBSCAN clustering and improved bilateral filtering algorithm in point cloud denoising," *Surveying and Mapping Bulletin*, vol. 12, no. 11, pp. 89–92, 2019.

[19] J. Wang, L. I. Huajian, and H. Wang, "Research on tunnel point cloud denoising based on centerline fitting," *Journal of Shenyang Jianzhu University(Natural Science)*, vol. 35, no. 2, pp. 676–683, 2019.

[20] P. H. Casajus, T. Ritschel, and T. Ropinski, "Total denoising: unsupervised learning of 3D point cloud cleaning," *IEEE*, vol. 8, no. 1, pp. 1–11, 2019.

[21] Y. Zhou, R. Chen, Y. Zhao, X. Ai, and G. Zhou, "Point cloud denoising using non-local collaborative projections," *Pattern Recognition*, vol. 120, no. 81, pp. 108128–108139, 2021.

[22] M. A. Irfan and E. Magli, "Exploiting color for graph-based 3D point cloud denoising," *Journal of Visual Communication and Image Representation*, vol. 75, no. 1, pp. 103027–103039, 2021.

[23] W. Hu, X. Gao, G. Cheung, and Z. Guo, "Feature graph learning for 3D point cloud denoising," *IEEE Transactions on Signal Processing*, vol. 68, no. 99, pp. 1–10, 2020.

[24] Y. Guo, S. Guo, K. Guo, and H. l. Zhou, "Seismic data denoising under the morphological component analysis framework by dictionary learning," *International Journal of Earth Sciences*, vol. 110, no. 5, pp. 963–978, 2021.

[25] L. Zhang, L. Han, A. Chang, and J. Fang, "Seismic data denoising using double sparsity dictionary and alternating direction method of multipliers," *Journal of Seismic Exploration*, vol. 29, no. 1, pp. 49–71, 2020.

[26] X. Wang and J. Ma, "Adaptive dictionary learning for blind seismic data denoising," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 7, pp. 1273–1277, 2020.

[27] L. Li, M. Cai, X. Guan, and D. Chu, "Piecewise adaptive-norm trend filtering method for ICESat/GLAS waveform data denoising," *IEEE Access*, vol. 8, no. 1, pp. 168965–168979, 2020.

[28] M. Gu, R. Xie, and L. Xiao, "A novel method for NMR data denoising based on discrete cosine transform and variable

length windows," *Journal of Petroleum Science and Engineering*, vol. 25, no. 1, pp. 1–12, 2021.

[29] Z. Peng, S. Peng, L. Fu et al., "A novel deep learning ensemble model with data denoising for short-term wind speed forecasting," *Energy Conversion and Management*, vol. 207, no. 6, pp. 112524–112535, 2020.

[30] J. W. Zheng and Z. Deli, "Seismic data denoising based on data-driven tight frame dictionary learning method," *Global Geology*, vol. 23, no. 04, pp. 45–50, 2020.

[31] M. Cao, L. Zheng, and X. Liu, "Single view 3D reconstruction based on improved RGB-D image," *IEEE Sensors Journal*, vol. 20, no. 20, pp. 12049–12056, 2020.

[32] S. Croix, P. Deckers, and S. M. Sindbæk, "Recasting a Viking warrior woman from Ribe: 3D digital image reconstruction compared," *Journal of Archaeological Science: Report*, vol. 32, no. 2, pp. 102455–102469, 2020.

[33] C. Liu and G. Liu, "Characterization of pore structure parameters of foam concrete by 3D reconstruction and image analysis," *Construction and Building Materials*, vol. 267, no. 1, pp. 120958–120972, 2020.

[34] M. Shahbazi, P. Ménard, G. Sohn, and J. Théau, "Unmanned aerial image dataset: ready for 3D reconstruction," *Data in Brief*, vol. 25, no. 1, pp. 103962–103975, 2019.

[35] H. X. Wu and Z. Zhang, "3D image reconstruction method based on absolute conic image," *Computer simulation*, vol. 38, no. 8, pp. 203–206, 2021.

[36] S. T. Wan, B. L. Zhang, and T. Yin, "Design of 3D reconstruction software for X-ray nondestructive testing image," *Chinese Journal of Construction Machinery*, vol. 18, no. 5, pp. 425–429, 2020.