Kinect-based 3D Indoor Environment Map Building

Yi ZHANG, Qi CHENa,*, Xiao-Lin ZHONG and Qiao-Zhen LIAO

National Engineering Research and Development Center for Information Accessibility, Chongqing University of Posts and Telecommunications, China

*Corresponding author

Keywords: Mobile robot, Kinect, Robustness, 3D map.

Abstract. For the problem of poor robustness and real-time performance of mobile robot during the process of 3D map building, a new 3D map building method which combines the improved EKF-based RANSAC algorithm and the G-ICP algorithm is proposed. First of all, a feature point extraction and matching process is carried out based on the RGB data collected by Kinect. During the process of matching, the improved RANSAC algorithm is employed to eliminate the false matching points. Then, in order to get the three-dimensional point cloud data of the indoor environment, a G-ICP algorithm which has better robustness and performance is employed to accomplish accurate image registration. Eventually, a 3D indoor environment map is created by the matching data and 3D point cloud data. The effectiveness of the proposed method is verified by the experiments.

Introduction

The establishment of a high-availability environment map is the basis of completing autonomous mobile robot localization, path planning and navigation [1]. In recent years, 3D environment map related research has received extensive attention. In literature [2], the 3D laser radar data is adopted to create a 3D map of the environment based on the improved ICP algorithm, thus generate a high-precision 3D map. But this approach is suboptimal for the lack of texture description of the environment and the 3D laser is too much expensive. Another method which uses ordinary camera and the particle filter algorithm to obtain the depth information of the environment is also hard to acquire high real-time performance for the time delay [3]. Henry etc. proposed an interactive 3D reconstruction system based on Kinect [4], but this system relies on the selection of key frames, thus it is difficult to apply to the online mobile robot 3D map generation for its large amount of calculation and high time complexity.

To solve these problems, this paper proposes a new method which combines the improved EKF-based RANSAC algorithm and G-ICP algorithm to generate a 3D map in the indoor environment. The overall framework of the proposed system is shown in Fig. 1.

![Figure 1. Overall framework of the proposed system.](image-url)
Feature Points Detection and Matching

Common feature point detection algorithms are: Harris, SUSAN, SIFT, SURF [5], etc. in many feature point detection algorithms, SURF algorithm has high efficiency and invariance attribute both in scale and rotation. Therefore, this paper selects SURF algorithm to extract the feature points.

a) Feature points detection: SURF detection algorithm is mainly based on Hessian matrix. Setting the integral image function \( f(x, y) \), Hessian matrix can be expressed as:

\[
H(x, y, \sigma) = \begin{bmatrix}
L_x(x, y, \sigma) & L_y(x, y, \sigma) \\
L_y(x, y, \sigma) & L_{yy}(x, y, \sigma)
\end{bmatrix}
\]

(1)

Where \( L_x \), \( L_y \) and \( L_{yy} \) are convolutions for an image \( f(x, y) \) with a Gaussian second order partial derivative of the point \((x, y)\), \( \sigma \) is the scale for Hessian matrix of the point \((x, y)\). Matrix discriminant is:

\[
\text{det}(H) = \left( \frac{\partial^2 f}{\partial x^2} \right) \left( \frac{\partial^2 f}{\partial y^2} \right) - \left( \frac{\partial^2 f}{\partial x \partial y} \right)^2
\]

(2)

The value of the formula (2) is the characteristic value of matrix \( H \), and whether that point is an extreme value point is determined according to the positive and negative attribute of the discriminant value.

b) Matching feature points: The main process of SURF feature points matching algorithm: First, use rapid indexing matching method for feature points to preliminary screening and then take the nearest neighbor matching, as shown in formula (3). Then, set a threshold value range and compare the nearest neighbor Euclidean distance with Euclidean distance of the last neighbor, if the ratio is in the range, then we consider it is a correct match, otherwise mismatch.

\[
\text{Dis}_g = \left[ \sum_{k=0}^{k=n} (X_{ik} - Y_{jk})^2 \right]^{1/2}
\]

(3)

Where \( n \) is the dimension of the feature vector, and \( X_{ik}, X_{jk} \) are the \( k \) th element of the \( i \) th and the \( j \) th feature point descriptor of the image to be matched. In SURF algorithm, the feature points matching is decided through the nearest neighbor matching threshold, usually exists many mismatches, thus further match process is needed to guarantee the expected accuracy rate.

Improved EKF-based RANSAC Algorithm

a) EKF iterative process: Extended Kalman filtering is capable of linearizing the mean and variance of the Kalman filter, its iterative process includes the prediction step and the filtering step [6]. During the prediction step, the motion model follows the standard Kalman filter state prediction and covariance estimation, which is expressed as follows:

\[
\hat{s}_{k|k} = f_k(\hat{s}_{k-1|k-1});
\]

(4)

\[
P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k;
\]

(5)

In the formula, \( F_k \) denotes the Jacobian matrix of the \( k \) th step of the state vector \( s_{k|k-1} \), and \( Q_k \) indicates the zero mean Gaussian noise covariance of the dynamic model.
During the filtering step, data are updated according to the traditional extended Kalman filter equations.

\[
\hat{s}_k = \hat{s}_{k|k-1} + K_k (z_k - H_k \hat{s}_{k|k-1});
\]

\[
P_{k|k-1} = (I - K_k H_k) P_{k|k-1};
\]

\[
K_k = P_{k|k-1} H_k^T S_k^{-1},
\]

(6)

Where, \( S_k \) is the covariance between measured values and predicted values, \( H_k \) is the Jacobian matrix of prediction model.

b) Improved algorithm RANSAC: In RANSAC algorithm, the effective sample data are calculated through the mathematical model parameters of the sample dataset which includes the abnormal data [7]. Assume that \( w \) indicates the probability of correct data point in the dataset of each iteration, and \( w^n \) denotes the probability that, in the iterative process, all the \( m \) selected data points are the correct data points, thus the probability that at least one abnormal data point in \( m \) data points can be denoted as \( 1 - w^n \), meanwhile these abnormal points may cause unreasonable models. Therefore, the iteration number \( n \) can be obtained as:

\[
n = \frac{\log(1 - p)}{\log(1 - w^n)}
\]

(7)

It shows that the iteration number \( n \) of the RANSAC algorithm increases in pace with \( m \), thus the complexity of the algorithm will increase accordingly. This paper introduced the improved RANSAC algorithm to make full use of the priori information obtained from the extended Kalman filter in the prediction step [8], thus reducing the size of the dataset, decreasing the iteration number of the RANSAC algorithm to simplify the computational complexity and improve the real-time property and robustness of the systems. The main steps of the algorithm are as follows:

1) Assume that \( Z \) is the data points set obtained via the matching process, and \( Z^{\text{inliers}} \) denotes the best matching part of the points set, other points set are indicated as \( Z^{\text{outliers}} \), EKF update step is carried out via \( Z^{\text{inliers}} \), the detailed processes are as follows:

\[
\hat{s}_{k|k} = \hat{s}_{k|k-1} + K_k (z^{\text{inliers}} - h(\hat{s}_{k|k}));
\]

\[
K_k = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1};
\]

\[
P_{k|k-1} = (I - K_k H_k) P_{k|k-1},
\]

(8)

In the formula, \( H_k = [H_1 \cdots H_n]^T \) represents Jacobian matrix of the measurement \( h(\hat{s}_{k|k-1}) \), and \( R_k \) represents the noise covariance.

2) Due to the long-range points are sensitive to the camera angle transformation in points set, and short distance points are sensitive to scale transformation, when a long-range point generate the point which is sensitive to the angular transformation and the scale transformation is not quite accurate, other long-range points will support the fitting model. However, due to the poor accuracy of scale transformation, the near points are possible to show high uncertainty. Therefore, after determining the fitting model, we need to remedy some interior points from the unstable points set. Partial EKF update processes of remedial points are as follows:

\[
\hat{s}_{k|k} = \hat{s}_{k|k-1} + K_k (z^{\text{inliers}} - h(\hat{s}_{k|k}));
\]
\begin{align*}
K_k &= P_{kk^{-1}}H_k^T (H_k P_{kk^{-1}}H_k^T + R_k)^{-1}; \\
P_{kk^{-1}} &= (I - K_k H_k) P_{kk^{-1}};
\end{align*}

After \( \omega_{inliers} \) is updated with EKF, most errors caused in the prediction step were corrected, and covariance are also reduced. Thus unstable interior points get remedied, as a result achieving the purpose of enhancing the system robustness.

**Accurate Registration Algorithm of G-ICP**

In the point cloud accurate registration step, we adopt the G-ICP algorithm [9], which is superior to the conventional ICP algorithm in the aspect of performance and robustness. Conventional ICP algorithm can be simply interpreted as: given two 3D data point sets to find the space conversion between the two point sets so that they can perform spatial matching.

G-ICP algorithm employs a probability model when the conventional ICP algorithm calculates the residual error \( E \), finding the transformation matrix \( T \) by minimizing the residual error function [10]. In the probabilistic model, assume that the existed point sets \( \hat{A} = \{ \hat{a}_i \}, \hat{B} = \{ \hat{b}_i \}, \) and \( A,B \) are subjected to the normally distribution \( \hat{a}_i \sim N(\hat{a}_i, C_i^A), \hat{b}_i \sim N(\hat{b}_i, C_i^B). \) \( \{ C_i^A \}, \{ C_i^B \} \) are the covariance matrix of the measured point sets, the transformation matrix \( T \) can be indicated as:

\[
T = \arg \max_T \sum_i \log(p(d_i^{TT})))
\] (10)

Simplify formula (10) to obtain the update formula of G-ICP algorithm:

\[
T = \arg \min_T \sum_i d_i^{T \gamma} (C_i^B + TC_i^A T^\gamma)^{-1} d_i^{T}\) (11)

When set \( C_i^B = I, C_i^A = 0 \), getting the update formula of the traditional ICP algorithm by the formula (11):

\[
T = \arg \min_T \sum_i d_i^{T \gamma} d_i^{T}\) (12)

G-ICP algorithm still uses the method of calculating the Euclidean distance to find the corresponding points in order to employ the k-d tree to accelerate, thereby reducing the complexity of the algorithm. Set \( T_0 \) as the initial transformation matrix, two point cloud sets are expressed as \( A = \{ a_i \}, B = \{ b_i \} \), respectively. Then the G-ICP algorithm can be summarized as follows:

1) Each point in the point set \( A \) is obtained by computing its closest point in point set \( B \), then employ k-d tree to accelerate, mark all points to the collection \( M \).
2) Rigid transformation matrix is obtained by calculation, and transform point set \( A \) to get a new transformation point set.
3) Iteratively calculate MLE (maximum likelihood estimation) and judge whether the minimum residual error function \( E \) is the minimum value. If it is true then stop computation, otherwise repeat step 2 until meet the requirement.
4) The obtained transformation matrix is the desired secondary registration parameter.

**Experimental Results and Analysis**

a) In this paper, the experimental platform consists of Pioneer3-DX robot, Kinect sensor and a laptop computer, using the controller of Hitachi H8S series. The experiments are conducted on ThinkPad E450C with 2.7GHz Intel Core i5-5200U Dual Core Processor, 4GB of RAM, running on Ubuntu
12.04 Linux operation system. To ensure high robustness, set the robot turning angular velocity as 0.52rad/s, and the moving speed is 0.2m/s.

Extract a series of environmental image information via Kinect, and then select two adjacent frame image from those images, after performing extraction and matching of the feature points, use the traditional RANSAC algorithm and the improved RANSAC algorithm to eliminate mismatch. The experimental results are shown in Fig. 2.

![Figure 2. Excluding the results of mismatching.](image)

The logarithm of mismatching points, the logarithm of matching points and running time are shown in Table. 1, It can be seen that the improved RANSAC algorithm can achieve better results in eliminating mismatches during the process of matching, and improve the accuracy of feature matching and the robustness of algorithm.

<table>
<thead>
<tr>
<th>Index</th>
<th>Traditional algorithms</th>
<th>Improved algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match point numbers</td>
<td>256</td>
<td>113</td>
</tr>
<tr>
<td>Mismatching points logarithmic</td>
<td>26</td>
<td>3</td>
</tr>
<tr>
<td>Running time (ms)</td>
<td>86</td>
<td>55</td>
</tr>
</tbody>
</table>

b) Fig.3 is the 3D map building result of a partial environment, here, Fig.3(a) is the experimental environment, Fig.3 (b) is the 3D point cloud information of the environment, Fig.3(c) is the front view of the constructed 3D map, and Fig.3(d) is the top view of the 2D map.

![Figure 3. 3D map of local environment.](image)

Fig. 4 is a 3D map constructed under a wide range viewing angle. Fig.4(a) is the experimental environment, Fig.4 (b) is the depth information of a frame image during the map construction process, Fig.4 (c) is the 3D scene under a partial perspective, Fig.4 (d) is the final construction result of a 3D wide range map. In this paper, the 3D point cloud acquired via G-ICP algorithm is much more
smooth and clear, as shown in Fig. 3(b). The color information of the environment captured via Kinect can be mapped onto the corresponding 3D point cloud model, and the created 3D map of the indoor environment is shown in Fig. 4(d). We can see that the 3D map of the indoor environment created by this method has clear and smooth outlines which are consistent with the real environment, that is to say the proposed method achieves better 3D reconstruction result.

Conclusions
This paper proposed a new method which combines the improved EKF-based RANSAC algorithm and G-ICP algorithm to create a 3D map of the indoor environment. Use Kinect to obtain the color information and depth information of the environment, effectively enhanced the real-time property and robustness of the system with the improved RANSAC algorithm and G-ICP algorithm. Experimental results demonstrate that the proposed method can quickly create a more accurate 3D map of the indoor environment which is consistent with the real environment. Future plans to improve the system performance can be summarized in two aspects: (1) make further efforts to the feature point detection algorithm and improve the effectiveness of the feature point detection algorithm. (2) Employ parallel GPU technology to further improve the efficiency of the system to acquire better system real-time performance.

References