Efficient Closed-Loop Multiple-View Registration

Xiaowei Shao, *Member, IEEE*, Yun Shi, Huijing Zhao, *Member, IEEE*,
Xuelong Li, *Fellow, IEEE*, and Ryosuke Shibasaki, *Member, IEEE*

***Abstract*—Registering multiple views is an essential and challenging problem for many intelligent transportation applications that employ a mobile sensing platform or consist of multiple stationary sensors. In this paper a novel algorithm is presented for multiple-view registration under a loop closure constraint. Different from most existing methods, which use general optimization techniques, our method studies the mechanism of adjusting the poses of views in a loop and provides a highly efficient and accurate solution. We prove that translation vectors can be decoupled if the same point set is used in each view to associate the previous and subsequent views, leading to our solution for such decouplable cases. If this condition does not hold, an exact solution of translation vectors is provided when rotation parameters are given, which results in our iterative solution for general cases by updating rotation and translation alternately. In our method, the effect of the accumulated pose error in a loop can be distributed to all views efficiently through loop factors, and only a few iterations are needed. Most important of all, in each iteration our method has linear computational complexity with respect to the number of views, which is much superior to that of state-of-the-art methods. A series of experiments was conducted, involving simulation of thousands of views and real vehicle-borne sensing data that include 65 371 point pairs in 352 views. Experimental results show that our proposed method is not only stable and highly efficient but also provides competitive accuracy relative to existing methods.**

***Index Terms*—Loop closure, matrix exponentiated gradient (MEG), multiple-view registration.**

Manuscript received July 17, 2013; revised December 1, 2013 and February 12, 2014; accepted April 12, 2014. This work was supported in part by the GRENE Program (Green Network of Excellence, 2011–2016) funded by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) in Japan, by the National Natural Science Foundation of China under Grant 61125106, and by the Key Laboratory of Agri-informatics, Ministry of Agriculture in China under Project 2012004. The Associate Editor for this paper was A. Amditis

X. Shao is with the Earth Observation Data Integration and Fusion Research Initiative, The University of Tokyo, Tokyo 113-8656, Japan, and also with the Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China (e-mail: shaoxw@iis.u-tokyo.ac.jp).

Y. Shi is with the Key Laboratory of Agri-informatics, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China, and also with the Center for Spatial Information Science, The University of Tokyo, Tokyo 113-0033, Japan (e-mail: shiyun@caas.cn).

H. Zhao is with the State Key Laboratory of Machine Perception, Peking University, Beijing 100871, China.

X. Li is with the Center for Optical Imagery Analysis and Learning (OPTIMAL), State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China.

R. Shibasaki is with the Center for Spatial Information Science, The University of Tokyo, Tokyo 113-0033, Japan.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2014.2319352

I. Introduction

In recent years, the rapid development of range sensors and stereovision techniques has greatly facilitated the acquisition of real-world 3-D object data. Compared with conventional 2-D vision systems, such data provide detailed and accurate spatial information of measured objects and this, in turn, has assisted the development of a wide spectrum of intelligent transportation systems and computer vision applications, ranging from intelligent vehicles and robots [1], [2], crowd surveillance and tracking [3], road safety [4], and 3-D environment modeling [5], to simultaneous localization and mapping (SLAM) systems [6]–[8].

In the past decade, there has been extensive research focusing on a single stationary sensor, and many mature systems have been developed. However, such systems have limited spatial coverage and can only collect data from one viewpoint. In real systems such as [9]–[11], it is necessary to deploy multiple sensors at different places or mount sensors on a mobile platform to increase the visibility of objects of interest, such as pedestrians near an intersection or inside a railway station. Therefore, it is a common problem to register point sets captured from multiple viewpoints. Traditional approaches for solving the two-view registration problem are the iterative closest point (ICP) [12] method and its variations [13], [14], which are popular owing to their simplicity and good performance. In addition, a number of methods have been proposed to address this issue from different aspects such as feature extraction, matching criteria, and numerical solutions. The reader may refer to [15], [16] for a detailed review and [17] for the open-source library of ICP variants and the performance evaluation on real-world 3-D point data sets.

In general, the problem of two-view registration [12]–[16] can be regarded as a nonlinear parameter estimation problem, by minimizing the cost function under some conditions. This framework can be directly extended to the case of multiple views, by defining the cost function as a sum of the matching error of associated point pairs between each of the two related views [18], [19]. An alternative way of solving this problem is to introduce additional estimates to the position of feature points [20]–[22], assuming that the ground truth of these feature points satisfies the pose constraint for the related views.

When only two views are involved, the explicit solution to this registration problem is available if the relation of associated point pairs is given. However, when addressing multiple views, in general, it is unlikely to give an explicit solution. Essentially, it is a problem of optimizing a complex cost function over large parameter space. Therefore, various numerical methods have been developed to solve this problem. Most existing methods such as [14], [18], and [19] have satisfactory performance