Research on Visual Positioning of a Roadheader and Construction of an Environment Map

Wentao Zhang 1,2, Guodong Zhai 1,2,* , Zhongwen Yue 2,3, Tao Pan 2,4,5 and Ran Cheng 1,2

Abstract: The autonomous positioning of tunneling equipment is the key to intellectualization and robotization of a tunneling face. In this paper, a method based on simultaneous localization and mapping (SLAM) to estimate the body pose of a roadheader and build a navigation map of a roadway is presented. In terms of pose estimation, an RGB-D camera is used to collect images, and a pose calculation model of a roadheader is established based on random sample consensus (RANSAC) and iterative closest point (ICP); constructing a pose graph optimization model with closed-loop constraints. An iterative equation based on Levenberg–Marquadt is derived, which can achieve the optimal estimation of the body pose. In terms of mapping, LiDAR is used to experimentally construct the grid map based on open-source algorithms, such as Gmapping, Cartographer, Karto, and Hector. A point cloud map, octree map, and compound map are experimentally constructed based on the open-source library RTAB-MAP. By setting parameters, such as the expansion radius of an obstacle and the updating frequency of the map, a cost map for the navigation of a roadheader is established. Combined with algorithms, such as Dijkstra and timed-elastic-band, simulation experiments show that the combination of octree map and cost map can support global path planning and local obstacle avoidance.

Keywords: roadheader; RGB-D camera; SLAM; pose solution; pose optimization; navigation map

1. Introduction

With the rapid development of the industrial Internet, big data, and artificial intelligence, the intellectualization and robotization of equipment have become an inevitable trend in industrial development. Based on literature data from the Web of Science database, a cluster analysis of hotspots in intelligent coal mine research, from 2010 to 2020, obtained using VOSviewer software, is shown in Figure 1. It reveals that the in-depth integration of intelligent technology and the development of the coal industry is inevitable. For China’s coal industry, the degree of automation and intelligence of fully mechanized mining faces has been significantly improved over years of innovative development, and has been applied practically in many mining areas [1,2]. However, the intelligent development of fully mechanized excavation faces is still in a bottleneck period due to the poor working environments and complex excavation processes; the imbalance between mining and tunneling seriously restricts the safety and efficiency of coal mine production. Therefore, it is urgent to carry out research on the intelligence of tunneling equipment and systems for a fully mechanized excavation face [3,4].
As the main equipment for coal roadway driving, autonomous positioning of roadheaders is a key link to realize intelligent tunneling, and many scholars have done a great deal of original work in this area. An indoor global position system (IGPS)-based pose measurement method for roadheaders was proposed by Tao et al. [5]. The position coordinates of a receiver can be obtained from a single transmitter station, which realized the autonomous measurement of the body pose in a multi-point time-sharing manner. A remote-controlled system for roadheaders was developed by Roman et al. [6]. The system had built-in active lighting infrared sensors, which could assist in the positioning of a roadheader without visual control by workers. Du et al. [7] studied the pose measurement method of roadheaders based on machine vision. A laser target was fixed on a roadheader to receive the laser beam. After the imaging features were collected using a monocular camera, the roadheader pose can accurately be solved by solving the pose model and performing error compensation. Mao et al. [8] explored a multi-source fusion measurement method based on ultrasound, inertial navigation, a laser, and geomagnetic information. The method could overcome the shortcoming of excessive accumulated errors in a single inertial navigation measurement and effectively monitor the roadheader pose. Fu et al. [9,10] developed a set of roadheader pose measurement systems based on an ultra-wideband. A time-of-arrival model was built after the distance between the positioning node of a roadheader and ultra-wideband (UWB) positioning base station was measured. After the positioning algorithm was solved, the roadheader’s pose information could be obtained. Yang et al. [11] presented a vision measurement method for the pose of roadheader’s cutting head, based on infrared LEDs. A 16-point infrared target was fixed on the roadheader, and then the target image was collected and processed using a monocular camera; the pose of the cutting head could then be estimated. Simulation and experimental results showed that the accuracy of pose measurements could meet the requirements of underground roadway construction. Yan et al. [12] investigated a roadheader pose measurement method based on multi-source information, such as a gyroscope, a total station, and an inertial measurement system, and built a positioning model. Experimental results showed that this method could meet precision requirements. Xue et al. [13] put forward a positioning method for roadheaders based on binocular vision technology. By collecting information with a binocular vision sensor, the coordinate model between a roadheader and the roadway space was established, and then the pose was calculated. Simulations and experiments
showed that the accuracy of the calculated positioning parameters was good and the established positioning model was reasonable. Yang et al. [14] proposed a vision measurement method for the pose of a roadheader body based on laser beams. The laser beams were used for the image feature, and a measurement system of roadheader body pose, based on monocular vision, was established. Experimental results showed that this method could effectively solve body pose information of roadheaders. A pose measurement method for a roadheader, combining total station and strapdown inertial navigation, was proposed by Zhang et al. [15]. Total station and inertial navigation technology were used to measure the pose of roadheaders, respectively, and the pose was determined using fusion data, which reduced the influence of dust in the mine.

The above research covers the main schemes of present pose measurement of roadheaders. Total station has high installation flexibility, but its light path is easily blocked, and its real-time performance is poor and cannot adapt to the dynamic working conditions of a roadheader. An inertial sensor is generally used as an auxiliary measurement method because of its large cumulative error. UWB technology has a strong real-time performance, but its signal is greatly affected by the complex electromagnetic environment of coal mines. IGPS technology has a large measurement range, but the cost of installation and construction is high. Visual measurements based on target imaging are susceptible to the influence of light beam scattering, and the installation of a light source emitter is limited by complicated roadway structures. The composition of the combined measurement system is complex, and the multi-source data are difficult to synchronize accurately in time and space; thus, the data reliability is not high. State perception, trajectory planning, and motion control are the three keys to realize intelligent and robotized roadheaders. As a core technology of autonomous positioning and navigation for mobile devices, simultaneous localization and mapping (SLAM) is widely used in mobile robot navigation, unmanned driving, and in other fields [16–18]. For instance, the open source library, Real-Time Appearance-Based Mapping (RTAB-MAP), provides a positioning and mapping solution that is independent of time and scale, which can be applied to various setups [19]. According to different sensor types, SLAM technology is mainly divided into laser-based and vision-based [20]. The point cloud obtained by LiDAR is sparse, a created map lacks semantic information, and the linearization error is large. In contrast, vision sensors can obtain far more status information than LiDARs. Furthermore, the RGB-D (red, green, blue, and depth) camera in a vision sensor is based on active ranging technology, and it has the advantages of not being affected by ambient lighting changes and textures compared with monocular and binocular cameras.

In summary, SLAM technology is a better solution to realize autonomous positioning and navigation of roadheaders and roadway mapping in a GPS-free roadway environment. At present, there are no reports on the application of visual SLAM technology to coal roadway roadheaders, and research on this technology is still lacking. Therefore, the positioning requirements of roadheaders and the objective conditions, such as tunnel environment and sensor attributes, are comprehensively considered, a method of body pose estimation for roadheaders, based on visual SLAM, is proposed, and multiple schemes of environmental map construction based on laser and visual SLAM are tested. These can provide information for pose correction, navigation of roadheaders, and three-dimensional reconstruction of roadways.

2. Positioning Foundation of the Roadheader and Modeling

2.1. Definition of Coordinate System and Pose

The positioning problem of a roadheader is transformed into a state estimation problem in this paper; that is, to solve the optimal pose. To accurately describe the pose of a roadheader in three-dimensional (3D) space, a series of coordinate systems need to be established, an earth coordinate system $O_eX_eY_eZ_e$, roadway coordinate system $O_tX_tY_tZ_t$, world coordinate system $O_wX_wY_wZ_w$, and camera coordinate system $O_cX_cY_cZ_c$. The $O_e$ system is fixed on the earth, which is the basis of the design of the roadway; the $O_t$ system
is the absolute reference of the location in the roadway, and its position is determined by
the structure of the roadway; the $O_w$ system is the reference frame for the pose description
of an airborne camera when it is moving, and the $O_e$ and $O_w$ systems are related to rotation
and translation.

As the input device of visual SLAM, an RGB-D camera can be installed in the sensing
system of a roadheader through intrinsic safety technology. Considering that the camera
is connected to the roadheader body, the pose of the roadheader’s body is expressed
by the pose of the RGB-D camera relative to the $O_w$ system. If the absolute pose of
the roadheader body in the $O_e$ system or the $O_e$ system is required, it can be obtained
using Euclidean coordinate transformation. In addition, the pose of the cutting head of
a cantilever roadheader can be solved according to the Denavit–Hartenberg model and
forward kinematics in robotics [21]. In sum, the pose acquisition of a roadheader body in
the $O_w$ system is the basis for realizing other functions.

At the initial moment, it is stipulated that the $O_e$ and $O_w$ systems are coincident. The
$X_w$ axis of the $O_w$ system is located in the horizontal plane and is perpendicular to the
roadway midline, pointing to the right of the heading direction. The $Y_w$ axis is located in
the horizontal plane and is parallel to the roadway midline, pointing to the heading
direction; the $Z_w$ axis is perpendicular to the horizontal plane and points to the roadway
roof. After the roadheader moves, the pose of the $O_e$ system changes relative to the $O_w$
system, and the pose of the roadheader body (camera) is described by axis deflection angles
$\alpha$, $\beta$, and $\gamma$, and axis displacements $t_x$, $t_y$, and $t_z$, as shown in Figure 2. $\alpha$ is defined as the
included angle between the projection of the axis in the tunneling direction on the $O_wX_wY_w$
plane and the $Y_w$ axis, namely the yaw angle (Figure 2a). $\beta$ is defined as the included angle
between the projection of the axis in the tunneling direction on the $O_wY_wZ_w$ plane and the
$Y_w$ axis, namely the pitch angle (Figure 2b). $\gamma$ is defined as the included angle between
the projection of the transverse axis of the roadheader on the $O_wX_wZ_w$ plane and the $X_w$
axis, namely the roll angle (Figure 2c); $t_x$, $t_y$, and $t_z$ are, respectively, the displacement of
the roadheader body in the direction of $X_w$, $Y_w$, and $Z_w$ axes (Figure 2d–f).

![Figure 2. Pose parameters of roadheader body (camera).](image-url)
According to rigid body kinematics, the motion of a roadheader can be decomposed into rotation and translation. The rotation of the body (camera) in the $O_w$ system is described by rotation matrix $R$; $R_{cw}$ means the rotation transformation from the $O_w$ system to the $O_c$ system. Specifically, point $P_w(x_w, y_w, z_w)$ in the $O_w$ system is rotated and transformed to point $P_c(x_c, y_c, z_c)$ in the $O_c$ system, which can be described as: (1) rotating at angle $\alpha$ along the $Z_w$ axis, as shown in Equation (1); (2) rotating at angle $\beta$ along the $X_w$ axis, as shown in Equation (2); and (3) rotating at angle $\gamma$ along the $Y_w$ axis, as shown in Equation (3). $R$ has a definite transformation relationship with the $\alpha$, $\beta$, and $\gamma$, as shown in Equation (4) [22].

$$
\begin{bmatrix}
x_c \\
y_c \\
z_c
\end{bmatrix} =
\begin{bmatrix}
\cos \alpha & -\sin \alpha & 0 \\
\sin \alpha & \cos \alpha & 0 \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
x_w \\
y_w \\
z_w
\end{bmatrix}
$$

(1)

$$
\begin{bmatrix}
x_c \\
y_c \\
z_c
\end{bmatrix} =
\begin{bmatrix}
1 & 0 & 0 \\
0 & \cos \beta & \sin \beta \\
0 & -\sin \beta & \cos \beta
\end{bmatrix}
\begin{bmatrix}
x_w \\
y_w \\
z_w
\end{bmatrix}
$$

(2)

$$
\begin{bmatrix}
x_c \\
y_c \\
z_c
\end{bmatrix} =
\begin{bmatrix}
\cos \gamma & 0 & -\sin \gamma \\
0 & 1 & 0 \\
\sin \gamma & 0 & \cos \gamma
\end{bmatrix}
\begin{bmatrix}
x_w \\
y_w \\
z_w
\end{bmatrix}
$$

(3)

$$
R_{cw} = \text{Rot}(Z_w, \alpha) \cdot \text{Rot}(X_w, \beta) \cdot \text{Rot}(Y_w, \gamma)
$$

$$
= \begin{bmatrix}
\cos \alpha \cos \gamma - \sin \alpha \sin \beta \sin \gamma & \cos \alpha \sin \beta + \sin \alpha \sin \beta \cos \gamma & \cos \alpha \sin \beta \cos \gamma \\
- \cos \beta \sin \gamma & \sin \beta & \cos \beta \cos \gamma
\end{bmatrix}
$$

(4)

The translation of the body (camera) is described by 3D vector $t$, and $t_{cw}$ represents the translation transformation from the $O_w$ system to the $O_c$ system, as shown in Equation (5). After homogeneous transformation, $R_{cw}$ and $t_{cw}$ can be unified into transformation matrix $T_{cw}$, as shown in Equation (6) [22]. In the absence of ambiguity, $R_{cw}$, $t_{cw}$, and $T_{cw}$ below are referred to as $R$, $t$, and $T$ without subscripts.

$$
t_{cw} = \begin{bmatrix} t_x & t_y & t_z \end{bmatrix}^T
$$

(5)

$$
T_{cw} =
\begin{bmatrix}
R_{cw} & t_{cw} \\
0^T & 1
\end{bmatrix}
$$

(6)

2.2. SLAM Mathematical Model of Roadheader

In the moving process of the roadheader, the airborne camera observes the environmental signposts and estimates its own motion. As shown in Figure 3, the SLAM process of a roadheader equipped with a camera can be visualized, and the pose state and signposts are represented by nodes, and processes of camera movement and observation are represented by edges. Further, at time $i$, let the pose variable of the airborne camera be $x_i$, the position variable of waypoint is $y_j$, the motion input is $u_i$, and the observation data of airborne camera are $z_{ij}$; then, the mathematical model of visual SLAM for the roadheader is shown as Equation (7).

$$
\begin{cases}
x_i = f_p(x_{i-1}, u_i, w_i) & i \in (1, \ldots, m) \\
z_{ij} = g_o(y_j, x_i, v_{ij}) & j \in (1, \ldots, n)
\end{cases}
$$

(7)

where $f_p$ and $g_o$ are the process function and observation function, respectively; $w_i$ and $v_{ij}$ are the motion noise and observation noise at time $i$, respectively.
3. Pose Estimation and Optimization of the Roadheader

3.1. Pose Estimation Scheme

The Pose estimation scheme of the roadheader (camera) is shown in Figure 4. Specifically, (1) the images of the roadway environment are collected by the airborne RGB-D camera and depth information is obtained; (2) feature extraction and matching are performed on the color images, and 3D matching pairs are formed after depth mapping; (3) combined with the random sample consensus (RANSAC) strategy and the iterative closest point (ICP) strategy, the pose of the roadheader is estimated; (4) the pose graph, based on keyframe and closed-loop constraints, is constructed, and the global pose optimization is carried out to obtain the optimal pose estimation of the roadheader.

Figure 3. SLAM motion model of the roadheader.

3.2. Image Feature Extraction and Matching

Images collected by the RGB-D camera need feature calculations and data associations, and commonly used methods include the direct method, optical flow method, and feature point method [23]. Both the direct method and the optical flow method rely too heavily on
the assumption of a constant gray level, so it is difficult for them to be directly used in a roadway environment; the feature point method, however, is less affected by motion and illumination and is relatively stable. In coal mine roadways, the forming quality of a section is affected by the real-time and accurate measurement of the roadheader pose. Compared with other features, ORB (oriented fast and rotated brief) features perform better in terms of real-time performance and accuracy, and ORB features are not sensitive to rotation and translation [24,25]. Consequently, ORB features are used for extraction and matching in this paper.

To verify the effect of ORB feature extraction and matching, two images from the public dataset, TUM-RGBD, were selected for the experiment. The experimental results are shown in Figure 5. Figure 5a shows the effect obtained after using the Brute-Force matcher and Figure 5b shows the effect after matching filtering, and the filtering basis is that the Hamming distance is less than twice the minimum distance.

![Original match](image1)

![Filtered match](image2)

**Figure 5.** Effect of ORB feature extraction and matching.

### 3.3. Pose Solution

Based on image feature extraction and matching, an initial motion transform with a high confidence is iteratively solved using the RANSAC method, which aims to provide reliable initial values for pose estimation based on ICP. The process is as follows: (1) According to the matched 3D point set, the initial model is generated by randomly sampling some point pairs; (2) the model is evaluated according to the given threshold by calculating the point-pair error; and (3) through calculating the number of interior point pairs and the average value of the error, the model is updated and iterated based on the given threshold. After several iterations, a relatively optimal motion transformation can be obtained [26,27].

After depth mapping and RANSAC iteration, a set of 3D point sets with exact matching is obtained, i.e., \( P = \{p_1, p_2, \ldots, p_n\} \), \( Q = \{q_1, q_2, \ldots, q_n\} \), the body pose transformation between the \( i \)-th pair of 3D matching points in the point sets can be written as:

\[
p_i = Rq_i + t
\]  

(8)
According to the SLAM mathematical model of a roadheader (Equation (7)), there are noises in the motion equation and observation equation of the airborne camera; thus, the state estimation is uncertain. The inter-frame error function, $e_{(3D-3D)i}$, is positioned as [28]:

$$e_{(3D-3D)i} = p_i - (Rq_i + t)$$

(9)

For $n$ pairs of 3D matching points in the point set, the least squares expression of the objective function, $F(R, t)$, can be expressed as [29,30]:

$$F(R, t) = \arg\min_{R, t} \frac{1}{2} \sum_{i=1}^{n} ||p_i - (Rq_i + t)||^2$$

(10)

To solve Equation (10), let the de-centroid coordinates of point sets $P$ and $Q$ be $p'_i$ and $q'_i$, respectively, that is:

$$\begin{cases}
    p'_i = p_i - \frac{1}{n} \sum_{i=1}^{n} p_i \\
    q'_i = q_i - \frac{1}{n} \sum_{i=1}^{n} q_i
  \end{cases}$$

(11)

Combining Equations (10) and (11), the de-centroid form can be obtained as:

$$F(R, t) = \arg\min_{R, t} \frac{1}{2} \sum_{i=1}^{n} \left(||p'_i - Rq'_i||^2 + \frac{1}{n} \sum_{i=1}^{n} p_i - \left(R \cdot \frac{1}{n} \sum_{i=1}^{n} q_i + t\right)||^2\right)$$

(12)

Equation (12) can be decomposed into two functions, $f_1(R)$ and $f_2(t)$, namely:

$$f_1(R) = \arg\min_{R} \frac{1}{2} \sum_{i=1}^{n} ||p'_i - Rq'_i||^2$$

(13)

$$f_2(t) = \arg\min_{t} \frac{1}{2} \sum_{i=1}^{n} \left(\frac{1}{n} \sum_{i=1}^{n} p_i - \left(R \cdot \frac{1}{n} \sum_{i=1}^{n} q_i + t\right)||^2\right)$$

(14)

Further, Equation (13) can be expanded and simplified as:

$$g_1(R) = \arg\min_{R} \left[-\text{tr}\left(R \sum_{i=1}^{n} q'_i (p'_i)^T\right)\right]$$

(15)

where, $\text{tr}()$ is the trace operator of matrix.

To understand the $R$ contained in Equation (15), a method based on singular value decomposition is adopted, that is, let [31,32]:

$$\sum_{i=1}^{n} q'_i (p'_i)^T = UDV^T$$

(16)

where $D$, $U$, and $V$ are all diagonal matrices, and matrix $D$ is composed of singular values.

According to the principle of Cauchy–Schwarz’s inequality, when the objective function $g_1(R)$ in Equation (15) takes the minimum value and $DUV^T$ is full rank, $R$ can be obtained, the expression is:

$$R = UV^T$$

(17)

After $R$ is solved, let $f_2(t)$ in Equation (14) be zero, and $t$ can be solved, i.e.:

$$t = \frac{1}{n} \sum_{i=1}^{n} p_i - R \cdot \frac{1}{n} \sum_{i=1}^{n} q_i$$

(18)
So far, through the construction of the least squares problem of pose solution and singular value decomposition, the pose parameters, \( R \) and \( t \), of the roadheader body can be solved.

### 3.4. Pose Optimization

Pose optimization must be carried out because of the accumulated errors in pose solutions. Waypoints are regarded as constraints of pose optimization rather than as optimization variables; the optimization model of a body pose graph is established by selecting the image key frame. Additionally, closed-loop constraints are added to carry out global pose optimization, so as to ensure the effectiveness and real-time performance of pose estimation, and then the deviation correction is completed in time.

#### 3.4.1. Pose Graph Construction and Closed-Loop Detection

The selection principle of key frames is based on the evaluation of the distance between frames. After \( R \) and \( t \) are given different weights, the motion transformation index representing the distance between frames can be calculated. If the index is less than a given threshold, the current frame is determined to be a key frame. After selecting the key frame, the pose graph of the motion transformation constraint can be constructed. To further modify the computed pose, closed-loop detection is needed and closed-loop constraints are added to the pose graph. The key of closed-loop detection lies in the judgment of the similarity between two frames of images, i.e., if the similarity between the current frame and the historical key frame exceeds a certain value of similarity between the current frame and the previous-moment key frame, and the consistency of space and time is satisfied, a closed loop can be determined.

Based on the SLAM motion model of the roadheader, shown in Figure 3, the pose graph optimization model, which eliminates the waypoint variables and adds closed-loop constraints, is shown in Figure 6. The figure shows two processes of motion transformation and closed-loop detection, \( T_{(i-1)} \) is the motion transformation constraint between pose nodes \( x_{i-1} \) and \( x_i \), and \( T_{(i+1)n} \) is the closed-loop constraint between pose nodes \( x_{i+1} \) and \( x_n \).

![Figure 6. Optimized model of the pose graph.](image)

#### 3.4.2. Pose Graph Optimization

According to the pose graph optimization model shown in Figure 6, let the inter-frame transformation matrix obtained in the pose solution stage or closed-loop detection be \( T_{ij} \), and the pose estimation error be \( e_{ij} \); then, \( e_{ij} \) can be expressed as [33]:

\[
e_{ij} = \ln \left( T_{ij}^{-1} T_{i-1}^{-1} T_i \right) ^\vee \tag{19}
\]

where the sign \( ^\vee \) means to transform the matrix into a vector.

To optimize the global pose graph, the least squares expression of the objective function, \( G(T) \), based on the error term is constructed, that is [34]:

\[
G(T) = \arg \min_T \sum_{ij} e_{ij}^T \Sigma_{ij}^{-1} e_{ij} \tag{20}
\]
where $\Sigma_{ij}$ is the covariance matrix of the pose measurement, which is used to represent the uncertainty of the measurement.

Equation (20) is nonlinear and is solved using an iterative method in this paper. Let $\Delta x$ be the iterative increment, $x^*$ be the initial value of the pose estimation, and then the Taylor expansion form of the error, $e_{ij}$, at $x^*$ can be expressed as:

$$e_{ij}(x^* + \Delta x) = e_{ij} + J_{ij}\Delta x$$  \hspace{1cm} (21)

where $J_{ij}$ is the Jacobian matrix of $e_{ij}$ at $x^*$.

Combined with Equations (20) and (21), Equation (22) can be obtained after approximate processing, that is:

$$
\begin{cases}
H\Delta x = -b \\
H = \sum J_{ij}^T\Sigma_{ij}^{-1}J_{ij} \\
b = \sum e_{ij}^T\Sigma_{ij}^{-1}J_{ij}
\end{cases}$$  \hspace{1cm} (22)

By introducing Lagrange multiplier $\rho$ on the basis of Equation (22), the increment equation based on the Levenberg–Marquadt method is shown in Equation (23):

$$(H + \rho I)\Delta x = -b$$  \hspace{1cm} (23)

where $I$ is the identity matrix.

After multiple iterations of Equation (23), when the termination condition of the iteration is satisfied, the optimized pose estimation can be obtained. Subsequently, a closed-loop control and deviation correction can be carried out for the traveling of the roadheader according to the optimized body pose, and trajectory planning of the cutting arm can be performed according to robotics to ensure the forming quality of the roadway section.

4. Map Construction and Test

4.1. Overview of Mapping Scheme

Although the acquired pose can provide a priori information for the moving of the roadheader, its navigation also depends on the environment map. Based on SLAM technology, a roadway map is established, which can help the roadheader identify obstacles and judge roadway edge shapes, and provide support for remote man–machine interactions, roadheader travel prediction, section forming control, and so on.

Common mapping schemes are shown in Table 1, including laser-based and vision-based schemes, and the real triangle symbol indicates that the algorithm supports the relevant interface. It can be seen from Table 1 that the RTAB-MAP algorithm can provide a variety of sensor input interfaces, and can output a dense 3D point cloud map, which is currently an excellent application-based mapping solution.

| Mapping Algorithm | Input | Output |
|-------------------|-------|--------|
|                   | Camera | Laser Radar | Odometer | Grid Map | Feature/Point Cloud Map |
|                   | Monocular | Binocular | RGB-D | 2D | 3D | 2D | 3D | Sparse | Dense |
| Gmapping [35]     | ▲       | ▲       | ▲ | ▲       | |
| Cartographe [36]  | ▲       | ▲       | ▲ | ▲       | |
| Karto [37]        | ▲       | ▲       | ▲ | ▲       | |
| Hector [38]       | ▲       | ▲       | ▲ | ▲       | |
| ORB-SLAM1 [39]    | ▲       | ▲       | ▲ | ▲       | |
| ORB-SLAM2 [40]    | ▲       | ▲       | ▲ | ▲       | |
| RTAB-MAP [19]     | ▲       | ▲       | ▲ | ▲       | ▲ | ▲ | ▲ | ▲ | ▲ | ▲ |
Comprehensively considering the underground environment of a coal mine as well as the existing conditions of a laboratory, a long corridor is adopted to simulate an underground roadway, and some mapping algorithms, listed in Table 1, are tested. The long corridor used for test, as shown in Figure 7, is as close as possible to the dimly lit, narrow and closed roadway environment.

Figure 7. Long corridor used for test.

4.2. Construction of Grid Map and Test

4.2.1. Principle of Laser Mapping

To construct an occupancy grid map based on laser SLAM, the key is to use the prior information to calculate the posterior probability of the map, and then identify the grid state, i.e., unknown, idle, or occupied [41,42]. The grid state schemas formed after two-dimensional (2D) LiDAR scanning are shown in Figure 8; O_sX_sY_s and O_bX_bY_b are the grid coordinate system and body coordinate system, respectively. The white area in the grid is idle, the black area in the grid is occupied, and the gray area in the grid is unknown, which is the target area for further exploration using SLAM technology.

Figure 8. Construction principle of occupied grid map.
The measured data after the laser beam is emitted by the 2D LiDAR is denoted as $z^L$, and is as shown in Equation (24). According to probability theory, if $s = 1$ that means that the grid is occupied, and if $s = 0$ that means that the grid is idle; then, the updated rule of state $S$ for any grid is shown in Equation (25).

$$z^L = \begin{bmatrix} d \\ \theta \end{bmatrix}^T$$

where $d$ represents the distance of the measuring point, and $\theta$ represents the included angle between the laser beam and the body orientation.

$$S_i = S_{i-1} + \log \frac{p(z | s = 1)}{p(z | s = 0)}$$

where $p()$ is the probability calculation expression.

The map constructed using laser SLAM can provide a priori information for navigation and obstacle avoidance of the roadheader, but it is not safe to move in the occupancy grid map using only Boolean values [42]. To realize global navigation of the roadheader and to avoid dynamic obstacles in real time, an obstacle map layer and expansion layer should be added, based on the static map layer constructed using SLAM; that is, a cost map for navigation should be constructed. The value principle of a cost map is shown in Figure 9, where each grid state is represented by a gray value from 0 to 255. The dashed line frame in the figure represents the contours of the roadheader, and the grid value of 254 indicates that the obstacles coincide with the center of the body, which will inevitably lead to a collision. The single dotted circle represents the inner tangent circle of the body contour line, the grid value in the inner tangent circle is 253, and collision is also inevitable in this area. The double-dot dotted circle represents the outer tangent circle of the body contour line, the grid value range in the outer tangential circle is 128–252, and collisions may occur in this area; the three-point scribbled circle represents the buffer area, and the grid value range is 1–128, which should be avoided. The grid value in the three-point double-dashed circle is 0, and the area is allowed free passage. The gray block with a grid value of 255 represents an unknown area, which needs to be updated according to the sensor information [43].

Figure 9. Value principle of cost map.
4.2.2. Mapping Test Based on Laser SLAM

Laser SLAM is divided into filtering-based and optimization-based, the Gmapping algorithm [35] in Table 1 belongs to the former, and the algorithms, Cartographer [36], Karto [37], and Hector [38], belong to the latter. In this paper, the four abovementioned laser SLAM algorithms are used to carry out mapping tests of the simulated roadway along the long corridor shown in Figure 7.

The test results of the four algorithms are shown in Figure 10. Figure 10a shows a map of simulated roadway, built based on the Gmapping algorithm, and the Gmapping algorithm relies on the Rao–Blackwellized particle filter, which takes laser data and odometer data as inputs; the map constructed based on the Cartographer algorithm is shown in Figure 10b. The Cartographer algorithm consists of two phases of local and global SLAM, with the first phase establishing and maintaining grid submaps, and the second phase eliminating cumulative errors and carrying out global optimization. Figure 10c shows a map built based on the Karto algorithm, which is based on the graph optimization theory, and uses the graph mean to represent the map, and the front-end and back-end are carried out by a single thread. The map constructed based on the Hector algorithm is shown in Figure 10d; the Hector algorithm only creates a grid map by matching laser data without the need for odometer data.

![Grid map of simulated roadway based on laser SLAM.](image)

Figure 10. Grid map of simulated roadway based on laser SLAM.

The Gmapping algorithm based on the filtering framework has a lower computational complexity, but it relies too heavily on odometer data. Although the Hector algorithm is not limited to the odometer, it requires a high laser frequency. In addition, the Cartographer algorithm is relatively resource intensive. Consequently, based on the 2D occupy grid map constructed using the Karto algorithm, the cost map for the navigation of the roadheader is constructed by introducing parameters, such as the expansion radius of obstacles and the updating frequency of map. The simulation test results are shown in Figure 11, the red arrow represents the body pose.
4.3. Construction of 3D Map and Test

4.3.1. Principle of Vision Mapping

According to Table 1, the map constructed using the ORB-SLAM1/2 algorithm is a sparse point cloud map, which cannot be used for navigation and obstacle avoidance of the roadheader. In addition, the algorithm can easily miss a target when the airborne camera rotates and it is sensitive to noise. Considering the complex environment of underground roadways and the requirements of tunneling work, the performance of running and being in real-time, the accuracy of algorithm should be given more priority. The RTAB-MAP framework provides a map construction scheme that is independent of scale and time. As shown in Figure 12 [19], its internal memory management is implemented through three memory nodes, i.e., WM (working memory), LTM (long-term memory), and STM (short-term memory). The weight of the registration point is updated in STM, and the closed-loop detection of the registration point is completed in WM. When the number of registration points exceeds a certain threshold, the registration point with the longest storage time is placed in WM from STM. To ensure real-time Bayesian estimation of the number of registration points, two functional nodes, namely fetch and transfer, are defined between WM and LTM during closed-loop detection in WM. When the closed loop is detected in the WM, the adjacent location points in the closed loop will be retrieved from the LTM and placed in the WM. When the processing time for closed-loop detection exceeds a certain threshold, the anchor point with the lowest weight is transferred to the LTM [19]. Based on the above memory management mechanism, the RTAB-MAP framework avoids the problem of having map data that are too large, which can meet the engineering needs of long-term, real-time, and online mapping in a large-scale environment.

![Figure 12. Memory management mechanism of RTAB-MAP.](image)

The RTAB-MAP algorithm can build a dense point cloud map. Considering the disadvantage of a large amount of data in a dense point cloud, a portable stored octree map can be formed after octree compression [44,45], which can be used for stereoscopic navigation of the roadheader. The data structure of the octree is shown in Figure 13; each parent node contains eight child nodes, among which the leaf layer node corresponds to the map with the highest resolution. In addition, each node points to a cube voxel, and

![Figure 13. Principle of octree compression.](image)
the probability of the \( n \)-th node corresponding to the voxel being occupied is calculated according to Bayesian theory, as shown in Equation (26) [45].

\[
p(n|z^c_1:t) = \left[ \frac{1}{1 + \frac{1 - p(n|z^c_1:t)}{p(n|z^c_1:t)} \cdot \frac{1 - p(n)}{p(n)}} \cdot \frac{1 - p(n)}{p(n)} \right]^{-1}
\]

where \( z^c \) represents the observation data of the airborne camera at the corresponding subscript time.

4.3.2. Mapping Test Based on Vision SLAM

Based on RTAB-MAP framework, the image data collected by RGB-D camera are input, the dense 3D point cloud map is constructed by mapping the test of the simulated roadway (shown in Figure 14), and the octree map is constructed via octree compression and is shown in Figure 15.

![Figure 13. Principle of octree compression.](image)

![Figure 14. Dense 3D point cloud map of simulated roadway.](image)
Octree map is convenient for navigation of the roadheader because of its portable characteristics. Based on the octree map shown in Figure 15, the multi-layer cost map for navigation is constructed by introducing parameters, such as the expansion radius of obstacles and the frequency of map updates, as shown in Figure 16. The red arrow in the figure represents the body pose, the darkly colored expansion area around the body is a local cost map, which is used for local path planning; the rest of the area is a global cost map, which is used for global path planning. In this test, the Dijskra algorithm is used for global path planning, and the red lines in the figure are the trajectory of the global path planning. The TEB (timed-elastic-band) algorithm is used for local path planning, and the blue line is the trajectory of local path planning.

Figure 16. Navigation and obstacle avoidance based on the octree map and cost map.
4.3.3. Mapping Test Based on Fusion of Laser and Vision

Although the mapping accuracy of laser SLAM is high and the computational complexity is low, the map constructed is a sparse point cloud, lacks semantic information, and has large linearization errors. In contrast, the environmental information obtained by the camera is much more than that obtained by LiDAR, and the semantic information is rich. Moreover, the RGB-D camera has the advantage of not being affected by the change in ambient light and texture, but its mapping accuracy is affected by the cumulative error. Based on the above considerations, the two mapping schemes are complementary. Specifically, in the dynamic environment of a coal mine roadway with strong interference, the mapping scheme based on the laser and vision fusion has a stronger robustness.

The RTAB-MAP framework provides an interface between the camera and LiDAR [19], as shown in Figure 17. First, the laser and visual data are entered, then the multi-source information is initialized and aligned, and finally a map of the environment is output.

![Figure 17. The RTAB-MAP mapping process based on vision and laser.](image)

The mapping result of the simulated roadway, based on the fusion of laser and visual data, is shown in Figure 18. The map is in a composite form, and includes a dense 3D point cloud and a 2D occupied grid. Compared with Figure 14, it can be seen that the SLAM mapping scheme that integrates laser and vision has more advantages in detail processing of the point cloud, and the map is more refined. Additionally, the point cloud can be used for 3D reconstruction of the environment after gridding and texture mapping in the later stage, and the 2D occupancy grid map can be used for roadheader navigation after being converted into a cost map.

![Figure 18. Composite map based on visual and laser data.](image)
5. Conclusions
In this paper, according to the development requirements of intelligent and robotization of fully mechanized excavation faces, the method of estimating roadheader poses and constructing navigation maps by using vision and laser SLAM technology is proposed. The main conclusions are as follows:
(1) The positioning method based on visual SLAM technology is proposed, the environment data are collected by an airborne RGB-D camera, and the RANSAC+ICP model is constructed to solve and realize the autonomous measurement of pose by establishing and solving the optimization model of pose graph with closed-loop constraints. The global consistent pose of the roadheader can be obtained, which can effectively solve the difficult positioning and orientation problem of roadheaders in a restricted space.
(2) A series of open-source algorithms, such as Gmapping, Cartographer, Karto, Hector, and RTAB-MAP, are applied to construct simulated roadway maps, including 2D occupied grid, 3D point cloud, and octree, based on a static map created by SLAM. The cost map for roadheader navigation is further constructed by setting parameters, such as expansion radius of obstacles and the updating frequency of map. Furthermore, obstacle avoidance tests based on the Djiskra+TEB algorithm show that a combination of octree map and cost map can support global path planning and local dynamic obstacle avoidance.
(3) Based on the open-source library RTAB-MAP framework, by fusing laser and vision data, a composite map composed of a dense 3D point cloud and a 2D occupancy grid is constructed, which could, not only be used for plane navigation of a roadheader, but also provide support for 3D reconstruction of an environment.

Author Contributions: Conceptualization, G.Z. and W.Z.; methodology, W.Z.; software, G.Z. and W.Z.; validation, W.Z., G.Z., and Z.Y.; formal analysis, G.Z., Z.Y., and T.P.; investigation, W.Z.; data curation, W.Z. and R.C.; writing—original draft preparation, W.Z.; writing—review and editing, W.Z. and G.Z.; visualization, W.Z. and R.C.; supervision, G.Z., Z.Y., and T.P.; funding acquisition, G.Z. and T.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Fundamental Research Funds for the Central Universities (Grant No. 2021YJSJD10); the National Key Research and Development Program of China (Grant No. 2017YFC0804307); Fujian Provincial Key Laboratory of Information Processing and Intelligent Control (Minjiang University) (Grant No. MJUKF-IPIC201905).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References
1. Wang, G.; Xu, Y.; Ren, H. Intelligent and ecological coal mining as well as clean utilization technology in China: Review and prospects. Int. J. Min. Sci. Technol. 2019, 29, 161–169. [CrossRef]
2. Wang, J.; Huang, Z. The Recent Technological Development of Intelligent Mining in China. Engineering 2017, 3, 439–444. [CrossRef]
3. Tian, J.; Wang, S.; Wu, M. Kinematic models and simulations for trajectory planning in the cutting of Spatially-Arbitrary crosssections by a robotic roadheader. Tunn. Undergr. Space Technol. 2018, 78, 115–123. [CrossRef]
4. Deshmukh, S.; Raina, A.K.; Murthy, V.M.S.R.; Trivedi, R.; Vajre, R. Roadheader—A comprehensive review. Tunn. Undergr. Space Technol. 2020, 95, 103148. [CrossRef]
5. Tao, Y.F.; Zong, K.; Zhang, M.J.; Jia, W.H.; Fu, S.C.; Yang, J.J.; Wu, M. Aposition and orientationmeasurement method of single-station, multipoint and time-sharing forroadheaderbody based on iGPS. J. China Coal Soc. 2015, 40, 2611–2616.
6. Roman, M.; Josif, B.; Nikolay, D. Development of Position System of a Roadheader on a Base of Active IR-sensor. In Proceedings of the 25th Daaam International Symposium on Intelligent Manufacturing and Automation, Oldest Technical Univ, Vienna, Australia, 26–29 November 2015; pp. 617–621.
36. Hess, W.; Kohler, D.; Rapp, H.; Andor, D. Real-time loop closure in 2D LIDAR SLAM. In Proceedings of the 2016 IEEE International Conference on Robotics and Automation (ICRA), Stockholm, Sweden, 16–21 May 2016; pp. 1271–1278.

37. Vincent, R.; Limketkai, B.; Eriksen, M. Comparison of indoor robot localization techniques in the absence of GPS. In Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XV; International Society for Optics and Photonics: Bellingham, WA, USA, 2010; p. 7664.

38. Kohlbrecher, S.; Von Stryk, O.; Meyer, J.; Klingauf, U. A flexible and scalable SLAM system with full 3D motion estimation. In Proceedings of the 2011 IEEE International Symposium on Safety, Security, and Rescue, Kyoto, Japan, 1–5 November 2011; pp. 155–160.

39. Mur-Artal, R.; Montiel, J.M.M.; Tardos, J.D. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. IEEE Trans. Robot. 2015, 31, 1147–1163. [CrossRef]

40. Mur-Artal, R.; Tardos, J.D. ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras. IEEE Trans. Robot. 2017, 33, 1255–1262. [CrossRef]

41. Qi, X.; Wang, W.; Yuan, M.; Wang, Y.; Li, M.; Xue, L.; Sun, Y. Building semantic grid maps for domestic robot navigation. Int. J. Adv. Robot. Syst. 2020, 17, 12. [CrossRef]

42. O’Meadhra, C.; Tabib, W.; Michael, N. Variable Resolution Occupancy Mapping Using Gaussian Mixture Models. IEEE Robot. Autom. Lett. 2018, 4, 2015–2022. [CrossRef]

43. Sun, J.; Chen, Z.H.; Wang, P.; Zhang, Q.B.; Bao, P. Multi-region coverage method based on cost map and minimal tree for mobile robot. Robot 2015, 37, 435–442.

44. Schön, B.; Mosa, A.S.M.; Laefer, D.F.; Bertolotto, M. Octree-based indexing for 3D pointclouds within an Oracle Spatial DBMS. Comput. Geosci. 2013, 51, 430–438. [CrossRef]

45. Jessup, J.; Givigi, S.N.; Beaulieu, A. Robust and Efficient Multirobot 3-D Mapping Merging with Octree-Based Occupancy Grids. IEEE Syst. J. 2015, 11, 1723–1732. [CrossRef]