Available online at www.sciencedirect.com

ScienceDirect

CAAI Transactions on Intelligence Technology 1 (2016) 90−103
http://www.journals.elsevier.com/caai-transactions-on-intelligence-technology/

Original article

A framework for multi-session RGBD SLAM in low dynamic workspace environment

Yue Wang [a], Shoudong Huang [b], Rong Xiong [a,*], Jun Wu [a]

[a] *State Key Laboratory of Industrial Control and Technology, Zhejiang University, Hangzhou, PR China*
[b] *Centre for Autonomous Systems, Faculty of Engineering and IT, University of Technology, Sydney, Australia*

Available online 4 June 2016

Abstract

Mapping in the dynamic environment is an important task for autonomous mobile robots due to the unavoidable changes in the workspace. In this paper, we propose a framework for RGBD SLAM in low dynamic environment, which can maintain a map keeping track of the latest environment. The main model describing the environment is a multi-session pose graph, which evolves over the multiple visits of the robot. The poses in the graph will be pruned when the 3D point scans corresponding to those poses are out of date. When the robot explores the new areas, its poses will be added to the graph. Thus the scans kept in the current graph will always give a map of the latest environment. The changes of the environment are detected by out-of-dated scans identification module through analyzing scans collected at different sessions. Besides, a redundant scans identification module is employed to further reduce the poses with redundant scans in order to keep the total number of poses in the graph with respect to the size of environment. In the experiments, the framework is first tuned and tested on data acquired by a Kinect from laboratory environment. Then the framework is applied to external dataset acquired by a Kinect II from a workspace of an industrial robot in another country, which is blind to the development phase, for further validation of the performance. After this two-step evaluation, the proposed framework is considered to be able to manage the map in date in dynamic or static environment with a noncumulative complexity and acceptable error level.

Keywords: Multi-session SLAM; RGBD sensor; Low dynamic mapping

1. Introduction

Simultaneous localization and mapping (SLAM) has been a core technique enabling the autonomy of robots, such as robot car [1] and autonomous underwater vehicle [2]. Compared to these high cost and large devices, the small to medium sized devices, such as mobile manipulator, flying robot and hand-hold devices began to raise the attention in recent years due to the high flexibility, low cost and thus the highly promising application. These devices also call for SLAM to achieve the capacity of long-term operation. The main challenge for such a solution includes three aspects: (1) the flying robot and hand-hold devices have a motion pattern with more frequent changes of orientation due to the free environment (no ground plane); (2) these devices aim on low cost, light-weighted and small scale, hence expensive or heavy sensors cannot be equipped; (3) these devices usually work periodically in a pre-defined human sharable workspace with low dynamics, which means objects in the workspace may be moved, added or removed across multiple sessions.

Consumer-level RGBD sensor has made it very convenient to collect both intensity and depth information at a low cost. For the first two challenges, we apply the RGBD sensor for perception, enabling the 3D pose estimation but also a dense environment map for subsequent navigation. The third challenge is to deal with the change of objects (move, add, remove) across multiple sessions. A quick solution is to build

* Corresponding author. State Key Laboratory of Industrial Control and Technology, Zhejiang University, 38 Zheda Road, Xihu District, Hangzhou, 310027, PR China.
E-mail address: rxiong@iipc.zju.edu.cn (R. Xiong).
Peer review under responsibility of Chongqing University of Technology

http://dx.doi.org/10.1016/j.trit.2016.03.009

the map at each session, but this method discards all history experience. Our solution is to manage the dynamics in a map. Specifically, the multi-session SLAM component was utilized to accumulate the map building. On the top of that, a map management component was proposed to keep the map compact and in track of the environment changing. With this framework, we are able to address all three challenges.

In the previous studies, various SLAM methods have been presented for mapping the environment with this kind of sensors. Existing RGBD mapping methods were mainly on single session and for relatively static environment or with high dynamics [3–5]. However, the low dynamics emerging in multi-session scenario did not draw much attention. Some methods [6–8] were proposed to deal with the challenge. They used vision or planar laser sensor, which captured limited dynamics and cannot be simply extended to that using RGBD sensor. The methods using vision sensor can tell whether a frame has a significant change in appearance as it is feature based. Since the RGBD sensor also provides depth information, we can capture the geometric change and know exactly what is changed in a frame. The methods using laser sensor usually took a 2D grid occupancy map as its map representation, which is not available in RGBD SLAM system due to the high complexity of 3D grid. Besides, the dynamics captured in 2D is only a slice of the 3D dynamics, which can be semantically insufficient.

To the best of our knowledge, our system may be the first one that build the map over the low dynamic environment using only a RGBD sensor in a scenario of 6 DoF multi-session SLAM. We proposed a framework that can build the map keeping track of the current environment, preventing the change of environment in previous sessions incorporated. Fig. 1 gives the comparison between the final map generated by multi-session SLAM system with and without considering low dynamics in a workspace in office environment. The objects (books, cans, boxes and so on) are added, removed and moved across the sessions. After 10 sessions of SLAM, the SLAM without considering the low dynamics mixed the current and out-of-dated information together, leading to a useless map with incorrect duplicated objects, while the proposed system considering the low dynamics, demonstrated the current environment in the map.

The main contributions of this paper include:

- A framework is proposed for multi-session RGBD SLAM in low dynamic environment consists of two components: multi-session SLAM and graph management. The multi-session SLAM component has a graph model with each node being a pose and each edge being a constraint, thus fusing the information from previous sessions and current session to keep the map in one global coordinates. The graph management component can keep the graph model in date and with non-accumulative complexity using the out-of-dated scan identification module and redundant scan identification module.
- An out-of-dated scan identification module is proposed to find the previous pose with RGBD scan on the environment which is changed in current session. The goal of this module can be explained by setting an example, a cup was on the desk in previous sessions, but is removed in the current session. Then, the poses observing that cup on the desk should be found and pruned to keep the map in track of the environment changes. Because the unavailability of grid occupancy model, our idea is to adopt camera projection model and connected component detection to find the difference between the maps generated by the scans in the previous sessions and that in the current session. With this method, the poses reserved are always with in-dated scans and robust to noise and holes occurred in RGBD sensor.
- A redundant scan identification module is proposed to find the pose with RGBD scan having large overlapping part with others. This module is to reduce the number of poses if the number of in-dated scan is higher than a pre-defined threshold, which enables the computational time of the SLAM relevant to the size of the map and one session SLAM, instead of the all sessions SLAM. The idea of our method is to find a subset of poses that can generate a map similar to the original one using all poses, in the measure of Kullback–Leibler divergence. By applying this method, when a robot executes multi-session SLAM in a fixed sized static region in low dynamic environment, the computation complexity will keep constant since poses with redundant scans have been pruned despite that they are in date.

To show the performance of the framework, we tune and test the algorithm in a 2-session and 10-session dataset on a workspace in office environment with multiple objects moved, added and removed across the sessions, which is collected by a hand hold Kinect sensor. The result in Fig. 1 validates the effectiveness of the proposed method. After that, the framework was applied on a 5-session external dataset captured in workspace of an industrial robot with various sized boxes manipulated across sessions, which is blind to the development phase, for evaluation of real performance. The remainder of the paper is organized as follows: In Section 2, related works on mapping dynamic environments and pose pruning will be discussed. In Section 3, the proposed framework for multi-session RGBD SLAM in low dynamic environment is introduced. In Sections 4 and 5, the proposed out-of-dated scans identification and redundant scans identification will be presented in detail. In Section 6, we will demonstrate the experimental results using the real world datasets. The conclusion and future work will be discussed in Section 7.

2. Related works

The multi-session SLAM in static environment is first presented in vision based methods [1,6,9]. These works formulate the basic concept that the robot cannot simply start a new mapping session without using the information in previous sessions, since the constraints in past sessions provide