

ORIGINAL ARTICLE

Open Access



Fast Estimation of Loader's Shovel Load Volume by 3D Reconstruction of Material Piles

Binyun Wu¹, Shaojie Wang^{1,2}, Haojing Lin¹, Shijiang Li¹ and Liang Hou^{1*}

Abstract

Fast and accurate measurement of the volume of earthmoving materials is of great significance for the real-time evaluation of loader operation efficiency and the realization of autonomous operation. Existing methods for volume measurement, such as total station-based methods, cannot measure the volume in real time, while the bucket-based method also has the disadvantage of poor universality. In this study, a fast estimation method for a loader's shovel load volume by 3D reconstruction of material piles is proposed. First, a dense stereo matching method (QORB–MAPM) was proposed by integrating the improved quadtree ORB algorithm (QORB) and the maximum a posteriori probability model (MAPM), which achieves fast matching of feature points and dense 3D reconstruction of material piles. Second, the 3D point cloud model of the material piles before and after shoveling was registered and segmented to obtain the 3D point cloud model of the shoveling area, and the Alpha-shape algorithm of Delaunay triangulation was used to estimate the volume of the 3D point cloud model. Finally, a shovel loading volume measurement experiment was conducted under loose-soil working conditions. The results show that the shovel loading volume estimation method (QORB–MAPM VE) proposed in this study has higher estimation accuracy and less calculation time in volume estimation and bucket fill factor estimation, and it has significant theoretical research and engineering application value.

Keywords Loader, Volume estimation, Binocular stereo vision, 3D terrain reconstruction, Point cloud registration and segmentation

1 Introduction

A loader's earthwork is a mechanized operation process that necessitates productivity planning, project progress evaluation, and labor fund release. It is important to quantify the amount of earthwork materials using various technologies and methodologies and then use it as a key parameter to evaluate loader operation and performance [1, 2]. For instance, according to the amount of material shoveled at a time, the time required to finish the entire earthwork operation is estimated and used as a reference

to estimate the labor compensation to be provided to operators [3, 4].

An entire measurement or a single shovel measurement of the earthwork stockpile can be used to determine the volume of the earthworks. The total station robot scanner is widely employed in the comprehensive measurement of earthworks as an efficient 3D modeling and measurement technology. The volume is estimated by employing the total station to reconstruct the 3D surface of the entire earthwork material surface while scanning the 3D coordinates of the relevant feature points [5]. Furthermore, a photogrammetric technique based on UAV aerial photography can be used to provide 3D information on the material surface, which is suitable for measuring the entire volume of earthwork [3, 6]. However, during the actual operation of earthwork, engineering construction managers often pay greater attention to

*Correspondence:

Liang Hou
hliang@xmu.edu.cn

¹ Pen-Tung Sah Institute of Micro-Nano Science and Technology, Xiamen University, Xiamen 361102, China

² Shenzhen Research Institute of Xiamen University, Shenzhen 518057, China



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

the project's ongoing progress throughout the operation process and assess the efficiency of the project operation. The entire measurement method is suitable for planning the productivity of the entire earthwork. However, it is challenging to assess the operational efficiency in real time, necessitating a single shovel measurement. Existing research on single shovel measurements mainly focuses on constructing a 3D point cloud model of the bucket and the materials in the bucket, estimating the volume of materials in the bucket, and evaluating the bucket filling rate and shovel loading efficiency [7–10]. However, buckets of different earth-moving machines have different specifications and sizes. The limitations of this method for estimating the bucket loading volume based on the bucket 3D point cloud model include its poor generality and high reliance on the bucket model. Therefore, it is critical to investigate the single shoveling volume measurement method of loaders based on the 3D reconstruction of material piles to evaluate the ongoing project progress and the operating efficiency of operators.

Nowadays, 3D laser scanning systems and visual sensors are commonly utilized to gather object volume information using non-contact measurements to build a 3D reconstruction of an object's surface [11, 12]. Yakar et al. [13] used a ground laser scanner to create a 3D point cloud and estimate the volume on the surface of shovel materials at a construction site, and then compared the results with photogrammetry. The results showed that the laser scanning method had a higher measurement accuracy. However, there are some limitations to using a laser scanning system for the 3D reconstruction of complex object surfaces. Because the amount of point cloud data generated using the laser scanning system for object scanning is significantly greater than that of the visual sensor, post-processing and 3D reconstruction of the 3D point cloud will be a more time-consuming task, and it is challenging to meet the requirements of real-time updating of object volume information. Additionally, the application of a 3D laser scanning system in earthwork is somewhat constrained by its limitations in terms of shape and size sensitivity, inability to gather color information, and cost.

To resolve the aforementioned issues, 3D reconstruction employing visual sensors offers an advantageous approach for the real-time estimation of single shovel loads in earthworks. The visual sensor can detect dynamic changes in the surrounding environment faster than laser radar because it can record the working site environment more quickly. Additionally, vision-based 3D reconstruction requires less computation time, is less expensive, and is lighter than laser-scanning-based 3D reconstruction. As a result, vision-based methods are increasingly being employed in the 3D modeling of

earthwork environments. Currently, the three primary types of vision-based 3D reconstruction technologies are structured light, time-of-flight, and stereo vision. Structured light 3D measurement technology has the advantages of simple hardware configuration, fast measurement speed, and high reconstruction accuracy. However, its applicability in outdoor earthwork is constrained by its failure to function in environments with strong light sources. The time-of-flight vision system works on the principle of calculating the depth information of an object in relation to the length of time it takes for a pulse of light (generally invisible light) to travel from the point of emission to the point of reception after being reflected by the object. However, owing to the fast propagation speed of light, the measurement accuracy is relatively low and the 3D reconstruction effect of the object surface is poor. Stereo vision has been extensively used as a typical passive measurement technique. Binocular stereo vision is one such method that employs two cameras to imitate human eyes to take pictures of the object being measured from various angles, find the corresponding matching points of the same spatial point in the left and right camera images based on image characteristics, and then compute and solve the spatial coordinate points using the parallax principle and spatial correspondence. This technique provides dense depth information quickly and at a low cost. It is suitable for 3D reconstruction of earthwork material piles [14, 15].

Earthmoving machinery such as loaders usually operates in an unstructured terrain environment that changes dynamically in real time. Their operating scenes are characterized by complex working conditions and harsh environments, which pose a significant challenge to earthwork volume estimation based on the 3D reconstruction of stereo vision. On the one hand, it is difficult to reconstruct earthwork material piles in 3D because of their irregular appearance, lack of texture on the surface, different soil softness, and poor lighting conditions. However, with the continuous operation of earthwork machinery, the local characteristics of the material pile shape change significantly, which puts forward higher requirements for the real-time update of earthwork volume estimation. Therefore, a fast estimation method for the load volume of a loader by 3D reconstruction of material piles using binocular stereo vision is proposed in this study. First, the improved quadtree ORB algorithm (QORB) was utilized to extract the feature points in the surface of material piles to improve the feature point detection uniformity, solving the issues of the traditional ORB algorithm, such as overly centralized and overlapping feature point detection and insufficient feature description ability. The homography matrix transformation was then applied

to accomplish feature point matching correction, and more sparse matching feature points were created to improve the accuracy of the 3D reconstruction of the material piles and shovel loading volume calculation. Subsequently, the robust sparse feature points and parallax matched by the QORB algorithm were used as prior knowledge. The dense parallax map was then calculated by creating a MAPM using Bayesian estimation, and the material piles were 3D reconstructed using the binocular vision stereo imaging principle. Subsequently, the ICP algorithm was used to register the 3D point cloud on the surface of the material piles, both before and after shoveling. The point cloud model of the shoveling area was then obtained according to the boundary constraint segmentation, and the volume of shoveling was estimated using the Alpha-shape algorithm of Delaunay triangulation. Finally, the shovel volume measurement experiment in loose soil conditions verifies the effectiveness of this research method.

This study provides the following contributions. Firstly, the authors propose a fast estimation method for the volume of a loader's shovel load through 3D reconstruction of the material pile surface before and after shoveling. It addresses the limitations of existing techniques, such as the inability of entire volume measurement methods using total stations or UAVs to assess operational efficiency in real-time. Additionally, it overcomes the strong reliance on bucket models and the limited applicability of single-shovel measurement methods for buckets of different specifications and sizes. While visual-based 3D reconstruction technology has been widely applied in various fields, this study represents a novel attempt to apply visual-based 3D reconstruction of the material surface before and after loading to estimate the volume of a loader's shovel load. Secondly, a dense stereo matching method based on the QORB-MAPM algorithm is proposed. This method offers significant advantages in terms of reconstruction accuracy and computational efficiency, providing robust technical support for the fast and accurate measurement of a loader's shovel load volume. Furthermore, the proposed novel method for estimating a loader's shovel load volume can also serve as a valuable reference and guide for addressing similar issues in other types of construction machinery.

The remainder of this paper is organized as follows. Section 2 presents related works on feature matching, 3D reconstruction, and volume estimation. Section 3 explains the methodology used in the proposed scheme. The experimental results and discussion are presented in Section 4. Finally, conclusions and prospects are presented in Section 5.

2 Related Work

2.1 Feature Matching and 3D Reconstruction

Image feature matching, a critical component of 3D reconstruction methods based on stereo vision, has a direct effect on the accuracy and computational efficiency of 3D reconstruction and volume estimates. Stereo matching algorithms are often divided into three categories: global stereo matching algorithm, semi-global stereo matching algorithm, and local stereo matching algorithm. Global stereo matching offers greater matching accuracy, but its applicability in 3D reconstruction is severely constrained by its high computational cost and memory usage. The semi-global stereo matching algorithm has been proposed and is widely used to increase the computational efficiency of the global stereo matching algorithm. Semi global block matching (SGBM) algorithm [16] is one well-known example. Its basic concept is to first determine the disparity of the image and then establish an energy generation value function related to the global image on the disparity map in accordance with the smoothing constraint criteria of the pre-selected scan. The method has higher matching accuracy and uses less computing power than the global stereo matching method. However, the algorithm still struggles to meet the requirements of real-time updating of earthwork volume estimation during loader shovel loading operations due to the complicated, real-time, and dynamic terrain environment. Local feature point stereo matching, in contrast to the approaches mentioned above, determines the local optimal matching cost by detecting a constrained set of pixels and is popular due to its superior computing efficiency.

The Seale invariance feature transform (SIFT) algorithm [17], speed up robust features (SURF) algorithm [18], and ORB algorithm [19] are the three primary traditional feature-matching algorithms. The improved oFAST and rBRIEF algorithms comprise the ORB algorithm [20, 21]. In this method, the key points in the image are identified using oFAST corners to detect pixels that differ considerably from the surrounding pixels. Subsequently, the rBRIEF description for each key-point is calculated. This algorithm, a classic example of a binary description technique, outperforms the SIFT and SURF algorithms in terms of computing efficiency. Deep learning-based image processing techniques are also widely used to address image matching issues in addition to traditional image matching methods based on manual labeling. A large number of deep learning networks have also been used to extract image features. For example, Chaudhuri et al. [22] proposed a new architecture based on deep neural networks to solve the cross-modal information retrieval problem in remote

sensing, which can learn a distinguishable shared feature space for all input modes, and is suitable for information retrieval with consistent semantics. Li et al. [23] proposed a large-scale remote sensing image retrieval method based on deep hash neural networks (DHNN), which can automatically learn feature extraction operations and feature hash mapping under the supervision of labeled samples. Furthermore, many deep learning-based feature matching methods [24–26] apply graph neural networks (GNNs) to aggregate neighborhood information and form a structured representation of nodes and edges, showing good matching performance.

Deep learning-based methods can extract subtle low-level features and abstract high-level features, which can describe features more accurately. Deep learning networks, on the other hand, typically have complex structures and high computing costs, and they cannot meet the needs of real-time computing without graphics processing unit (GPU accelerators). Traditional feature description methods based on manual labeling, such as the ORB algorithm, have significant advantages in computing speed and cost, making them more suitable for 3D reconstruction of earthwork material piles and shovel loading volume estimation in real-time. However, this method has the following drawbacks: 1) The ORB algorithm employs the oFAST algorithm to detect feature points by comparing a pixel's gray value to its surroundings. This comparison allows the program to identify feature points in the image that are more noticeable than surrounding pixels. However, the algorithm employs an arbitrarily selected global fixed threshold, causing the detected feature points to be excessively clustered or even overlapped in areas with obvious features and excessively sparse or even empty areas without obvious features, resulting in a loss of information in some image areas. 2) The ORB algorithm describes the feature points using an improved rBRIEF algorithm. Due to the low feature description ability of the binary description method, it is difficult to obtain more correct matching feature points in the feature point matching phase, which seriously affects the accuracy of 3D reconstruction and volume estimation of the material pile.

In summary, it is necessary to study the feature matching method with higher computing efficiency and stronger robustness to realize the 3D real-time reconstruction of the material piles in the earthwork operation of the loader in the face of the characteristics of the material piles with an irregular shape and no surface texture, as well as the real-time dynamic change of the material pile surface characteristics during the earthwork operation.

2.2 Volume Estimation

Stereo vision technology has been widely used for the 3D measurement of objects in different fields, such as industrial manufacturing [27–29], agriculture and fishing [30, 31], and aerospace [32, 33]. Research on earthwork volume measurements using visual techniques is gaining increasing interest. Bügler et al. [3] estimated the production efficiency of engineering activities by determining the entire amount of earth that had been dug using photogrammetric technology and combining it with video data. Borthwick et al. [34] installed a stereo camera on the large arm of a mechanical excavator to shoot the truck bucket at a distance and calculated the load volume by constructing a 3D surface model before and after the truck loaded materials. Yakar et al. [35] rectified the point cloud using the placed control points, then, based on the corrected point cloud, constructed a grid 3D model of the gravel pile and computed the volume. The aforementioned study may be used to plan the overall productivity of earthwork; however, it is challenging to evaluate operational efficiency in real time. Anwar et al. [7] proposed the use of stereo vision to estimate the volume of materials in an excavator bucket and verified the effectiveness of the method through simulation and field tests. Guevara et al. [8] used a binocular stereo camera to construct a 3D point cloud on the material surface of a bulldozer bucket, and used the Alpha-shape algorithm of Delaunay triangulation to estimate the effective shovel load of the bucket. Lu et al. [9, 10] developed a new perception system based on the stereo vision perception method, as well as advanced technologies such as point cloud registration, splicing, and surface interpolation, to realize the 3D point cloud reconstruction of materials in the loader bucket and accurate estimation of shovel loading. The aforementioned study produced an accurate assessment of the volume of materials in a single bucket during earthwork, offering reliable assurance for real-time estimation of earthwork volume and real-time evaluation of operational efficiency. However, the majority of current investigations are restricted to measuring the amount of materials in containers, such as buckets. This approach makes it difficult to employ when the bucket size varies because it necessitates the construction of an in-advance 3D model of an empty bucket. Fu et al. [36] suggested a fast estimation method for landslide deposit volumes by the 3D reconstruction of smartphone images and interpolation of the bottom surface considering the steepest gradient. This method resolves the issue of estimating the volume of irregular landslide deposits. However, it is not appropriate for the dynamic change in earthwork that occurs before and after loading in real time.

In summary, the method of single-shovel volume measurement by constructing a 3D point cloud model of the bucket and the material in the bucket needs to build a 3D model of the empty bucket in advance, which has a strong dependence on the bucket model and poor universality, and is difficult to apply to a situation where the bucket has different specifications and sizes. As a result, research on the measurement method of loader single-shoveling volume based on 3D reconstruction of material piles provides a novel idea for real-time project progress and operational efficiency evaluation.

3 Methodology

3.1 Overall Framework

The overall research idea of this study, which mainly includes data acquisition and processing, feature matching and correction based on QORB algorithm, dense 3D reconstruction of material piles based on MAPM, and volume estimation of the shovel load, is shown in Figure 1. First, the camera was calibrated according to the acquired binocular image to obtain the internal and external parameters of the camera, and then the image was calibrated on the same polar line. Second, a feature matching and correction method based on QORB algorithm is proposed to improve the feature point detection uniformity and obtain more correct matching feature points. Third, sparse 3D point clouds are generated by employing 2D Delaunay triangulation on the matching feature points in accordance with the determined camera calibration parameters. MAPM is built using the sparse feature points generated via matching as prior knowledge

to estimate the parallax of the residual pixel points and achieve the densification of sparse point clouds. Finally, reference and matching models were constructed using 3D point clouds on the surface of the material piles before and after shoveling, respectively. The matching and reference models are then registered using the ICP algorithm. The point cloud was segmented according to the shovel loading boundary, and the shovel load volume was estimated using the Alpha-shape algorithm of Delaunay triangulation.

3.2 Fast Matching of Feature Points and Dense 3D Reconstruction of Material Piles Based on QORB-MAPM Algorithm

To achieve a fast estimation of the loader’s shovel load volume, this paper presents a novel dense stereo matching algorithm, namely, QORB-MAPM, by integrating the QORB algorithm and MAPM. This algorithm is combined with the principle of binocular stereo imaging to realize the fast matching of feature points and dense 3D reconstruction of material piles. Figure 2 shows the flow-chart of this method, which involves three stages.

Step1: Fast matching of feature points based on the QORB algorithm. The QORB algorithm is used instead of the traditional ORB algorithm for feature point detection and matching to improve feature detection uniformity, solving the issue that feature points detected by the traditional ORB algorithm tend to be too clustered or even overlapped in areas with obvious features and are too sparse or even holes in areas without obvious features. Then, feature matching and

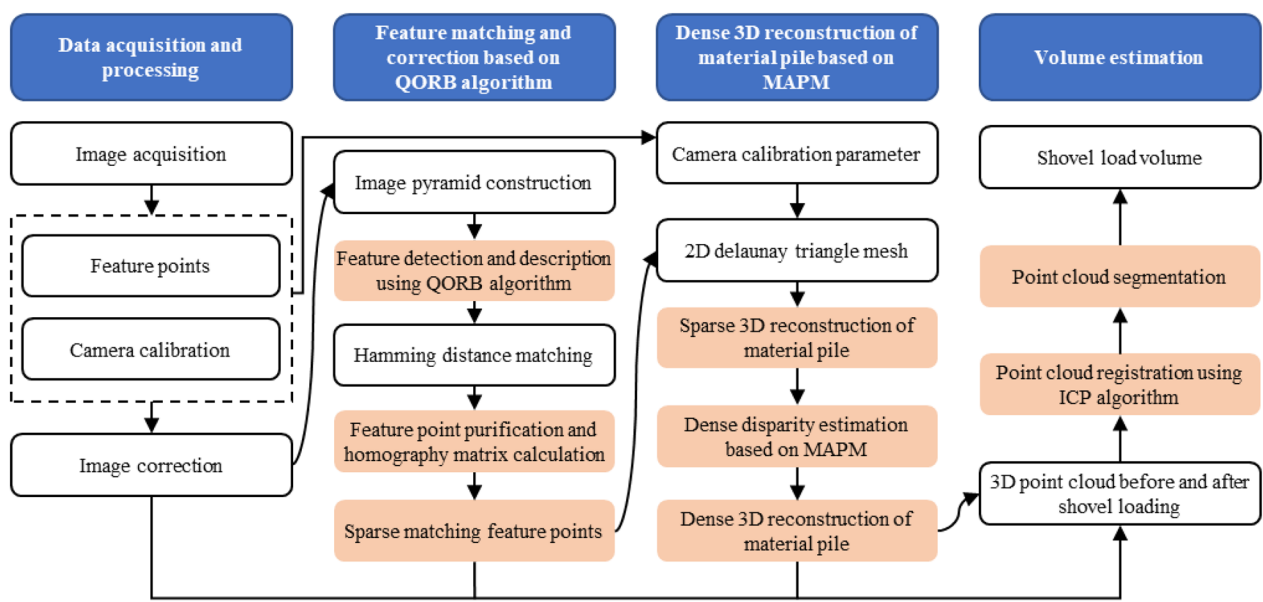


Figure 1 Overall framework

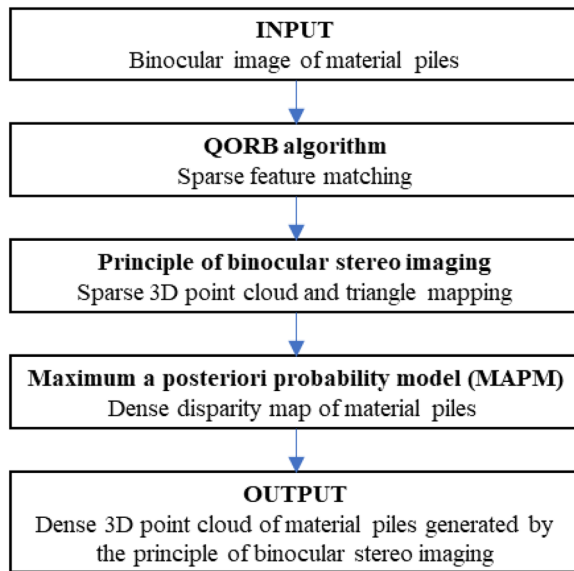


Figure 2 Flowchart of the proposed QORB-MAPM algorithm for dense stereo matching and 3D reconstruction of material piles

correction are performed through feature matching selection based on the RANSAC algorithm and feature point mapping based on the homography matrix to obtain more sparse matching feature points that improve the accuracy of 3D reconstruction of material piles and volume estimation of shovel loads.

Step2: Reconstruction of sparse 3D point cloud based on binocular stereo imaging principle. Based on the binocular stereo imaging principle, a sparse 3D point cloud and triangle mapping are generated by combining camera calibration parameters and matching sparse feature points to provide prior information for building MAPM.

Step3: Dense disparity estimation based on MAPM. The dense disparity map of the material pile is estimated by building a MAPM using the obtained sparse 3D point cloud and triangle mapping as prior information, which ensures the accuracy of disparity estimation, reduces the range of disparity search, and effectively improves the computational efficiency of 3D reconstruction and volume estimation of material piles.

The detailed approach of each stage is discussed as follows.

3.2.1 Fast Matching of Feature Points of Material Piles Based on QORB Algorithm

By calculating the gray difference between the feature point and the surrounding pixels and comparing it with the predetermined threshold, the traditional ORB algorithm employs the oFAST algorithm to choose the most

prominent pixel in the image as the feature point. The feature point detection procedure of traditional ORB are as follows:

Step 1: Construct an image pyramid according to the specified number of layers and scale factor.

Step 2: Calculate the number of feature points that need to be detected in each image layer according to the specified number of feature points and pyramid layers. The number of feature points N_i of the i^{th} image layer is calculated as follows:

$$N_i = \frac{N \cdot (1 - q)}{1 - q^n} \cdot q^{i-1}, i \geq 1, \quad (1)$$

where, n , q and N represent the number of layers of the image pyramid constructed, the scale factor between layers of the image pyramid, and the total number of feature points detected, respectively.

Step 3: Use oFAST to detect feature points from i^{th} image layer, sort them in descending order according to response values, and select the top N_i feature points are taken as the feature points extracted in the i^{th} image layer.

Step 4: Use rBRIEF to create descriptors of feature points in each layer of image.

For the surface features of the material pile with irregular surface morphology and no texture, the feature points detected by the traditional ORB algorithm are tend to be too clustered or even overlapped in areas with obvious features and are too sparse or even holes in areas without obvious features, which is not conducive to the 3D reconstruction of the earthwork material piles. Artal et al. [37, 38] used a quadtree algorithm to manage the detected feature points, which improved the uniformity of ORB algorithm feature point detection. However, this method is prone to excessive splitting of quadtree nodes because a quadtree has an unlimited depth, which increases the computational burden. Therefore, in this study, a novel algorithm is proposed, namely QORB, in which different node split depths are set according to the feature points allocated at different pyramid levels to prevent excessive node splitting, reduce redundant feature calculations, improve the uniformity of feature point detection, and ensure calculation efficiency. Compared to traditional ORB algorithm, **Step 3** is improved in QORB algorithm, as shown in Figure 3, **Step 3** in QORB algorithm is described as follows:

Step 3.1: Divide the i^{th} layer image into regular grids, with the size of each grid cell being 30×30 pixels, and use oFAST to extract feature points from each grid cell image block. Assuming that N_i is the number of desired feature points allocated to the current pyramid image layer, $D_{max} (4^{D_{max}} \geq N_i)$ is set as the maximum depth of the image layer. Consider the image layer as the initial

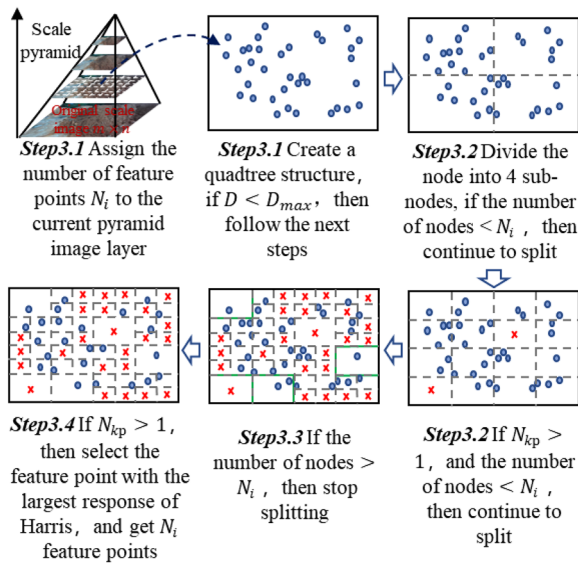


Figure 3 Feature point detection procedure of the Step 3 in QORB algorithm

node in the quadtree splitting process, which creates a fundamental quadtree structure. Assume that the current pyramid image layer's splitting depth is D , if $D < D_{max}$, then complete **Step 3.2**.

Step 3.2: Assume that N_{kp} is the total number of feature points included in the node. If $N_{kp} = 0$, then the node is deleted. If $N_{kp} = 1$, then the node is not split. If $N_{kp} > 1$, then the node is divided into 4 child nodes.

Step 3.3: Repeat **Step 3.2** until the required number of feature points is reached by the number of nodes, at which time the splitting stops.

Step 3.4: If $N_{kp} > 1$, the feature point with the largest Harris response value in each node is selected as the current feature point to obtain N_{set} feature points.

The procedure of feature points extraction and fast matching of material piles based on QORB algorithm is shown in Figure 4. First, the image pyramid is constructed for the material pile image, the regular grid is generated for the current image layer, and the oFAST algorithm is used for feature point detection. The improved quadtree algorithm is performed for the management of extracted feature points to improve the uniformity of feature point detection, reduce the influence of the characteristics of the material piles without obvious features and no surface texture, and improve the accuracy of 3D reconstruction and shovel loading volume estimation of the material piles. The rBRIEF algorithm is then used to calculate the binary descriptors of the feature points.

Second, after the binary descriptor of each feature point is obtained, a matching threshold is set (the ratio

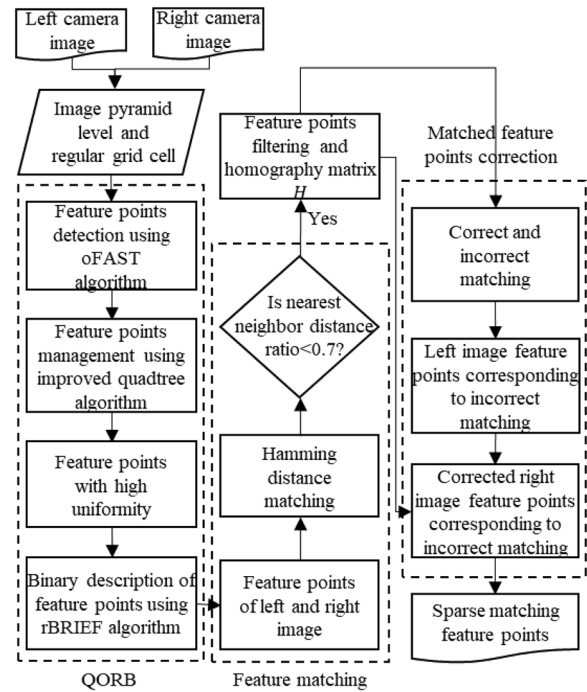


Figure 4 Fast matching and correction of feature points based on QORB algorithm

of the two nearest neighbor distances is less than 0.7), and the Hamming distance is used to judge the feature points one by one. When the similarity of the corresponding feature points on the two images meets the conditions, the matching point with the highest similarity is selected. Calculate the homography matrix H according to the matching feature points, and assume that the projection points of a point P in the 3D space on the left and right images are respectively $p_1(u_1, v_1)$ and $p_2(u_2, v_2)$. The formula is as follows:

$$p_2 = Hp_1 = K \left(R - t \cdot \frac{n^T}{d} \right) K^{-1} p_1, \quad (2)$$

where K is the camera internal parameter, R is the camera rotation matrix, t is the camera translation column vector, and (n, d) represents the coordinates of a plane in the world coordinate system. Eq. (2) can be further expanded to obtain Eq. (3). When there are four pairs of matching feature points, the homography matrix H can be calculated by linear transformation according to Eq. (3).

$$\begin{pmatrix} u_2 \\ v_2 \\ 1 \end{pmatrix} = H \begin{pmatrix} u_1 \\ v_1 \\ 1 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} u_1 \\ v_1 \\ 1 \end{pmatrix}. \quad (3)$$

Finally, RANSAC algorithm is used to filter matching feature points to obtain correct matching and incorrect matching. The left image feature points in the incorrect matching are calculated by homography matrix H to obtain the corresponding right image mapping points. The calculated results are used as the corrected right image feature points and replace the right image feature points in the incorrect matching to obtain more sparse matching feature points.

3.2.2 Principle of Binocular Stereo Imaging and Sparse 3D Reconstruction

Camera calibration and image correction for the binocular camera should be performed first to reconstruct a 3D point cloud from the images. Camera calibration was used to determine the internal and external parameters of the camera. The internal camera parameters, such as the projection position coordinates (u_0, v_0) and focal length f of the camera lens optical axis in the pixel coordinate system, are related to the optical properties of the camera. The mapping relationship between spatial points and pixel points may be constructed using internal parameters, which enables any 3D coordinates under the camera coordinate system to be mapped to the pixel coordinate system. The external parameters of the camera include a rotation matrix R and translation matrix T , through which the transformation between the camera coordinate system and world coordinate system can be performed. The transformation relationship between the left and right camera coordinate systems can be established once the camera calibration parameters have been determined. The imaging planes of the left and right cameras can be transformed into the same plane according to the transformation relationship. Finally, sparse 3D point cloud reconstruction was performed utilizing the obtained camera calibration parameters and the corrected left and right images in accordance with the binocular camera stereo imaging principle, as illustrated in Figure 5.

Assume that $P(X_c, Y_c, Z_c)$ is a point in 3D space. In the pinhole camera model, coordinate transformation is required to finally become a pixel $p(u, v)$ on the 2D image. Transform point $P(X_c, Y_c, Z_c)$ from the world coordinate system to point $p(x, y)$ of the camera coordinate system through the external parameters of the camera according to Eq. (4). According to Eq. (5), $p(x, y)$ is transformed from the camera coordinate system to point $p(u, v)$ of the pixel coordinate system through the camera internal parameter matrix.

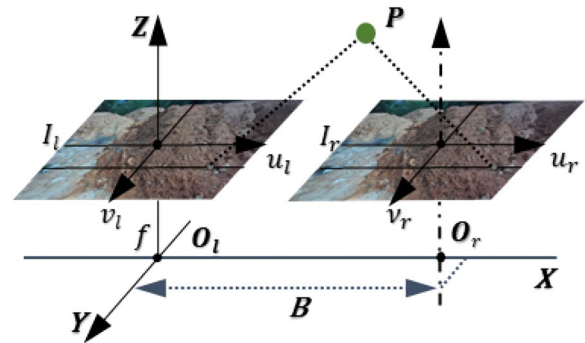


Figure 5 Principle of binocular stereo imaging

$$Z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}, \tag{4}$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1/d_x & 0 & u_0 \\ 0 & 1/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}. \tag{5}$$

For simultaneous Eq. (4) and Eq. (5), Eq. (6) can be obtained as

$$\begin{cases} X_c = (u - u_0) * \frac{Z_c}{f_x}, \\ Y_c = (v - v_0) * \frac{Z_c}{f_y}, \\ Z_c = f * \frac{B}{u_l - u_r}, \end{cases} \tag{6}$$

where $f_x = f/d_x$ and $f_y = f/d_y$ indicate that the focal length f of the camera is transformed into a pixel measurement in x and y directions, respectively. B represents the baseline length of the binocular camera. Therefore, the parallax of the matching points and sparse 3D point cloud can be calculated according to Eq. (6).

The generated sparse 3D point cloud was first triangulated using Delaunay triangulation to achieve dense 3D reconstruction of the material piles, and the triangulation mapping parameter $\mu_i(o_n^{(l)})$ was computed using Eq. (7).

$$\mu_i(o_n^{(l)}) = a_i u_n + b_i v_n + c_i, \tag{7}$$

where n represents a triangular serial number containing the pixel point $o_n^{(l)} = (u_n, v_n)$. The triangle plane parameters (a_i, b_i, c_i) can be obtained by solving the linear equation for the three vertices of each triangle. Triangulation mapping provides an accurate initial disparity value for the subsequent construction of the probability model,

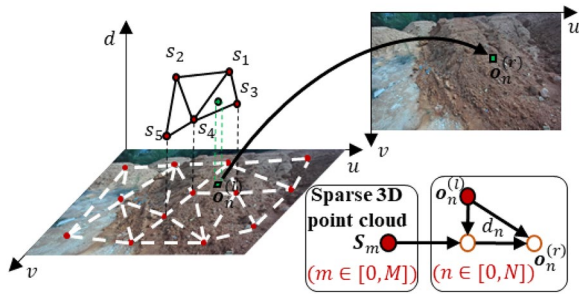


Figure 6 Probability estimation model

which reduces the disparity search range of the residual pixels of the input image and significantly improves computational efficiency.

3.2.3 Maximum a Posteriori Probability Model (MAPM) and Dense Disparity Estimation

The maximum a posteriori probability model (MAP) is constructed to estimate the optimal disparity of the residual pixel points in accordance with the obtained sparse 3D point clouds and triangular mapping in order to achieve the reconstruction of dense 3D point clouds. Figure 6 illustrates the fundamental principle, whereas Eq. (8) represents the probability estimation model.

$$\hat{d}_n = \operatorname{argmax} P(d_n | \mathbf{o}_n^{(l)}, \mathbf{o}_1^{(r)}, \dots, \mathbf{o}_N^{(r)}, \mathbf{S}), \quad (8)$$

where $\mathbf{S} = (s_1, s_2, \dots, s_M)$ represents the sparse 3D point cloud constructed previously; each point is represented as $s_m = (u_m, v_m, d_m)$; d_m is the parallax corresponding to the point (u_m, v_m) . $\mathbf{o}_1^{(r)}, \dots, \mathbf{o}_N^{(r)}$ represent all the pixels in the right image that have the same horizontal polar line as point $\mathbf{o}_n^{(l)}$. If the disparity d_n of a point $\mathbf{o}_n^{(l)}$ in the left image is regarded as a random variable to be solved, the posterior probability can be expressed as the product of a prior probability and likelihood probability, as shown in Eq. (9).

$$P(d_n | \mathbf{o}_n^{(l)}, \mathbf{o}_1^{(r)}, \dots, \mathbf{o}_N^{(r)}, \mathbf{S}) \propto P(d_n | \mathbf{S}, \mathbf{o}_n^{(l)}) * P(\mathbf{o}_1^{(r)}, \dots, \mathbf{o}_N^{(r)} | \mathbf{o}_n^{(l)}, d_n) \quad (9)$$

Assume that the prior probability is proportional to the Gaussian distribution, as shown in Eq. (10).

$$P(d_n | \mathbf{S}, \mathbf{o}_n^{(l)}) \propto \begin{cases} \exp\left(-\frac{(d_n - \mu(\mathbf{S}, \mathbf{o}_n^{(l)}))^2}{2\sigma^2}\right) & \text{if } |d_n - \mu| < 3\sigma, \\ 0, & \text{otherwise,} \end{cases} \quad (10)$$

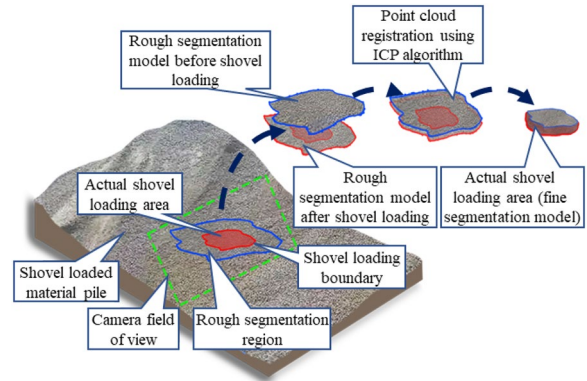


Figure 7 Point cloud registration and segmentation considering the boundary restriction of shovel loading

where $\mu(\mathbf{S}, \mathbf{o}_n^{(l)})$ represents a triangle mapping consisting of a sparse 3D point cloud \mathbf{S} containing pixels $\mathbf{o}_n^{(l)} = (u_n, v_n)$. The likelihood probability can be expressed as a Laplace distribution, as shown in Eq. (11).

$$P(\mathbf{o}_n^{(r)} | \mathbf{o}_n^{(l)}, d_n) \propto \begin{cases} \exp(-\|D_n^{(l)} - D_n^{(r)}\|) & \text{if } \begin{pmatrix} u_n^{(l)} \\ v_n^{(l)} \end{pmatrix} = \begin{pmatrix} u_n^{(r)} + d_n \\ v_n^{(r)} \end{pmatrix}, \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

where, $D_n^{(l)}$ and $D_n^{(r)}$ represent the feature description vectors of the n th pixel in the left image and the n th pixel in the right image, respectively. The if condition in Eq. (11) ensures that this constraint is satisfied because the corresponding points of the left and right images must appear on the same horizontal polar line, because the left and right images obtained by the binocular camera have been rectified in advance. Therefore, the likelihood probability model derived from Eq. (11) is given by Eq. (12).

$$P(\mathbf{o}_1^{(r)}, \dots, \mathbf{o}_N^{(r)} | \mathbf{o}_n^{(l)}, d_n) \propto \sum_{i=1}^N P(\mathbf{o}_i^{(r)} | \mathbf{o}_n^{(l)}, d_n). \quad (12)$$

It is worth noting that the disparity range of the remaining pixel points is limited to $|d_n - \mu(\mathbf{S}, \mathbf{o}_n^{(l)})| < 3\sigma$ in the process of disparity estimation, which ensures the precision of disparity estimation, reduces the range of disparity search, effectively improves the calculation efficiency, and

provides strong support for the dense 3D reconstruction of earthwork material piles and the accurate estimation of shovel volume.

3.3 Point Cloud Segmentation and Volume Estimation Considering Boundary Constraints

3.3.1 Point Cloud Registration and Segmentation

Constructing a dense 3D point cloud model on the surface of the material piles, both before and after shoveling, is required to calculate the shoveled volume. The actual shoveling 3D point cloud model was generated by matching and segmenting the point cloud models before and after shoveling, and the volume was calculated using the point cloud model, as illustrated in Figure 7. Binocular images of the material pile surface taken before and after shoveling as well as the corresponding 3D point cloud model's position and attitude change arbitrarily during the actual earthwork operation of engineering vehicles. This affects the accuracy of the subsequent point cloud processing and shoveling volume calculation. Therefore, the point cloud after shoveling is aligned with the reference point cloud model before shoveling using a set of suitable point cloud change parameters, such as translation, rotation, and yaw, determined using the ICP point cloud registration algorithm.

For example, with an image resolution of 1920×1080 , there can be up to 1 million 3D point clouds on the surface of material piles in a single reconstruction owing to the large amounts of dense point cloud data that have been reconstructed. The voxel-mesh method is used to first realize point cloud down-sampling before registration to improve computing efficiency while maintaining calculation accuracy. Following downsampling, the point cloud is roughly segmented in a certain area where the boundary of the shovel loading area is extended outward. Rough segmentation can improve the calculation efficiency, reduce the impact of the non-shovel loading area, and improve the accuracy of the volume estimation. The ICP algorithm is used to register the 3D point cloud model of the material pile surface after rough segmentation. The ICP algorithm is described as follows: 1) Find the point in the target point cloud closest to each point in the source point cloud. 2) The root mean square of the point-to-point or point-to-plane distance metric is minimized to determine the transformation parameters. 3) Transform the point cloud according to the transformation parameters. 4) Repeat until the matching requirements are satisfied. After ICP algorithm registration, the position and pose of the 3D point cloud on the surface of the material piles before and after shoveling are highly consistent.

After rough segmentation, the point cloud contains not only the effective shovel loading area but also many unrelated areas. Therefore, it is necessary to further

fine-segment the registered point cloud to obtain the point cloud model of the actual shovel loading area. The point cloud was segmented according to the boundary constraints of the shovel loading area during the point cloud segmentation process to optimize the segmentation of the effective shovel loading area. The 3D point cloud model on the surface of the material piles before shoveling after segmentation is represented by the blue point cloud in the figure, whereas the model on the surface of the material piles after shoveling after segmentation is represented by the red point cloud. The space volume enclosed by the two models was the shoveling volume to be estimated.

3.3.2 Volume Estimation

A point cloud model of the shovel loading area was obtained after the point cloud registration and segmentation. To estimate the volume of shoveled materials, the Alpha-shape algorithm of Delaunay triangulation was used for calculation. The shovel loading area point cloud model was first Delaunay triangulated, and the point cloud contour envelope was created using parametric fitting of the point cloud. After performing Delaunay triangulation on the point cloud model in the shovel loading area, the contour envelope was created by parametric fitting of the point cloud. The alpha-radius parameter value, which is used to regulate the fineness of the produced contour, enables the customization of the envelope surface around the point set. The created contour envelope becomes convex when the parameter value is too large. If the parameter value is too small, it will result in an overestimation of the shovel volume, and the contour envelope may have holes. Therefore, in the estimation of actual shovel load volume, with the increase of shovel load volume, the parameter value should be adjusted reasonably to generate a complete envelope of point cloud contour.

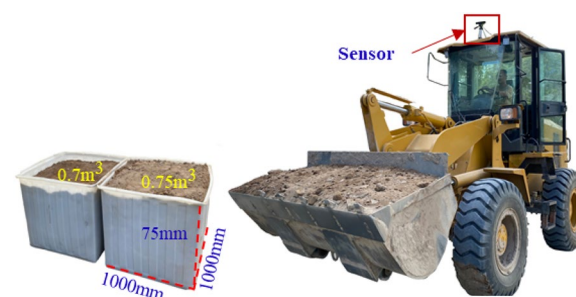


Figure 8 Loader and measuring tool for experiment

4 Experiment Results and Discussions

To verify the effectiveness of this research method for estimating the shovel load volume, we conducted a shovel load volume measurement experiment based on an XG931K loader (rated volume of the loader bucket is 1.8 m^3). The vision sensor was fixed on top of the loader cab to obtain better image information of the material piles. At the same time, in order to verify the accuracy of the volume estimation method in this study, a measuring tool is used to measure the actual shoveled volume, as shown in Figure 8. In this study, 8 shoveling experiments on loose soil were carried out at the terrain experimental site of a well-known construction machinery manufacturer. During the experiment, images of the material piles in various shoveling processes were acquired, and 3D reconstruction and shovel load volume estimation of the material piles were performed. Table 1 lists the system and development platforms used in this experiment.

4.1 3D Reconstruction of Material Piles

According to this research method, it is necessary to reconstruct the surface model of material piles before estimating the shovel load volume. To verify the effectiveness of the dense stereo matching and 3D reconstruction method of the material piles proposed in this study based on QORB–MAPM algorithm, some image sets of the material piles were selected from the 8 shoveling tests, which were before shoveling, after two shovels, after four shovels, and after 6 shovels, as shown in Figure 9. In addition, in the experimental process, we also compared 3 current methods that have good performance and similar characteristics to this research method: the block matching method (BM), semi-global block matching method (SGBM) and traditional ORB-based dense matching and 3D reconstruction method (ORB–MAPM). To evaluate the performance of this research method in detail, in addition to the difference in feature point detection and

matching, a comparison of traditional ORB and QORB algorithm ensures the consistency of other parts, such as feature matching and dense 3D reconstruction. For the sake of generality, this experiment used the VC++ language and the mainstream open-source computer vision library (OpenCV) for algorithm development.

First, feature points were detected. In this process, the traditional ORB algorithm and QORB algorithm in this study used oFAST to detect feature points. In this study, the improved ORB algorithm first divides the 2D space of each pyramid image into image blocks using the improved quadtree algorithm and then employs the oFAST algorithm to detect feature points for each image block. During the experiment, the number of feature points to be detected was set to 5000. The effects of feature point detection using the different methods are shown in Figure 10. It can be seen from the figure that the feature points extracted by the traditional ORB algorithms are too concentrated and overlapped in some areas with prominent edges and are too sparse or even appear as empty holes in areas with insignificant edges, resulting in the loss of local image feature information. The feature points extracted using the QORB algorithm combined with the improved quadtree algorithm in this study were more evenly distributed. The surface of the earthwork material piles with irregular surface morphology and no texture can more fully reflect the shape characteristics of the material piles, which is conducive to their 3D reconstruction of the material piles.

After feature point detection, feature point description and matching are required. In the feature point description, the traditional ORB algorithm and the QORB algorithm in this study use binary descriptors, therefore, the matching similarity is calculated using the Hamming distance to preliminarily select the matching points with the highest similarity. In the process of feature point matching, there may be incorrect matches of keypoints matched through specific similarity measurement relationships and corresponding search strategies. Therefore, the RANSAC algorithm was further introduced to purify feature points to filter out correct matches and incorrect matches. According to this research method, the homography matrix H is calculated for the filtered matching feature points. The left image feature points in the incorrect matching are calculated through the homography matrix H to obtain the corresponding right image mapping points. The right image mapping point can be taken as the corrected right image feature point to replace the right image feature point in the incorrect matching to obtain the final sparse matching feature points. The more sparse matching feature points can be obtained through feature point filtering based on the RANSAC algorithm and feature point mapping based on the homography

Table 1 Development platform

Items	Specification
ZED2 Sensor	Video resolutions: 1920×1080 Frame rate: 30 frames/s Interface: USB 3.0 Base line: 12 cm
Computer	64-bit Win10 platform 3.30 GHz processor and 32 GB memory
Program language	VC++ and Matlab
OpenCV library	Version: 3.4.1
Point cloud library	Version: 1.8.1

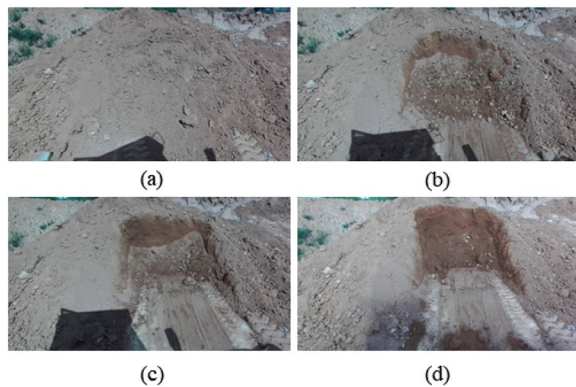


Figure 9 Left camera images under different shovel loading conditions, (a) Image I (before shoveling), (b) Image II (after 2 shovels), (c) Image III (after 4 shovels), (d) Image IV (after 6 shovels)

matrix H to improve the accuracy of 3D reconstruction and shovel load volume estimation of the material piles. The matching effects of the two methods are shown in Figure 11. Only the correctly matched feature point pairs are shown in the figure, as indicated by the green line. It can be seen from the figure that compared with the traditional ORB algorithms, the feature matching and correction method based on QORB algorithm in this study can obtain more correct matching feature point pairs, and the matching point pairs are more evenly distributed.

The obtained correctly matched feature points are then further generated into a sparse 3D point cloud in accordance with the camera calibration parameters and binocular camera stereo imaging principle. A dense 3D point cloud of material piles may be constructed by building a probability model in accordance with the obtained sparse 3D point cloud and triangular mapping. To verify the effectiveness of the 3D reconstruction of material piles based on QORB-MAPM algorithm proposed in this study, we compared it with three different methods: BM, SGBM, and ORB-MAPM. The 3D point clouds of the material piles reconstructed by different methods are shown in Figure 12. The figure shows how the ORB-MAPM, and QORB-MAPM methods in this study were able to reconstruct dense 3D point clouds that closely represent the actual material pile morphology on all test images. However, there were several noise points and holes in the 3D point cloud created by the BM and SGBM algorithms. This is because the pixel intensities are compared by BM and SGBM to calculate the generation value between the corresponding matching points. This pixel block-based descriptor performs has weak description performance, which increases the error in matching corresponding points and reduces the effectiveness of dense 3D reconstruction. However,

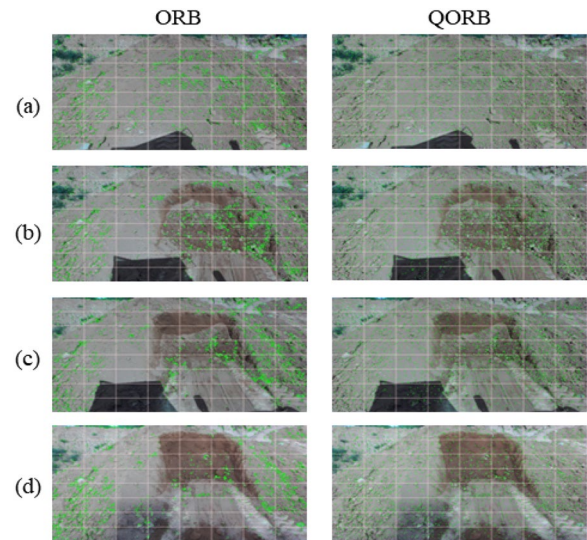


Figure 10 Comparison of different algorithms for feature point detection: (a) Image I (before shoveling), (b) Image II (after 2 shovels), (c) Image III (after 4 shovels), (d) Image IV (after 6 shovels)

in the dense 3D reconstruction method based on the MAPM proposed in this study, the triangular mesh calculated from matched support points as a priori information provides an accurate initial depth value for dense 3D reconstruction. Compared with pixel-block-based dense reconstruction methods, such as BM and SGBM, this method improves the robustness of the 3D point cloud. In addition, the 3D point clouds reconstructed by the ORB-MAPM algorithms have some point cloud holes and missing edge point clouds. Compared to this algorithm, the QORB-MAPM algorithm in this study retains more abundant material piles information, and the reconstructed 3D material piles point cloud model is more accurate.

The local features of the morphology of the material piles vary dramatically while the operation task proceeds in the actual application of earthwork, which raises expectations for the real-time reconstruction of the material piles and earthwork volume estimation. Therefore, we consider 5 groups of images before shoveling as an example to compare the average calculation time of different methods for the 3D reconstruction of material piles, as shown in Table 2. Figure 12 and Table 2 show that while the 3D material piles reconstructed by the BM algorithm include significant noise points, their computation time is the shortest. In contrast, the SGBM algorithm not only performs poorly in terms of 3D point cloud reconstruction, but also takes the longest time to complete. It is challenging to use these two methods in actual earthwork. Compared with the traditional ORB-MAPM method,

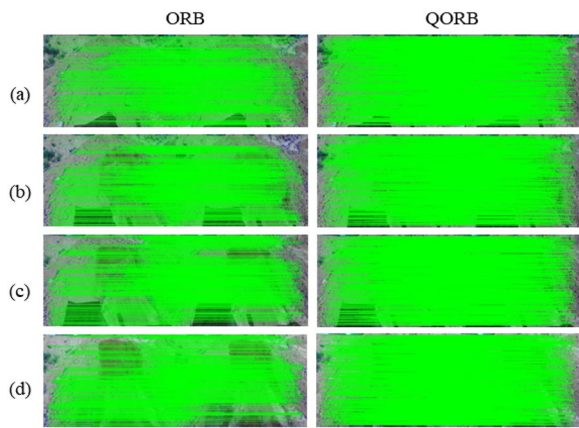


Figure 11 Comparison of feature matching of different algorithms: (a) Image I (before shoveling), (b) Image II (after 2 shovels), (c) Image III (after 4 shovels), (d) Image IV (after 6 shovels)

the calculation time of the QORB–MAPM method in this study was slightly increased, but the reconstructed

3D point cloud was more accurate, and the overall performance was better.

4.2 Volume Estimation of Material Piles

The 3D point cloud model of the material piles before and after shovel loading is not consistent because of the movement and changes in the loader’s position and attitude during the operation process. Therefore, the ICP algorithm is used to register the 3D point cloud model of the material pile surface before and after the shovel loading. Before the point cloud model registration, the dense point cloud was first down-sampled using the voxel-mesh method, and then the point cloud was roughly segmented at the boundary of the shovel loading area, extending outward to a certain range, as shown in Figure 13(a). This reduces the computational burden in the point cloud registration process and increases the computational efficiency of the algorithm. The blue and red point clouds in the image represent the 3D point clouds on the surface of the material piles before and after shoveling, respectively. It can be clearly seen from the figure that, owing

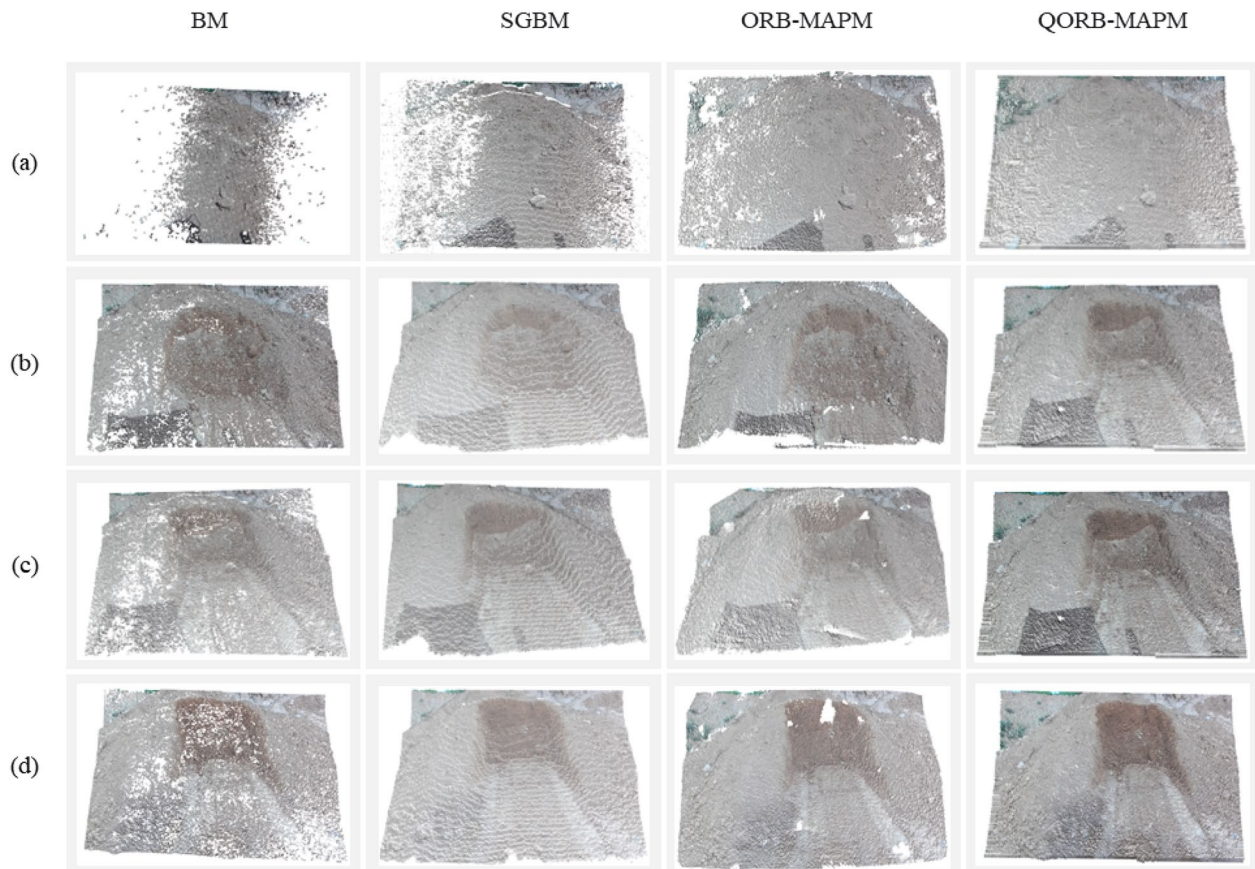


Figure 12 Comparison of dense 3D reconstruction of different algorithms: (a) Image I (before shoveling), (b) Image II (after 2 shovels), (c) Image III (after 4 shovels), (d) Image IV (after 6 shovels)

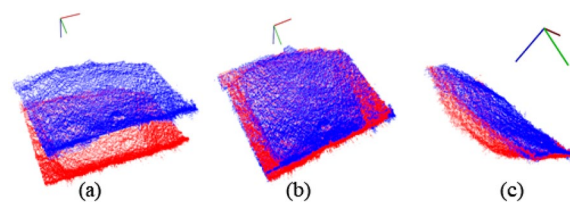
Table 2 Calculation time of each picture used for 3D reconstruction of material piles (ms)

Items	BM	SGBM	ORB-MAPM	QORB-MAPM
Keypoint detection	--	--	1036.2	1544.41
Descriptor computation	--	--	351.61	430.60
Feature match	10537.65	84748.45	2223.63	2350.94
Dense 3D reconstruction	4694.55	6213.15	5994.16	5941.05
Total	15232.2	90961.6	9605.6	10267.0

Note: The calculation time consumed for Feature match in the BM and SGBM algorithms includes the total time required for Keypoint detection, Descriptor computation, and Feature match.

to the movement and change of the loader, there is a large difference in the position and attitude of the point cloud models built before and after shoveling. The point cloud registered using the ICP algorithm is shown in Figure 13(b). The figure clearly shows that the position and pose of the 3D point cloud on the surface of the material piles before and after shoveling, registered by the ICP algorithm, are highly consistent. Finally, the point cloud model was finely segmented after registration according to the boundary of the shoveling area to retain the point cloud of the shoveling area, as shown in Figure 13(c).

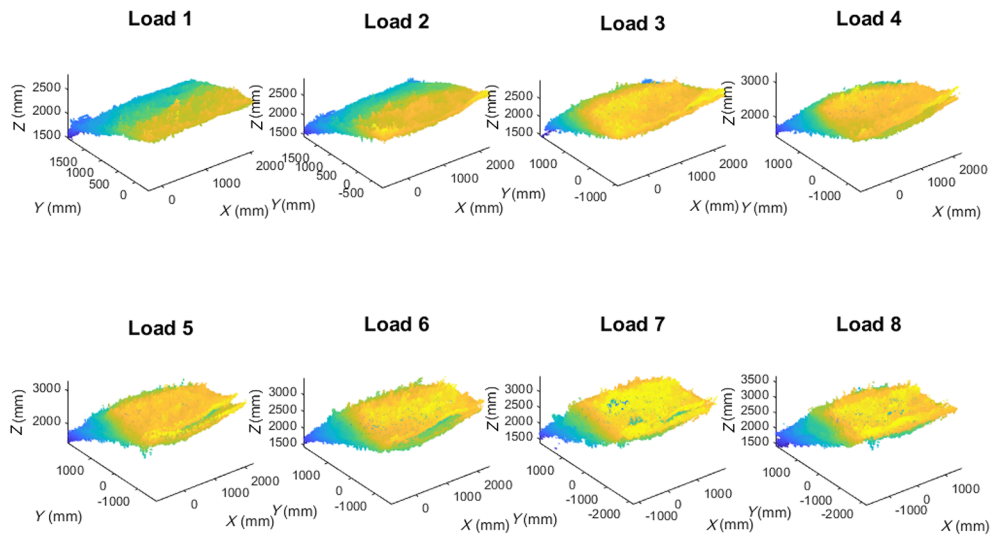
To verify the effectiveness of the proposed method in estimating the shoveling volume, the 3D point cloud of the material piles before shoveling was used as the reference point cloud model in this experiment, and the 3D point cloud of the material piles after each shoveling was used as the registration point cloud model in 8 shoveling tests. The point cloud model of the shoveling area was obtained through point cloud registration and segmentation. Finally, the Alpha-shape algorithm of Delaunay triangulation in the MATLAB toolbox was used to calculate the shoveling volume. The alpha-radius parameter values for each of the 8 shoveling tests during the experiment were set to 0.36, 0.36, 0.36, 0.32, 0.32, 0.38, 0.38, and 0.38, respectively. Each shoveling test consisted of 5 trials, and the shoveling volume was measured using the measuring tool shown in Figure 8. The segmented point cloud model and corresponding volume calculation results of the shoveling area after 8 shoveling tests are shown in Figure 14(a) and (b), respectively. The experimental values are listed in Table 3. It can be seen from the table that the maximum relative error and standard deviation of the estimated volume of the 8 shoveling tests are 10.19 % and 0.2395 m³, respectively. A relatively large error occurs after the first and second shoveling, which is due to the small amount of shoveling that makes the change of material piles in the shoveling area insignificant, resulting

**Figure 13** Point cloud registration and segmentation (the coordinate axis represents the camera position, where red, green and blue represent the x axis, y axis and z axis respectively)

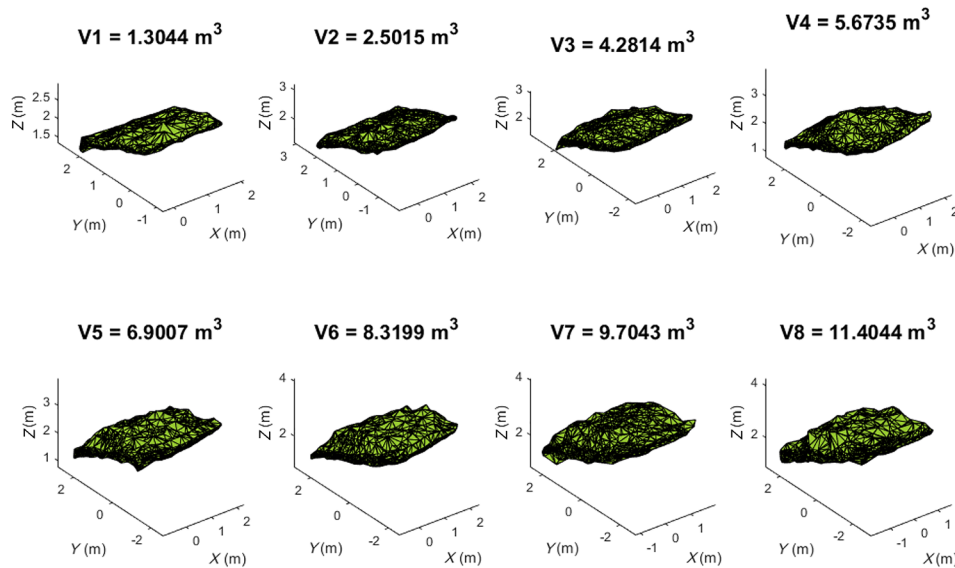
in inaccurate point cloud model segmentation. In general, the shovel loading volume estimation method proposed in this study is highly accurate and can meet the requirements of accurate estimation of shovel loading volume in practical earthwork applications.

The volume estimation methods based on BM (BM VE), SGBM (SGBM VE), and ORB-MAPM (ORB-MAPM VE), which have good performance and similar characteristics to the proposed method, were compared and verified to further verify the effectiveness of the volume estimation method based on QORB-MAPM (QORB-MAPM VE) in this study. 4 methods were used to estimate the volume of 8 shoveling tests, with 5 shoveling trials conducted for each test. The average of the 5 volume estimates was calculated and compared with the actual value (represented by the magenta solid line) measured by the measuring tool, as shown in Figure 15. The figure shows that as the shovel loading volume increases, the volume estimation values from the different methods follow the same trend of change as the actual values. This strongly confirms the reliability of the volume estimation method for the loader's shovel load based on the 3D reconstruction of the material piles proposed in this study. The test results are presented in Table 4. The maximum relative error and standard deviation of the estimated results of the 4 methods are as follows: BM VE (20.53% and 0.2511 m³), SGBM VE (14.31% and 0.6526 m³), ORB-MAPM VE (17.07% and 0.4768 m³), and QORB-MAPM VE (10.19% and 0.2395 m³). Therefore, the QORB-MAPM VE volume estimation method proposed in this study has higher estimation accuracy.

In practical earthwork applications, the bucket fill factor, which is an important parameter for evaluating the performance and operating efficiency of earthwork machinery, is often estimated using the ratio of the volume of materials in a single bucket to the rated volume of the bucket. Using the 8 shoveling test images from above, one can indirectly estimate the volume of materials loaded into the bucket in a single shovel by constructing a three-dimensional point cloud model of the material piles



(a)



(b)

Figure 14 Point cloud segmentation and volume estimation in shovel loading area

before and after a single shovel. The volume of 8 shoveling tests was estimated throughout the test using 4 different methods, and 5 trials were performed for each shoveling test. As illustrated in Figure 16, the average value of the 5 volume estimations was calculated and compared with the actual value (represented by the magenta solid line) measured by the measuring tool. It can be seen from the figure that the estimation results of the different methods

are quite different. The volume estimation results of the QORB–MAPM VE method in this study were consistent with the actual values. The detailed results are presented in Table 5. It can be seen from the table that the maximum relative error and standard deviation of the estimated results of the 4 methods are: BM VE (61.42 % and 0.2515 m³), SGBM VE (43.68 % and 0.6463 m³), ORB–MAPM VE (33.82 % and 0.4959 m³) and QORB–MAPM

Table 3 Volume estimation test results

Items	Load 1	Load 2	Load 3	Load 4	Load 5	Load 6	Load 7	Load 8
Test 1	1.304	2.502	4.281	5.674	6.901	8.320	9.704	11.404
Test 2	1.276	2.512	4.229	5.412	6.810	8.322	9.980	11.441
Test 3	1.328	2.625	4.264	5.594	7.299	8.139	9.809	11.195
Test 4	1.407	2.562	4.211	5.561	7.400	8.361	9.812	11.424
Test 5	1.428	2.688	4.260	5.515	7.306	8.281	9.811	11.190
Alpha-radius	0.36	0.36	0.36	0.32	0.32	0.38	0.38	0.38
Average volume estimation (m ³)	1.349	2.578	4.249	5.551	7.143	8.285	9.823	11.331
Actual volume of single bucket (m ³)	1.45	1.42	1.40	1.43	1.35	1.35	1.42	1.40
Actual volume of shovel load(m ³)	1.45	2.87	4.27	5.70	7.05	8.40	9.82	11.22
Relative error (%)	-6.99	-10.19	-0.49	-2.61	1.32	-1.37	0.04	0.99
Standard deviation (m ³)	0.0589	0.0703	0.0254	0.0866	0.2395	0.0770	0.0886	0.1133

VE (15.08 % and 0.2705 m³) proposed in this study. Therefore, the QORB-MAPM VE volume estimation

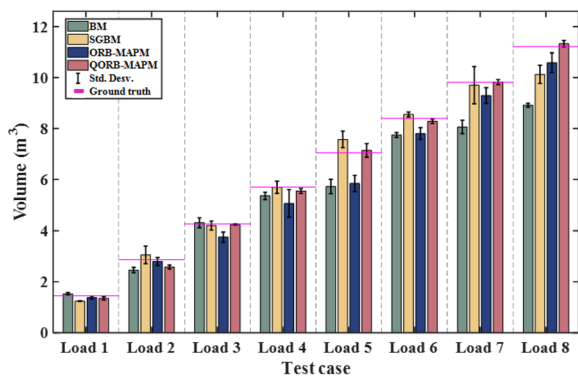


Figure 15 Comparison of volume estimation results of different methods

method in this study also has a higher estimation accuracy in bucket fill factor estimation and has important engineering application value.

The average calculation times of the different methods from the 3 main stages of 3D reconstruction of material piles, point cloud registration, and Alpha-shape algorithm volume calculation are compared, as shown in Table 6, and calculate the total time consumption. It can be observed from the table that the 3D reconstruction of material piles and point cloud registration account for most of the total calculation time, and point cloud registration takes the longest time. The SGBM VE method requires the highest computation time, whereas the ORB-MAPM VE and QORB-MAPM VE methods used in this study require the least computation time. It is worth noting that in the earthwork operation process, loaders, excavators, scrapers, and other earthwork machinery usually have a relatively determined operation mode. Taking loaders as an example, completing an

Table 4 Volume estimation test results of different methods

Methods	Items	Load 1	Load 2	Load 3	Load 4	Load 5	Load 6	Load 7	Load 8
BM VE	Average volume estimation (m ³)	1.525	2.462	4.317	5.362	5.726	7.751	8.065	8.916
	Relative error (%)	5.15	-14.22	11.07	-5.92	-18.78	-7.73	-17.87	-20.53
	Standard deviation (m ³)	0.0469	0.0942	0.1742	0.1234	0.2511	0.0851	0.2352	0.0676
SGBM VE	Average volume estimation (m ³)	1.242	3.056	4.206	5.693	7.573	8.555	9.703	10.125
	Relative error (%)	-14.31	6.48	-1.50	-0.13	7.41	1.84	-1.19	-9.76
	Standard deviation (m ³)	0.0124	0.3087	0.1579	0.2127	0.2909	0.0853	0.6526	0.3202
ORB-MAPM VE	Average volume estimation (m ³)	1.381	2.798	3.752	5.075	5.847	7.805	9.297	10.584
	Relative error (%)	-4.76	-2.51	-12.14	-10.97	-17.07	-7.08	-5.32	-5.66
	Standard deviation (m ³)	0.0497	0.1400	0.1802	0.4768	0.2833	0.2030	0.2752	0.3475
QORB-MAPM VE	Average volume estimation (m ³)	1.349	2.578	4.249	5.551	7.143	8.285	9.823	11.331
	Relative error (%)	-6.99	-10.19	-0.49	-2.61	1.32	-1.37	0.04	0.99
	Standard deviation (m ³)	0.0589	0.0703	0.0254	0.0866	0.2395	0.0770	0.0886	0.1133

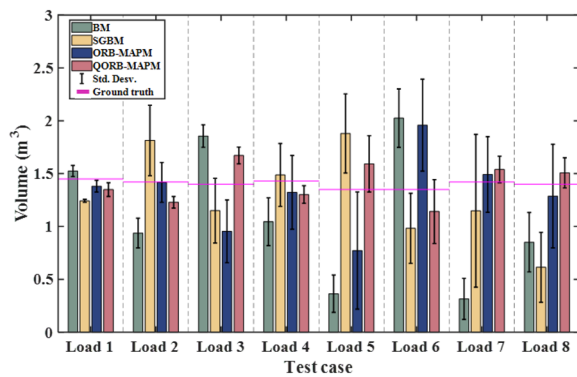


Figure 16 Comparison of bucket fill factor estimation results of different methods

operation cycle includes three main operational stages: shoveling, transportation, and unloading. According to many practical experiences and test statistics, it usually takes 40–50 s to complete an operational cycle. Therefore, the BM VE ORB–MAPM VE and the QORB–MAPM VE method in this study can satisfy the requirements of shovel loading volume estimation in real-time operation. In summary, considering the volume estimation accuracy and total calculation time, the QORB–MAPM VE

method in this study has wider application prospects in earthwork shovel loading volume estimation.

5 Conclusions and Prospects

To achieve fast and accurate measurement of the volume of earthmoving materials, a fast estimation method for a loader’s shovel load volume based on 3D reconstruction of material piles was proposed in this study. The main conclusions are as follows.

- (1) A feature point matching and correction method based on QORB algorithm was proposed to improve the uniformity of feature point detection and obtain more sparse matching feature points. And a dense stereo matching algorithm, namely, QORB–MAPM, was proposed by integrating the QORB algorithm and the maximum a posteriori probability model (MAPM) to achieve fast matching and dense 3D reconstruction of feature points of material piles. Compared with BM, SGBM, and ORB–MAPM 3D reconstruction methods, the QORB–MAPM method proposed in this study has significant advantages in reconstruction accuracy and calculation efficiency, which provides strong

Table 5 Bucket fill factor estimation results of different methods

Methods	Items	Load 1	Load 2	Load 3	Load 4	Load 5	Load 6	Load 7	Load 8
BM VE	Average volume estimation (m ³)	1.525	0.9375	1.8555	1.0455	0.3635	2.025	0.3145	0.8515
	Relative error (%)	4.15	−26.82	25.29	−21.38	−54.81	37.49	−61.42	−30.49
	Standard deviation (m ³)	0.0469	0.1254	0.0949	0.2022	0.1578	0.2473	0.1738	0.2515
SGBM VE	Average volume estimation (m ³)	1.242	1.814	1.150	1.487	1.880	0.982	1.148	0.614
	Relative error (%)	−11.53	21.86	−13.90	3.17	29.44	−20.44	−15.12	−43.68
	Standard deviation (m ³)	0.0123	0.2978	0.2736	0.2657	0.3344	0.2965	0.6463	0.2959
ORB–MAPM VE	Average volume estimation (m ³)	1.381	1.417	0.954	1.323	0.772	1.959	1.492	1.287
	Relative error (%)	−3.84	−0.17	−24.79	−5.96	−32.11	33.82	4.00	−6.27
	Standard deviation (m ³)	0.0497	0.1681	0.2657	0.3118	0.4959	0.3889	0.3205	0.4381
QORB–MAPM VE	Average volume estimation (m ³)	1.349	1.229	1.6719	1.302	1.5929	1.1419	1.539	1.5079
	Relative error (%)	−5.63	−10.62	15.08	−7.11	13.45	−11.59	6.60	5.97
	Standard deviation (m ³)	0.0589	0.0489	0.0712	0.0743	0.2385	0.2705	0.1122	0.1275

Table 6 Calculation time of volume estimation (ms)

Calculation time (ms)	BM VE	SGBM VE	ORB–MAPM VE	QORB–MAPM VE
3D reconstruction of material piles	15232.2	90961.6	9605.6	10267.0
Point cloud registration (Iterations: 3)	18584.5	22635.2	23597.7	23343.0
Volume estimated by Alpha-shape algorithm	2340.3	2568.2	2567.7	2387.8
Total	36157.0	116165.0	35771.0	35997.8

technical support for fast and accurate measurement of the loader's shovel load volume.

- (2) A volume estimation method based on a 3D point cloud on the surface of material piles before and after shoveling was proposed to solve the problem that existing methods for volume measurement, such as total station-based methods, cannot measure the volume in real time, while the bucket-based method also has the disadvantage of poor universality. The proposed method includes 5 stages: point cloud down-sampling, point cloud rough segmentation, point cloud registration, point cloud fine segmentation, and volume calculation. This method provides a new idea and method for measuring the loader's shovel load volume.
- (3) The volume estimation and bucket fill factor estimation were tested using 8 shoveling experiments under loose soil conditions and were compared with BM VE, SGBM VE, and ORB-MAPM VE. The test results show that the maximum relative error and standard deviation of the 4 methods are BM VE (20.53% and 0.2511 m³), SGBM VE (14.31% and 0.6526 m³), ORB-MAPM VE (17.07% and 0.4768 m³), and QORB-MAPM VE (10.19% and 0.2395 m³) respectively in volume estimation. In bucket fill factor estimation, the corresponding test results of the 4 methods are BM VE (61.42% and 0.2515 m³), SGBM VE (43.68% and 0.6463 m³), ORB-MAPM VE (33.82% and 0.4959 m³) and QORB-MAPM VE (15.08% and 0.2705 m³). Therefore, the QORB-MAPM VE volume estimation method in this study has higher estimation accuracy and lower calculation time consumption in volume estimation and bucket fill factor estimation and provides reliable technical support for real-time evaluation of loader operation efficiency and unmanned autonomous operation. In addition, this research method is applicable to the estimation of the shovel load volume of other earthmoving machinery and has important theoretical research and engineering application value.

It should be pointed out that the research in this study needs to be further deepened and expanded.

- (1) Owing to the limitation of test conditions, this study only takes the shovel loading test of loose soil as an example, considering the complexity of the actual operating environment of earthmoving machinery, such as the different softness of soil and the rolling and collapse of small granular soil, such as fine sand before and after shovel loading, which

creates great difficulties in the registration and segmentation of three-dimensional point clouds. Shovel loading experiments in more challenging terrain environments will be performed later to verify the effectiveness of the proposed method.

- (2) It can be seen from the calculation time of the shoveling volume estimation in this study that the point cloud registration takes the longest time. Therefore, the focus of future research will be to study point cloud registration algorithms with higher computational efficiency to meet the needs of real-time shoveling operations. In addition, there is room for improvement in algorithm program optimization.

Acknowledgements

Not applicable.

Author Contributions

BW performed the data analyses, the experiment and wrote the manuscript; SW contributed significantly to analysis and manuscript preparation; HL helped perform the experiment; SL helped revised the manuscript. LH contributed to the conception of the study. All authors read and approved the final manuscript.

Authors' Information

Binyun Wu born in 1993, is currently a doctoral candidate in the *Pen-Tung Sah Institute of Micro-Nano Science and Technology, Xiamen University, China*. His research interests include intelligent optimization and autonomous operation of construction machinery.

Shaojie Wang born in 1985, is currently an assistant professor at *Pen-Tung Sah Institute of Micro-Nano Science and Technology, Xiamen University, China*. He received his doctor degree from *Xiamen University, China* in 2016. His research interests include artificial intelligence system and industrial big data analysis.

Haojing Lin born in 1989, is currently a doctoral candidate in the *Pen-Tung Sah Institute of Micro-Nano Science and Technology, Xiamen University, China*. Her research interests include artificial intelligence system and industrial big data analysis.

Shijiang Li born in 1999, is currently a doctoral candidate in the *Pen-Tung Sah Institute of Micro-Nano Science and Technology, Xiamen University, China*. His research interests include innovative design of complex equipment.

Liang Hou born in 1974, is currently a professor at *Pen-Tung Sah Institute of Micro-Nano Science and Technology, Xiamen University, China*. He received his doctor degree from *Tianjin University, China*, in 2002. His research interests include innovative design, optimization and intelligence of equipment products.

Funding

Supported by National Key R&D Program of China (Grant Nos. 2020YFB1709901 and 2020YFB1709904), National Natural Science Foundation of China (Grant Nos. 51975495 and 51905460), Guangdong Provincial Basic and Applied Basic Research Foundation (Grant No. 2021A1515012286), and Guiding Funds of Central Government for Supporting the Development of the Local Science and Technology (Grant No. 2022L3049).

Declarations

Competing Interests

The authors declare no competing financial interests.

Received: 26 December 2022 Revised: 28 August 2023 Accepted: 1 September 2023

Published online: 13 October 2023

References

- [1] S Dadhich, U Bodin, U Andersson. Key challenges in automation of earth-moving machines. *Automation in Construction*, 2016, 68: 212–222.
- [2] D Pratt. *Fundamentals of construction estimating*. Boston: Cengage Learning, 2010.
- [3] M Bügler, A Borrmann, G Ogunmakin, et al. Fusion of photogrammetry and video analysis for productivity assessment of earthwork processes. *Computer-Aided Civil and Infrastructure Engineering*, 2017, 32(2): 107–123.
- [4] M Savia, H N Koivo. Neural-network-based payload determination of a moving loader. *Control Engineering Practice*, 2004, 12(5): 555–561.
- [5] M Yakar, H M Yilmaz, Ö Mutluoğlu. Close range photogrammetry and robotic total station in volume calculation. *International Journal of the Physical Sciences*, 2010, 5(2): 86–96.
- [6] H He, T Chen, H Zeng, et al. Ground control point-free unmanned aerial vehicle-based photogrammetry for volume estimation of stockpiles carried on barges. *Sensors*, 2019, 19(16): 3534.
- [7] H Anwar, S M Abbas, A Muhammad, et al. Volumetric estimation of contained soil using 3D sensors. *Commercial Vehicle Technology Symposium*, 2014: 11–13.
- [8] J Guevara, T Arevalo-Ramirez, F Yandun, et al. Point cloud-based estimation of effective payload volume for earthmoving loaders. *Automation in Construction*, 2020, 117: 103207.
- [9] J X Lu, Q S Bi, Y N Li, et al. Estimation of fill factor for earth-moving machines based on 3D point clouds. *Measurement*, 2020, 165: 108114.
- [10] J X Lu, Z W Yao, Q S Bi, et al. A neural network-based approach for fill factor estimation and bucket detection on construction vehicles. *Computer-Aided Civil and Infrastructure Engineering*, 2021, 36: 1600–1618.
- [11] Y Arayici. An approach for real world data modelling with the 3D terrestrial laser scanner for built environment. *Automation in Construction*, 2007, 16(6): 816–829.
- [12] M Golparvar-Fard, J Bohn, J Teizer, et al. Evaluation of image-based modeling and laser scanning accuracy for emerging automated performance monitoring techniques. *Automation in Construction*, 2011, 20(8): 1143–1155.
- [13] M Yakar, H M Yilmaz, O Mutluoglu. Performance of photogrammetric and terrestrial laser scanning methods in volume computing of excavation and filling areas. *Arabian Journal for Science & Engineering*, 2013, 39(1): 387–394.
- [14] Z Ma, S Liu. A review of 3D reconstruction techniques in civil engineering and their applications. *Advanced Engineering Informatics*, 2018, 37, 163–174.
- [15] C Sung, P Y Kim. 3D terrain reconstruction of construction sites using a stereo camera. *Automation in Construction*, 2016, 64: 65–77.
- [16] H Hirschmuller. Stereo processing by semiglobal matching and mutual information. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2008, 30(2): 328–341.
- [17] D G Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004, 60(2): 91–110.
- [18] H Bay, A Ess, T Tuytelaars, et al. Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 2008, 110(3): 346–359.
- [19] E Rublee, V Rabaud, K Konolige, et al. ORB: an efficient alternative to SIFT or SURF. *IEEE International Conference on Computer Vision*, 2011: 2564–2571.
- [20] J Jiao, B Zhao, S Wu. A speed-up and robust image registration algorithm based on FAST. *IEEE International Conference on Computer Science & Automation Engineering*. 2011: 10–12.
- [21] M Calonder, V Lepetit, C Strecha, et al. BRIEF: Binary robust independent elementary features. *Proceedings of the 11th European Conference on Computer Vision*. Heraklion, Crece. 2010, 6314: 778–792.
- [22] U Chaudhuri, B Banerjee, A Bhattacharya, et al. CMIR-NET: A deep learning based model for cross-modal retrieval in remote sensing. *Pattern Recognition Letters*, 2010, 131: 456–462.
- [23] Y S Li, Y J Zhang, X Huang, et al. Large-scale remote sensing image retrieval by deep hashing neural networks. *IEEE Transaction on Geoscience and Remote Sensing*, 2017, 56(2): 950–965.
- [24] R Z Wang, J C Yan, X K Yang. Learning combinatorial embedding networks for deep graph matching. in: *IEEE International Conference on Computer Vision*, 2019: 3056–3065.
- [25] B Jiang, P F Sun, B Luo. GLMNet: Graph learning-matching convolutional networks for feature matching. *Pattern Recognition*, 2022, 121: 108167.
- [26] P Sarlin, D DeTone, T Malisiewicz, et al. Superglue: learning feature matching with graph neural networks. *IEEE Conference on Computer Vision and Pattern Recognition*, 2020: 4937–4946.
- [27] C Liguori, A Paolillo, A Pietrosanto. An on-line stereo-vision system for dimensional measurements of rubber extrusions. *Measurement*, 2004, 35(3): 221–231.
- [28] T Zhang, J H Liu, S L Liu, et al. A 3D reconstruction method for pipeline inspection based on multi-vision. *Measurement*, 2017, 98: 35–48.
- [29] G F Xiao, Y T Li, Q X Xia, et al. Research on the on-line dimensional accuracy measurement method of conical spun workpieces based on machine vision technology. *Measurement*, 2019, 148: 106881.
- [30] J Miller, J Morgenroth, C Gomez. 3D modelling of individual trees using a handheld camera: Accuracy of height, diameter and volume estimates. *Urban Forestry & Urban Greening*, 2015, 14(4): 932–940.
- [31] P Muñoz-Benavent, G Andreu-García, J M Valiente-González, et al. Enhanced fish bending model for automatic tuna sizing using computer vision. *Computers and Electronics in Agriculture*, 2018, 150: 52–61.
- [32] S Barone, A Paoli, A V Rationale. Shape measurement by a multi-view methodology based on the remote tracking of a 3D optical scanner. *Optics and Lasers in Engineering*, 2012, 50(3): 380–390.
- [33] M R Yao. *Research on 3D vision measurement technology of aeroengine blade profile*. Harbin: Harbin Institute of Technology, 2019. (in Chinese)
- [34] J R Borthwick. *Mining haul truck pose estimation and load profiling using stereo vision*. Vancouver: University of British Columbia, 2009.
- [35] M Yakar, H M Yilmaz, O Mutluoglu. Performance of photogrammetric and terrestrial laser scanning methods in volume computing of excavation and filling areas. *Arabian Journal for Science & Engineering*, 2014, 39(1): 387–394.
- [36] L Fu, J Zhu, W L Li, et al. Fast estimation method of volumes of landslide deposit by the 3D reconstruction of smartphone images. *Landslides*, 2021, 18(9): 3269–3278.
- [37] R Mur-Artal, J M M Montiel, J D Tardos. ORB-SLAM: A versatile and accurate monocular SLAM system. *IEEE Transactions on Robotics*, 2015, 31(5): 1147–1163.
- [38] R Mur-Artal, J D Tardos. ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras. *IEEE Transactions on Robotics*, 2017, 33(5): 1255–1262.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)