Automatic Camera Pose Estimation Based on Textured 3D Map Information

Yonghoon Ji[*1], Atsushi Yamashita[*1], and Hajime Asama[*1]
[*1] Graduate School of Engineering, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656 Japan

To construct an intelligent space which has a distributed camera network, pre-calibration of all cameras (i.e., determining the absolute poses of each camera) is an essential task that is extremely tedious. This paper deals with automatic calibration method for the distributed camera based on 3D map information of an environment. Parameterized line features that are extracted from both a distributed camera image and the map information are transformed to Hough space and utilized for matching process in particle filter-based estimation. We evaluate the proposed method in a simulation environment with a virtual camera.

1 Introduction

In order to obtain reliable information from a camera network system, pre-calibration of distributed cameras (i.e., determining absolute positions and orientations of each camera) is an essential task that is extremely tedious. In this respect, automatic calibration methods for the camera network has been intensively studied in the past [1][2]. These approaches, however, mobile agents are needed, so that these cannot be applied where the mobile agents cannot use. To this end, we propose a novel approach to realize automatic calibration of external parameters using only map information of the environment. As shown in Fig. 1, textured 3D map information can be utilized to generate virtual 2D images from arbitrary viewpoints (i.e., arbitrary 6DOF camera pose) using 3D projective geometry when the internal camera parameters are known; thus, the external camera parameters can be determined by matching between virtual images generated at every viewpoint and a real image from the camera network. However, it is impossible to search a 6DOF solution without strong constraints because of a myriad number of local minimum solutions. Here, the 3D map information is very useful to reduce the solution space since the cameras are generally installed on the occupied region such as interior wall because of space limitations. We use line features which extracted from both the textured 3D map information and the real image data to make constraints.

Proposed calibration scheme is divided into two steps: ‘parameterization of 3D geometrical lines’ and ‘particle filter-based calibration’ step. During parameterization step, 3D geometrical line parameters for the environment are automatically generated from the textured 3D map information. The calibration step determines the 6DOF camera parameters by matching between generated 3D geometrical line parameters and 2D photometrical line parameters that are generated from the real image from the camera sensor network. We focus on the consideration of expression of the line features and the design of its new measurement model to apply particle filter algorithm.

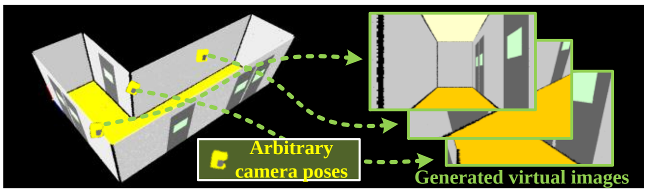
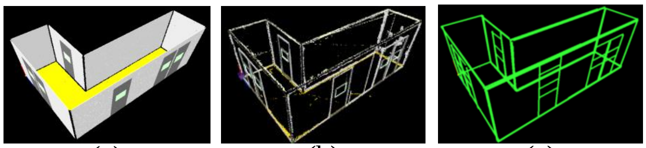


Fig. 1 Arbitrary camera poses and generated virtual images from textured 3D map information.



(a) (b) (c)

Fig. 2 Parameterization of 3D geometrical lines: (a) textured 3D map information, (b) extracted point cloud on line segments, and (c) learned robust line parameters.

2 Parameterization of 3D geometrical lines

The parameterization process of the 3D line segments for the entire environment consists of two major steps: ‘extraction of point cloud on line segments’ and ‘learning robust line parameters’ as illustrated in Fig. 2. The ‘extraction of point cloud on line segments’ step involves generating point cloud on the 3D geometric line segments as shown in Fig. 2 (b). To generate the point cloud, 2D photometrical lines from virtual images from arbitrary camera poses are extracted and they are back-projected onto the 3D map information. The ‘learning robust line parameters’ step learns coefficients of the 3D geometrical line segments based on random sample consensus (RANSAC) algorithm. The parameterized 3D geometrical line segments can be represented only using 6 parameters; thus, the computational burden for particle filter-based matching process can be reduced significantly.

3 Particle filter-based parameter calibration

Particle filter is used for parameter calibration in this study. It is one of the popular methods to implement Bayesian filter that can estimate the probability distribution using a set of random particles. At each iteration, the probabilities of particles are updated using a prediction model and a measurement model, and then particles are resampled. The state (i.e., 6DOF camera pose in this study) is represented by the weighted sum of all particles.

To estimate camera pose based on textured 3D map information, a novel feature comparison model (i.e., measurement model in particle filter) is required for the matching process. To this end, a novel descriptor based on quantized line parameters in Hough space (QLH) which represents the distribution of slopes and distances from the origin of the 2D image plane is proposed in this study, as shown in Fig. 3. The QLH descriptor is robust to illumination changes since they do not require any intensity information and they are always available as long as the edge information is detected.

2D photometrical lines in the image plane are uniquely determined by two properties: a slope α and a distance ρ from the origin as shown in Fig. 3 (a). They can be transformed into Hough space as shown in Fig. 3 (b). This signature can be a robust descriptor for line segment-based matching. The QLH descriptor for the real camera image is simply generated from the image including 2D photometrical lines that are extracted by Hough transform as shown in Fig. 3 (a). On the other hand, the predicted QLH descriptors from the arbitrary camera poses in the 3D map information can be obtained from 2D photometrical line segments that are generated by projecting parameterized 3D geometrical lines onto the virtual 2D image plane.

Bayesian filter in this study determines the posterior probability of the 6DOF camera pose $p(\boldsymbol{x} \mid \boldsymbol{z}, \boldsymbol{m})$ based on the measurement data from the camera image and the map information as follow:

$$p(\boldsymbol{x} \mid \boldsymbol{z}, \boldsymbol{m}) = \eta p(\boldsymbol{z} \mid \boldsymbol{x}, \boldsymbol{m})\, p(\boldsymbol{x} \mid \boldsymbol{m}), \tag{1}$$

where $\boldsymbol{x}=[x\ y\ z\ \psi\ \theta\ \varphi]^{\mathrm{T}}$ is 6DOF camera parameters that should be estimated and η is a normalization constant. $p(\boldsymbol{x} \mid \boldsymbol{m})$ refers to the prior distribution which means the map information $\boldsymbol{m}$ can provides constraints to the camera parameters $\boldsymbol{x}$. $p(\boldsymbol{z} \mid \boldsymbol{x}, \boldsymbol{m})$ denotes the measurement model which is defined by the likelihood function that should be newly designed for the QLH image descriptor in this study. Here, the proposed QLH descriptor can be exploited for criteria of similarity comparison. Thus, we design a novel measurement model as follow:

$$p(\boldsymbol{z} \mid \boldsymbol{x}, \boldsymbol{m}) = \frac{1}{\sigma_{\mathrm{EMD}}\sqrt{2\pi}} \exp\left(-\frac{\mathrm{EMD}(\boldsymbol{z}, h(\boldsymbol{x}))^2}{2\sigma_{\mathrm{EMD}}{}^2} \right), \tag{2}$$

where $\boldsymbol{z}$ and $h(\boldsymbol{x})$ refer to QLH descriptors extracted from a real camera image and predicted from an arbitrary camera pose $\boldsymbol{x}$, respectively. σ_{EMD} means the standard deviation associated with an uncertainty of the QLH descriptor. EMD(·) denotes earth mover's distance between two sets of the QLH descriptor. Earth mover's distance is the most appropriate criteria for this problem on the ground that it calculates the similarity between lines' distributions.

4 Experimental result

We simulated the proposed calibration method in a simulation environment (Fig. 2 (a)) with a virtual camera image (Fig. 3 (a)). The simulation environment includes

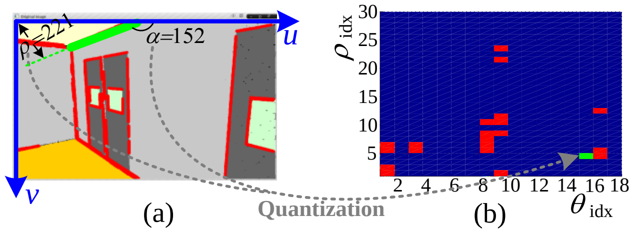


Fig. 3 Generation of QLH descriptor: (a) extracted 2D photometrical lines from image data and (b) QLH descriptor consisting of quantized 2D lines in Hough space.

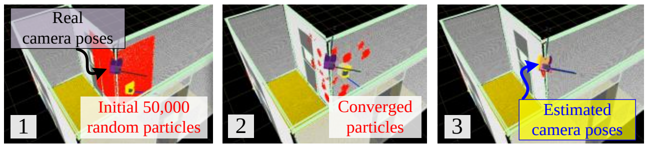


Fig. 4 Experiment result for particle-filter based camera parameter estimation.

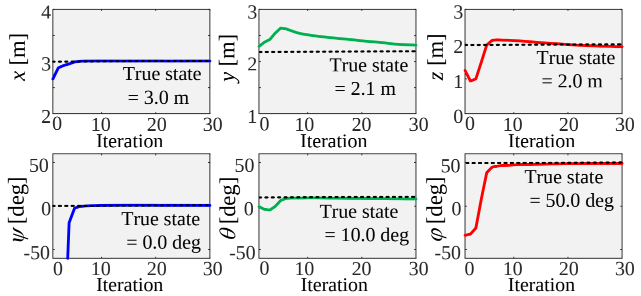


Fig. 5 Convergence process of camera state variables.

various line features located on the sides of the walls, doors, and windows. As shown in Fig. 4, particles are initialized based on human observation and the prior information (i.e., constraints from the map information). The QLH descriptor-based similarities between the 2D photometrical line segments from the real camera image and those from all particles were computed. The results were then used in the probability update of each particle.

The several stages of particle filter iterations and convergence process for the camera parameters are illustrated in Fig. 4 and Fig. 5, respectively. The simulation result shows that the complete 6DOF external parameters are estimated very accurately.

5 Conclusion

In this paper, an automatic calibration scheme was developed for the camera sensor network system. In order to realize complete 6DOF external parameter estimation, line segment-based QLH descriptor was designed for particle filter-based matching process and 3D map information was used for prior information.

The proposed approach was demonstrated in a simulation of a virtual environment with a distributed camera system. We showed that our system is able to estimate 6DOF camera pose accurately.

Acknowledgement

This work was in part supported by Tough Robotics Challenge, ImPACT Program (Impulsing Paradigm Change through Disruptive Technologies Program).

References

[1] Huiying Chen, Kohsei Matsumoto, Jun Ota, and Tamio Arai, "Self-calibration of Environmental Camera for Mobile Robot Navigation", *Robotics and Autonomous Systems*, Vol. 55, No. 3, pp. 177–190, 2007.

[2] Yonghoon Ji, Atsushi Yamashita, and Hajime Asama, "Automatic Calibration and Trajectory Reconstruction of Mobile Robot in Camera Sensor Network", *Proceedings of the 11th Annual IEEE International Conference on Automation Science and Engineering (CASE2015)*, pp. 206–211, 2015.