MVA2013 IAPR International Conference on Machine Vision Applications, May 20-23, 2013, Kyoto, JAPAN

Scan-matching Based 6DOF SLAM Using Omnidirectional Stereo

Hiroshi Koyasu
Graduate School of Science and Engineering, Saitama University
Simo-Okubo 255, Saitama, Japan
koyasu@cda.ics.saitama-u.ac.jp

Hitoshi Maekawa
Graduate School of Science and Engineering, Saitama University

Hiroshi Kawasaki
Graduate School of Science and Engineering, Kagoshima University

Shintaro Ono
Institute of Industrial Science, The University of Tokyo

Katsushi Ikeuchi
Institute of Industrial Science, The University of Tokyo

Abstract

Simultaneous localization and mapping (SLAM) of working environments is one of the important problems in mobile robotics. For the problem, we present a new method using an omnidirectional stereo system. By using 3D range information obtained from the omnidirectional stereo, we construct scan-matching based extended Kalman filter SLAM(EKF-SLAM) system. In the system, the output of the piecewise iterative closest point(ICP) algorithm is used as the observation of the EKF. The global consistency of the robot trajectory is kept by a loop closing method based on the Kalman smoothing. Experimental results of the method are shown for a simulated environment.

1 Introduction

3D simultaneous localization and mapping(SLAM) is one of the important problems in the area of mobile robotics (*e.g.*[1]). To estimate robot pose and its working environment, the robot should obtain 3D information of the environment. Laser range finder (LRF) is one of commonly used sensors that obtains the 3D information. The LRF can measure depth information of a target point accurately by laser beam projection, however additional scanning mechanism is required to reconstruct 3D surface structure[2]. Lider is also hopeful solution, however, its maximum vertical resolution is currently 32.

The other possible approach is a vision based one which uses conventional cameras to form stereo view. By this approach, 3D surface structure can be obtained together with its texture. However, conventional camera has still limited view area to cover whole surrounding environment in which the robot should navigate and work. To overcome this drawback of vision based approach, we employ an omnidirectional stereo vision system which consists of hyperboloidal mirror and CCD camera.

In our method, accuracy of each stereo 3D-information is improved by sub-pixel estimation of disparity. After this pre-process, these 3D data are aligned by the Iterative Closest Point (ICP) algorithm[3] at each step. Then, the estimation is improved by re-calculating the error of the observations by Kalman smoothing when the robot detects loop closure of its trajectory.

The main contributions of this paper are as follows:

1. Extended Kalman filter SLAM(EKF-SLAM) formulation for a system which uses an output of the scan-matching as an EKF observation is proposed. Similar method was proposed by Cole and Newman[2], however, their method is complete SLAM; which keeps history of all previous robot pose. This causes too much updating poses which do not relate to the current observation; this requires meaning-less computation. On the other hand, our method is formulated as partial SLAM; which keeps only the current robot pose.

2. SLAM system with an omnidirectional stereo vision with a sub-pixel level disparity estimation, that is effective in both estimating 3D positions and reconstructing environments accurately, is constructed; to the best of our knowledge, there are no range based localization method that uses an omnidirectional stereo with sub-pixel level disparity estimation.

The most significant advantage of an omnidirectional camera is its wide field of view. Therefore, almost any part of view is shared with a different view points; this leads to stable localization using ICP. Furthermore, the required number of observations to reconstruct whole environment decreases.

The rest of this paper is structured as follows: In Section 2, we present related work. In Section 3, we describe scan-based EKF-SLAM formulation. In Section 4, we describe implementation of our EKF-SLAM using omnidirectional stereo. Experimental results are shown in Section 5. Finally, we conclude and present plans for future work in Section 6.

2 Related work

Our work is positioned as a part of SLAM. Many studies have been carried out in this domain. Most of the researches use Bayesian filter to formulate the problem. To implement the filter, some of these use particle filter, and the others use extended Kalman filter(EKF).

Extended Kalman filter is widely used[4, 5]. Many method is based on feature tracking. Such methods uses only a salient point of an observation. This may reduce computation, however, most of the observation is wasted. For dense 3D reconstruction of environments, it is required to use many data. Threfore, scan-matching is suitable for dense 3D reconstruction.

Scan-matching methods can divided into two categories: one is matching with current range observation with generated map, and the other is matching with the current and a previous range observation.

The advantage of former is comparing with rich information. Thus it is sutable for poor resolution range sensor (*e.g.*sonar). However, the quality of the global map affects the localization results, *i.e.*the previous localization errors degrade the accuracy of consequent localization. Nevertheless, correcting previous localize error in the map is difficult. One of the solution of this problem is FastSLAM[6]. The method keeps various map hypothesis in particles and paticle filter selects most likely maps. However, the number of particle increases exponentially as the dimension of the state increases.

The latter requires a sensor which has wide feild of view with enough resolution. So the number of methods in this category is not many.

Cole and Newman[2] propose scan-matching based EKF-SLAM. Checchin *et al.* [7] apply the method to the rader-based system with improving matching process. Their method is formulated as complete SLAM; which keeps history of all previous robot pose. This causes too much updating poses which do not relate to the current observation.

Borrmann *et al.* [8] propose another approach; which is based on Graph-SLAM[9]. Their method optimize the robot trajectory by using piecewise ICP results and loop information. The optimization process is based on relaxation; which requires much computation. We think the method based on a back-propagation such as Kalman smoothing is enough to optimize.

3 Scan-matching based EKF-SLAM

In this section, we describe the framework of the scan-based extended Kalman filter SLAM(EKF-SLAM). Note that our EKF formulation is not depend on the sensor model, so it can apply to the system using other sensor such as lider.

The state vector includes only the current robot position and orientation, *i.e.*not including history of the previous robot poses or map elements. By using the current robot's position $(x_t, y_t, z_t)^T$ and rotation quaternion $\boldsymbol{q}_t$, the state vector at time t is denoted as follows:

$$\boldsymbol{x}_t = (x_t, y_t, z_t, \boldsymbol{q}_t^T)^T,$$

where T denotes transpose. The state transition equation is defined as follows by a function $\boldsymbol{F}$ to update state vector by odometer inputs:

$$\boldsymbol{x}_{t+1} = \boldsymbol{F}(\boldsymbol{x}_t, \boldsymbol{u}_{t+1}) + \boldsymbol{v}_{t+1}, \tag{1}$$

where $\boldsymbol{u}$ represents the inputs from odometer, $\boldsymbol{v}$ represents white noise vector, respectively. We use a robot driven by the wheel, so we model $\boldsymbol{F}$ as moving a planer surface, then the system noise vector compensates others; such as entering slope.

In our EKF formulation, the uncertainty of the state $\boldsymbol{P}_{t+1}$ is calculated by

$$\boldsymbol{P}_{t+1} = \Delta_{\boldsymbol{x}}\boldsymbol{F}\boldsymbol{P}_t\Delta_{\boldsymbol{x}}\boldsymbol{F}^T + \Delta_{\boldsymbol{u}}\boldsymbol{F}\boldsymbol{\Sigma}_{\boldsymbol{u}_{t+1}}\Delta_{\boldsymbol{u}}\boldsymbol{F}^T + Q_{t+1}, \tag{2}$$

where $\Delta_{\boldsymbol{x}}\boldsymbol{F}$ and $\Delta_{\boldsymbol{u}}\boldsymbol{F}$ represent Jacobian of $\boldsymbol{F}$ differentiated with respect to $\boldsymbol{x}$ and $\boldsymbol{u}$, $\boldsymbol{\Sigma}_{\boldsymbol{u}_{t+1}}$ represents the error covariance of the $\boldsymbol{u}_{t+1}$, Q represents the covariance of $\boldsymbol{v}$, respectively.

A scan-matching method gives us a relative displacement and rotation between the current and a previous robot pose. Therefore, the observation equation is formed as follows:

$$z_{t,t-k} = \boldsymbol{H}(\boldsymbol{x}_t, \boldsymbol{x}_{t-k}) + \boldsymbol{w}_{t,t-k}, \tag{3}$$

where $z_{t,t-k}$ represents the relative displacement and rotation between $\boldsymbol{x}_t$ and $\boldsymbol{x}_{t-k}$, and $\boldsymbol{w}_{t,t-k}$ is an error of the result of the ICP, respectively. We assume that the covariance of the $\boldsymbol{w}$, denoted as R, is constant with respect to any pair of $\boldsymbol{x}$. $\boldsymbol{H}$ is the function to calculate relative pose of $\boldsymbol{x}_t$ with respect to $\boldsymbol{x}_{t-k}$. It includes rotation, so it is a non-lenear function. Note that Eq. (3) can use ICP matching between current and any previous range data.

Both $\boldsymbol{x}_t$ and $\boldsymbol{x}_{t-k}$ have some uncertainty, we modify the Kalman filter formulation to deal with the uncertainties. We assume that each error covariance of $\boldsymbol{x}$ is independent, the innovation covariance S is calculated as follows:

$$\begin{aligned}\boldsymbol{S}_{t,t-k} \;=\; & \Delta_{\boldsymbol{x}_t}\boldsymbol{H}P_t\Delta_{\boldsymbol{x}_t}\boldsymbol{H}^T + \Delta_{\boldsymbol{x}_{t-k}}\boldsymbol{H}P_{t-k}\Delta_{\boldsymbol{x}_{t-k}}\boldsymbol{H}^T \\ & + R_{t,t-k}. \end{aligned} \tag{4}$$

By using this innovation covariance, the Kalman gain $K_{t,t-k}$ is calculated as same as the linear Kalman filter. Updating state is also same. The error covariance of the state is done by using following:

$$\tilde{\boldsymbol{P}}_t = (I - K_{t,t-k}\Delta_{\boldsymbol{x}_t}\boldsymbol{H})\hat{P}_t - K_{t,t-k}\Delta_{\boldsymbol{x}_{t-k}}\boldsymbol{H}P_{t-k}, \tag{5}$$

where I represents the identity matrix.

Usually, we use $k = 1$, so the error of the above successive ICP estimation is accumulated along with robot's navigation. To reduce this accumulated error, we apply a closing loop method based on Kalman smoothing[10]. When a loop is detected by the method described in Section 4.3, the k is set to the size of the loop, then the newest pose is updated. The result is propagated by using Kalman smoothing.

4 Implementation using omnidirectional stereo

4.1 Omnidirectional stereo

Our omnidirectional stereo is composed out of two omnidirectional cameras. The omnidirectional camera

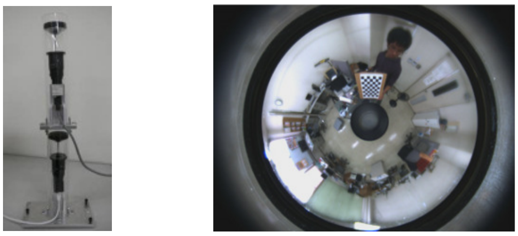
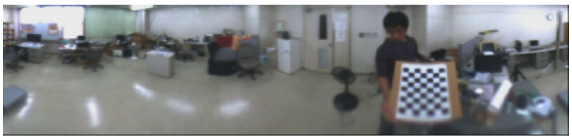


Figure 1. Omnidirectional stereo camera(left), and its image(right).

(a)



(b)

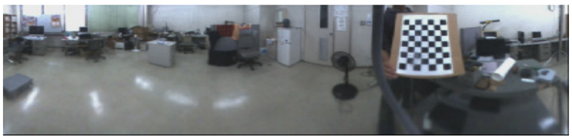


Figure 2. Panoramic image pair converted from omnidirectional image shown in the Fig. 1. (a): Upper image, (b): Lower image.

that we used is based on a hyperboloidal mirror[11], as shown in the Fig.1(left). It can acquire omnidirectional images as shown in the Fig.1(right).

To apply omnidirectional stereo, we align two omnidirectional cameras vertically same as Gluckman *et al.* [12] By converting input omnidirectional images to panoramic images, we obtain image pairs that have parallel and vertically aligned epipolar lines. An example of the panoramic image pair is shown in the Fig. 2.

To calculate disparities, we apply a SAD (Sum of Absolute Difference) based block-matching method to search corresponding points between a stereo image pair. Then, we estimate sub-pixel level disparities by using a corresponding evaluation function as proposed by Arai *et al.*[13].

4.2 Scan-matching using ICP

We use Iterative Closest Point (ICP)[3] algorithm in order to align range data between two successive observations. We emphasize that such piecewise ICP is advantageous for computational cost, however, precision is not guaranteed, and the drawback can be eliminated by the *closing loop* method explained in the next section.

The ICP aligns two sets of range data, which are observed at different positions, by estimating their positional relation. The method repeats following steps unless mean-square error falls less than a threshold: i) search for correspondence between points in the two range data sets. ii) calculate the translation and rotation which minimize sum of distances between corresponding points.

To avoid converging local minima, the initial value of the ego-motion should be chosen carefully. We use an odometry information of the robot as the initial value.

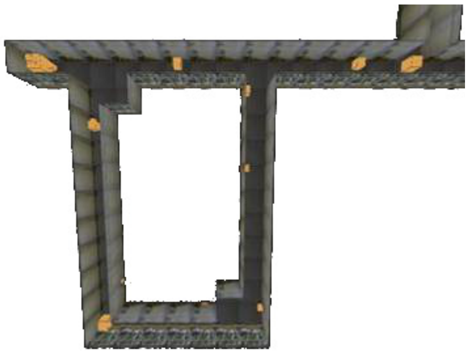


Figure 3. A corridor environment for simulation(top view).

The uncertainty of the ICP result is currently given as a constant value determined experimentally.

4.3 Loop-closure detection

The error of the successive ego-motion estimation is accumulated. To fix this accumulated error, we apply a closing loop method based on Kalman smoothing[10].

To detect a loop-closure, we use the residual of the ICP. When a previous position is inside the 99% probability ellipsoid of the current position, this pair of the position may be loop-closure. Then, we calculate relative motion between the pair by the ICP. If the residual of the ICP is less than a threshold, the loop is detected.

5 Experiments

To evaluate our method, we conducted a simulation experiment. We simulated an indoor environment, which consists of planer objects such as walls, boxes; each object has a texture made from a photograph. The simulated environment is shown in Fig. 3. The robot moved the environment with taking pairs of panoramic stereo image and odometer values.

The robot started from the bottom-left corner of the Fig. 3 to top-left corner straightly, then turned 30° right 3 times for the sake of going to the next corner, finally went back to the bottom-left corner through the top-left corner and the top-right corner. While moving, the robot obtained stereo image pairs and odometer values with every 1m moving and 30° turning. The robot moved totally 64m and obtained 74 observations.

Fig. 4 shows the estimated trajectory before applying the closing loop. In the figure, purple and green lines indicate true and estimated trajectories, respectively. Black circles indicate the estimated position where robot observed. Red ellipsoids indicate 99% probability ellipsoid of the robot position. The positional error of the final position were about 1.1m on x-y plane and 0.24m along z axis. The root mean square error was 0.51m.

At the final positon, since the initial position was inside the 99% probability ellipsoid, the robot ran ICP matching with range data obtained at initial and final position. Then the residual was less than threshold, the robot detected a loop and processed the Kalman smoothing.

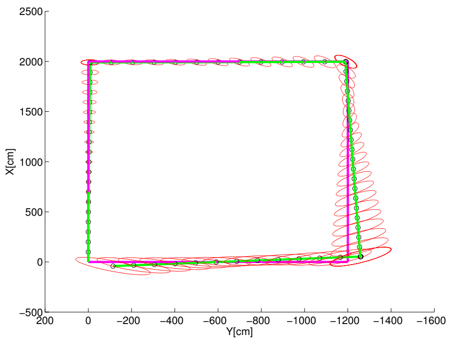


Figure 4. Estimated trajectory of the robot before applying the closing loop. Purple and green lines indicate true and estimated trajectories, respectively. Black circles indicate the estimated position where robot observed. Red ellipsoids indicate 99% probability ellipsoid of the robot position.

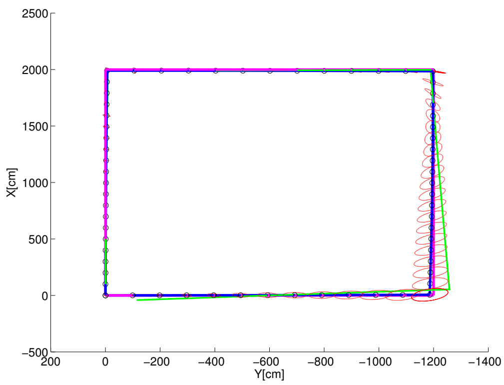


Figure 5. Estimated trajectory of the robot after applying the closing loop. Two trajectories (purple and green lines) are the same as Fig. 4. Blue line indicates the estimated trajectory after applying the closing loop. The estimated positions (black circle) and their 99% probability ellipsoids (red ellipsoid) are on the trajectory after applying the closing loop.

Fig. 5 shows the estimated trajectory after applying the closing loop. In the figure, the estimated trajectory after applying the closing loop (blue line) and the same trajectory as Fig. 4 are drawn. Also the estimated position (black circle) and its 99% probability ellipsoid (red ellipsoid) after the closing loop are drawn. Accumulating error significantly decreased by the closing loop. The root mean square error was reduced to 0.14m.

The reconstruction results before closing the loop and after closing the loop are shown in the figures 6 and 7, respectively. Before closing the loop, a misalignment is seen on the bottom-left corner. But after closing the loop, the corner is correctly lapped over.

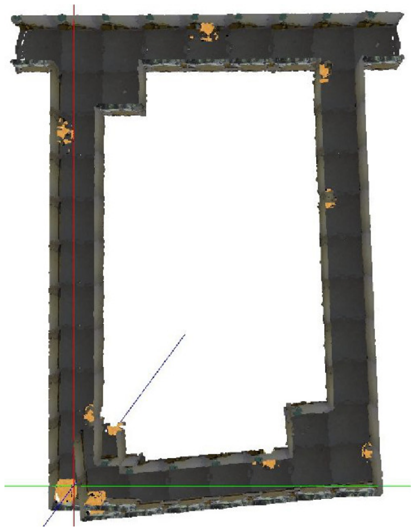


Figure 6. Reconstruction result before closing loop.

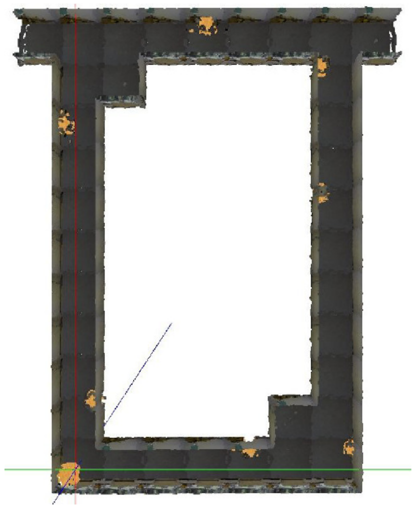


Figure 7. Reconstruction result after closing loop.

6 Conclusion

We propose a new 3D simultaneous localization and mapping method that uses scan-matching of range data obtained by the omnidirectional stereo with a sub-pixel level disparity. We apply novel formulation of EKF-SLAM and Kalman smoothing-based simple closing loop. The simulation experiment suggests that our extended Kalman filter formulation works effectively.

One of our future work is to develop a method to handle the uncertainty of the ego-motion calculated by the ICP, which is treated as a constant value in this paper.

Acknowledgements

This work was partly supported by JSPS KAKENHI Grant Number 23611006.

References

[1] Yu-Shin Chou and Jing-Sin Liu. A robotic indoor 3d mapping system using a 2d laser range finder mounted

on a rotating four-bar linkage of a mobile platform. *International Journal of Advanced Robotic Systems*, 10, 2013.

[2] D. M. Cole and P. M. Newman. Using Laser Range Data for 3D SLAM in Outdoor Environments. In *Proc. IEEE Int'l Conf. on Robotics and Automation*, pages 1556–1563, 2006.

[3] P. J. Besl and N. D. McKay. A Method for Registration of 3-D Shapes. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(2):239–256, 1992.

[4] A. J. Davison and N. Kita. 3D Simultaneous Localisation and Map-Building Using Active Vision for a Robot Moving on Undulation Terrain. In *Proc. of IEEE Int'l Conf. on Computer Vision and Pattern Recognition*, pages 384–391, 2001.

[5] A. Davison, I. D. Reid, N. D. Molton, and O. Sasse. MonoSLAM: Real-Time Single Camera SLAM. *IEEE Trans. on PAMI*, 29(6):1052–1067, 2007.

[6] D. Hähnel, D. Schulz, and W. Burgard. An Efficient FastSLAM Algorithm for Generating Maps of Large-Scale Cyclic Environments from Raw Laser Range Mearsurements. In *Proc. of 2003 IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 206–211, 2003.

[7] P. Checchin, F. Grossier, C. Blanc, R. Chapuis, and L. Trassoudaine. Radar scan matching slam using the fourier-mellin transform. In *7th Int'l Conf. on Field and Service Robotics*, 2009.

[8] D. Borrmann, J. Elseberg, K. Lingemann, A. Nüchter, and J. Hertzberg. Globally consistent 3d mapping with scan matching. *Robotics and Autonomous Systems*, 56(2):130–142, 2008.

[9] F. Lu and E. Milios. Globally Consistent Range Scan Alignment for Environment Mapping. *Autonomous Robots*, 4(4):333–349, 1997.

[10] H. Baltzakis and P. Trahanias. Closing Multiple Loops while Mapping Features in Cyclic Environments. In *Proc. of 2003 IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 717–722, 2003.

[11] K. Yamazawa, Y. Yagi, and M. Yachida. Omnidirectional Imaging with Hyperboloidal Projection. In *Proc. of 1993 IEEE/RSJ Int'l Conf. on Intelligent Robots and Systems*, pages 1029–1034, 1993.

[12] J. Gluckman, S. K. Nayar, and K. J. Thoresz. Real-Time Omnidirectional and Panoramic Stereo. In *Proc. of Image Understanding Workshop*, volume 1, pages 299–303, 1998.

[13] M. Arai, K. Sumi, and T. Matsuyama. Optimization of Correlation Function and Sub-Pixel Estimation Method on Block Matching. In *CVIM-144-5, Information Processing Society of Japan*, pages 33–40, 2004. (in Japanese).