

Ladar-Based Detection and Tracking of Moving Objects from a Ground Vehicle at High Speeds

Chieh-Chih Wang, Charles Thorpe and Arne Suppe

Robotics Institute
Carnegie Mellon University
Pittsburgh, PA, 15213, USA
Email: {bobwang, cet, suppe}@cs.cmu.edu

Abstract

Detection and tracking of moving objects (DATMO) in crowded urban areas from a ground vehicle at high speeds is difficult because of a wide variety of targets and uncertain pose estimation from odometry and GPS/DGPS. In this paper we present a solution of the simultaneous localization and mapping (SLAM) with DATMO problem to accomplish this task using ladar sensors and odometry. With a precise pose estimate and a surrounding map from SLAM, moving objects are detected without a priori knowledge of the targets. The interacting multiple model (IMM) estimation algorithm is used for modeling the motion of a moving object and to predict its future location. The multiple hypothesis tracking (MHT) method is applied to refine detection and data association. Experimental results demonstrate that our algorithm is reliable and robust to detect and track pedestrians and different types of moving vehicles in urban areas.

1 Introduction

Detection and tracking of moving objects (DATMO) is one of the most important and challenging problems for driving assistance and autonomous driving. Although the DATMO problem has been extensively studied for several decades [1, 2, 3, 4, 9, 10], it is still very difficult to accomplish DATMO in crowded urban environments from a ground vehicle at high speeds.

One of the most difficult issues is to separate moving objects and stationary objects. In indoor environments, the most important targets are people. If cameras are used to detect people, the appearance-based approaches are widely used and people can be detected no matter if they are moving or not. If laser scanners are used, the feature-based approaches are usually the preferred solutions [5, 6, 7]. Both appearance-based and feature-based methods rely on *a priori* knowledge of targets. In urban areas, because there are many kinds of moving objects such as pedestrians, animals, wheelchairs, bicycles, motorcycles, cars, buses, trucks, trailers, etc., it is very difficult to define features or appearances by using laser scanners.

In order to accomplish DATMO from a moving platform, a precise localization system is essential [7, 8, 14]. Unfortunately, a good inertial measurement system is very expensive and it is known that GPS and DGPS often fail in the urban areas because of the urban canyon effects. In the past decade, the simultaneous localization and mapping (SLAM) problem has received substantial interest in robotics and AI literature [11], which provides a more precise pose estimate than inertial measurement systems; and a global consistent surrounding map without *a priori* map and without access to independent position information. However, most of the published work on SLAM assumes that the environment is static.

In [12], we presented an approach to tackle the SLAM problem and the DATMO problem at once. SLAM provides more accurate pose estimation and a surrounding map, which are used to detect moving objects reliably. SLAM can be more accurate because moving objects are filtered out of the SLAM process thanks to the moving object location prediction from DATMO. SLAM and DATMO are mutually beneficial, as shown in Figure 1. In [13], we derived the Bayesian formula of the SLAM with DATMO problem, which provides a solid basis for understanding and solving this problem. Our solution of the SLAM with DATMO problem satisfies both the safety and navigation demands of the driving assistant and autonomous driving systems by using laser scanners and odometry.

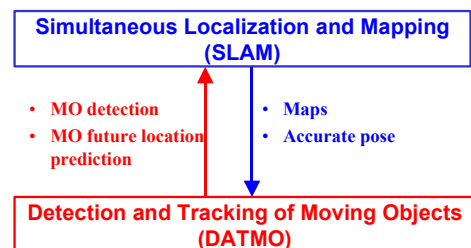


Figure 1: SLAM with DATMO

Since SLAM in urban and suburban areas was addressed in [12], in this paper the SLAM part of the whole problem is treated as a black box, which provides a

surrounding map and a better pose estimate than odometry. In order to establish the basic terminology used throughout this paper, the Bayesian formula of the SLAM with DATMO problem is briefly introduced and its meaning is explained in Section 2. Section 3 addresses our DATMO algorithms step by step. The results of experiments, carried out with the CMU Navlab11 vehicle in crowded urban areas, are shown in Section 4, and the conclusion is in Section 5.

2 SLAM with DATMO

The SLAM with DATMO problem is not only to solve the SLAM problem in dynamic environments but also to detect and track these dynamic objects. More specifically, what we discuss is how to estimate the pose of the robot, build a map, detect other moving objects and to predict their motions, given odometry and laser scanner measurements from the robot. In this section, the Bayesian formula of the SLAM with DATMO problem is introduced and its meaning is explained. In addition, local SLAM with DATMO is addressed.

2.1 Notation

We denote the discrete time index by the variable k , the vector describing an odometry measurement from time $k-1$ to time k by the variable u_k , a laser scanner measurement from the vehicle at time k by the variable z_k , the state vector describing the true location of the vehicle at time k by the variable x_k , and the stochastic map which contains l features by the variable $M = \{m^1, \dots, m^l\}$. $Y_k = \{y_k^1, \dots, y_k^n\}$ are the locations of moving objects, of which there are n moving objects that appeared inside the sensor's range at time k . In addition, we define the following set to refer data leading up to time k .

$$U_k = \{u_0, u_1, \dots, u_k\} = \{U_{k-1}, u_k\} // \text{Odometry} \quad (1)$$

$$Z_k = \{z_0, z_1, \dots, z_k\} = \{Z_{k-1}, z_k\} // \text{Ladar} \quad (2)$$

$$X_k = \{x_0, x_1, \dots, x_k\} = \{X_{k-1}, x_k\} // \text{True.Location} \quad (3)$$

where the initial location of the vehicle x_0 is assumed known.

2.2 Bayesian Formulation

Because the environment contains not only static objects but also dynamic objects, the measurements from sensors contain information from both static and dynamic objects. The general recursive formula for SLAM with DATMO can be expressed as:

$$p(x_k, Y_k, M | Z_k, U_k) \quad (4)$$

We start with the following assumptions:

- The vehicle motion model is Markov.
- The sensor measurement can be separated into moving parts and stationary parts and that they are independent:

$$z_k = z_k^s + z_k^m \quad \text{and hence} \quad Z_k = Z_k^s + Z_k^m \quad (5)$$

Here the sensor measurement belonging to stationary objects is denoted by the variable z_k^s and the sensor measurement belonging to moving objects is denoted by the variable z_k^m .

- When estimating the posterior over the map and the vehicle pose, the measurements of moving objects carry no information, neither do their location Y_k .

Then the general recursive Bayesian formula of (4) can be derived and expressed as (See [13] for the derivation of this equation):

$$\begin{aligned} p(Y_k, M, x_k | Z_k, U_k) &\propto \quad (6) \\ \frac{p(z_k^m | Y_k, x_k) p(Y_k | Z_{k-1}, U_k)}{\text{DATMO}} &\frac{p(z_k^s | M, x_k) p(M, x_k | Z_{k-1}^s, U_k)}{\text{SLAM}} \\ = \frac{p(z_k^m | Y_k, x_k)}{\text{Update}} &\frac{\int p(Y_k | Y_{k-1}) p(Y_{k-1} | Z_{k-1}, U_{k-1}) dY_{k-1}}{\text{Prediction}} \\ \frac{p(z_k^s | M, x_k)}{\text{Update}} &\frac{\int p(x_k | u_k, x_{k-1}) p(x_{k-1}, M | Z_{k-1}^s, U_{k-1}) dx_{k-1}}{\text{Prediction}} \end{aligned}$$

2.3 Solving the SLAM with DATMO Problem

From (6), input to the SLAM with DATMO filter are two separate posteriors, one of the conventional SLAM form, $p(x_{k-1}, M | Z_{k-1}^s, U_{k-1})$, and a separate one for DATMO, $p(Y_{k-1} | Z_{k-1}, U_{k-1})$. The posterior of the SLAM part is recovered by:

$$p(x_k, M | Z_k^s, U_k) = \int p(Y_k, M, x_k | Z_k, U_k) dY_k \propto \quad (7)$$

$$p(z_k^s | M, x_k) \int p(x_k | u_k, x_{k-1}) p(x_{k-1}, M | Z_{k-1}^s, U_{k-1}) dx_{k-1}$$

and the posterior of the DATMO part is computed by:

$$\begin{aligned} p(Y_k | Z_k, U_k) &= \iint p(Y_k, M, x_k | Z_k, U_k) dM dx_k \\ &\propto \int [p(z_k^m | Y_k, x_k) \int p(Y_k | Y_{k-1}) p(Y_{k-1} | Z_{k-1}, U_{k-1}) dY_{k-1}] \\ &\quad p(x_k | Z_k^s, U_k) dx_k \quad (8) \end{aligned}$$

Equation (8) shows that DATMO should take account of the uncertainty in the pose estimate of the robot because the laser scanner measurements are directly from the robot. Figure 2 illustrates the procedures for solving the SLAM with DATMO problem. Figure 2(a) shows the estimations and the corresponding distributions of the robot's pose, a detected moving object's pose (in green),

and the map at time $k - 1$. In Figure 2(b), the robot moves and gets a measurement u_k from odometry. The moving object also moves but there is no measurement directly associated with its motion. Unlike using u_k and the robot's motion model to predict the robot's pose, only previous measurements associated with this moving object are used to model the moving object's motion and predict its pose. Figure 2(c) shows that the robot gets a new measurement z_k at the new location. Here z_k contains information from both stationary objects and moving objects. In Figure 2(d), the estimations and the corresponding distributions of the robot's pose and the map at time k are updated using information only associated with stationary objects. Finally Figure 2(e) shows the pose estimate and the corresponding distribution of the moving object are updated using more precise information from SLAM.

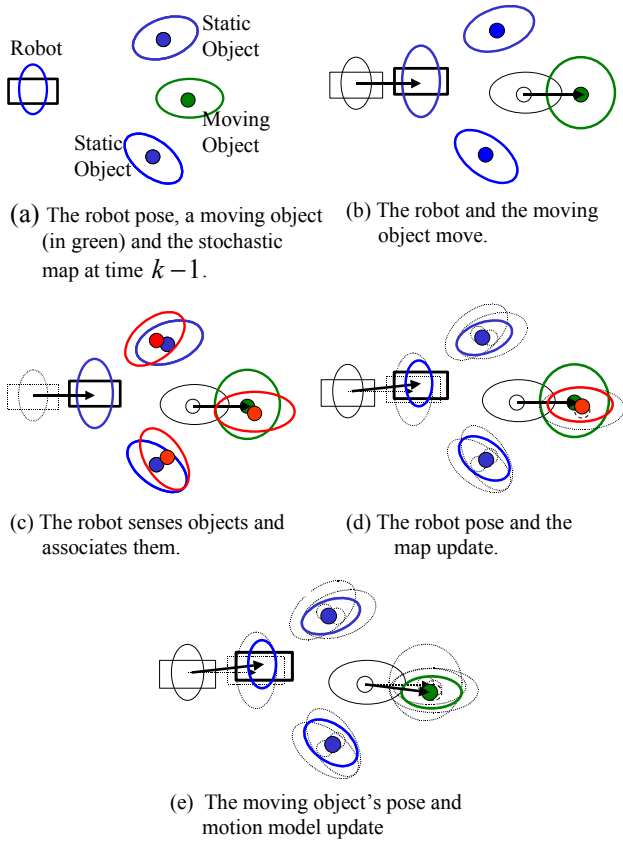


Figure 2: SLAM with DATMO procedures

Although the distributions are shown by ellipses in Figure 2, the Bayesian formula does not assume that the estimations are Gaussian distributions.

2.2 Local SLAM with DATMO

The goal of SLAM is to build a globally consistent map, so the uncertainty estimate of the pose of the robot from

SLAM is with respect to the global coordinate frame. Equation (8) shows that DATMO has to process more uncertain information because DATMO has to take into account the pose estimate uncertainty from SLAM, which makes data association and tracking of moving objects more difficult (See Figure 3). Fortunately for applications such as driving assistant systems, it is not necessary to solve DATMO globally. The absolute positions and velocities of moving objects can be estimated with respect to the temporary global coordinate frame instead of the true global coordinate frame. Figure 3 and Figure 4 illustrate the differences between global SLAM with DATMO and local SLAM with DATMO. Once a moving object is detected at the first time $k - 3$, the test vehicle frame at $k - 3$ (O' in Figure 4) would be assigned as the temporary global frame for tracking this moving object. This transforms relative error in a global frame into the equivalent, but more convenient, representation of absolute error in a local frame.

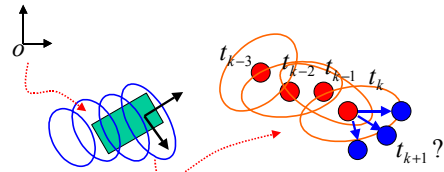


Figure 3: Global SLAM with DATMO

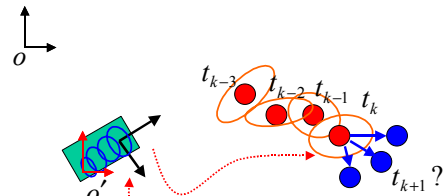


Figure 4: Local SLAM with DATMO

3 DATMO Implementation

Using a precise pose estimate and a surrounding map from SLAM, the proposed DATMO algorithm solves the problems in the following manner. First, a new scan is segmented into several groups using a simple distance criterion. With the surrounding map and the pose estimate from SLAM, moving objects (groups) are detected by finding inconsistencies between the new scan and the map. Using the data associated with a moving object, the Interacting Multiple Model (IMM) estimation algorithm [2, 3, 4] tracks and predicts the motion of this moving object with the constant velocity model and the constant acceleration model. The multiple hypothesis tracking (MHT) [9, 3] method is applied to refine detection and data association.

3.1 Scan Segmentation

Scan segmentation is the first step of the DATMO algorithm. Because the laser scans in our applications are not dense and the targets we want to detect and track do not have specific sizes and shapes, we use a simple distance criterion, namely that the distance between points in two groups must be longer than 1 meter, to segment measurement points into several groups. Figure 5 shows one example of scan segmentation.

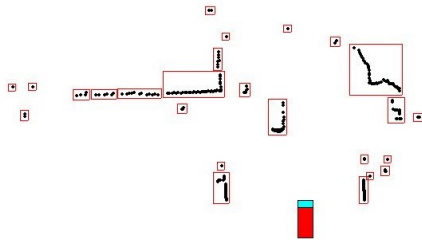


Figure 5: Scan Segmentation

3.2 Moving Object Detection

Intuitively, any inconsistent part from SLAM should belong to moving objects. But the idea isn't totally correct. There are two cases for detecting moving objects:

Case 1:

From previous scans or the map, we know some space is not occupied. If we find any object in this space, this object must be moving. In Figure 6, object A must be a moving object.

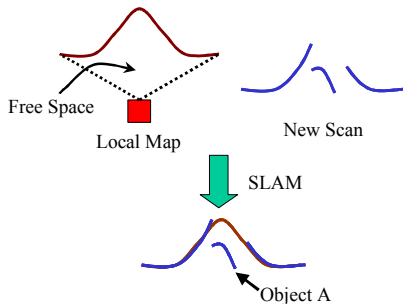


Figure 6: Detection Case 1.

Case 2:

In Figure 7, we can't say that object B is a moving object. Object B may be a new stationary object because Object B may have been occluded by object C. What we are sure is that object C is a moving object. Although we can't tell if object B is moving or not by registering only two scans, the previous information does help us to guess the characteristics of object B.

The moving object detection algorithm consists of two parts: the first is to detect moving points; the second is to combine the results from segmentation and moving point detection for deciding which groups are potential moving objects. Given a new scan, the local surrounding map, and the pose estimate from local SLAM, we first transform the local surrounding map to the coordinate frame of the

current laser scanner, and then convert the map from a rectangular coordinate system to a polar coordinate system. Now it is easy to detect moving points by comparing values along the range axis of the polar coordinate system. A group is identified as a potential moving object if the ratio of the number of moving points to the number of total points is greater than 0.5. Figure 8 shows the results of moving object detection and a red box indicates a moving car recognized by our motion-based detector.

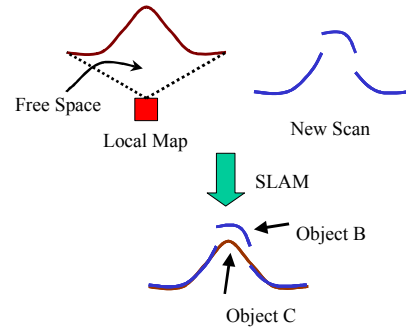


Figure 7: Detection Case 2.

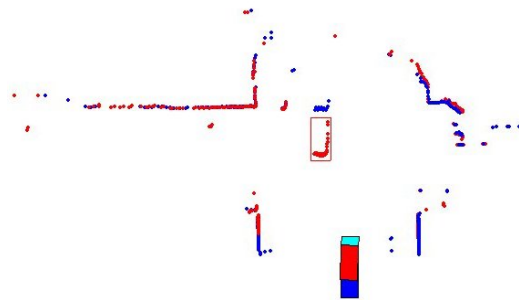


Figure 8: Moving object detection

3.3 Slow Moving Object Detection

Detection of pedestrians at very low speeds is difficult but possible by including information from the map. From our experimental data, we found that the data associated with a pedestrian is very small, generally 2-4 points. Also, the motion of a pedestrian can be too slow to be detected by the motion-based detector. Because the map contains information from previous moving objects, we can say that if a blob is in an area that was previously occupied by moving objects, this object can be recognized as a moving object.

3.4 The Interacting Multiple Model (IMM) algorithm

In SLAM, we can use odometry and the robot motion model to predict the future location of the robot. However, in DATMO neither *a priori* knowledge of moving objects' motion models nor odometry measurements about moving objects is available. Therefore we applied the IMM algorithm with the constant velocity (CV) model

and the constant acceleration (CA) model to model the motion of a moving object and to predict its future location. Here we use the Piecewise Constant White Acceleration Model from the CV model and the Piecewise Constant Wiener Process Acceleration Model for the CA model [2, 3]. The prediction and estimation of a moving object is from the mixing (combination) of these two motion models.

3.5 Multiple Hypothesis Tracking

Moving object detection is often erroneous because of measurement errors and pitch/roll motions of the test vehicle. We apply the multiple-hypothesis-tracking (MHT) method to filter out wrong detections and refine data association. Local SLAM is not run for every new scan because of computational power limitations and because odometry is good enough locally. Moving object detection initializes the new tracks and then track association is accomplished using spatial and shape information. The motion patterns of the hypothesis tree's branches are used to prune the hypothesis tree. At the next time of moving object detection the hypothesis tree will be confirmed and new moving objects will be detected.

4 Experimental Results

One SICK LMS221 and two SICK LMS291 laser scanners are mounted in various positions on Navlab11 (Figure 10), doing horizontal profiling. The range data were collected at 37.5 Hz with 0.5 degree resolution. The maximum measurement range of the scanners is 80 m. The images from our three-camera system are for visualization. Figure 11 shows the raw range images from the ladar mounting on the front of the test vehicle.

The results of local SLAM with DATMO are shown in Figure 12, Figure 13 and Figure 14. Instead of discarding information from moving objects, a stationary object map (SO-map) and a moving object map (MO-map) are created to store information from stationary objects and moving objects which are shown in black and red respectively. The solid blue rectangle indicates the location of the test vehicle. The blue boxes represent the tracked objects. In Figure 12, one car (Object A) and one trailer (Object B) are detected and tracked, but a pedestrian (Object C) is missed because the motion of the pedestrian is too small to detect at the beginning. After accumulating information in the MO-map, this pedestrian is successfully detected and tracked, which is shown in Figure 13. Figure 14 shows that our algorithm segments a trailer (Object A) into three groups and tracks these groups individually. An algorithm to analyze the relationships of groups for merging multiple tracks as a single object or splitting a track into multiple objects is ongoing.



Figure 10: The Navlab11 testbed

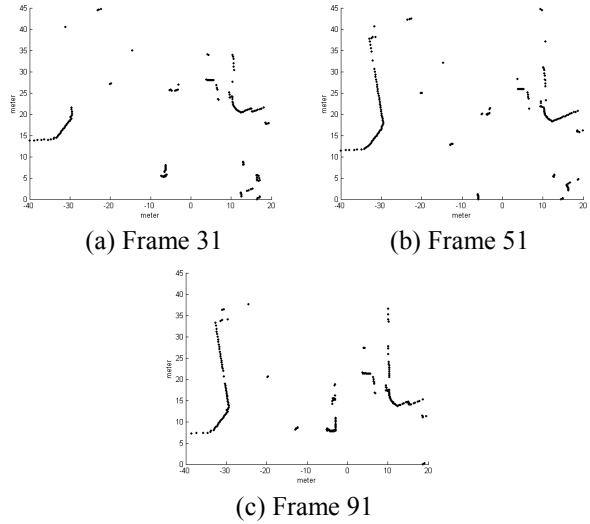


Figure 11: Raw range images from SICK LMS221.

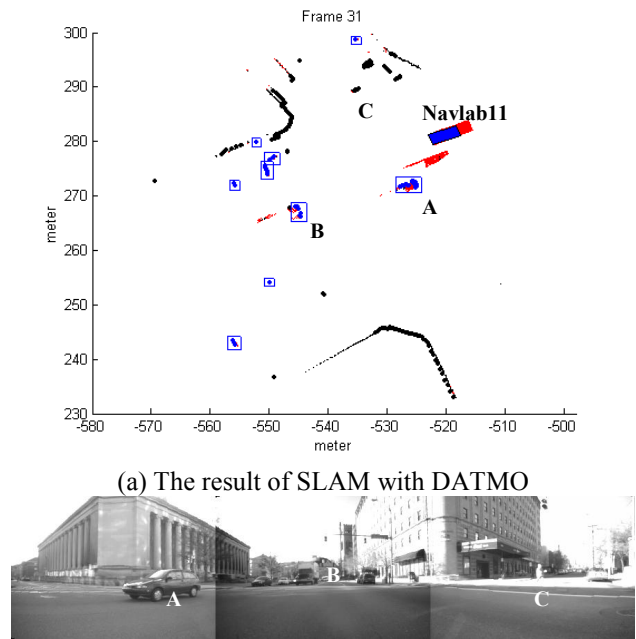


Figure 12: Ground Vehicle Detection and Tracking

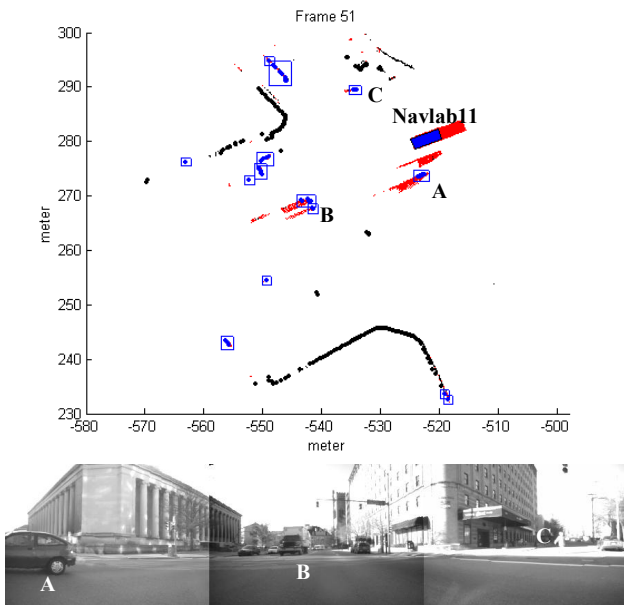


Figure 13: Pedestrian Detection and Tracking

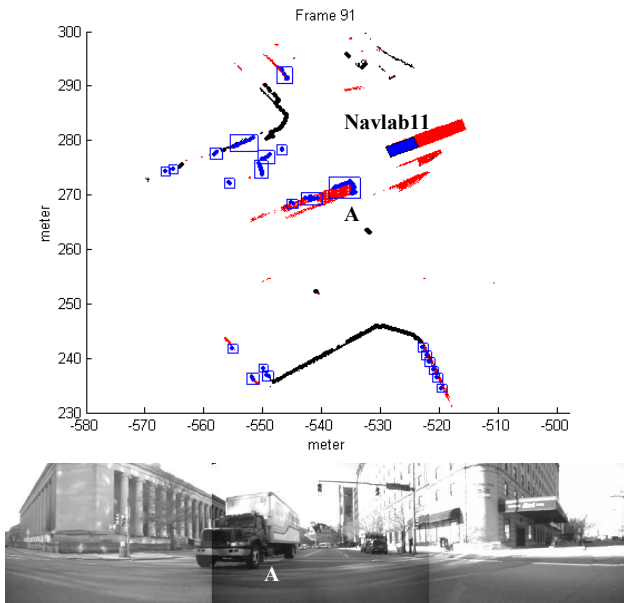


Figure 14

5 Conclusion

We have presented a method to accomplish detection and tracking of moving objects from a moving platform based on the Bayesian formula of SLAM with DATMO. Experimental results have shown that DATMO in crowded urban area from a ground vehicle at high speeds is feasible using odometry and laser scanners.

Acknowledgements

We thank people in the Navlab group for their help with the experiments. This work is partially supported by the Federal Transit Administration through the Pennsylvania Department of Transportation Agreement Number TEC-116 under the Federal Assistance Program

References

- [1] Y. Bar-Shalom, Tracking Methods in a Multitarget Environment, *IEEE Trans. On Automatic Control*, Vol. 23, No. 4, Aug. 1978.
- [2] Y. Bar-Shalom and X.-R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*, YBS, Danvers, MA, 1995.
- [3] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*, Artech House, MA, 1999
- [4] H. A. P. Blom and Y. Bar-Shalom, The Interacting Multiple Model Algorithm for Systems with Markovian Switching Coefficients, *IEEE Trans. On Automatic Control*, Vol. 33, No. 8, Aug. 1988.
- [5] A. Fod, A. Howard, and M. J Mataric, A Laser-Based People Tracker, *IEEE Int. Conf. on Robotics and Automation*, May, 2002.
- [6] B. Kluge, C. Kohler and E. Prassler, Fast and Robust Tracking of Multiple Objects with a Laser Range Finder, *IEEE Int. Conf. on Robotics and Automation*, pp. 1683-88, 2001.
- [7] M. Lindstrom and J.-O. Eklundh, Detecting and Tracking Moving Objects from a Mobile Platform using a Laser Range Scanner. *Proc. Int. Conf. On intelligent Robots and Systems*, Oct. 2001.
- [8] E. Prassler, J. Scholz, and P. Fiorini, A Robotic Wheelchair for Crowded Public Environments, *IEEE Robotics & Automation Magazine*, March, 2001.
- [9] D. B. Reid, An Algorithm for Tracking Multiple Targets, *IEEE Trans. On Automatic Control*, vol. 24, no. 6, December 1979.
- [10] D. Schulz, W. Burgard, D. Fox and A. B. Cremers. Tracking Multiple Moving Targets with a Mobile Robot using Particle Filters and Statistical Data Association, *IEEE Int. Conf. on Robotics and Automation*, pp. 1665-70, 2001.
- [11] Summer School on SLAM 2002, <http://www.cas.kth.se/SLAM/>
- [12] C.-C. Wang and C. Thorpe, Simultaneous Localization And Mapping with Detection And Tracking of Moving Objects, *IEEE Int. Conf. on Robotics and Automation*, May, 2002.
- [13] C.-C. Wang, C. Thorpe and S. Thrun, Online Simultaneous Localization and Mapping with Detection and Tracking of Moving Objects: Theory and Results from a Ground Vehicle in Crowded Urban Areas, *IEEE Int. Conf. on Robotics and Automation*, May, 2003.
- [14] L. Zhao and C. Thorpe, Qualitative and Quantitative Car Tracking from a Range Image Sequence, *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, 1998.