Reducing Effect of Outliers in Landmark-based Spatial Localization using MLESAC

Sunglok Choi ∗ Jong-Hwan Kim ∗ ∗

∗ U-Robot Research Division, Electronics and Telecommunications Research Institute, 138, Gajeong-dong, Yuseong-Gu, Daejeon, 305-700, Republic of Korea (e-mail: sunglok@etri.re.kr)

∗ ∗ Department of Electrical Engineering and Computer Science, KAIST, 373-1, Guseong-dong, Yuseong-gu, Daejeon, 305-701, Republic of Korea (e-mail: johkim@rit.kaist.ac.kr)

Abstract: In the landmark-based localization problem, movement and ambiguity of landmarks and imperfect identification process make measurements of the landmarks completely different from its true value. The incorrect measured data have degraded existing localization methods in the practical applications. This paper proposes a framework to improve accuracy of the existing landmark-based localization methods regardless of such incorrect measured data. The framework is based on Maximum Likelihood Estimation Sample Consensus (MLESAC). It samples a set of measured data randomly to estimate position and orientation, and the estimated pose is evaluated through likelihood of whole measured data with respect to the result. Iterations of sampling, estimation, and evaluation are performed to find the best result to maximize the likelihood. Simulation results demonstrate that the proposed framework improved the the existing localization methods. Analysis using a concept of loss functions also explains that the framework is superior compared to previous researches such as Random Sample Consensus (RANSAC).

1. INTRODUCTION

Localization is to identify a robot’s pose (position and orientation) using measured data and prior knowledge such as a map or the initial position. It is known as one of the most important tasks of mobile robots because it is a basic step to perform more complex tasks such as cleaning, serving, and guidance.

Localization methods are typically categorized into local methods and global methods. Local methods try to estimate relative displacement from the initial position while minimizing odometry error, so they need the initial position as prior knowledge. They are also called as dead-reckoning or incremental positioning. However, without knowing the initial position, global methods can estimate the absolute position using a map and measured data from the environment. Therefore, they can overcome a serious position error and a kidnapping problem.

Sources of data are useful to understand characteristics of localization methods: a motion, an environment, and colleagues (Fig. 1). Fox et al. [2000] also mentioned a similar categorization. Motion data can be measured by an encoder (for wheeled robots), footstep (for legged robots), and inertia sensors. Approaches to use only motion data are typical local methods, which accumulate the data in temporal axis. Environmental data are obtained through a vision system, sonar sensor, or laser range finder. These data originate from landmarks or appearance of the given environment. Proximity to electrical tags or stations are the most important tasks of mobile robots because it is a basic step to perform more complex tasks such as cleaning, serving, and guidance.

Many global localization methods are based on landmarks because the landmarks describe the environment as a simple and clear manner such as a set of absolute positions of landmarks. There are also many commercial localization modules which use landmarks (e.g. GPS, NorthStar by Evolution Robotics, and StarLITE by ETRI). Measurements from landmark can be a distance from robot to the landmark or a bearing from robot's heading direction to the landmark. These measurements have errors resulted from noise and quantization, but they are near their true values. The measurements of this class are referred as inliers. However, measurements can be totally different from their true values because of movement and ambiguity.
Fig. 1. Localization methods in the view of sources of data of landmarks and imperfect identification process. The measurements of this class are referred as outliers.

For applications in the real world, a landmark-based localization method should overcome errors resulted from outliers. However, there are few researches related to a framework which investigated outliers, and their approaches were based on RANSAC or their own frameworks similar with RANSAC (Jaulin et al. [2002], Se et al. [2004], and Yuen and MacDonald [2005]). Even though RANSAC can deal with outliers and make more accurate results than the existing methods without it, it needs to adjust parameter which is sensitive to the distribution of measured data.

In this paper, a novel robust framework is proposed to improve the accuracy of existing landmark-based localization methods through MLESAC. MLESAC is initially suggested by Torr and Zisserman [2000] to estimate fundamental matrices and image homography in the computer vision community. In MLESAC, random sampling picks up a set of measured data to estimate a robot’s position and orientation. Error of measurements with respect to the estimation is modeled as a mixture of Gaussian and uniform distribution, and likelihood of measurements is derived from the error model. Iterations of sampling, estimation, and evaluation are performed to find the best position and orientation to maximize the likelihood.

The remainder of this paper is organized as follows. Section 2 formulates the landmark-based localization problem and examines the previous researches. Section 3 introduces the robust landmark-based localization framework based on MLESAC. Section 4 demonstrates experimental results and the comparison to the existing methods. Analysis using a concept of loss functions is included as well. Finally, Section 5 contains discussion and further works.

2. REVIEW OF LANDMARK-BASED LOCALIZATION

2.1 Problem Statements

Landmark-based localization problem is to estimate a robot’s position and orientation, \( \Theta = (p, \theta)^T \), from a map, \( M \), and measured data of observed landmarks, \( D \). The map is described as a set of absolute positions of landmarks in a given environment: \( M = (m_1, m_2, \ldots, m_n)^T \), where \( n \) is the number of landmarks. The measured data can be a set of displacements (vector) or distances (scalar) from robot to each landmark. Moreover, it can be a set of bearings from robot’s heading direction to each landmark or relative bearings between a landmark and its adjacent landmark: \( D = (d_1, d_2, \ldots, d_k)^T \), where \( k \) is the number of observed landmarks. Some landmarks can be unobservable due to occlusion or limited sensing range. In other words, the number of measured data \( k \) is less than or equal to the number of landmarks \( n \). Among \( k \) measured data, inliers are near their true value, and outliers are completely irrelevant to their true value.

2.2 Landmark-based Localization Methods

Landmark can be an artificial patch or a conventional object such as a door, television, and refrigerator. Wireless station can be landmark because it can provide its distance from the robot through signal strength. Local image feature such as Harris corner and SIFT can also be landmark. As an example, Se et al. [2004] used huge amounts of SIFT image features as landmarks.

Landmark-based localization methods can be classified by the types of measurement: displacement, distance, bearing, and relative bearing (Table. 1). A displacement of landmark can be regarded as the combination of distance and bearing of landmark. According to available measurements, it is necessary to apply a suitable localization method. Se et al. [2004] estimated a robot’s position and orientation from two displacements measured through a stereo-vision system. Calabrese and Indiveri [2005] utilized two pairs of distances and bearings measured through observed size of a landmark and omnidirectional vision system. Thomas and Ros [2005] offered a trilateration method which used three distance measurements to identify a robot’s position in three-dimensional space. They analyzed three kinds of errors which caused inliers: station location error, range measurement error, and numerical error. Betke and Gurvits [1997] proposed a bearing-based method, which represents a position of landmark as complex number, and Shimshoni [2002] also suggested a bearing-based algebraic method, which used Direct Linear Transformation (DLT). Briechle and Hanebeck [2004] developed a recursive method, which used three relative bearings and nonlinear filtering. They considered unknown but bounded error which caused inliers. Note that there were various works to investigate the landmark-based localization problem, but they did not take into account of outliers except Se et al. [2004].

2.3 Previous Works

Se et al. [2004] tackled the outlier problem through Hough transform and RANSAC, and showed the efficiency of RANSAC when SIFT image features were used. RANSAC, proposed by Fischler and Bolles [1981], is a popular model fitting framework in the computer vision community. Its procedure is iteration of generating of a hypothesis and its verification (Fig. 2). It selects a hypothesis which maximize the number of measured data within predefined tolerance, \( \delta \). In other words, its result is to maximize the number of data classified as inliers. Jaulin et al. [2002] solved the outlier problem through their framework, Outlier Minimal Number Estimator (OMNE). OMNE tries to find the position and orientation to minimize the number of data classified as outliers, which is the same objective with RANSAC.
Table 1. Recent works for the landmark-based localization

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Recent Works</th>
<th>Results</th>
<th>The Num. of Landmarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Displacement</td>
<td>Se et al. [2004], Calabrese and Indiveri [2005]</td>
<td>Location, Direction</td>
<td>2</td>
</tr>
<tr>
<td>Distance</td>
<td>Thomas and Rios [2005]</td>
<td>Location</td>
<td>3</td>
</tr>
<tr>
<td>Bearing</td>
<td>Betke and Gurvits [1997], Shimshoni [2002]</td>
<td>Location, Direction</td>
<td>3 ~ (Least Square Method)</td>
</tr>
<tr>
<td>Relative Bearing</td>
<td>Brieche and Hanebeck [2004]</td>
<td>Location</td>
<td>3</td>
</tr>
</tbody>
</table>

3. THE ROBUST LOCALIZATION FRAMEWORK

3.1 The Error Model and Maximum Likelihood Estimation

Torr and Zisserman [2000] proposed a probabilistic error model of inlier and outlier. They suggested the pdf (probability distribution function) of error by an inlier as unbiased Gaussian distribution and the pdf of error by an outlier as uniform distribution, which is described as follows:

\[ p_{in}(e) = \left( \frac{1}{\sqrt{2\pi\sigma}} \right)^s \exp\left(-\frac{e^2}{2\sigma^2}\right) \quad (1) \]

\[ p_{out}(e) = \frac{1}{\nu} \quad (2) \]

where \( \sigma \) is the standard deviation of Gaussian distribution, \( s \) is the dimension of measured data, and \( \nu \) is the area of the given environment. If the measurement is a displacement and the size of environment is \( 10 \times 18 \), then \( s = 2 \) and \( \nu = 180 \).

Without knowledge whether a measurement is an inlier or an outlier, but the ratio of inliers to whole data \( \gamma \) and the standard deviation \( \sigma \) are known, likelihood of the measurement \( d \) becomes a mixture of \( p_{in}(e) \) and \( p_{out}(e) \) as follows:

\[ p(d \mid \Theta) = \gamma \left( \frac{1}{\sqrt{2\pi\sigma}} \right)^s \exp\left(-\frac{e^2}{2\sigma^2}\right) + (1 - \gamma) \frac{1}{\nu} \quad (3) \]

where \( e \) is error between a measurement \( d \) and pose \( \Theta \). Fig. 3 shows the likelihood with various inlier ratios.

Under an assumption that \( k \) measured data are independent, a likelihood of whole data \( D \) is obtained as

\[ p(D \mid \Theta) = \prod_{i=1}^{k} p(d_i \mid \Theta). \quad (4) \]

Generally, the likelihood \( p(D \mid \Theta) \) becomes a quite small real value. For computational feasibility, the negative log likelihood is used as follows:

\[ NLL(\Theta) = -\log p(D \mid \Theta) \]

\[ = -\sum_{i=1}^{k} \log p(d_i \mid \theta). \quad (5) \]

The most probable \( \Theta \) is to maximize the likelihood or to minimize the negative log likelihood. Maximum Likelihood estimation is follows:

\[ \Theta = \arg \min_{\Theta} NLL(\Theta). \quad (6) \]
3.2 The Parameter Adaptation

The shape of error model are described by two parameters, σ and γ. To calculate (5), two parameters should be available. In spite of the fact that σ needs to be tuned by hands, Torr and Zisserman [2000] offered the method to adapt γ through Expectation Maximization (EM) procedure. If data class μ is 1 when the corresponding data is an inlier, and 0 when the data is an outlier, then the inlier ratio γ can be estimated as the mean of data class μ:

\[ γ = \frac{1}{k} \sum_{i=1}^{k} P(μ_i = 1 | γ) = \frac{1}{k} \sum_{i=1}^{k} p_i + p_o \]  

where

\[ p_i = γ \left( \frac{1}{\sqrt{2π}} \right)^d \exp \left( -\frac{e_i^2}{2σ^2} \right) \]

\[ p_o = (1 - γ) \frac{1}{p_o} \]  

EM is an iterative procedure to find locally optimal value. In this problem, EM tries to find locally optimal value of γ through update equation. At first, guess the initial value of γ as 0.5, then update γ through (7), and repeat this update step until its convergence. Since five iterations are enough for convergence experimentally, this paper uses this value for adapting γ.

3.3 The Overall Procedure

Now, the localization problem is formulated as (6). To search Θ, MLESAC utilizes the iterative probabilistic sampling. Therefore, the overall procedure is an iteration of three stages: selecting data, estimating Θ from the selected data, and evaluating it using (5). It is described in Fig. 4.

It is difficult to determine the number of iteration t by hands. Fischler and Bolles [1981] offered the following relationship,

\[ P_{fail} = (1 - \hat{γ})^t, \]

where \( P_{fail} \) is the probability to fail in selecting t inliers during t repetitions, and \( \hat{γ} \) is a guessed inlier ratio before running an algorithm. From the above equation, t is derived as follows:

\[ t = \frac{\log P_{fail}}{\log (1 - \hat{γ})^t}. \]

4. SIMULATION RESULTS

4.1 Simulation Configurations

Simulation was performed in a space whose size was 1000 × 1000 and contained 20 landmarks. The robot was placed at \( p = (329, 82)^T \) and its orientation is \( θ = 0.314 \) [rad]. Two parameters, α and β, control observation condition of landmarks. α is a probability of being an outlier, that is, it is the rate of outliers to whole landmarks. β is a magnitude of observation noise. The simulator used additive Gaussian noise \( N(0, β^2) \) to generate the observation noise.

Four existing methods were employed for the experiment. Betke97, Shimshoni02, Thomas05, and Se04(RANSAC) were the existing methods in Betke and Gurvits [1997], Shimshoni [2002], Thomas and Ros [2005], and Se et al. [2004]. Betke97+MLESAC, Shimshoni02+MLESAC, and Thomas05+MLESAC were methods incorporated with MLESAC, and Se04+MLESAC was a method fused with MLESAC, not RANSAC. Table 1 shows their characteristics, and Table 2 shows their error definitions and parameters used in the simulation. In Table 2, \( m_j \) is an absolute position of a landmark matched with measurement \( d_j \), \( R(θ) \) is the clockwise rotation matrix, and \( AD \) is the function to calculate the difference of two angles as follows:

\[ AD(φ, ϕ) = \min \left( |φ - ϕ - 2π|, |φ - ϕ + 2π| \right). \]

Two experiments aimed to examine accuracy with respect to various situations. Experiment #1 was performed in a set of various outlier ratios α, and Experiment #2 was performed in a set of various magnitude of observation noises β. For a statistically representative result, 1000 runs were evaluated.

4.2 Results and Discussion

- **Experiment #1**  Fig. 5 shows that the proposed framework made the existing methods more accurate in various outlier rates. Below analysis explains why the methods with MLESAC are more accurate than methods without it. Se04+MLESAC shows slightly better accuracy in high outlier ratios compared to Se04(RANSAC). Methods with MLESAC were about 100 times more accurate than methods without MLESAC until α = 0.5. However, methods with MLESAC showed significant degradation of accuracy after α = 0.5. Since the experiment used the parameter t using \( \hat{γ} = 0.5 \), the number of trial, t, is not sufficient to overcome bigger outlier ratio than 0.5. When t was increased (or \( \hat{γ} \) was decreased), the accuracy became better but it needs more computation time. RANSAC also has the same problem.

- **Experiment #2**  Fig. 6 shows that the proposed framework improve the accuracy of existing methods. Se04+MLESAC shows 30 ~ 40% better accuracy than
\( e_i = \text{Err}(d_i; \Theta, m_j) \) where \( \Theta = (p, \theta)^T \)

\( \text{Betke97+MLESAC} \)

Bearing \( \text{Err}(d_i; \Theta, m_j) = AD(\text{atan}(m_j - p) - \theta, d_i) \)

\( l = 1 \), \( t(\text{Refer (11)}) = 72 \)

\( \sigma = 0.0175 \)

\( \text{Shimshoni02+MLESAC} \)

Bearing \( \text{Err}(d_i; \Theta, m_j) = AD(\text{atan}(m_j - p) - \theta, d_i) \)

\( l = 1 \), \( t(\text{Refer (11)}) = 72 \)

\( \sigma = 0.0175 \)

\( \text{Thomas05+MLESAC} \)

Distance \( \text{Err}(d_i; \Theta, m_j) = \| m_j - p \| - d_i \)

\( l = 1 \), \( t = 35 \)

\( \sigma = 4.566 \)

\( \text{Se05(RANSAC)} \)

Displacement \( \text{Err}(d_i; \Theta, m_j) = \| R(\theta)(m_j - p) - d_i \| \)

\( l = 2 \), \( t = 17 \)

\( \delta = 11.086 \)

\( \text{Se05+MLESAC} \)

Displacement \( \text{Err}(d_i; \Theta, m_j) = \| R(\theta)(m_j - p) - d_i \| \)

\( l = 2 \), \( t = 17 \)

\( \delta = 4.566 \)

Table 2. Parameters used in the simulation (common parameters: \( P_{\text{fail}} = 0.01, \hat{\gamma} = 0.5 \))

\[ L(e_i) = \begin{cases} 0 & |e_i| < \delta \\ \text{const} & \text{otherwise} \end{cases}, \quad (15) \]

and a loss function of MLESAC can be derived as

\[ L(e_i) = -\log \left( \frac{1}{\sqrt{2\pi}\sigma} \exp\left( -\frac{e_i^2}{2\sigma^2} \right) + (1-\gamma)^{\frac{1}{2}} \right). \quad (16) \]

Three loss functions are plotted in Fig. 7.

It is possible to understand why LSM does not overcome outliers from Fig. 7. Serious error by an outlier causes big loss enough to make the result of LSM wrong. In contrast, RANSAC and MLESAC have a constant loss out of defined range so they are more robust to outliers than LSM. The


