Robotics and Autonomous Systems 79 (2016) 132–146

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

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

Simulation study of a constant time hybrid approach for large scale terrain mapping using satellite stereo imagery

A. Sarkar[a,b,*], R. Reiger[c], D. Chatterjee[d], S. Patranabis[e], H. Singh[e], P. Mukherjee[e]

[a] Department of Mathematics, IIT Kharagpur, India
[b] Techno India University, Kolkata, India
[c] Airbus Group HQ HWD1, Germany
[d] ISI, Kolkata, India
[e] Department of Computer Science, IIT Kharagpur, India

HIGHLIGHTS

- This paper proposes an innovative use of particles for pose estimation in the SLAM problem.
- The proposed approach achieves constant time complexity despite the use of EKF filters.
- The performance of the system is independent of the deployment environment.
- The paper provides a unique approach to combine strengths of multiple SLAM techniques.

ARTICLE INFO

Article history:
Received 20 August 2014
Received in revised form
30 December 2015
Accepted 15 January 2016
Available online 8 February 2016

Keywords:
Bundle adjustment
SURF
Topographic labelling
Particle filters
DEM mapping

ABSTRACT

This work addresses the problem of performing large scale SLAM (Simultaneous Localization And Mapping) with satellite stereo imagery for terrain mapping, using a constant time estimation approach. The approach adopts the relative bundle adjustment approach (RBA) and integrates with it a particle-based framework to obtain a constant time probabilistic pose estimation model. The approach further uses a concept of fuzzy landmark-based similarity between poses to make common landmark identification across poses easier, especially when landmarks are sparsely encountered. In order to achieve robustness under varying environmental conditions, we use Speeded Up Robust Features (SURF) for computing spatial and temporal landmark correspondences across time steps. Finally, we use a fast loop closure approach to reduce drifts and obtain global pose estimates. For simulation study, the robot images are cropped from stereo-pair satellite images at different time steps incorporating errors in the robot's control information. Extensive experimentation has been carried out to study the robot trajectories and the determination of Digital Elevation Model (DEM), with encouraging findings. We have also compared our work with 6D FastSLAM 2.0 (Thrun et al. (2005)) as well as Relative SLAM (RSLAM) due to Mei et al. (2010).

1. Introduction

Recent work in the field of obtaining real time solutions to estimating robot's poses rely on approaches that scale with the environment without using any probabilistic framework, such as RSLAM. On the other hand, approaches like FastSLAM that are based purely on particle filters and other similar approaches using EKF filters involve parameters that grow with time. However, the filter based approach gives us a distribution of each pose rather than a single point in the global environment. In this work, we attempt to merge the advantages of both approaches by combining the particle based framework with a continuous relative pose representation. The proposed approach involves factorization of the SLAM posterior over the robot's path, in which each individual particle follows a constant time stereo SLAM approach and the particle distribution is harnessed by the algorithm to estimate the optimal trajectory. The result is a bundle distribution instead of a single bundle, which is then adjusted to estimate the current pose. For accurate bundle adjustment, especially when landmarks are sparsely encountered, fuzzy pose similarity technique is leveraged. We further ensure that the robot works in external environment condition using an appropriate loop closure mechanism.

* Corresponding author at: Department of Mathematics, IIT Kharagpur, India.
E-mail address: anjan@maths.iitkgp.ernet.in (A. Sarkar).

http://dx.doi.org/10.1016/j.robot.2016.01.004
0921-8890/

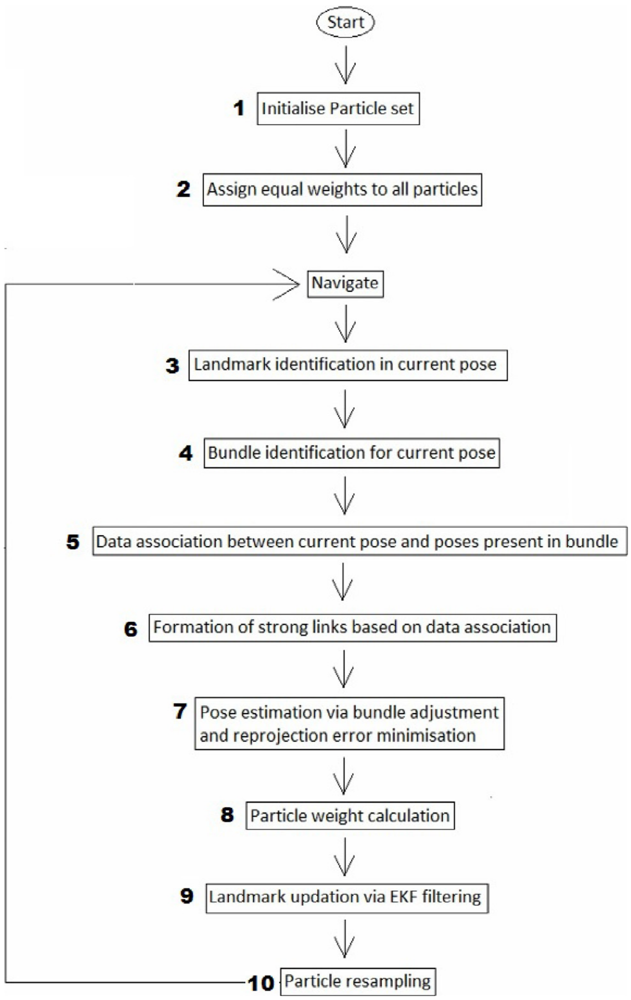

Fig. 1. Flow diagram of the different stages.

To avoid problems with scale propagation, we use a stereo pair of images at each time step which are used to extract stereo features via SURF. This also avoids extensive computations for feature depth estimation that a monocular approach would otherwise demand. The contribution of the paper lies in the innovative use of particles for constant time pose estimation with robust data association phase across frames. Unlike other EKF filter based approaches, where the time complexity worsens due to growing parameters at each time step, the proposed work achieves constant time complexity for each time step. Further the complexity of the current approach is independent of environment size making it suitable for on line SLAM applications for extremely large environments. Fig. 1 describes the basic work flow of the methodology.

2. Related work

In this section, we review some of the major advances in visual SLAM that use stereo systems. Surveys on SLAM approaches are available in [1], [2, part (i)] and [3, part (ii)]. Fast SLAM 2.0 comprises Rao–Blackwellized particle filters for pose estimation and Extended Kalman Filter (EKF) for landmark estimation. The method has a major limitation — the number of particles necessary for a specific environment is difficult to estimate and affects the computational complexity. Zhu et al. [4] among others has attempted to solve this problem by proposing a method to estimate the required number of particles according to the uncertainty of sensors with some success.

However, these methods are still prone to inconsistent estimates, leading to the recent trend of local [5] or global [6–9] bundle adjustment as the underlying estimator. More recent approaches intend to provide real time solutions using stereo pair of images. This includes the work of Nister et al. [10] that harnesses local bundle adjustment using global pose estimates. However the work does not address the problem of loop closure which is essential for drift reduction. The work of Jeong et al. [11] may be noted for many facet of bundle adjustment. Special mention must be made of the FrameSLAM system [12] and constant time RSLAM framework [13] that use bundle adjustment for constant time pose estimation. While the former aims for reduction of large-scale solution complexity, the latter focuses on locally accurate map and trajectory using relative bundle adjustment.

The innovation in our work is that it combines the relative bundle adjustment technique of RSLAM with the concept of Rao–Blackwellized particle filters. We also leverage on the concept of fuzzy pose similarity for bundle adjustment which is particularly useful when landmarks are sparsely encountered.

3. Problem formulation

3.1. Simulation environment

We simulate the robot's path and test our methodology on stereo-pair satellite image environment to demonstrate that such an UAV with stereo-pair sensor aboard can be planned in practice. Two large stereo-pair satellite images of size 12000×12000 (covering area of 30 KM by 30 KM), are considered from Panchromatic cameras AFT (after ward) and FORE (forward), which are at angles -5° and $+26^\circ$, respectively, along the track with respect to nadir and are very helpful in deriving the depth information of the terrain imaged [14]. A series of sub-scenes corresponding to the robot's sensors' intake at different time steps is cropped from these images as if the robot is in motion and is taking such images at different time steps. The robots' path is also simulated by incorporating errors in the input control information.

Initially all the particles are at the same position as in stage 1 of Fig. 1. But as they move through odometry, they encounter different landmark distribution, and thus they entail different hypothesis. Each particle carries equal weight as described in stage 2 of Fig. 1. A continuous relative representation (CRR) is used to represent the environment [13]. The advantage of this approach is that it incorporates standard bundle adjustment and does not require complex map-merging for loop closure. However, it aims mostly to provide stable motion estimates without laying too much stress on the mapping aspect. In our work, we maintain multiple copies of the same map and at each step choose the best map to represent the environment.

3.2. Representation of the environment

In our representation, the set of poses are treated as nodes of a graph. Each new pose is treated as a newly added vertex in the already existing graph. The new pose is connected to other existing poses by two kinds of links:

1. Each adjacent pair of poses is connected by an odometric link (or a weak link) which represents the required transformation to travel from one pose to another.
2. Each pair of poses that have matching features will have another additional link called a matching link (or a strong link). This link essentially represents a set of constraints imposed on the relative location of the poses, manifested in the form of common stereo features between them.