Research Article

An Improved Extended Information Filter SLAM Algorithm Based on Omnidirectional Vision

Jingchuan Wang and Weidong Chen

Department of Automation, Key Laboratory of System Control and Information Processing, Shanghai Jiao Tong University, Ministry of Education of China, Shanghai 200240, China

Correspondence should be addressed to Jingchuan Wang; jchwang@sjtu.edu.cn

In the SLAM application, omnidirectional vision extracts wide scale information and more features from environments. Traditional algorithms bring enormous computational complexity to omnidirectional vision SLAM. An improved extended information filter SLAM algorithm based on omnidirectional vision is presented in this paper. Based on the analysis of structure a characteristics of the information matrix, this algorithm improves computational efficiency. Considering the characteristics of omnidirectional images, an improved sparsification rule is also proposed. The sparse observation information has been utilized and the strongest global correlation has been maintained. So the accuracy of the estimated result is ensured by using proper sparsification of the information matrix. Then, through the error analysis, the error caused by sparsification can be eliminated by a relocation method. The results of experiments show that this method makes full use of the characteristic of repeated observations for landmarks in omnidirectional vision and maintains great efficiency and high reliability in mapping and localization.

1. Introduction

The first problem of a mobile robot is that it has to deal with a complicated environment when the robot fulfills its assignments. Simultaneous localization and mapping (SLAM) is one of the key enabling technologies for a mobile robot's autonomous ability. SLAM addresses the problem of building up a map within an unknown environment, while at the same time keeping track of their current location. Many popular SLAM implementations use a distance sensor such as a laser range finder or sonar to explore the environments [1–3]. We propose to extend this approach to an omnidirectional vision-based system for economic reasons.

Many vision sensor applications have been developed in recent years. Vision sensors can provide continuous image data. Until now, the research work about vSLAM mainly used stereo vision or monocular vision [4, 5]. Paper [6] presents a vSLAM method based on an encoder integrated system and applies a distributed particle filter approach. A vSLAM method based on a tracking approach for "2.5D space" is proposed in paper [7]. A semantic structure of environments is built based on object recognition and laser scan [8]. A calibrated monocular camera is used in a vehicle SLAM system [9]. 3D reconstruction of the environment is built through the matching interest points.

However, there are limitations in the view angle of ordinary vision sensors. In contrast, omnidirectional vision has a 360 degree field of view and it has been widely applied in robot navigation, video conferencing, and surveillance [10]. Omnidirectional vision based robot localization has been researched and has been popular for years. Paper [11] develops a topological localization method, where the matching algorithm of omnidirectional images based on a color histogram is proposed. It obtains relatively complete environmental information and strengthens the capability of tracking a target; this makes up for deficiencies in the study of vSLAM.

Omnidirectional vision not only provides opportunities for studying vSLAM problems, but it also provides several challenges. First, the large amount of information will increase the computational complexity of the SLAM algorithm; second, the distortion of the omnidirectional

image is relatively large, which is very difficult to extract and match features directly. As a result, vSLAM methods based on omnidirectional vision have seldom been applied in the case of localization and mapping for an unknown environment. Therefore, the omnidirectional vision based SLAM is not only able to promote vSLAM technology, and provide a new idea, but it also can broaden the application field of omnidirectional vision (e.g., environment exploration and rescue).

The extended information filter SLAM (EIF-SLAM) algorithm [3, 12] describes the uncertainty of SLAM by information matrix and vector. However, the efficiency of traditional EIF-SLAM is very low due to the enormous computation of mean vector restoration from matrix and vector information. Papers [12, 13] improve the accuracy of EIF-SLAM through sparsification of the information matrix by constructing a reasonable topology structure of Bayes network. Paper [14] eliminates elements by using a new sparsification approach in EIF-SLAM and maintains SLAM consistent. Because there are more features in the omnidirectional images, these methods cannot reduce the computational complexity due to the large landmarks classification. In this paper, based on an analysis of structural characteristics of the information matrix, an improved extended information filter SLAM algorithm (IEIF-vSLAM) and sparsification rule are presented. The analysis of sparsification error and the verification of experiments show that this algorithm not only improves computational efficiency, but also maintains the accuracy of the estimated result by using proper sparsification of the information matrix. It can be used in a mobile robot and fulfill map building tasks though omnidirectional vision in indoor environments.

2. IEIF-vSLAM Algorithm

2.1. Algorithm Structure. The structural diagram of the proposed IEIF-vSLAM algorithm is shown in Figure 1. It includes 3 parts: features detection, IEIF, and robot control.

In the module of features detection, a Harris-SIFT feature extraction method based on SIFT [10] is used to match and track features [15]. The KLT operator [16] is used to track Harris-SIFT feature points. The RANSAC method [17] is used for affine transform and removing the wrong match points which are caused by blocking. If a new landmark has been found, it will be recorded in the feature database.

In the IEIF module, the information variables based on the observation data will be updated in blocks of observation update. According to the movement of the robot, the state variables will be estimated in the block of motion update. In fact, the computational burden of IEIF-vSLAM is mainly concentrated on both of the processes of observation and of motion update. Due to the wide scale of information and a huge amount of omnidirectional image feature, the dimension of information variables rapidly increases. This situation will heat up if the environment becomes larger [18]. Therefore, it is reasonable for an improved IEIF-vSLAM algorithm—based on an analysis of structural characteristics of the information matrix—to improve computational efficiency and maintain accuracy of the state estimated.
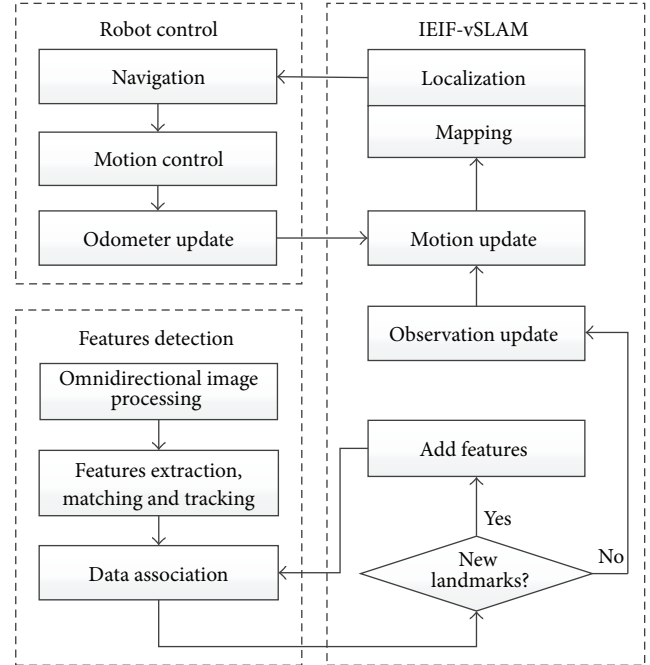

Figure 1: IEIF-vSLAM algorithm structure diagram.

2.2. Information Matrix of SLAM. At time t, the state estimation of robot pose and landmark position satisfies Gaussian distribution:

$$
\begin{aligned}
p\left(x_t, M \mid z^t, u^t\right) &= N\left(\begin{bmatrix} \mu_{x_t} \\ \mu_M \end{bmatrix}, \begin{bmatrix} \Sigma_{x_t x_t} & \Sigma_{x_t M} \\ \Sigma_{M x_t} & \Sigma_{MM} \end{bmatrix}\right) \\
&= N^{-1}\left(\begin{bmatrix} \eta_{x_t} \\ \eta_M \end{bmatrix}, \begin{bmatrix} \Lambda_{x_t x_t} & \Lambda_{x_t M} \\ \Lambda_{M x_t} & \Lambda_{MM} \end{bmatrix}\right) \quad (1) \\
&\doteq N\left(\mu_t, \Sigma_t\right) = N^{-1}\left(\eta_t, \Lambda_t\right).
\end{aligned}
$$

x_t: State of robot, including the pose of the robot.
M: Set of landmarks.
z^t: The observation of landmarks.
u^t: The input vector of robot control.

As shown in (1), if u^t and z^t are given, the joint conditional probability density of x_t and M satisfies normal distribution in time domain, whose mean value is μ_t and covariance matrix is Σ_t. It also satisfies inverse normal distribution, whose mean value is η_t and covariance matrix is Λ_t in the information domain.

The IEIF-vSLAM algorithm estimates the robot pose and landmarks positions according to the update of information vector η_t and information matrix Λ_t. However, η_t and Λ_t have no obvious physical meanings, so they should be converted to μ_t and Σ_t. The conversion formula is

$$
\Lambda_t = \Sigma_t^{-1} \qquad \eta_t = \Lambda_t \mu_t. \tag{2}
$$

As shown in Figure 1, observation will be updated at every motion cycle through movements. The motion update,

observation update, and add features are expressed as follows [13]:

Motion Update

$$p\left(x_{t+1}, Mz^t, u^{t+1}\right) = N^{-1}\left(\eta_{t+1}, \Lambda_{t+1}\right),$$

$$\Lambda_{t+1} = \begin{bmatrix} \Psi & Q^{-1}F\Omega^{-1}\Lambda_{x_t M} \\ \Lambda_{Mx_t}\Omega^{-1}F^T Q^{-1} & \Lambda_{MM} - \Lambda_{Mx_t}\Omega^{-1}\Lambda_{x_t M} \end{bmatrix}, \quad (3)$$

$$\eta_{t+1} = \begin{bmatrix} Q^{-1}F\Omega^{-1}\eta_{x_t} + \Psi u \\ \eta_M - \Lambda_{Mx_t}\Omega^{-1}\left(\eta_{x_t} - F^T Q^{-1} u\right) \end{bmatrix},$$

where

$$\Psi = \left(Q + F\Lambda_{x_t x_t}^{-1}F^T\right)^{-1},$$

$$\Omega = \left(\Lambda_{x_t x_t} + F^T Q^{-1}F\right), \quad (4)$$

$$u = F^T Q^{-1}\left(f\left(\mu_{x_t}, u_{t+1}\right) - F\mu_{x_t}\right)$$

$\Lambda_{x_t x_t}$ expresses the robot poses' own correlation.

F is the Jacobian matrix of the mobile robot's motion model.

Add Features

$$p\left(x_{t+1}, x_t, Mz^t, u^{t+1}\right)$$

$$= N^{-1}\left(\eta_{t+1}, \Lambda_{t+1}\right),$$

$$\Lambda_{t+1} = \begin{bmatrix} \Lambda_{x_t x_t} + G^T R^{-1}G & \Lambda_{x_t M} & -G^T R^{-1} \\ \Lambda_{Mx_t} & \Lambda_{MM} & 0 \\ -R^{-1}G & 0 & R^{-1} \end{bmatrix},$$

$$\quad (5)$$

$$\eta_{t+1} = \begin{bmatrix} \eta_{x_t} - G^T R^{-1}u \\ \eta_M \\ R^{-1}u \end{bmatrix}$$

$$u = g\left(\mu_{x_t}, z_t\right) - G|_\mu \mu_{x_t}$$

$\Lambda_{x_t M}$ expresses the correlation between robot pose and landmark.

G is the Jacobian matrix of omnidirectional vision's observation model.

Observation Update

$$p\left(x_t, M \mid z^{t+1}, u^t\right) = N^{-1}\left(\eta_{t+1}, \Lambda_{t+1}\right),$$

$$\Lambda_{t+1} = \Lambda_t + H^T R^{-1}H,$$

$$\quad (6)$$

$$\eta_{t+1} = \eta_t + H^T R^{-1}u,$$

$$u = z_t - h\left(\mu\right) + H|_\mu \mu.$$

In (6), H is the sparsification matrix. Given the main processes of EIF-SLAM, the question of how to reduce the computation complexity of motion update, observation update, and add features to the database is important to EIF-SLAM's application. The performance of enhancing efficiency by sparsification should be reasonable. Prior to discussion of sparsification, the observation model of omnidirectional vision is analyzed.

2.3. Observation Model of Omnidirectional Vision. In an omnidirectional vision system, z_{F_i} is the observation of features (landmarks), its position in image pixel coordinate $(x_{F_i}, y_{F_i})^T$ is $(u_{F_i}, v_{F_i})^T$. $w(k)$ is its Gaussian noise. $R(k)$ is its covariance matrix. The observation measurements, $(r_{F_i}, \theta_{F_i})^T$ which are similar to those of a laser sensor, are used to replace $(u_{F_i}, v_{F_i})^T$. r_{F_i} is the distance between the robot and features and θ_{F_i} is the head angle for features in robot local coordinates. This allows us to describe the derivation process of transformation.

Figure 2 is the omnidirectional vision schematic plot, which is fixed on top of the robot. The height of the lens from the floor is h. Omnidirectional vision's hyperboloid mirror is designed as $z^2/a^2 - (x^2 + y^2)/b^2 = 1$. Its focal point is $C(u_c, v_c)^T$ (one of the foci of the mirror) and the focal distance is f. The imaging pixel of features Fi is $(u_{F_i}, v_{F_i})^T$ [19].

Then,

$$\theta_{Fi} = \text{atan2}\left(u_{F_i} - u_c, v_{F_i} - v_c\right). \quad (7)$$

As shown in Figure 2, the 2D coordinates are built based on the other focus point of the mirror. Then, the function of the mirror is changed to

$$\frac{(z+e)^2}{a^2} - \frac{x^2}{b^2} = 1, \quad e = \sqrt{a^2 + b^2}. \quad (8)$$

The reflected light goes through the center point of the lens: $C = (0, -2e)^T$, its slope is

$$k = \frac{f}{\sqrt{\left(u_c - u_{F_i}\right)^2 + \left(v_c - v_{F_i}\right)^2}}. \quad (9)$$

Then the intersection point $(x_m, z_m)^T$ between the incident light ray and the mirror will be calculated as

$$\frac{(z_m + e)^2}{a^2} - \frac{x_m^2}{b^2} = 1$$

$$z_m = kx_m - 2e$$

$$x_m = \frac{b^2}{ke - \sqrt{a^2\left(1 + k^2\right)}} \quad (10)$$

$$\implies$$

$$z_m = \frac{kb^2}{ke - \sqrt{a^2\left(1 + k^2\right)}} - 2e \cdot$$

According to the geometrical relation [19]:

$$\frac{r_{F_i}}{2e + h} = -\frac{x_m}{z_m} \implies r_{F_i} = \frac{b^2\left(2e + h\right)}{k\left(a^2 + e^2\right) - 2ea\sqrt{1 + k^2}}. \quad (11)$$

So, the new observations $(r_{F_i}, \theta_{F_i})^T$ are obtained through (7) and (11). The observation of landmarks in different distance for omnidirectional vision has the same gauss distribution $R(k)$. However, the uncertainty of landmarks observation $P_{F_i}(k)$ is different and is determined by calibration [20]. In hyperboloid mirror, $a = 29.0949$, $b = 23.4125$, and $h = 1650$ mm. The uncertainty of one radius axis of omnidirectional image (640×640) is obtained by calibration with the interval of 5 pixels. According to the characteristics of center symmetry of omnidirectional image, the uncertainty in other direction of whole image will also be determined and recorded in a table for query.

According to (6), we get

$$
\begin{aligned}
u &= z_t - h(\mu) + H|_\mu \mu \\
&= \begin{bmatrix} r \\ 0 \end{bmatrix} - \begin{bmatrix} \sqrt{(\mu_{fx_i} - \mu_x)^2 + (\mu_{fy_i} - \mu_y)^2} \\ \mathrm{atan2}\left(\dfrac{\mu_{fy_i} - \mu_y}{\mu_{fx_i} - \mu_x}\right) - \mu_\theta \end{bmatrix} \\
&\quad + \begin{bmatrix} -\dfrac{\mu_{fx_i} - \mu_x}{r} & -\dfrac{\mu_{fy_i} - \mu_y}{r} & 0 & \cdots & \dfrac{\mu_{fx_i} - \mu_x}{r} & \dfrac{\mu_{fy_i} - \mu_y}{r} & \cdots \\[2mm] \dfrac{\mu_{fy_i} - \mu_y}{r^2} & -\dfrac{\mu_{fx_i} - \mu_x}{r^2} & -1 & \cdots & -\dfrac{\mu_{fy_i} - \mu_y}{r^2} & \dfrac{\mu_{fx_i} - \mu_x}{r^2} & \cdots \end{bmatrix} \begin{bmatrix} \mu_x \\ \mu_y \\ \mu_\theta \\ \vdots \\ \mu_{fx_i} \\ \mu_{fy_i} \\ \vdots \end{bmatrix}.
\end{aligned}
\tag{12}
$$

Therefore, H in (6) is described in a sparsification matrix expression [13].

3. Sparsification of Information Matrix

3.1. Structure Characteristics Analysis. According to the process description of motion update, observation update and add features, the structural characteristics [14] of the information matrix based on the EIF formula can be concluded as follows:

(1) information matrix is Hermit matrix of symmetric positive definite [12, 13];

(2) according to (3), the element of Λ_{MM} will be decreased during movement;

(3) according to (6), H is the sparsification matrix. Then, in $H^T R^{-1} H$, only the elements which are related to robot pose and positions of unupdated landmarks are nonzero. Furthermore, observation update strengthens the expression of the unupdated landmarks.

Then, the structural characteristics of information matrix are obtained as follows:

(1) two elements in the diagonal line denote the link strength of two notes;

(2) if the shortest link between two nodes is longer, then that from the main diagonal in the information matrix of the relevant elements is farther;

(3) the endpoints of vice diagonal denote the link strength between the current state x_t and landmark $(r_{F_i}, \theta_{F_i})^T$, which is relevant to x_t. So, if the robot observes the landmark again, the link strength becomes larger.

From the above analysis, most of the elements in the information matrix of omnidirectional vision SLAM are nearly zero. It is reasonable for the structure of the information matrix to be sparsed.

Because of th0e required omnidirectional vision's observation model linearization and data association, the information matrix should be calculated in these processes (motion update, observation update, and add features). Among them, the main computation is concentrated on solving (2). In order to analyze this conveniently, it has been transformed into standard linear equations to solve the problem:

$$
\begin{aligned}
\Lambda \Sigma &= I, \\
\Lambda \mu &= \eta,
\end{aligned}
\tag{13}
$$

where, both I and Λ are the unit matrix with the same dimension.

With an increase of landmarks is in the map, the computation cost of solving linear equations will increase greatly. As mentioned above, the information matrix is an almost sparse matrix. If we can reasonably use the sparse characteristic of the information matrix, the efficiency of the calculation of extended information filtering will be enhanced.

Compared with the dense matrix, the efficiency of linear equations solving through a sparse matrix will be improved significantly. The sparse matrix operation method [21] is used in this paper. Due to the special operation rule of the sparse matrix, it can greatly reduce the computation cost. If the sparse degree is unchanged, the inversion cost for sparse and
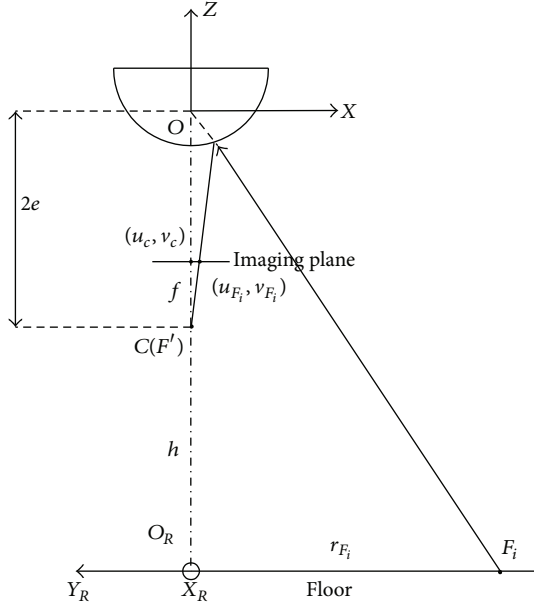
FIGURE 2: Omnidirectional vision schematic plot.

---

while $t > N$

$$\Lambda_t(i,j) = \begin{cases} 0 & |\Lambda_t(i,j)| \le \dfrac{1}{\lambda \dim(\Lambda_t)^2 \, \|\Lambda_{t-1}^{-1}\|} \\ \Lambda_t(i,j) & \text{others} \end{cases}$$

$\forall i,j \quad 1 \le i \le r(\Lambda_t) \quad 1 \le j \le c(\Lambda_t)$

end

ALGORITHM 1: Sparsification rule for information matrix.

---

dense matrix has an exponential relationship with the matrix dimension increase.

3.2. Improved Sparsification Rule. The improved sparsification rule is proposed in this chapter to reduce the computation cost of EIF-SLAM. According to the analysis of the omnidirectional vision observation model and the structural characteristics of the information matrix, this rule is described as shown in Algorithm 1.

Where, $0.5 \le \lambda \le 1$ denotes the conservative coefficient, $N > 1$ is the initial step of sparsification, Λ_t is the information matrix at time t, $\dim(\Lambda_t)$ denotes the dimension of Λ_t, $r(\Lambda_t)$ and $c(\Lambda_t)$ denotes the number of rows and columns of the information matrix. The larger N is, the higher the precision of EIF-SLAM will be and the poorer the real time performance will be. The larger λ is, the more conservative the initial sparsification is.

Only motion update will make the information matrix dense [22]. Therefore, the process of sparsification should be carried out once in every time cycle. Under the above analysis, the proposed sparsification algorithm will lead to higher computational efficiency. However, the proposed sparsification algorithm will bring an error to the information matrix and decrease the accuracy of state estimation simultaneously.

In SLAM application, the error caused by sparsification can be eliminated by a loop closure method. A closed loop environment should be designed for a robot to navigate; however it is unrealistic in practice. In information matrix, the correlations are expressed as the elements of matrix. The difference of different correlations is large due to the distortion and wide range view of omnidirectional vision. So, the sparsification rule has been improved in this chapter. The sparse observation information has been utilized and the strongest global correlation has been maintained.

Based on the improved sparsification rule, a relocation method is used to eliminate the error. Relocation refers to the robot finding landmarks which have been recorded in its database (this area has been explored before).

In the process of sparsification, environmental features map are divided into three independent parts [22]:

$$M = M^0 + M^+ + M^-. \tag{14}$$

Among that, M^0 are the passive features (they have some correlation with robot pose) and are converted from active features (they have no correlations with robot pose), M^+ are the features which are still maintained active, and M^- are the features which are still maintained passive.

In SLAM, the posterior probability of state vector is $P(X_t, M^0, M^+, M^-)$. According to Bayes theory, it will break down as follows:

$$P\left(X_t, M^0, M^+, M^-\right)$$
$$= P\left(X_t \mid M^0, M^+, M^-\right) P\left(M^0, M^+, M^-\right) \tag{15}$$
$$= P\left(X_t M^0, M^+, M^- = \mu^-\right) P\left(M^0, M^+, M^-\right).$$

In (15), if M^+ and M^0 are given, it is said that the correlation between robot pose and the passive feature is independent. So, the passive feature can be set as an arbitrary value, such as its estimate value μ^-.

In the traditional sparsification rule, M^0 is mandatorily converted to passive. Like M^-, M^0 is also independent of robot post. So, the posterior probability of state vector in (15) is expressed as follows:

$$P_{\text{SEIF}}\left(X_t, M^0, M^+, M^-\right)$$
$$= P\left(X_t M^+, M^- = \mu^-\right) P\left(M^0, M^+, M^-\right). \tag{16}$$

M^0 has been divided into two parts:

$$M^0 = M^0_{\text{keep}} + M^0_{\text{observed}}. \tag{17}$$

Among that, M^0_{keep} is the feature which is mandatorily maintained passive during prediction. M^0_{observed} is the feature which has been observed again and maintained active. In order to limit the amount of active features under a boundary, M^+_{deactive} as the features with weaker correlation should be selected from M^+. It has the same amount as M^0_{observed}. Then, M^+ is divided into two parts:

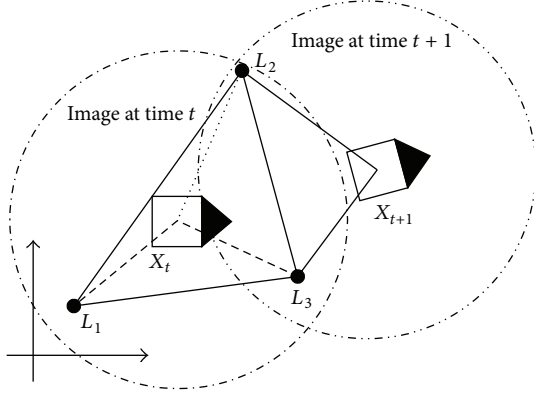$$M^+ = M^+_{\text{keep}} + M^+_{\text{deactive}}. \tag{18}$$

FIGURE 3: Correlation description in omnidirectional vision SLAM.

Among that, M_{keep}^+ is the feature which is maintained active during prediction. M_{deactive}^+ is the feature which has been mandatorily converted to passive. In the improved sparsification rule, M_{new}^0 expresses the feature which has been converted to passive from active, and M_{new}^+ is the feature which is maintained actively. Then, we can say that

$$M^0 = \left\{ M_{\text{new}}^0, M_{\text{new}}^+, M^- \right\},$$
$$M_{\text{new}}^0 = M_{\text{keep}}^0 + M_{\text{deactive}}^+, \qquad (19)$$
$$M_{\text{new}}^+ = M_{\text{observed}}^0 + M_{\text{keep}}^+.$$

As shown in Figure 3, in time t, there are 3 landmarks (L_1, L_2, L_3) in the omnidirectional image of the robot. The active feature M^+ is expressed as $M^+ = \{L_1, L_2, L_3\}$. Under the above analysis, the feature with a weaker correlation will be removed during the prediction of X_{t+1}. According to the traditional rules, the correlation between the robot and L_2 whichever is farthest from X_t will be removed. So, $M^0 = L_2$. However, in fact, (L_2, L_3) will be observed in time $k + 1$. The correlations between both robot and L_2, robot and L_3 will become stronger. According to the improved rule, L_1 will be removed. Then, $M_{\text{new}}^0 = L_1$. Due to the correlation of L_2 that has been maintained in the sparsification matrix, the repeated observation of L_2 is defined as relocation, it will decrease the uncertainty and the error of localization.

According to the extended information filtering algorithm [12, 13], the positions of landmarks which have been observed repeatedly will be modified. Positions of landmarks without repeated observations but relevant to the relocation area will also be revised. The overall error of state estimated results will be greatly reduced. In practice, omnidirectional vision has a wide range of view, as shown in Figure 4. There is a large amount of repeated landmark information in two continuous images, if the robot is placed in an environment with obvious features. Compared with the previous sparsification approach [13], the proposed sparsification algorithm does not need to classify landmarks. The relocation method can effectively restrain estimation errors caused by sparse matrix information.
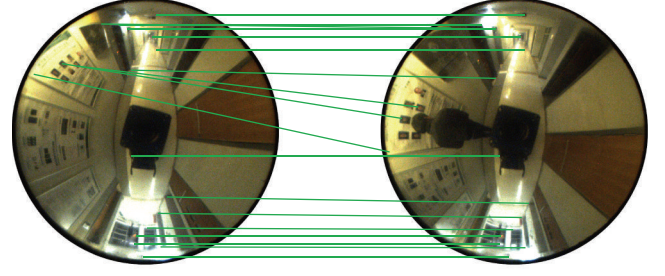

FIGURE 4: Repeated landmarks information in two continuous images.

4. Experimental Results and Discussions

4.1. Experiment Platform. Intelligent wheelchair-JiaoLong is used to evaluate the proposed IEIF-vSLAM algorithm. It is developed based on a commercial powered wheelchair and equipped with two encoders (one encoder for one driven wheel), a smart motion controller, an on-board PC, a laser range finder, and an omnidirectional vision system [23]. In sensor layer, laser range finder and omnidirectional vision are used to gather environment information. The sensor is processed and IEIF-vSLAM algorithm is running in on-board PC of controller layer which is equipped with a 1.8 GHz i5 CPU. In powered wheelchair layer, the wheelchair base is driven by two differential wheels and is powered by a group of 12 V batteries. The users can operate the wheelchair through the joystick. Figure 5 provides an overview of Jiaolong's system architecture. During verification, Harris-SIFT and IEIF-vSLAM algorithms also run in the on-board PC.

The omnidirectional vision system which has a Point Grey Chameleon camera with an image resolution of 1296(H) $\times$ 964(V) pixels is used in this experiment. The camera's frame rate is 10 Hz.

4.2. Platform Model Analysis. As a differential-driven mobile robot, JiaoLong's motion model is expressed as

$$f = \begin{bmatrix} x_v(k+1) \\ y_v(k+1) \\ \theta_v(k+1) \end{bmatrix} = \begin{bmatrix} x_v(k) + Vdt \cos(\theta_v(k)) \\ y_v(k) + Vdt \sin(\theta_v(k)) \\ \theta_v(k) + \omega dt \end{bmatrix} + \xi, \quad (20)$$

where, $[x_v(k) \ \ y_v(k) \ \ \theta_v(k)]^T$ and $[x_v(k+1) \ \ y_v(k+1) \ \ \theta_v(k+1)]^T$ are the robot pose at time k and $k + 1$. V and ω are linear and angular velocity. dt is the control period. ξ is a white Gaussian noise with mean value of 0 and variance of σ^2. So, the Jacobian Matrix F of the mobile robot's motion model in (3) will be expressed as

$$F = \frac{\partial f}{\partial x_v} = \begin{bmatrix} 1 & 0 & -Vdt \sin(\theta_v(k)) \\ 0 & 1 & Vdt \cos(\theta_v(k)) \\ 0 & 0 & 1 \end{bmatrix}. \qquad (21)$$

The observation model of omnidirectional vision is expressed as

$$g = \begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} x_v + r \cos(\theta + \theta_v) \\ y_v + r \sin(\theta + \theta_v) \end{bmatrix} + \eta, \qquad (22)$$
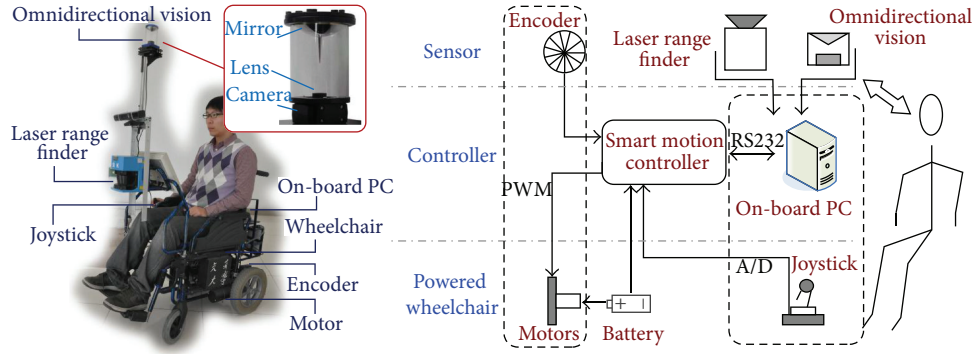
FIGURE 5: Prototype and hardware structure of the JiaoLong wheelchair.

where, $[x_i \quad y_i]^T$ are the global coordinates of features. $[x_v \quad y_v \quad \theta_v]^T$ is the current pose of robot. $[r \quad \theta]^T$ is the measurement (distance and angle) of omnidirectional vision. η is a white Gaussian noise with mean value of 0 and variance of σ^2. So, the Jacobian Matrix G of observation model in (5) will be expressed as

$$G = \frac{\partial g}{\partial x_v} = \begin{bmatrix} 1 & 0 & -r\sin(\theta + \theta_v) \\ 0 & 1 & r\cos(\theta + \theta_v) \end{bmatrix}. \tag{23}$$

In the proposed IEIF-vSLAM algorithm, the symmetric form of observation model is also used:

$$h = \begin{bmatrix} r \\ 0 \end{bmatrix} = \begin{bmatrix} \sqrt{(x_i - x_v(k))^2 + (y_i - y_v(k))^2} \\ \text{atan2}\left(\dfrac{y_i - y_v(k)}{x_i - x_v(k)}\right) - \theta_v \end{bmatrix}. \tag{24}$$

Then, H in (6) will be expressed as

$$H = \frac{\partial h}{\partial \varepsilon} = \begin{bmatrix} -\dfrac{x_i - x_v(k)}{r} & -\dfrac{y_i - y_v(k)}{r} & 0 & \cdots & \dfrac{x_i - x_v(k)}{r} & \dfrac{y_i - y_v(k)}{r} & \cdots \\ \dfrac{y_i - y_v(k)}{r^2} & -\dfrac{x_i - x_v(k)}{r^2} & -1 & \cdots & -\dfrac{y_i - y_v(k)}{r^2} & \dfrac{x_i - x_v(k)}{r^2} & \cdots \end{bmatrix}. \tag{25}$$

4.3. Map Building Experiments.

The wheelchair is placed in an environment with a room, a door, and a long corridor. A map built by a laser range finder [23] is superimposed for comparison with the map built by proposed IEIF-vSLAM. The trajectory followed by JiaoLong in the experiment is shown in Figure 6. JiaoLong is driven by a user through a joystick to follow this trajectory, starting inside room number 200, travelling through the door, entering two atriums and corridors, and travelling back to room number 200. This process has been repeated for 3 cycles. During the driving process, the data are transmitted and recorded in the on-board PC. The data contains time t, odometer data, data of laser range finder, and omnidirectional images. In order to avoid significant data redundancy, a lower threshold is set. If the wheelchair does not move by more than 4 cm or turn by more than 1.5°, the image will be discarded. The map building and localization experiments are carried out offline to verify the proposed omnidirectional vision SLAM algorithm. The robot's "real" pose (as the ground truth compared with the pose obtained by IEIF-vSLAM algorithm, the accuracy of which has been limited to within ±4 cm) is calculated by scan-match method [24] based on odometer data and data of laser range finder. Based on time t which is recorded in each frame of data during acquisition, the one-to-one relationship between the "real" pose and the pose obtained by

IEIF-vSLAM is built. The data acquisition experiment cost 10 minutes and the total distance of 3 cycles is 360 m.

As shown in Figure 6, the blue dotted "■" indicates the planned trajectory of wheelchair, which is controlled through the joystick. This trajectory is obtained by scan-match method in the 1st circle. The Sparsification index is set as $N = 2$, $\lambda = 0.8$.

In Figure 7, some images of omnidirectional vision and feature map are shown. Room number 200 is an example of a complex indoor environment. There is rich furniture, such as tables, chairs, and various objects. The atrium had some exhibition boards, doors, and windows along the wall. The corridor, by contrast, has white walls with very few features available. In Figure 8, the black rectangle "■" indicate the localization results and the green solid line indicates the odometer trajectory for the 1st circle from starting point to destination. During the initial period, the feature points are extracted, meanwhile the map was built based on omnidirectional vision's observation. Although the odometer data have small accumulative errors, there is no obvious distinction between the trajectories of odometer and plan trajectory (Figure 6). However, when the wheelchair navigates out of Room number 200, odometer errors accumulate away from the planned trajectory. Near the destination, odometer errors have accumulated on a large scale, which
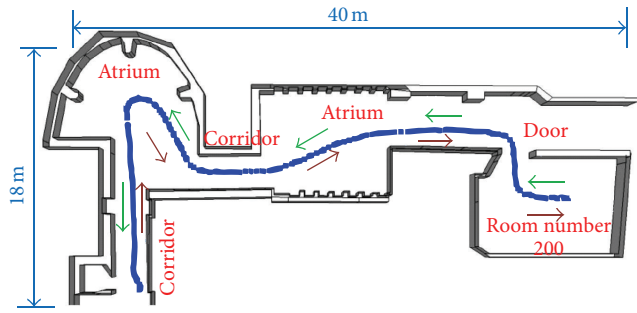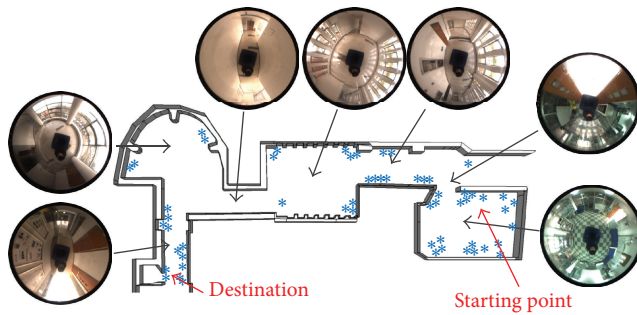
FIGURE 6: Environment and trajectory for experiment.

FIGURE 8: Odometer trajectory and localization of IEIF-vSLAM.

∗∗∗ Positions of landmarks

FIGURE 7: Feature map and omnidirectional image.

does not reflect the wheelchair's navigation position exactly. As shown in Figure 8, based on modification of the IEIF-vSLAM algorithm, the deviation of IEIF-vSLAM localization is still relatively small from start to end during navigation of three circles.

4.4. Localization Analysis. In order to analyze the localization accuracy of IEIF-vSLAM, as shown in Figure 8, the robot trajectory is divided into 3 sections:

  (i) Section I: from the starting point to the door exit;

  (ii) Section II: from the door exit to the entrance of the last corridor;

  (iii) Section III: from the entrance of the last corridor to the destination.

In Figure 8, the blue star "∗" indicates the position of landmark features. There are rich feature points on both sides of the path in Section I and their distribution is more dispersed; there are fewer feature points on both sides of the path in Section II and they are relatively concentrated; the feature points of Section III are rich and more concentrated. Table 1 shows the localization error (compared with the results of scan-match, Max., and Avg. are the maximum and average value of all absolute errors) during the three sections.

In Section I, it is a complex indoor environment. There is rich furniture, such as tables, chairs, and various objects. There are more features with more dispersed distribution. Therefore, some more accurate results are obtained.
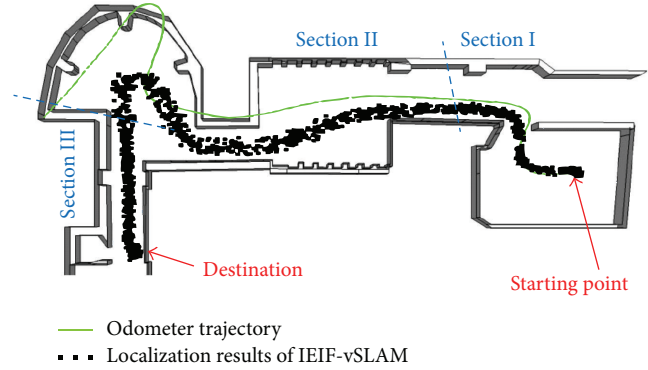
In Section II, much of the environment is a white wall with windows with similar features. There are fewer features. It leads to inaccurate localization which is caused by lack of relocation.

In Section III, there are more feature points and their distribution is relatively concentrated, which easily leads to miss-match. Therefore, positioning errors are some bigger than that of Section II.

As shown in Table 1, the position of the robot become more accurate in the 3rd circle of all 3 Sections. In the 3rd circle, the average error is 0.246 m in x-direction, 0.198 m in y-direction, and 1.6° in heading angle. There are more repeated observations for landmarks which lead to more relocations; the estimation errors of robot pose have been effectively restrained.

4.5. Comparison Experiments. In order to verify the validity of the proposed IEIF-vSLAM algorithm, the EIF [13] algorithm is selected to be compared with the proposed algorithm. The EIF algorithm combined with Harris-SIFT and SLAM algorithm also runs on the same platform which described in Section 4.1. The comparison experiment is also repeated for 3 cycles in same scenario. Only the Max. and Avg. in the 3rd circle are calculated and recorded. t as the algorithm's average of time consuming (including of EIF or IEIF, excluding of image processing, feature detecting, matching, and tracking) of each algorithm cycle has also been recorded.

Through the results in Table 2, compared with EIF algorithm, the proposed EIF-vSLAM algorithm is better in accuracy and the time consumption of the proposed EIF-vSLAM algorithm is only 1/2 of that of EIF algorithm. The advantage of computation efficiency of IEIF-vSLAM algorithm is obvious.

5. Conclusion

An IEIF-vSLAM method based on omnidirectional vision has been proposed in this paper. Both of the characteristics of information matrix structure and omnidirectional vision's repeated observations for landmarks are analyzed. Based on these analyses, the sparsification rule has been improved.

Table 1: Localization error of IEIF-vSLAM.

	Section I		Section II		Section III	
	Max.	Avg.	Max.	Avg.	Max.	Avg.
1st circle						
X error (m)	0.432	0.296	0.633	0.437	0.546	0.349
Y error (m)	0.395	0.247	0.578	0.432	0.430	0.271
Angle error (°)	3.7	1.9	4.3	2.5	3.6	2.2
2nd circle						
X error (m)	0.329	0.267	0.541	0.363	0.451	0.273
Y error (m)	0.318	0.231	0.481	0.365	0.318	0.231
Angle error (°)	2.9	1.6	3.5	2.4	2.5	1.5
3rd Circle						
X error (m)	0.401	0.235	0.502	0.253	0.471	0.245
Y error (m)	0.264	0.172	0.467	0.236	0.286	0.217
Angle error (°)	2.4	1.1	3.2	1.8	2.3	1.2

Table 2: Quantitative results of different algorithms.

	EIF algorithm		IEIF-vSLAM	
	Max.	Avg.	Max.	Avg.
X error (m)	0.714	0.351	0.502	0.246
Y error (m)	0.571	0.249	0.467	0.198
Angle error (°)	3.4	1.9	3.2	1.6
t (s)		0.487		0.279

The sparse observation information has been utilized and the strongest global correlation has been maintained. Both the computation efficiency and accuracy of the estimated results have been improved by using proper information matrix sparsification. Before real platform experiments, through the error analysis, the error caused by sparsification can be eliminated by proposed method. The results of the experiments show that this method which uses omnidirectional vision's characteristic of repeated observations for landmarks can be used for mobile robot map building and localization.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

References

[1] M. W. M. G. Dissanayake, P. Newman, S. Clark, H. F. Durrant-Whyte, and M. Csorba, "A solution to the simultaneous localization and map building (SLAM) problem," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 3, pp. 229–241, 2001.

[2] J. D. Tardós, J. Neira, P. M. Newman, and J. J. Leonard, "Robust mapping and localization in indoor environments using sonar data," *International Journal of Robotics Research*, vol. 21, no. 4, pp. 311–330, 2002.

[3] S. Thrun, Y. Liu, D. Koller, A. Y. Ng, Z. Ghahramani, and H. Durrant-Whyte, "Simultaneous localization and mapping with sparse extended information filters," *International Journal of Robotics Research*, vol. 23, no. 7-8, pp. 693–716, 2004.

[4] S. Se, D. G. Lowe, and J. J. Little, "Vision-based global localization and mapping for mobile robots," *IEEE Transactions on Robotics*, vol. 21, no. 3, pp. 364–375, 2005.

[5] A. J. Davison, "Real-time simultaneous localisation and mapping with a single camera," in *Proceedings of the 9th International Conference on Computer Vision*, pp. 1403–1410, Nice, France, October 2003.

[6] Y. J. Lee and S. Sung, "Vision based SLAM for mobile robot navigation using distributed filters," in *Mobile Robots Navigation*, A. Barrera, Ed., InTech, 2010.

[7] D. Burschka and G. D. Hager, "V-GPS(SLAM): vision-based inertial system for mobile robots," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 409–415, May 2004.

[8] S. Ekvall, P. Jensfelt, and D. Kragic, "Integrating active mobile robot object recognition and SLAM in natural environments," in *Proceedings of the IEEE International Conference on Intelligent Robots and Systems (IROS '06)*, pp. 5792–5797, Beijing, China, October 2006.

[9] E. Mouragnon, M. Lhuillier, M. Dhome, F. Dekeyser, and P. Sayd, "Monocular vision based SLAM for mobile robots," in *Proceedings of the 18th International Conference on Pattern Recognition (ICPR '06)*, vol. 3, pp. 1027–1031, Hong Kong, August 2006.

[10] H. Ishiguro, "Development of low-cost compact omnidirectional vision sensors and their applications," in *Proceedings of International Conference on Information Systems, Analysis and Synthesis*, pp. 433–439, Boston, Mass, USA, 1998.

[11] P. Blaer and P. Allen, "Topological mobile robot localization using fast vision techniques," in *Proceedings of the IEEE International Conference on Robotics and Automation*, Washington, DC, USA, May 2002.

[12] M. R. Walter, R. M. Eustice, and J. J. Leonard, "Exactly sparse extended information filters for feature-based SLAM," *International Journal of Robotics Research*, vol. 26, no. 4, pp. 335–359, 2007.

[13] R. M. Eustice, H. Singh, and J. J. Leonard, "Exactly sparse delayed-state filters for view-based SLAM," *IEEE Transactions on Robotics*, vol. 22, no. 6, pp. 1100–1114, 2006.

[14] H. Dong, J. Tang, W. Chen, A. Nagano, and Z. Luo, "Novel information matrix sparsification approach for practical implementation of simultaneous localization and mapping," *Advanced Robotics*, vol. 24, no. 5-6, pp. 819–838, 2010.

[15] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[16] J. Shi and C. Tomasi, "Good features to track," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 593–600, June 1994.

[17] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *International Journal of Computer Vision*, vol. 60, no. 1, pp. 63–86, 2004.

[18] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: Part I," *IEEE Robotics and Automation Magazine*, vol. 13, no. 2, pp. 99–110, 2006.

[19] K. Yamazawa, Y. Yagi, and M. Yachida, "Omnidirectional imaging with hyperboloidal projection," in *Proceedings of the International Conference on Intelligent Robots and Systems*, pp. 1029–1034, Tokyo, Japan, July 1993.

[20] J. Y. Xu, J. C. Wang, and W. Chen, "Omni-vision-based simultaneous localization and mapping of mobile robots," *Robots*, vol. 30, no. 4, pp. 289–297, 2008 (Chinese).

[21] J. Neira and J. D. Tardós, "Data association in stochastic mapping using the joint compatibility test," *IEEE Transactions on Robotics and Automation*, vol. 17, no. 6, pp. 890–897, 2001.

[22] R. Eustice, M. Walter, and J. Leonard, "Sparse extended information filters: insights into sparsification," in *Proceedings of the IEEE IRS/RSJ International Conference on Intelligent Robots and Systems (IROS '05)*, pp. 3281–3288, August 2005.

[23] Y. Wang and W. D. Chen, "Hybrid map-based navigation for intelligent wheelchair," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA '11)*, pp. 637–642, Shanghai, China, May 2011.

[24] D. Hähnel, W. Burgard, and D. Fox, "An efficient FastSLAM algorithm for generating maps of large-scale cyclic environments from raw laser range measurements," in *Proceeding of the International Conference on Intelligent Robots and Systems*, pp. 206–211, October 2003.