Real-time Grid Map Construction Based on ORB-SLAM2 (Postprint)

Authors: Lu Jianwei, Wang Yaoli

Date: 2018-08-13T00:00:00+00:00

Abstract

To address the limitation of current visual SLAM systems that can only output camera motion trajectories but cannot generate maps for path planning and navigation, this paper proposes a real-time grid map construction algorithm based on ORB-SLAM2. First, an inverse sensor model (ISM) suitable for visual SLAM is established. Second, the construction mechanism of the grid map algorithm is reorganized specifically for the ISM model, and a detailed derivation is presented. Finally, the specific implementation scheme for ORB-SLAM2 grid map construction is introduced. Through experiments, both the ISM model and the grid map model are analyzed to ensure algorithm feasibility. Real-time experiments conducted with monocular and RGB-D depth cameras achieve real-time grid map construction that clearly reveals obstacle positions, thereby validating the effectiveness of the proposed algorithm.

Full Text

Real-time Occupancy Grid Mapping Using ORB-SLAM2

Lu Jianwei, Wang Yaoli†
(College of Information & Computer Science, Taiyuan University of Technology, Taiyuan 030024, China)

Abstract

Current visual SLAM systems can only output camera motion trajectories but cannot generate maps suitable for path planning and navigation. To address this limitation, this paper proposes a real-time grid map construction algorithm based on ORB-SLAM2. First, we establish an inverse sensor model (ISM) tailored for visual SLAM. Second, we restructure the grid map algorithm' s construction mechanism for the ISM model and derive it in detail. Finally, we present the specific implementation scheme for ORB-SLAM2 grid map construction. Through experimental analysis of the ISM model and grid map model, we verify the algorithm' s feasibility. Real-time experiments using both monocular and RGB-D cameras achieve real-time grid map construction that clearly reveals obstacle positions, demonstrating the effectiveness of the proposed algorithm.

Key words: ORB-SLAM2; inverse sensor model; grid map; RGB-D camera

0 Introduction

Simultaneous Localization and Mapping (SLAM) represents a critical domain in robotics and autonomous navigation. Operating without prior environmental information, SLAM systems build environmental models while estimating their own motion using onboard sensors. From simple vacuum cleaning robots to complex autonomous vehicles, SLAM serves as an indispensable component across intelligent applications. Whether for autonomous quadrotors exploring harsh environments or self-driving cars, automatic navigation and mapping have become foundational supports for numerous applications. Consequently, efficient robot navigation and map construction have remained persistent objectives for researchers, particularly in recent years as this field has witnessed rapid development and significant advances.

While LiDAR-based 3D SLAM systems currently operate without challenge in indoor environments, their high cost restricts deployment in low-cost systems. This has led to the emergence of Visual SLAM (V-SLAM). Popular V-SLAM systems include MSCKF [?], ROVIO [?], OKVIS [?], ORB-SLAM2 [?], and VINS-mono [?]. However, these systems do not truly achieve mapping functionality, merely outputting simple camera motion trajectory graphs that cannot be used for robot navigation and obstacle avoidance in practice.

Santana et al. [?] introduced a monocular camera-based method for constructing 2D maps using planar information (homography matrices) in 2011. This approach rapidly determines obstacle presence by assessing whether image feature points lie on the same plane during segmentation. However, due to the lack of scale information in monocular vision, the resulting 2D grid maps are unsuitable for real robot localization and navigation. Dia et al. [?] presented a novel inverse sensor model for occupancy grid mapping, though without SLAM applications. Hull [?] described real-time occupancy grid mapping based on LSD-SLAM, similar to [?], implementing monocular SLAM for occupancy grid map construction through point cloud projection. Gregorio et al. [?] introduced an efficient robot navigation mapping framework using 3D occupancy grid mapping, but its computational demands are too high for SLAM-based robot systems where resource consumption must be minimized.

ORB-SLAM2 is a popular V-SLAM system capable of creating relatively simple 3D point cloud maps, yet these point clouds cannot be used for robot path planning and navigation. Therefore, to address ORB-SLAM2' s limitation of being unable to generate maps for navigation and path planning, we utilize keyframes, map points, and pose x to generate 2D occupancy grid maps suitable for robot path planning and navigation.

1 Establishing the Inverse Sensor Model

A mobile robot operating in an unknown environment estimates its external environment by establishing a probability distribution model that considers noise and uncertainty factors through its sensors. This model is called the sensor model (SM). The model that builds an external environment map using observed data is called the inverse sensor model (ISM).

To better explain the ISM model, we first consider an ideal sensor. Assuming its probability distribution function is $\delta(z-r)$, where δ is the Kronecker delta, we can obtain the ideal ISM model function $g(r)$ as a closed-form solution to equation (2). Here, r represents the sensor observation range, and L is the buffer zone width of the sensor observation area, with L sized as the diagonal distance between two corners of a grid cell.

In the ideal ISM model function $g(r)$, a probability value of 1 is not assigned at r but rather across a band of width L near r. This ensures that cells containing obstacles receive a value of 1. The practical ISM model is obtained by convolving the ideal inverse sensor model with Gaussian noise. The convolution can be performed piecewise since $g(r)$ is a piecewise function. The convolution equation can be rewritten as:

$$g(r) * f(r) = \int_{-\infty}^{\infty} g(\tau) f(r-\tau) d\tau$$

where $f(r)$ represents the Gaussian noise model. For computational convenience, we use the error function to express the result. The error function $\text{erf}(x)$ is defined as:

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

The final practical ISM model after convolution with Gaussian noise is given by:

$$g(r) * f(r) = F(z+L, z-L) - F(z+L, z+L) + F(z-L, z-L)$$

where $F(a, b)$ represents the convolution integral result over the interval $[b, a]$.

2 Establishing the Grid Map Model

Grid maps divide the environment into numerous equal-scale cells, representing partial environments through grid cells and storing information about whether each cell contains obstacles—that is, whether the cell state is occupied, free, or unknown. In the occupied case, it can represent obstacle boundaries for path planning and robot navigation. The free case indicates no obstacles, while unknown means the sensor cannot observe the cell state. Grid map scale is represented through resolution, such as 1 cm representing 5 m.

Since sensor observations introduce noise and uncertainty, these factors must be incorporated into the grid map model. If m_i represents the i-th cell in the map, its state can be represented by a binary occupancy variable (occupied or free). Ideally, with sensor observations $z_{1:t}$ and robot poses $x_{1:t}$ known from time 1 to t, the posterior probability for map cell m_i can be expressed as:

$$p(m_i|z_{1:t}, x_{1:t})$$

If the map m is divided into i equal-scale grids with index i, we have $m = \{m_i\}$. Based on whether grid cells are occupied, $m_i = 1$ indicates occupancy, $m_i = 0$ indicates free space, and $m_i = 0.5$ represents unknown state.

For 1,000 cells, the map would have 2^{1000} possible configurations, making this approach impractical for real mapping. Based on the Markov assumption that each cell’ s probability $p(m_i)$ is statistically independent, the probability of N independent grid cells can be expressed as:

$$p(m) = \prod_{i=1}^{N} p(m_i)$$

Over time, the probability of whether the i-th cell is occupied is incorporated into the map probability distribution, enabling continuous map updates. To achieve this mechanism, we employ the binary Bayes filter algorithm. For static environments where sensor observations are independent of map cells, applying Bayes’ rule yields:

$$p(m_i|z_{1:t}) = \frac{p(z_t|m_i)p(m_i|z_{1:t-1})}{p(z_t)}$$

To avoid numerical instability near probabilities of 0 or 1, we typically apply the logit function to obtain:

$$L(m_i|z_{1:t}) = L(m_i|z_{1:t-1}) + \log \frac{p(z_t|m_i)}{p(z_t|\neg m_i)}$$

where $L(m_i) = \log \frac{p(m_i)}{1-p(m_i)}$ is the logit function. This can be rewritten more clearly as:

$$l_{t,i} = l_{t-1,i} + \mathrm{ISM}(m_i, z_t, x_t) - l_0$$

where $l_{t,i}$ represents the log-odds of cell i at time t, and l_0 is the prior.

3 ORB-SLAM2 Generating Grid Map Model

ORB-SLAM2 generates grid maps by publishing keyframes, map points, and robot poses x in ROS, then subscribing to these publications as sensor observation data z and robot poses x for the ISM model to generate the required grid map through the grid map model. The algorithm flow is shown in Figure 2 [Figure 2: see original paper].

The specific implementation involves: 1. Initializing all grid cells with prior occupancy probabilities 2. For each time step t from 1 to T, processing all grid cells within the perceptual range 3. Updating occupancy probabilities using the logit function and ISM model 4. Recovering occupancy probabilities from log-odds values

At time t, with robot pose x_t, sensor observation z_t, and all grid cells $\{m_i\}$, each cell' s posterior probability is $p(m_i|z_{1:t}, x_{1:t})$. According to maximum a posteriori (MAP) mapping:

$$m^* = \arg\max_m p(m|z_{1:t}, x_{1:t})$$

4 Experimental Results and Analysis

4.1 Selection of L Value

When discussing the impact of L on the ISM model, we employ a fixed-variable strategy. First, we fix the z value and Gaussian noise model, then observe ISM model changes by varying L. As shown in Figure 3 [Figure 3: see original paper], as L increases, the ISM model' s grid cell probability continuously increases, indicating more cells will be judged as occupied. Conversely, free cells in the grid map decrease. When L exceeds a certain threshold, significant misjudgment occurs—areas without obstacles in reality are incorrectly judged as occupied. Experiments show that when L equals the diagonal length of a grid cell, the misjudgment probability is minimized.

4.2 Impact of Uncertainty Factors and Observation Range z

Since sensor observations introduce noise and uncertainty factors, these also affect map construction. The influence of the uncertainty parameter σ on the ISM model is shown in Figure 4 [Figure 4: see original paper]. As σ increases, indicating greater influence from uncertainty and noise, the ISM model' s posterior probability for grid cells decreases, meaning more cells will be judged as free space, which does not reflect reality.

Different observation ranges z also affect the ISM model (Figure 5 [Figure 5: see original paper]). Analysis reveals that as z increases, the robot' s observation range expands, but judgment accuracy for map grid cells decreases accordingly.

4.3 Indoor Real-time Experiments

Indoor real-time testing was conducted in a laboratory and corridor using an XBOX360 depth camera and a Logitech USB camera.

Monocular Camera Experiments: Using a USB camera in the laboratory, the generated grid map showed insufficiently clear obstacle annotation. Obstacles like desks along the central axis could not be labeled. This occurs because monocular cameras lack scale information, causing keyframes and feature points to be incorrectly assigned to the laboratory edges during point cloud map generation, resulting in failure to annotate central axis obstacles and producing erroneous grid maps. The real indoor environment and monocular-generated grid map are shown in Figure 6 [Figure 6: see original paper].

RGB-D Camera Experiments: To overcome monocular scale limitations, RGB-D cameras with scale information were used. Compared to monocular results, RGB-D cameras offer two advantages: (a) They possess scale information, enabling correct grid map generation that can annotate central axis obstacles; (b) They produce clearer grid maps that distinctly display obstacles and their relative positions, unlike monocular cameras that normalize all obstacles to a single plane. Figure 7 [Figure 7: see original paper] shows the RGB-D camera’ s motion trajectory point cloud, displaying the sparse point cloud map of the indoor environment after one complete camera rotation. Figure 8 [Figure 8: see original paper] shows the grid map generated during RGB-D camera motion. The irregular grid pattern reflects real-world irregular obstacles like tables and chairs.

Large-scale Corridor Experiments: After small-scale indoor testing, large-scale corridor experiments were conducted with the RGB-D camera (Figure 9 [Figure 9: see original paper]). The generated grid map clearly shows corridor walls, laboratory thresholds, and prominently marked elevator entrances, effectively reflecting the corridor environment and annotating obstacles (Figure 10 [Figure 10: see original paper]).

5 Conclusion

Addressing the limitation that previous visual SLAM systems cannot generate maps for robot navigation and path planning, this paper proposes a real-time grid map construction algorithm based on ORB-SLAM2. The algorithm establishes an inverse sensor model (ISM) suitable for visual SLAM, qualitatively analyzes the model to determine a practical ISM, derives the grid map construction mechanism in detail, and clearly explains the ORB-SLAM2 grid map generation process through algorithm flowcharts. Experimental analysis of monocular and RGB-D camera results demonstrates the effectiveness of the ISM model and grid map algorithm, verifying that the proposed algorithm has strong practicality and can generate real-time grid maps for robot navigation and path planning.

References

[1] Cadena C, Carlone L, Carrillo H, et al. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age [J]. IEEE Trans on Robotics, 2016, 32 (6): 1309-1332.

[2] Mourikis A I, Roumeliotis S I. A multi-state constraint Kalman filter for vision-aided inertial navigation [C]// Proc of International Conference on Robotics and Automation. Roma, Italy: Institute of Electrical and Electronics Engineers Inc, 2007: 3565-3572.

[3] Bloesch M, Omari S, Hutter M, et al. Robust visual inertial odometry using a direct EKF-based approach [C]// Proc of IEEE//RSJ International Conference on Intelligent Robots and Systems. Hamburg, Germany: Institute of Electrical and Electronics Engineers Inc, 2015: 298-304.

[4] Bloesch M, Burri M, Omari S, et al. Iterated extended Kalman filter based visual-inertial odometry using direct photometric feedback [J]. The International Journal of Robotics Research, 2017, 36 (10): 1053-1072.

[5] Leutenegger S, Lynen S, Bosse M, et al. Keyframe-based visual–inertial odometry using nonlinear optimization [J]. The International Journal of Robotics Research, 2015, 34 (3): 314-334.

[6] Mur-Artal R, Montiel J M M, Tardos J D. ORB-SLAM: a versatile and accurate monocular SLAM system [J]. IEEE Trans on Robotics, 2015, 31 (5): 1147-1163.

[7] Mur-Artal R, Tardos J D. Orb-slam2: an open-source slam system for monocular, stereo, and rgb-d cameras [J]. IEEE Trans on Robotics, 2017, 33 (5): 1255-1262.

[8] Qin Tong, Shen Shaojie. Robust initialization of monocular visual-inertial estimation on aerial robots [C]// Proc of the IEEE//RSJ International Conference on Intelligent Robots and Systems. Vancouver, BC, Canada: Institute of Electrical and Electronics Engineers Inc, 2017: 24-28.

[9] Santana A M, Aires K R T, Veras R M S, et al. An approach for 2D visual occupancy grid map using monocular vision [J]. Electronic Notes in Theoretical Computer Science, 2011, 281: 175-191.

[10] Dia R, Mottin J, Rakotovao T, et al. Evaluation of occupancy grid resolution through a novel approach for inverse sensor modeling [J]. IFAC-Papers OnLine, 2017, 50 (1): 13841-13847.

[11] Hull G. Real-time occupancy grid mapping using LSD-SLAM [D]. [S. l.]: Stellenbosch University, 2017.

[12] De Gregorio D, Di Stefano L. SkiMap: an efficient mapping framework for robot navigation [C]// Proc of International Conference on Robotics and Automation. Singapore: Institute of Electrical and Electronics Engineers Inc, 2017: 2569-2576.

[13] Levine M P. System and method of drawing anti-aliased lines using a modified bresenham line-drawing algorithm: U. S. Patent 5, 815, 162 [P]. 1998-09-29.

[14] Hou Rongbo, Wei Wu, Huang Ting, et al. Indoor robot localization and 3D dense mapping based on ORB-SLAM [J]. Journal of Computer Applications, 2017, 37 (5): 1439-1444.

[15] Wang Quan, Hu Yueli, Zhang He. Mobile robot simultaneous localization and mapping based on ROS and Kinect [J]. Application Research of Computers, 2017, 34 (10): 3184-3187.

[16] Wang Wei, Pu Yunming, Li Wang. A multi-robot 2D map fusion algorithm based on CI factor graph [J]. Robot, 2017, 39 (6): 872-878.

[17] Luo Yuan, Fu Youli, Cheng Tiefeng. Simultaneous localization and mapping implementation based on the improved rao-blackwellized particle filter [J]. Control Theory & Applications, 2015, 32 (2): 267.

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.