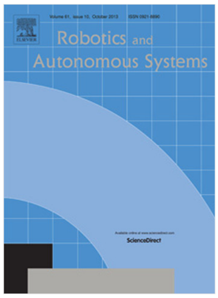
Accepted Manuscript



Motion removal for reliable RGB-D SLAM in dynamic environments

Yuxiang Sun, Ming Liu, Max Q.-H. Meng

PII: S0921-8890(18)30069-1
DOI: https://doi.org/10.1016/j.robot.2018.07.002
Reference: ROBOT 3056

To appear in: *Robotics and Autonomous Systems*

Received date : 28 January 2018
Revised date : 27 June 2018
Accepted date : 8 July 2018

Please cite this article as: Y. Sun, M. Liu, M.Q.-H. Meng, Motion removal for reliable RGB-D SLAM in dynamic environments, *Robotics and Autonomous Systems* (2018), https://doi.org/10.1016/j.robot.2018.07.002

Motion Removal for Reliable RGB-D SLAM in Dynamic Environments

Yuxiang Sun[a], Ming Liu[a], Max Q.-H. Meng[b,*]

[a]*Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China.*
[b]*Department of Electronic Engineering, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong, China.*

Abstract

RGB-D data-based Simultaneous Localization and Mapping (RGB-D SLAM) aims to concurrently estimate robot poses and reconstruct traversed environments using RGB-D sensors. Many effective and impressive RGB-D SLAM algorithms have been proposed over the past years. However, virtually all the RGB-D SLAM systems developed so far rely on the static-world assumption. This is because the SLAM performance is prone to be degraded by the moving objects in dynamic environments. In this paper, we propose a novel RGB-D data-based motion removal approach to address this problem. The approach is on-line and does not require prior-known moving-object information, such as semantics or visual appearances. We integrate the approach into the front end of an RGB-D SLAM system. It acts as a pre-processing stage to filter out data that are associated with moving objects. Experimental results demonstrate that our approach is able to improve RGB-D SLAM in various challenging scenarios.

Keywords: Motion Removal, Codebook Model, Dynamic Environments, RGB-D SLAM

1. Introduction

Simultaneous Localization and Mapping (SLAM) is a fundamental step for many robotic applications. It concurrently estimates robot poses and reconstructs traversed environment models. Many effective SLAM algorithms using visual sensors, such as monocular cameras [1], stereo cameras [2] and RGB-D cameras [3], have been proposed over the past years. Related technologies, such as the augmented reality [4] and autonomous driving [5], have benefited from the development of the SLAM technology. It is worth noting that the advent of the RGB-D cameras has changed the computer vision world [6]. They provide coloured point clouds with real-scale distance information, which greatly benefits the dense 3-D environment reconstruction. Many impressive RGB-D SLAM systems have been developed [7, 8, 9, 10, 11, 12] in recent years, and most of them adopt the graph optimization framework. We refer readers to this survey [13] to get an overview of the progress on the graph SLAM algorithm.

However, virtually all the current RGB-D SLAM algorithms are proposed under the *static-world assumption*. It requires that there is no moving object existing in the environment during the traversal of robots. The data associations in the SLAM front end can be hindered by moving objects. With incorrect data associations fed into the SLAM back end, the graph optimization process could be severely jeopardized, which finally leads to a catastrophic failure for the localization and mapping processes. Thus, the technology of RGB-D SLAM is still vulnerable in dynamic environments.

The data associations in the SLAM front end consists of two components, namely, the short-term data association and the long-term data association [13]. The short-term data association determines adjacent pose estimations, while the long-term one has an impact on the loop detection. Take as an example the sparse feature-based RGB-D SLAM. Standard robust estimators, such as the RANdom SAmple Consensus (RANSAC) algorithm [14], are usually employed in the SLAM front end to reject outlier feature associations. However, it is hard to reliably reject outliers when moving objects are not trivial in the camera field of view. In such a case, the outliers are unavoidably used for computing the robot poses, which makes the pose estimation erroneous. When a robot returns to a previously visited place where moving objects have gone away, the loop detection would be confused by matching the same scene but with different visual appearances. If we eliminate moving objects at the first exploration of a place, we can compare the image frames using just the static feature points, which would lead to a much more reliable loop detection result. Moreover, we can merely use the static feature points to get accurate robot pose estimations. Therefore, eliminating moving objects is able to reduce the incorrect data associations, which is critical to improve the SLAM performance.

In this paper, we develop a novel RGB-D data-based motion removal approach to address the problem of RGB-D SLAM in dynamic environments. We refer to the dense pixel-wisely moving-object segmentation as *motion removal*. Our approach serves as a pre-processing stage to filter out data that are associated with moving objects. With our approach, incorrect data associations can be greatly reduced in the SLAM front end.

Fig. 1 qualitatively compares the resulting point-cloud maps produced by the `ORB-SLAM` system [11] and the system integrated with our motion removal approach in a dynamic envi-

*Corresponding author.
Email addresses: eeyxsun@ust.hk (Yuxiang Sun), eelium@ust.hk (Ming Liu), max.meng@cuhk.edu.hk (Max Q.-H. Meng)