

Semantic Map Building Based on Object Detection for Indoor Navigation

Regular Paper

Jinfu Yang^{1,3}, Jizhao Zhang^{1*}, Guanghui Wang² and Mingai Li^{1,3}

¹ Department of Control Science and Engineering, Beijing University of Technology, Beijing, PR China

² Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence, USA

³ Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing, PR China

*Corresponding author(s) E-mail: aaltonen@126.com

Received 12 March 2015; Accepted 23 November 2015

DOI: 10.5772/62040

© 2015 Author(s). Licensee InTech. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Building a map of the environment is a prerequisite for mobile robot navigation. In this paper, we present a semantic map building method for indoor navigation of a robot using only the image sequence acquired by a monocular camera installed on the robot. First, a topological map of the environment is created, where each key frame forms a node of the map represented as visual words (VWs). The edges between two adjacent nodes are built from relative poses obtained by performing a novel pose estimation approach, called one-point RANSAC camera pose estimation (ORPE). Then, taking advantage of an improved deformable part model (iDPM) for object detection, the topological map is extended by assigning semantic attributes to the nodes. Extensive experimental evaluations demonstrate the effectiveness of the proposed monocular SLAM method.

Keywords Semantic Map, Topological Map, Object Detection, Deformable Part Model, Monocular SLAM

1. Introduction

Map building of the environment is a prerequisite for mobile robot navigation. In order to accurately plan a path

in an environment, a mobile robot needs to know where it is and how to get to a specified destination from an initial location. The process of simultaneously tracking the position of a mobile robot and building a map of the environment is known as simultaneous localization and mapping (SLAM). SLAM is one of the most important topics in the robotics community. There are two major aspects that need to be considered when employing a SLAM system. One is pose estimation based on sensor data and the other is reconstruction of the surrounding environment, such as objects, or other featured landmarks. In recent years, numerous SLAM algorithms have been performed. Davison et al. [1] proposed a real-time algorithm, called monocular SLAM, which can recover the 3D trajectory of a monocular camera. In the system, they utilized an extended Kalman filter (EKF) to update the estimation of the robot poses as well as the map of the landmarks in the environments recursively, achieving real time but drift-free performance. Clemente et al. [2] used a hierarchical map technology to create a SLAM system for wide-range environments and the approach was tested at the Oxford University campus. In the work carried out in [3], a GPU-based method was employed for feature detection, description, and matching procedures in order to attain a higher processing speed. Firstly, an 8-point algorithm was adopted to initialize the coordinate system. Then, the 3D coordinates of the feature points were