ISPRS Journal of Photogrammetry and Remote Sensing 105 (2015) 1–12

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Particle filtering methods for georeferencing panoramic image sequence in complex urban scenes

Shunping Ji [a], Yun Shi [b,c,*], Jie Shan [d], Xiaowei Shao [e,b], Zhongchao Shi [f], Xiuxiao Yuan [a], Peng Yang [b], Wenbin Wu [b], Huajun Tang [b], Ryosuke Shibasaki [c]

[a] School of Remote Sensing and Information Engineering, Wuhan University, 129 Luoyu Road, Wuhan 430079, China
[b] Key Laboratory of Agri-informatics, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, 12 Southern Street of Zhongguanchun, Haidian, Beijing 10008, China
[c] Center for Spatial Information Science, University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8568, Japan
[d] Lyles School of Civil Engineering, Purdue University, West Lafayette, IN 47907, USA
[e] Earth Observation Data Integration and Fusion Research Initiative, University of Tokyo, Tokyo 153-8505, Japan
[f] Department of Restoration Ecology and Built Environment, Faculty of Environmental Studies, Tokyo City University, 3-3-1 Ushikubo-nishi, Tuzuki-ku, Yokohama, Kanagawa 224-8551, Japan

ARTICLE INFO

Article history:
Received 23 April 2014
Received in revised form 10 March 2015
Accepted 14 March 2015
Available online 13 April 2015

Keywords:
Mobile mapping
Particle filtering
Monte Carlo
Kalman filtering
Image matching
Gross errors

ABSTRACT

Georeferencing image sequences is critical for mobile mapping systems. Traditional methods such as bundle adjustment need adequate and well-distributed ground control points (GCP) when accurate GPS data are not available in complex urban scenes. For applications of large areas, automatic extraction of GCPs by matching vehicle-born image sequences with geo-referenced ortho-images will be a better choice than intensive GCP collection with field surveying. However, such image matching generated GCP's are highly noisy, especially in complex urban street environments due to shadows, occlusions and moving objects in the ortho images. This study presents a probabilistic solution that integrates matching and localization under one framework. First, a probabilistic and global localization model is formulated based on the Bayes' rules and Markov chain. Unlike many conventional methods, our model can accommodate non-Gaussian observation. In the next step, a particle filtering method is applied to determine this model under highly noisy GCP's. Owing to the multiple hypotheses tracking represented by diverse particles, the method can balance the strength of geometric and radiometric constraints, i.e., drifted motion models and noisy GCP's, and guarantee an approximately optimal trajectory. Carried out tests are with thousands of mobile panoramic images and aerial ortho-images. Comparing with the conventional extended Kalman filtering and a global registration method, the proposed approach can succeed even under more than 80% gross errors in GCP's and reach a good accuracy equivalent to the traditional bundle adjustment with dense and precise control.

1. Introduction

Image or sensor georeferencing is essential for a vehicle-borne mobile mapping system. There are different solutions. GPS is the most common device that provides direct georeferencing; however, it often suffers from canyon effect in urban streets (Cui and Ge, 2003). Direct georeferencing with an integrated GPS/IMU system can provide better localization than GPS alone at a higher cost. The conventional solution needs to survey a number of ground control points (GCPs) that are then used in a global optimization process, such as bundle adjustment (BA). However, an extensive GCP distribution is difficult or costly in some situations. Long traverse in urban streets requires many GCPs for satisfactory localization because the baselines of vehicle-mounted images are very short. The third solution is automatic image-to-image matching (Kümmerle et al., 2010), which globally georeferences vehicle-borne image series to existing georeferenced

* Corresponding author at: Key Laboratory of Agri-informatics, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, 12 Southern Street of Zhongguanchun, Haidian, Beijing 10008, China.

E-mail addresses: jishunping@whu.edu.cn (S. Ji), shiyun@caas.cn (Y. Shi), jshan@ecn.purdue.edu (J. Shan), shaoxw@iis.u-tokyo.ac.jp (X. Shao), shizc@tcu.ac.jp (Z. Shi), yuanxx@whu.edu.cn (X. Yuan), yangpeng@caas.cn (P. Yang), wuwenbin@caas.cn (W. Wu), tanghuajun@caas.cn (H. Tang), shiba@csis.u-tokyo.ac.jp (R. Shibasaki).

http://dx.doi.org/10.1016/j.isprsjprs.2015.03.005
0924-2716/

imagery. Similar ideas include image-to-map (Alonso et al., 2012; Najjar and Bonnifait, 2005) and image-to-landmark, both of which detect correspondences automatically by computer perception. This paper focuses on the image-to-image approach for geo-referencing of mobile sensors. Specially, we study the geo-localization of monocular panoramic image sequences of complex street environments with reference to given aerial ortho-images.

The success of an image-to-image approach relies on acquisition of reliable control information. Unlike measurements in field surveying, GCPs or GCP patches acquired from matching image sequences to ortho-images may consist of many gross errors, especially for complex and dynamic scenes, which in turn would fail the bundle adjustment. Fig. 1 illustrates a panoramic image (a) taken by a vehicle-mounted camera and the difficulty of matching its ortho-rectified patches (b–d) with an aerial ortho-image (e) in an urban street scene. Shadows, moving cars, zebra crossing changes, and rectification errors all lead to an imperfect matching. Following problems have been noticed: (1) the false match between (b) and (e) introduces a gross error; (2) three-peak matches between (c) and (e) are all wrong; (3) match candidate between (d) and (e) is missing. As a result, the corresponding GCPs found through image matching do not have a consistent quality and their errors may not follow a Gaussian distribution with a small variance; the existence of blunders would deteriorate the optimal properties of bundle adjustment and even make it fail.

There existed alternative solutions to bundle adjustment. Dynamic or autonomous localization, which commonly addresses the current pose, often uses filtering methods such as extended Kalman filtering (EKF, Barrios and Motai, 2011; Solà et al., 2011) or particle filtering (PF, Montemerlo, 2003; Sim et al., 2005; Thrun et al., 2001). Both EKF and PF also could provide global optimization as bundle adjustment does, however, the latter was more accurate than filtering methods if good initials and precise control were available (Strasdat et al., 2012). Another optimization strategy was global registration that matches the whole trajectories extracted from mobile sensors to the corresponding trajectories extracted from geo-referenced photographs (Jaud et al., 2013). GPS observations and visual odometry were integrated first and then combined with map-matching observations to guarantee accurate localization under an EKF framework (Alonso et al., 2012). Kümmerle et al. (2010) introduced aerial images as constraints to the graph-based simultaneous localization and mapping (SLAM) method to achieve global consistency. In this approach, matching between aerial images and ground laser points was under favorable conditions and could supply good geo-referenced observations for a least squares solution. In (Jaud et al., 2013), a Fourier–Mellin registration method based on the global similarity between a radar map and an ortho-photograph was used for radar map geo-referencing and achieved good accuracy. However, its performance depended heavily on a preprocessing step that extracted reliable road features in both the radar map and the ortho-photograph. In (Marks et al., 2009), visual odometry of stereo camera (as predictions) and height information in grid cells (as updates) were utilized for outdoor SLAM, which was further aligned to georeferenced coordinates based on a noisy GPS signal. In this study, the error of GPS observations was assumed proportional to the known standard deviation of the GPS reading errors, simplifying the transform between the map and GPS to be a translation and rotation. Brubaker et al. (2013) proposed a probabilistic model as well as an efficient inference to cope with uncertainty due to noisy visual odometry and ambiguities in road maps. However, the inference was based on known road structure and the ambiguities were not specifically analyzed. (Najjar and Bonnifait, 2005) addressed more explicitly the non-linearity and multi-hypotheses of road map observations. They selected only the most likely one from the candidate set for an EKF solution because multi-hypothesis implementation in EKF was very complicated. In spite of the above related studies the presence of a large percentage of gross errors was not sufficiently addressed.

The purpose of this paper is to seek effective mobile localization solutions under the presence of significant image matching errors, both in terms of their magnitude and their percentage with respect to the total number of observations. As one implementation of the Monte Carlo methods, particle filtering uses discrete samples (particles) instead of continuous posterior distributions when dealing with the sequential localization problem. Theoretically, there may be no need for a separate process to detect a large number of gross errors, since PF itself simulates the arbitrary probability distributions of all observations, including gross errors. This is exactly the property we pursue that may satisfy a challenging localization problem under high noises. PF methods have been widely used for dynamic navigation (Thrun et al., 2001) and global localization

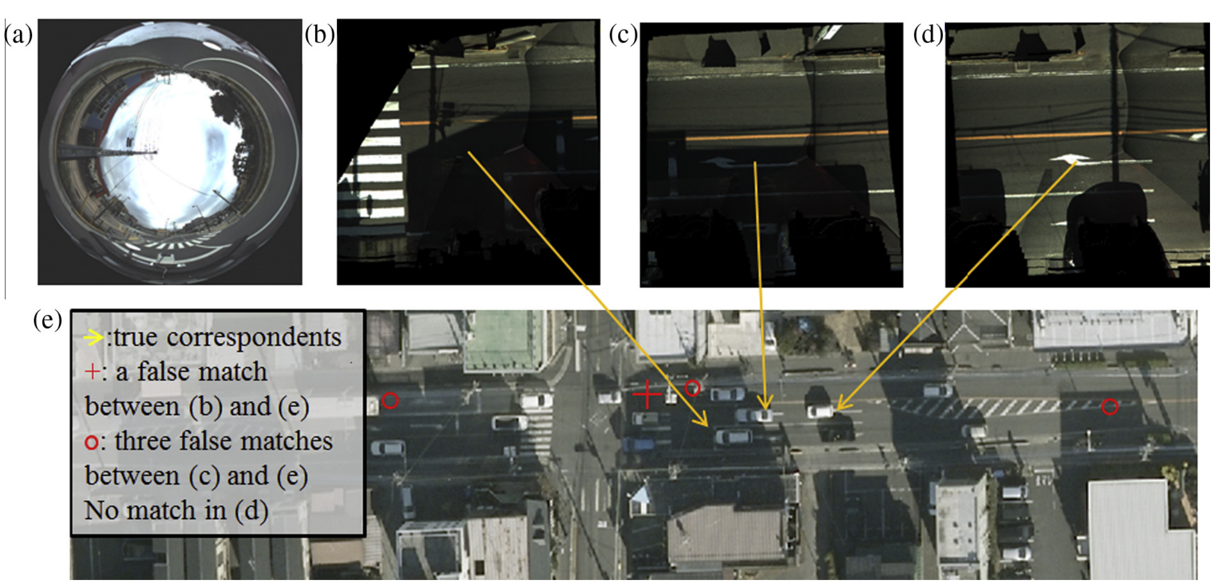

Fig. 1. Scenarios of matches between panoramic images and an aerial ortho-image. (a) Original panoramic image; (b–d) ortho-rectified and mosaicked panoramic image patches; (e) geo-referenced aerial ortho-image.