A Bayesian Approach to Learning 3D Representations of Dynamic Environments

Ralf Kästner, Nikolas Engelhard, Rudolph Triebel, and Roland Siegwart

Abstract. We propose a novel probabilistic approach to learning spatial representations of dynamic environments from 3D laser range measurements. Whilst most of the previous techniques developed in robotics address this problem by computationally expensive tracking frameworks, our method performs in real-time even in the presence of large amounts of dynamic objects. The computer vision community has provided comparable methods for learning foreground activity patterns in images. However, these methods generally do not account well for the uncertainty involved in the sensing process. In this paper, we show that the problem of detecting occurrences of non-stationary objects in range readings can be solved online under the assumption of a consistent Bayesian framework. Whilst the model underlying our framework naturally scales with the complexity and the noise characteristics of the environment, all parameters involved in the detection process obey a clean probabilistic interpretation. When applied to real-world urban settings, the results produced by our approach appear promising and may directly be applied to solve map building, localization, or robot navigation problems.

1 Introduction and Related Work

Understanding dynamic properties of the world has become an increasingly popular research topic in mobile robotics. The motivations for this popularity are manifold. The occurrence of moving objects in the robot’s sensor range may for example corrupt the localization or map building process [12, 1, 3]. On the other hand, novel
planning approaches aim at navigating platforms through highly dynamic environments \cite{11,6}. They therefore strongly rely on robust motion parameter estimates for objects that may potentially interfere with the robot’s trajectory.

A widely common group of methods addressing motion estimation is committed to tracking the displacement of entire point clusters \cite{10}. Whilst such approaches succeed in obtaining a parametric description of the cluster motion, they usually take strong assumptions about the size or shape of objects. In the above scenarios, however, we generally do not want to constrain ourselves with a limited number of object classes. In fact, we seek to detect motion rather than to find explicit motion parameters. Consider therefore a sensor reading that has been introduced by a dynamic object. If, at any later point in time, we acquire another reading that matches the previous observation, we may not care to also answer the difficult question of identity. That is, no matter if the measurement originates from a single or two different objects, we would still want to classify it as being dynamic.

In this paper, we propose a novel approach to the problem of learning 3D representations of dynamic environments from range data. In strong analogy to background modeling in computer vision \cite{8}, this problem constitutes a binary classification task. That is, for a series of range observations we seek to estimate whether single measurements originate from a static or a dynamic object. We therefore represent correspondences between measurements and objects using Gaussian mixture distributions \cite{14,13}.

Where standard methods for learning Gaussian mixtures fail due to the non-stationary nature of a dynamic world model \cite{5}, we propose an alternative on-line solution that does not make any assumptions about the number of Gaussians in the mixture model but efficiently scales with the complexity of the environment. Even in highly populated settings, our approach is thus capable of distinguishing dynamic from static objects in real-time.

The emphasis of this work strongly lies on the Bayesian formalization of all steps involved in the learning process \cite{8}. In fact, following techniques used in probabilistic change detection \cite{7} our method is strictly governed by the laws of probability, and each effective parameter comes with a clear probabilistic interpretation.

We demonstrate the practicability of our approach in simulation and by experiments involving several urban outdoor scenarios with a diversity of static structures and dynamic objects.

2 Probabilistic Formulation

Our algorithm to learning dynamic environment representations operates on range readings acquired with a nodding 2D laser range scanner that pitches up and down during data acquisition to produce 3D point clouds. This setup has been used frequently in the literature (see, e.g. \cite{13}) and is usually known as the nodding laser scanner configuration. Throughout this paper, we define a sensor measurement \( z_t \) as the tuple \((r_t, \theta_t, \phi_t)\), where \( r_t \) is the measured range, and \( \theta_t \) and \( \phi_t \) are the pitch and yaw angles of the laser beam at time \( t \) of the data acquisition. We assume that \( r_t \) is