Online Simultaneous Localization And Mapping with Detection And Tracking of Moving Objects: Theory and Results from a Ground Vehicle in Crowded Urban Areas

Chieh-Chih Wang, Charles Thorpe and Sebastian Thrun
Robotics Institute, Carnegie Mellon University
Pittsburgh, PA, 15213, USA
Email: [bobwang, cet, thrun]@cs.cmu.edu

***Abstract* – The simultaneous localization and mapping (SLAM) with detection and tracking of moving objects (DATMO) problem is not only to solve the SLAM problem in dynamic environments but also to detect and track these dynamic objects. In this paper, we derive the Bayesian formula of the SLAM with DATMO problem, which provides a solid basis for understanding and solving this problem. In addition, we provide a practical algorithm for performing DATMO from a moving platform equipped with range sensors. The probabilistic approach to solve the whole problem has been implemented with the Navlab11 vehicle. More than 100 miles of experiments in crowded urban areas indicated that SLAM with DATMO is indeed feasible.**

I. INTRODUCTION

The simultaneous localization and mapping (SLAM) problem has attracted immense attention in the mobile robotics literature [17], and SLAM techniques are at the core of many successful robot systems. Most researchers on SLAM assume that the unknown environment is static, containing only rigid, non-moving objects. In [20], we presented a method to solve the SLAM problem and the detection and tracking of moving objects (DATMO) problem concurrently and showed that the initial results of SLAM with DATMO are dramatically better than SLAM without DATMO in crowded urban environments. But at that moment we did not present a theoretic framework for solving the SLAM with DATMO problem; the tracking of moving objects also had not been fully developed. In this paper, we extend the Bayesian formula of the SLAM problem to the SLAM with DATMO problem. In order to supplement our previous paper, we also present the approach for solving the DATMO problem in detail.

The new focus of the Navlab group at Carnegie Mellon University is on short-range sensing, to look all around the vehicle for improving driving safety and preventing traffic injuries caused by human factors such as speeding, or distraction. We believe that being able to detect and track every stationary object and every moving object, to reason about the dynamic traffic scene, to detect and predict every critical situation, and to warn and assist drivers in advance, is essential to prevent these kinds of accidents.

Fig. 1: NAVLAB 11 testbed

In order to perform DATMO by using sensors mounted on a moving ground vehicle at high speeds, a precise localization system is essential. It is known that GPS and DGPS often fail in the urban areas because of urban canyon effects; and a good IMU system is very expensive. Our solution of the SLAM with DATMO problem satisfies both the safety and navigation demands by using laser scanners and odometry. SLAM with DATMO can provide a better estimation of the vehicle's location and provide information of the dynamic environments, which are critical to driving assistance and autonomous driving.

If we can have a stationary object map in advance, the SLAM problem reduces to a localization problem with a known map, which is easier solved than the full SLAM problem. Unfortunately, it is difficult to build a usable stationary object map because of *temporary stationary objects* such as parked cars. Even though we can filter moving objects out, the stationary object maps of the same scene built from different times could still be different, which means that we still have to do online map building for updating the current stationary object map. For driving assistance applications, basically a globally consistent metric stationary object map is not necessary. As a result, we include a digital map in our system and accomplish global localization in a topological way.

The DATMO problem has been extensively studied for several decades [1, 2]. It is not easy to solve the DATMO problem in crowded urban environments from a moving ground vehicle at high speeds. There are many kinds of moving objects, such as pedestrians, animals, wheelchairs, bicycles, motorcycles, cars, buses, trucks, trailers, etc., which means that targets have a wide range of sizes and velocities. The range of the velocities is from under 5mph (such as the pedestrian's movement) to 50mph. When using laser scanners, the features of moving objects can change significantly from scan to scan. The observation of a single object such as a trailer may be shown as several objects; multiple objects such as pedestrians may also be shown as a single object and moving objects may disappear and reappear. Besides, the vehicle may have extreme roll and pitch motions. To solve these difficulties, we presented a motion-based detector to detect different kinds of moving objects in [20]. A hypothesis tree is managed for data association and moving object merging/removal. The results show that our DATMO algorithm can be run in the crowded urban areas robustly and efficiently.

Both SLAM and DATMO have been studied in isolation. However, when driving in crowded urban environments composed of stationary and moving entities, neither of them is sufficient. The contribution of this paper is to establish a mathematical framework that integrates both, SLAM and DATMO. The paper provides ample experimental results that show that performing both at the same time is superior to doing just one or the other.

The rest of this paper is arranged as follows: In Section II, the Bayesian formula for the SLAM with DATMO problem is introduced; Section III reviews our approach of SLAM in outdoor environments briefly; Section IV introduces the algorithms to solve the DATMO problem in detail. The experimental results are in Section V, and the conclusion and future work are in Section VI.

II. SLAM WITH DATMO PROBLEM DEFINITION

In this section the Bayesian formulation of the SLAM with DATMO problem is introduced.

A. 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 $M = \{m^1, \cdots, m^l\}$. In addition, we define the following set to refer data leading up to time *k*.

$$U_k = \{u_0, u_1, ..., u_k\} = \{U_{k-1}, u_k\} \quad (1)$$

$$Z_k = \{z_0, z_1, ..., z_k\} = \{Z_{k-1}, z_k\} \quad (2)$$

$$X_k = \{x_0, x_1, ..., x_k\} = \{X_{k-1}, x_k\} \quad (3)$$

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

B. Bayesian Formulation of the SLAM Problem

Before formulating the SLAM with DATMO problem, we briefly introduce the formula of the SLAM problem. The SLAM problem is to determine the robot location X_k and the stationary object map M from laser scanner measurements Z_k and odometry measurements U_k. The general probabilistic formula for SLAM is expressed as:

$$p(X_k, M \mid Z_k, U_k) \quad (4)$$

Although the batch methods can provide an optimal estimation, the recursive methods are considered and used because of the online requirement and the computational power limitation in most of applications. According to the Bayes' theorem and the assumptions that the vehicle motion model is Markov and the objects in the environment are stationary, the general recursive Bayesian formula for SLAM can be derived and expressed as: (See [19, 10] for more details.)

$$p(x_k, M \mid Z_k, U_k) = \eta \cdot p(z_k \mid x_k, M) \int p(x_k \mid x_{k-1}, u_k)\, p(x_{k-1}, M \mid Z_{k-1}, U_{k-1})\, dx_{k-1} \quad (5)$$

where $p(x_{k-1}, M \mid Z_{k-1}, U_{k-1})$ is the posterior probability at time $k-1$, $p(x_k, M \mid Z_k, U_k)$ is the posterior probability at time k, $p(x_k \mid x_{k-1}, u_k)$ is the motion model, $p(z_k \mid x_k, M)$ is the update stage which can be inferred as the sensor (measurement, perceptual) model, and η is a normalizing constant. The motion model is calculated according to the vehicle kinematics/dynamics. The sensor model can be represented and calculated by different ways, such as feature/landmark based and occupancy-grid-map based approaches.

C. Bayesian Formulation of SLAM with DATMO

The SLAM with DATMO problem is not only to accomplish SLAM in dynamic environments but also to detect and track these dynamic events. Since the unknown

environment is dynamic and contains moving objects, the general recursive probabilistic formula for SLAM with DATMO can be expressed as:

$$p(x_k, Y_k, M \mid Z_k, U_k) \tag{6}$$

Here $Y_k = \{y_k^1, \cdots y_k^n\}$ are the locations of moving objects, of which there are n moving objects that appeared inside the sensor's range at time k.

C.1 Assumptions

Before introducing the derivation, the assumptions we made are addressed.

Assumption 1: Measurements can be decomposed into measurements of stationary and moving objects:

$$z_k = z_k^s + z_k^m \quad \text{and hence} \quad Z_k = Z_k^s + Z_k^m \tag{7}$$

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. In particular this implies the following conditional independence

$$\begin{aligned} p(z_k \mid Y_k, M, x_k) &= p(z_k^s \mid Y_k, M, x_k) p(z_k^m \mid Y_k, M, x_k) \\ &= p(z_k^s \mid M, x_k) p(z_k^m \mid Y_k, x_k) \end{aligned} \tag{8}$$

Assumption 2: 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:

$$p(M, x_k \mid Y_k, Z_k, U_k) = p(M, x_k \mid Z_k^s, U_k) \tag{9}$$

This is correct if we have no information whatsoever about the speed at which objects move. Here it is an approximation, but one that reduces the complexity of SLAM with moving objects enormously.

C.2 Derivation

We begin by factoring out the most recent measurement:

$$\begin{aligned} p(Y_k, M, x_k \mid Z_k, U_k) \propto\ & p(z_k \mid Y_k, M, x_k, Z_{k-1}, U_k) \\ & p(Y_k, M, x_k \mid Z_{k-1}, U_k) \end{aligned} \tag{10}$$

Observing the standard Markov assumption, we note that $p(z_k \mid Y_k, M, x_k, Z_{k-1}, U_k)$ does not depend on Z_{k-1} and U_k. Furthermore, we can now partition the measurement into stationary and moving, and obtain by exploiting Assumption 1 and (8):

$$\begin{aligned} & p(Y_k, M, x_k \mid Z_k, U_k) \propto \\ & \quad p(z_k^s \mid M, x_k) p(z_k^m \mid Y_k, x_k) p(Y_k, M, x_k \mid Z_{k-1}, U_k) \end{aligned} \tag{11}$$

The rightmost term $p(Y_k, M, x_k \mid Z_{k-1}, U_k)$ can now be further developed, exploiting Assumption 2:

$$\begin{aligned} & p(Y_k, M, x_k \mid Z_{k-1}, U_k) \\ & \quad = p(Y_k \mid Z_{k-1}, U_k) p(M, x_k \mid Y_k, Z_{k-1}, U_k) \\ & \quad = p(Y_k \mid Z_{k-1}, U_k) p(M, x_k \mid Z_{k-1}^s, U_k) \end{aligned} \tag{12}$$

Hence we get our desired posterior

$$\begin{aligned} & p(Y_k, M, x_k \mid Z_k, U_k) \propto \\ & \underline{\underline{p(z_k^s \mid M, x_k) p(M, x_k \mid Z_{k-1}^s, U_k)}} \\ & \quad \textit{SLAM} \qquad \underline{\underline{p(z_k^m \mid Y_k, x_k) p(Y_k \mid Z_{k-1}, U_k)}} \\ & \qquad\qquad\qquad\qquad \textit{DATMO} \end{aligned} \tag{13}$$

The term $p(Y_k \mid Z_{k-1}, U_k)$ resolves to the following predictions

$$\begin{aligned} & p(Y_k \mid Z_{k-1}, U_k) \\ & \quad = \int p(Y_k \mid Z_{k-1}, U_k, Y_{k-1}) p(Y_{k-1} \mid Z_{k-1}, U_k) dY_{k-1} \\ & \quad = \int p(Y_k \mid Y_{k-1}) p(Y_{k-1} \mid Z_{k-1}, U_{k-1}) dY_{k-1} \end{aligned} \tag{14}$$

Finally, the term $p(M, x_k \mid Z_{k-1}^s, U_k)$ in (13) is obtained by the following step:

$$\begin{aligned} & p(M, x_k \mid Z_{k-1}^s, U_k) = p(x_k \mid Z_{k-1}^s, U_k, M) p(M \mid Z_{k-1}^s, U_k) \\ & = \int p(x_k \mid Z_{k-1}^s, U_k, M, x_{k-1}) p(x_{k-1} \mid Z_{k-1}^s, U_k, M) \\ & \qquad\qquad p(M \mid Z_{k-1}^s, U_k) dx_{k-1} \\ & = \int p(x_k \mid u_k, x_{k-1}) p(x_{k-1}, M \mid Z_{k-1}^s, U_{k-1}) dx_{k-1} \end{aligned} \tag{15}$$

which is the familiar SLAM prediction step. Putting everything back into (13) we now obtain the final filter equation:

$$\begin{aligned} & p(Y_k, M, x_k \mid Z_k, U_k) \propto \\ & \underbrace{p(z_k^m \mid Y_k, x_k)}_{\textit{Update}} \underbrace{\int p(Y_k \mid Y_{k-1}) p(Y_{k-1} \mid Z_{k-1}, U_{k-1}) dY_{k-1}}_{\textit{Prediction}} \\ & \underbrace{p(z_k^s \mid M, x_k)}_{\textit{Update}} \underbrace{\int p(x_k \mid u_k, x_{k-1}) p(x_{k-1}, M \mid Z_{k-1}^s, U_{k-1}) dx_{k-1}}_{\textit{Prediction}} \end{aligned} \tag{16}$$

D. Solving the SLAM with DATMO problem

From (16), input to the SLAM with DATMO filter are two separate posteriors, one of the conventional SLAM form, $p(x_{k-1}, M \mid Z^s_{k-1}, U_{k-1})$, and a separate one for DATMO, $p(Y_{k-1} \mid Z_{k-1}, U_{k-1})$.

The remaining question is now how to recover those posterior at time *k*. For the SLAM part, the recovery is simple.

$$p(x_k, M \mid Z^s_k, U_k) = \int p(Y_k, M, x_k \mid Z_k, U_k) dY_k \propto \quad (17)$$
$$p(z^s_k \mid M, x_k) \int p(x_k \mid u_k, x_{k-1}) p(x_{k-1}, M \mid Z^s_{k-1}, U_{k-1}) dx_{k-1}$$

For DATMO, we get

$$p(Y_k \mid Z_k, U_k) = \iint p(Y_k, M, x_k \mid Z_k, U_k) dM dx_k$$
$$\propto \int [p(z^m_k \mid Y_k, x_k) \int p(Y_k \mid Y_{k-1}) p(Y_{k-1} \mid Z_{k-1}, U_{k-1}) dY_{k-1}]$$
$$p(x_k \mid Z^s_k, U_k) dx_k \quad (18)$$

where the posterior over the pose $p(x_k \mid Z^s_k, U_k)$ is simply the marginal of the joint posterior calculated in (17):

$$p(x_k \mid Z^s_k, U_k) = \int p(x_k, M \mid Z^s_k, U_k) dM \quad (19)$$

Equation (18) 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.

There are a number of possible methods for solving the SLAM problem such as the Particle Filter, the Extended Kalman Filter (EKF), the Unscented Kalman Filter, and the Sum-of-Gaussian method [10]. The SLAM with DATMO problem can be solved by these methods as well. The main differences of these methods are the representations of the joint posterior density. Given enough particles (sample points), the Particle filter can provide a complete representation of the joint posterior density, which is the key to deal with non-linearity and non-Gaussianity.

Currently, because the computational power of our system is not enough to run particle filter-based algorithms, the whole SLAM and DATMO problem is solved through the use of the EKF. We assume that the sensor model and motion model are Gaussian so that all these integrals of the formula are easily carried out in closed form. Since SLAM can get a better result if the moving objects are filtered out in advance, the whole procedure can be operated iteratively in order to get a more accurate result.

It should be noted that SLAM with DATMO could be handled by calculating a joint posterior over all features (map, robot pose, moving objects). Such an approach would be similar to existing SLAM algorithms, but with additional structure to allow for motion of the moving objects. Our choice to decompose the estimation problem into two separate estimators is motivated by two observations: First, moving features are highly unpredictable and including them in the vehicle localization (as would be the case for the single-filter solution) would have a negative effect on the vehicle's localization. Second, by maintaining separate posteriors for the static map variables and the moving features, the resulting estimation problems are much lower dimensional than the joint estimation problem. This makes it possible to update both filters in real-time.

III. SLAM IMPLEMENTATION IN URBAN AND SUBURBAN AREAS

Extracting features robustly and correctly in outdoor environments is difficult. Whenever a feature is extracted, an error from feature extraction will occur. The error analysis of feature extraction is not yet rigorously studied. Instead of feature-based approaches, our system applies a scan matching technique, the Iterative Closest Point (ICP) algorithm [22], and uses a grid-map to represent the environments. The map updating in our system is similar to the approach presented in [18]. Unlike other mapping methods, the map in our system contains information not only from stationary objects but also from moving objects. Checking the consistency of both the moving object map and the stationary object map provides important information of the SLAM with DATMO algorithm performance. The main problem of the scan-matching-based approaches is that there is no efficient and good way to estimate the uncertainty of the location estimation. This does decrease the likelihood of the successful large loop closing and DATMO. The study for solving this problem is ongoing.

In order to globally localize the vehicle online, a digital map is included into our system and the global localization is accomplished in a topological way. Online global topological SLAM using a digital map without GPS will be presented in another paper.

IV. DATMO IMPLEMENTATION

Basically, an algorithm for solving DATMO problems has to address the following issues:

- Detection and initiation of new moving objects;
- Moving object motion modeling;
- Data association;

- Merging moving objects when two or more moving objects coalesce;
- Removal of moving objects that have moved outside the sensor's range;
- Occlusion;
- Adaptation of the false measurements;
- Algorithm can be shown to work robustly over long sequences of data.

Our DATMO algorithm solves above problems in the following manner. The motion-based detector is used to detect different kinds of moving objects in crowded urban environments. Then the Multiple Hypothesis Tracking (MHT) method [3, 15] is applied to accomplish data association. Given the data associated with a moving object, the motion of this moving object is modeled and this model is used to predict the future motion. In this section, the motion-based detector presented in [20] is briefly introduced. The approaches of data association and motion modeling are addressed. In the end of this section, we will show that our DATMO algorithm has the ability to adapt the false measurements.

A. Motion-based Detection

In indoor environments, the most important targets of the DATMO problem are people [6, 8, 9, 11, 16]. 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. Both appearance-based and feature-based methods rely on the prior knowledge of the targets. In our application, because of the variety of our targets, it is very difficult to define features or appearances by using laser scanners. Also, the task is to detect *moving* objects, not temporary *stationary* objects such as parked cars, which are still useful information for the online SLAM.

Other than appearance-based and feature-based approaches, we presented a *motion*-based detection approach in [20] for both indoor and outdoor environments. As long as an object is moving, our approach can detect it. Although this method cannot detect stationary cars and pedestrians, these temporary stationary objects actually do not have to be dealt with, because their stationary state will not cause any critical threat that the driver has to be aware of, therefore this drawback is tolerable. Fig. 2 shows the results of SLAM with moving vehicle detection by our motion-based approach.

The detection of moving people at very low speeds is difficult but it is possible by fusing information from the moving object map. From Fig. 3, we found that the data associated with a pedestrian is very small, e.g. 2-4 points. Also, the motion of a pedestrian can be too slow to be detected by the motion-based detector. Since the map also contains information from previous moving objects, we can say if this blob is in an area that was previously occupied by moving objects, this object can be recognized as a moving object. In the bottom of Fig. 3, even if an object has no motion, this object is defined as a moving object according to the information from the map.

B. Moving Object Initiation and Data Association

Once a new moving object is detected, our algorithm initializes a new track for this object, such as assigning an initial state and a motion model to this new moving object. By using laser scanners, we can only get the position but not the velocity and orientation, therefore our algorithm uses the data from different times and then accomplishes data association in order to initialize a new track.

Data association and tracking problems have been extensively studied and a number of statistical data association techniques have been developed, such as the Joint Probabilistic Data Association Filter (JPDAF) [1] and the Multiple Hypothesis Tracking (MHT). Our system applies the MHT method, which maintains a hypothesis tree and can revise its decisions while getting new information. This delayed decision approach is more robust than other approaches. The main disadvantage of the MHT method is its exponential complexity. If the hypothesis tree is too big, it will not be feasible to search the whole hypotheses to get the most likely set of matching. Fortunately, the number of moving objects in our application is usually less than twenty and most of the moving objects only appear for a short period of time. Also, useful information about moving objects from laser scanners, such as location, size, shape, and velocity, is used for updating the confidence for pruning and merging hypotheses. In practice, the hypothesis tree is always managed in a reasonable size.

C. Motion Modeling and Tracking

Given the data associated with a moving object, the goal of motion modeling is to find the motion model of this moving object. Without any prior knowledge, the procedure of getting a motion model is complicated and it needs enough data and time in order to get the correct model. The Interacting Multiple Model (IMM) estimator is a sub-optimal hybrid filter that has been successful implemented in various target-tracking applications. For instance, in [21] an extended Interacting Multiple Model (IMM) algorithm was demonstrated on the Navlab5 vehicle for tracking of moving cars on highways.

In order to analyze the interactions between the vehicle and other moving objects, precise motion models of

moving objects and a long period of observation are necessary. Intuitively and experimentally, it is a good approximation that moving cars have three behavior modes: the constant-velocity mode, the constant-acceleration mode, and the turning mode. Regarding other moving objects such as moving people, selecting good and efficient motion models is challenging. But if the task is to predict the future motions of moving objects, we found that the results are satisfying by assuming that moving objects have only one behavior mode, which is the constant-velocity mode.

Tracking algorithms estimate the state of moving objects according to motion models, sensor models, and measurements. Since the whole SLAM and DATMO problem is solved through the use of the EKF in this paper, motion models of moving objects are described in terms of a constant velocity model subject to zero mean Gaussian errors.

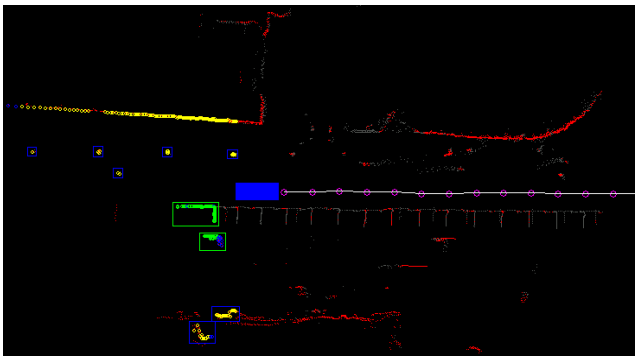


Fig. 2: The result of SLAM with Detection using a laser scanner mounting on the front of the vehicle. The solid blue rectangle represents the vehicle itself, whose width is 2m and length is 5m; the faint chain of magenta circles trailing from the vehicle is the list of positions from which the vehicle made its observations; the light red dots and light gray dots respectively belong to stationary parts and moving parts of the map; yellow, green, and blue circles represent the current scan; yellow small circles mean *stationary*; blue small circles are *unidentified* since we don't have enough information to tell if they are moving or stationary; small green circles are belonged to *moving* objects; green boxes are the moving objects recognized by our motion-based detector.

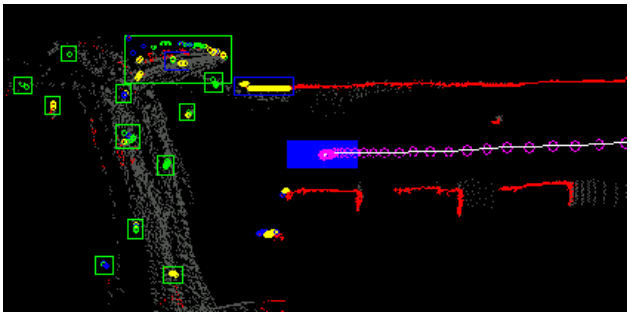


Fig. 3: Moving People Detection

D. Adaptation of the False Measurements

False measurements are often observed in our experiments. One is due to roll and pitch motions of the vehicle, which are unavoidable because of making turns at high speeds or sudden stops or starts. These motions may cause false measurements such as wrong scan data from the ground instead of other objects. Another cause of false measurements due to the fact that outdoors, the ground is not always flat; hence the flat world assumption in SLAM is only approximately valid; therefore uphill environments may cause false measurements because the scan may hit the ground as well. The DATMO algorithm adapts these false measurements implicitly without other pitch and roll measurement. First, the false measurements are detected and initialized as new moving objects by our motion-based detector. After data association and tracking are applied to these measurements, the shape and motion inconsistency is detected by our algorithm, which shows that these are false measurements. Also these false measurements will disappear immediately once the motion of the vehicle is back to normal.

V. EXPERIMENTAL RESULTS

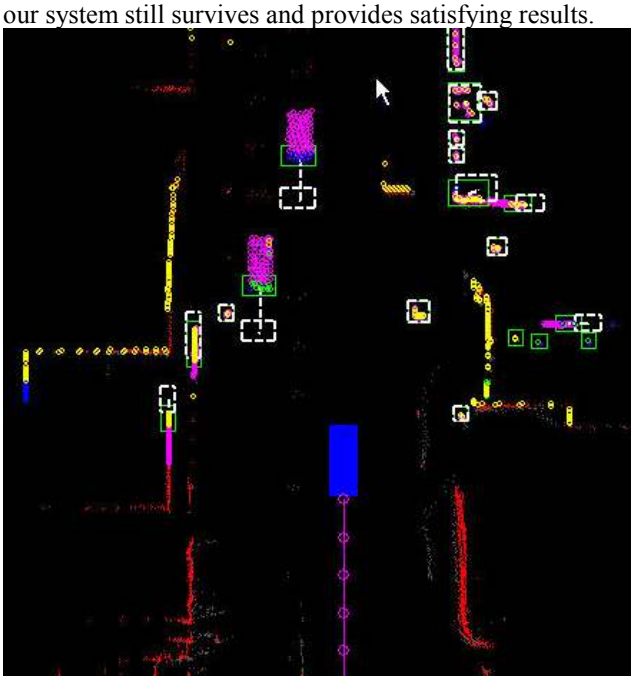
Currently the Navlab11 vehicle (See Fig. 1) is equipped with motion sensors (IMU, GPS, differential odometry, compass, inclinometer, angular gyro), video sensors (five video cameras, an omni-directional camera), three SICK single-axis scanning rangefinders, a light-stripe rangefinder, and five 500-MHz Pentium computers. The results in this paper use data only from SICK laser scanners and the odometry. The images from the omni-directional camera are for visualization.

One SICK LMS221 and two SICK LMS291 laser scanners were mounted in various positions on Navlab11, doing horizontal or vertical profiling. Navlab11 was driven through the Carnegie Mellon University campus and around nearby streets. The range data were collected at 37.5 Hz with 0.5 degree resolution. The maximum measurement range of the scanners is 80 m. Fig. 4 and Fig. 5 shows the results of SLAM with DATMO. The white dashed boxes present the predicted locations of the tracked moving objects from our tracking algorithm. The magenta points are previous data associated with the moving objects. In Fig. 4, two cars are detected and tracked. The speeds of these two cars are similar to the vehicle. Several false detections occurred at the top-right corner of the figure because online SLAM is performed and the confidence of that area is still low. But because the tracking algorithm indicted that the speeds of these objects are slow, the false detections will be removed and the map will be modified. In Fig. 5, two moving pedestrians are detected. The blur region in the unwarped

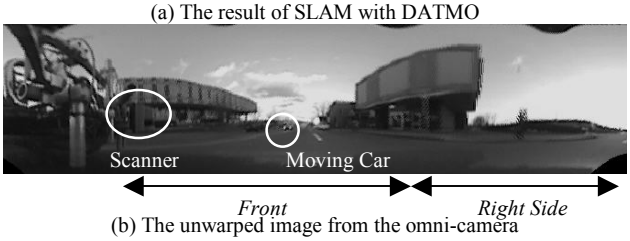
image shows the motion and location of these two moving pedestrians. Also on the right side of the vehicle, there is an intersection, which is uphill. A lot of magenta points are associated with this false detection. Fig. 6 also shows the effects of violent roll and pitch motions. The results show that our algorithm can survive under these critical situations. Fig. 7 shows the quality of our SLAM with DATMO algorithm. Another scanner was mounted on the top of the vehicle and performs the vertical profiling. A high quality 3D model is produced in a minute. More results and videos are available at http://www.cs.cmu.edu/~bobwang.

VI. CONCLUSION AND FUTURE WROK

We derived a Bayesian formula of the SLAM with DATMO problem. The formula provides a rigorous foundation to understand and solve the SLAM with DATMO problem. A probabilistic approach to this problem was implemented through the use of the Extend Kalman Filter. In the course of this research, our approach was exposed to data acquired over a total distance of 100 miles. Even the assumption that the ground is flat is not valid and the vehicle has extreme roll and pitch motions, our system still survives and provides satisfying results.



(a) The result of SLAM with DATMO



(b) The unwarped image from the omni-camera

Fig. 4: Detection and tracking of moving cars. The SLAM result shows that the vehicle was moving at 22.8 mph.

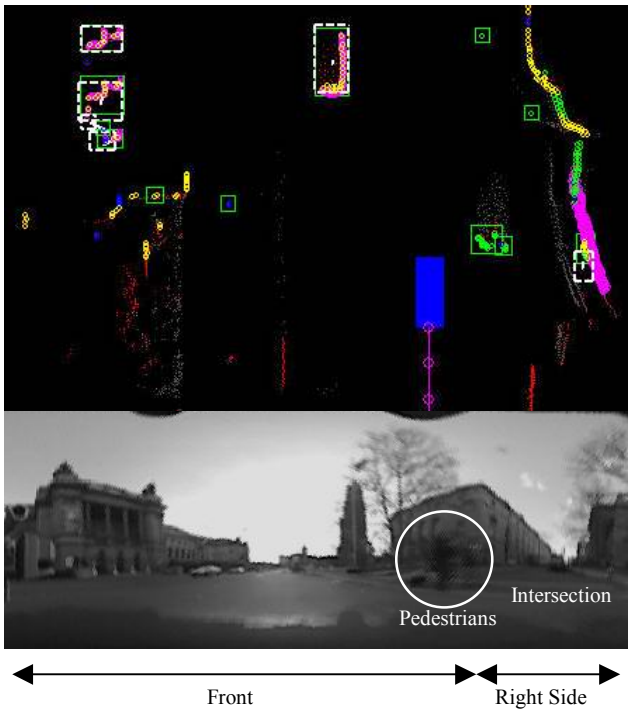


Fig. 5: Pedestrian Detection. The SLAM result shows that the test vehicle was moving at 21.0 mph.

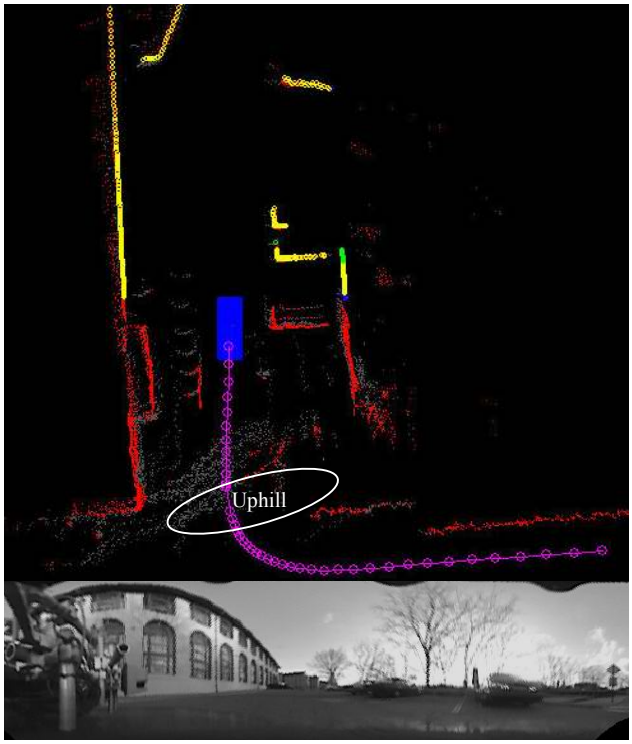


Fig. 6: The effects of roll and pitch on SLAM with DATMO. The SLAM result shows that the vehicle was moving at 12.3 mph.

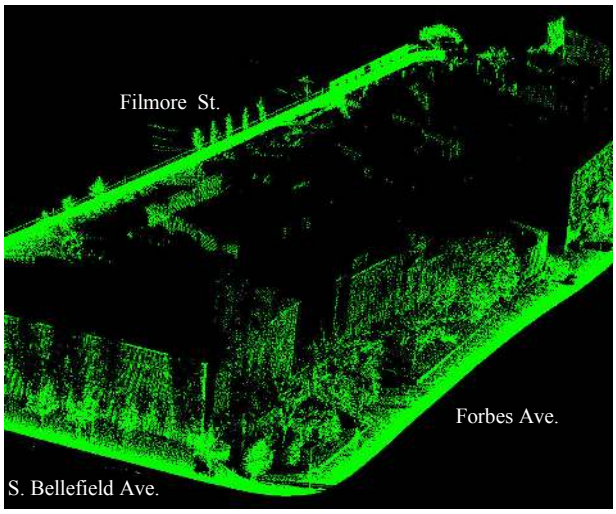


Fig. 7: 3D City Mapping

This paper raises several interesting topics. Maps of SLAM not only can be used for detecting moving objects but also can be constraints to improve the tracking performance. According to our experiment data, using cameras to detect moving objects is harder than using laser scanners. On the other hand, cameras provide very rich information about the global localization, which could benefit SLAM greatly. More experiments will be conducted using multiple laser scanners in different weather conditions for testing our algorithm.

VII. ACKNOWLEDGEMENTS

We thank people in the 3D Computer Vision Group of CMU and at NASA Ames Research Center for fruitful discussions. This work is partially supported by the Army Research Laboratories under Collaborative Technology Alliance program; by the Federal Transit Administration; by Bosch Corporation; and by SAIC Inc.

VIII. REFERENCES

[1] Y. Bar-Shalom and X.-R. Li, *Estimation and Tracking: Principles, Techniques, and Software*, YBS, Danvers, MA, 1988.

[2] Y. Bar-Shalom and X.-R. Li, *Multitarget-Multisensor Tracking: Principles and Techniques*, YBS, Danvers, MA, 1995.

[3] I. J. Cox and S. L. Hingorani, An Efficient Implemenation of Reid's Multiple Hypothesis Tracking Algorithm and its Evaluation for the Purpose of Visual Tracking, *IEEE Trans. on Pattern Analysis and Machine intelligence,* Vol. 18, No. 2, Feb., 1996.

[4] L. Davis, V. Philomin and R. Duraiswami. Tracking humans from a moving platform. *15th Int. Conf. on Pattern Recognition*. Vol. 4, pp. 171-78, 2000.

[5] M. Dietl, J.-S. Gutmann and B. Nebel, Cooperative Sensing in Dynamic Environments, *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, 2001.

[6] A. Fod, A. Howard, and M. J Mataric, A Laser-Based People Tracker, *IEEE Int. Conf. on Robotics and Automation*, May, 2002.

[7] H. Gauvrit, J. P LE Cadre and C. Jauffret, A formulation of Multitarget Tracking as an Incomplete Data Problem, *IEEE Tran. On Aerospace and Electronic Systems*, Vol. 33, No. 4, October, 1997

[8] 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.

[9] B. Kluge, D. Bank, and E. Prassler, Motion Coordination in Dynamic Environments: Reaching a Moving Goal while Avoiding Moving Obstacles, *IEEE Int. Workshop on Robot and Human Interactive Communication*, September, 2002.

[10] S. Majumder, H. Durrant-Whyte, S. Thrun and M. de Battista, An Approximate Bayesian Method for Simultaneous Localization and Mapping, submitted to *IEEE Trans. Robotics and Automation*.

[11] M. Montemerlo, S. Thrun, and W. Whittaker, Conditional Particle Filters for Simultaneous Mobile Robot Localization and People-Tracking, *IEEE Int. Conf. on Robotics and Automation*, May, 2002.

[12] D. M. Lane, M. J. Chantler, and D. Dai, Robust Tracking of Multiple Objects in Sector-Scan Sonar Images Sequences Using Optical Flow Motion Estimation, *IEEE Journal of Ocean Engineering*, Vol.23, No. 1, Jan., 1998.

[13] 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.

[14] E. Prassler, J. Scholz, and P. Fiorini, A Robotic Wheelchair for Crowded Public Environments, *IEEE Robotics & Automation Magazine,* March, 2001.

[15] D. B. Reid, An Algorithm for Tracking Multiple Targets, *IEEE Trans. On Automatic Control*, vol. 24, no. 6, Dec., 1979.

[16] 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.

[17] Summer School on SLAM 2002, http://www.cas.kth.se/SLAM/

[18] S. Thrun, W. Burgard, and D. Fox, A Real-time Algorithm for Mobile Robot Mapping with Applications to Multi-robot and 3D Mapping, *IEEE Int. Conf. on Robotics and Automation*, 2000.

[19] S. Thrun, D. Fox, and W. Burgard, Monte Carlo localization with mixture proposal distribution, *AAAI Natl. Conf. on Artificial Intelligence*, 2000.

[20] 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.

[21] 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*, pp. 496-501, 1998.

[22] P. Besl and N. McKay. A method for registration of 3D shapes. *Trans. Pattern Analysis and Machine Intelligence*, Vol. 12, No. 2, pp. 239-256, 1992.