Extraction of Dynamic Objects from 3D Point Clouds using Multimodal Detection and Tracking

Rudolph Triebel, Luciano Spinello, and Roland Siegwart

Autonomous Systems Lab, ETH Zurich, Switzerland
email: {rudolph.triebel, luciano.spinello, roland.siegwart}@mavt.ethz.ch

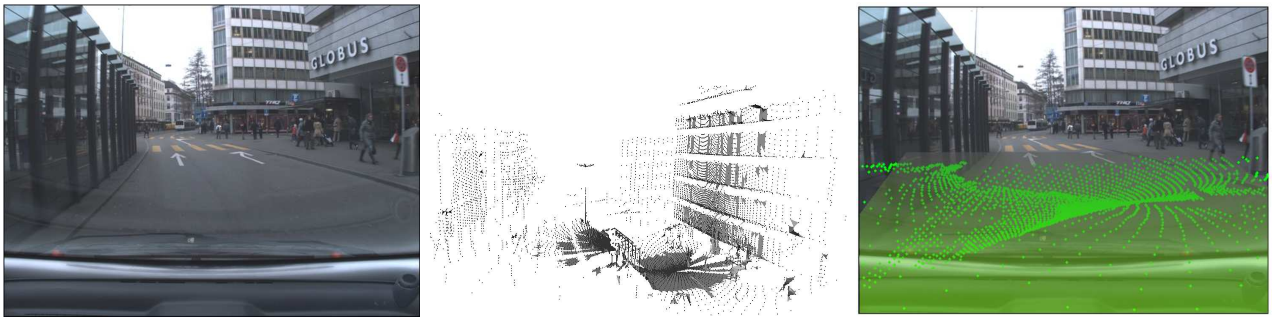

Fig. 1. 3D ground plane extraction. **Left:** Camera image as seen from the inside of the vehicle. **Middle:** Triangulated 3D point cloud of the same scene (seen from above). **Right:** Camera image with the points of the extracted ground plane overlaid.

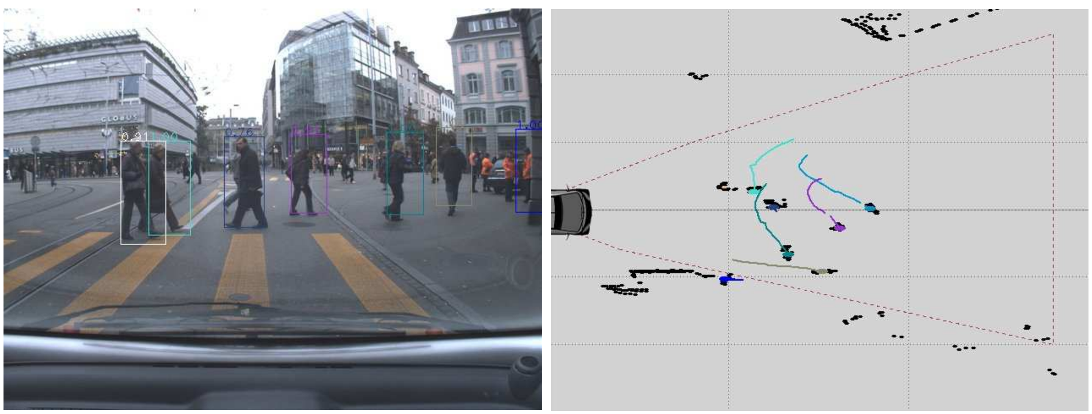

Fig. 2. Tracking results from a pedestrian crossing. The colored boxes in the image describe different tracks and probability levels; the size of the filled circle in the tracking figure is proportional to pedestrian detection confidence.

In this talk, we present a technique to detect and track dynamic objects – in particular pedestrians – in outdoor environments. The aim is to utilize the knowledge of dynamic objects when creating 3D maps. For example, the information about dynamic objects can be used to annotate 3D maps with motion vectors in areas where objects were in motion during the data acquisition process. This results in a better scene understanding where highly frequented areas such as zebra crossings can be distinguished from less ”dangerous” regions with no moving objects. Furthermore it enables to create 3D maps in challenging environments where moving objects are encountered at any time.

Our detection method is based on a combined multimodal sensor approach that utilizes 2D and 3D laser range and camera data. Laser data points are clustered and classified with a set of geometrical features using an SVM AdaBoost method. The clusters define a region of interest in the image that is adjusted using the ground plane information extracted from the 3D laser (see Fig. 1). In this areas a novel vision based people detector based on the Implicit Shape Model (ISM) is applied. Each detected person is tracked using a greedy data association technique and multiple Extended Kalman Filters, that use different motion models (see Fig. 2). This way, the filter can cope with a variety of different motion patterns. The tracker is asynchronously updated by the detections from the laser and the camera data. Experiments conducted in real-world outdoor scenarios with crowds of pedestrians demonstrate the usefulness of our approach.