Proc. CVPR'98, Santa Barbara, CA, June 23-25, pp. 496-501, 1998 

# Qualitative and Quantitative Car Tracking from a Range Image Sequence

Liang Zhao and Chuck Thorpe
Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213
E-mail: {lzhao, cet}@ri.cmu.edu

**Abstract**

*In this paper, we present a car tracking system which provides quantitative and qualitative motion estimates of the tracked car simultaneously from a moving observer. First, we construct three motion models (constant velocity, constant acceleration, and turning) to describe the qualitative motion of a moving car. Then the models are incorporated into the Extended Kalman Filters to perform quantitative tracking. Finally, we develop an Extended Interacting Multiple Model (EIMM) algorithm to manage the switching between models and to output both qualitative and quantitative motion estimates of the tracked car. Accurate motion modeling and efficient model management result in a high performance tracking system. The experimental results on simulated and real data demonstrate that our tracking system is reliable and robust, and runs in real-time. The multiple motion representations make the system useful in various autonomous driving tasks.*

## 1 Introduction

Vehicle tracking is an important application of computer vision. In an automated driving system, both numerical or quantitative tracking and symbolic or qualitative tracking are the prerequisites for the success of other autonomous driving tasks such as motion planning, obstacle detection, and path planning. Accurate motion estimation allows safe and efficient motion planning, while symbolic motion interpretation simplifies the reasoning procedure by presenting useful information in an efficient manner. So it becomes attractive to integrate quantitative and qualitative motion analysis in a single tracking system, enabling an automated car to react quickly and correctly to the rapid maneuvers of other vehicles.

This paper presents a real-time vehicle tracking system which provides quantitative and qualitative motion estimates simultaneously from a moving platform and using a laser rangefinder. The system was developed in the context of the Automated Highway System (AHS) project [1] and is intended to provide accurate motion estimation, motion classification, and reliable maneuver detection. To our knowledge, this is the first work that incorporates both quantitative and qualitative motion estimates into a single algorithm. Currently, most research has been focused on how to quantitatively estimate the motion parameters [2, 3, 4]. On the other hand, qualitative approaches [5, 6] have been investigated in the context where numerical motion estimates are unavailable, unstable, or unnecessary. Little attempt has been made to pursue quantitative and qualitative tracking simultaneously. Our work, however, will contribute to this direction.

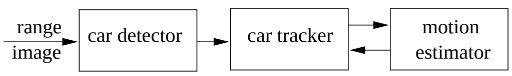


**Figure 1.** Block diagram of our car tracking system

Our tracking system consists of three parts (see Figure 1): car detection, car tracking, and quantitative/qualitative motion estimations. The fact that we can derive depth information directly from a laser rangefinder makes it a popular tool in many motion analysis systems [7, 8, 9]. In this work, we use a 2D laser scanner to obtain range data which greatly simplifies the car detection procedure. Thus, in this paper, we focus on the motion analysis part. First, we classify the vehicle motion into three qualitative kinematic modes: constant velocity mode, constant acceleration mode, and turning mode. Then, we derive the motion model of each mode and incorporate each model into an Extended Kalman Filter (EKF) [10] to track vehicles separately. The advantage of employing multiple motion models for tracking is that we can detect and predict the car maneuver as a separate event, and we can obtain more accurate motion interpretation in both quantitative and qualitative senses. The Interacting Multiple Model (IMM) algorithm first proposed by Blom [11] is a superior technique for multiple model management and maneuvering target tracking [12, 13]. Here we extend the algorithm to output both quantitative and qualitative motion estimates, and we call the extended IMM algorithm the EIMM algorithm. Accurate motion modeling and efficient model management result in a high performance tracking system. The experimental results on simulated and real data demonstrate that our tracking system is reliable and robust, and runs in real-time.

The remainder of the paper is organized as follows. In Section 2, we briefly explain the range image acquisition and car detection/tracking procedures. In Section 3, we describe the qualitative and quantitative motion representations. In Section 4, we present the motion analysis based on the EIMM algorithm. Experimental results are given in Section 5, followed by conclusions and future work in the last section.

## 2 Range Image Acquisition and Car Detection/Tracking

### 2.1 Range Image Acquisition

Range images are obtained from a single-line laser rangefinder mounted on the front bumper of a moving vehicle. The laser beam is controlled by a mirror that rotates about the vertical axis, providing a 180° field of view. Figure 2(a) illustrates a range image obtained from this range scanner – two barriers on the roadsides and two cars on the road. The simple range data allow fast car detection and tracking.

### 2.2 Car Detection/Tracking

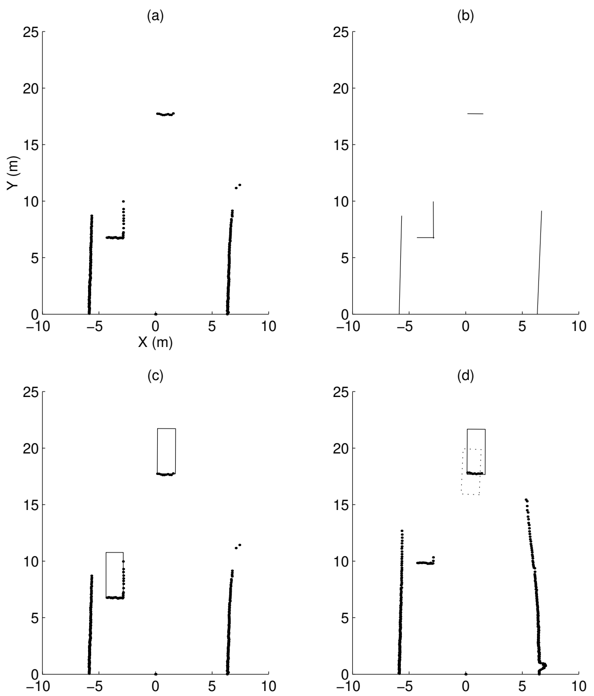


**Figure 2.** (a) raw range image (in the observer coordinate system) (b) line extraction (c) car detection (d) car tracking

Car detection is based on the prior knowledge of the rectangular shape, the limited size of vehicles, and the ground plane assumption. In a 2D range image (Figure 2(a)), vehicles consist of one to two line segments. The obvious features to detect are straight lines; this is done with the Hough Transform (Figure 2(b)). To group the extracted lines into cars, we first exclude the line segments belonging to barriers by using the length of the line segments and their poses with respect to the laser scanner. Two lines are grouped to a single car if they are close to each other and perpendicular; otherwise they are identified as two distinct cars (Figure 2(c)).

The steps listed above constitute the procedure for detecting cars in the first frame, requiring a search for cars across the entire image. In the following frames, the car is tracked by searching within a small area around the position predicted by the motion estimator (Figure 2(d): dotted box). This can be considered as soft gaze control, which is much simpler and faster than hard gaze control [14]. If there are two cars present in the area, the one nearest to the predicted position is chosen. From the car detector and tracker, the position and the heading of the tracked car are obtained and sent to the motion estimator for further analysis.

## 3 Qualitative and Quantitative Motion Representations

Similar to other areas in computer vision, motion representation has a great impact on the generality, reliability, accuracy, and efficiency of motion analysis algorithms. Both quantitative and qualitative representations as discussed by Thompson and Kearney [15] have advantages and disadvantages. Qualitative representation captures the significant characteristics of motion, but it has a low degree of precision. Quantitative representation, on the other hand, provides detailed, numerical description of motion parameters, but because of noisy measurements, problems of numerical instability and estimation errors of the motion parameters often occur. However, the two representations can complement each other and produce both stable and accurate motion interpretation. So the best strategy is to employ both qualitative and quantitative representations in the tracking system, and this is what we have done in this work.

### 3.1 Qualitative motion representation

**Table 1.** Motion Classification

| $V_{cv}$ | $V_{ca}$ | $V_t$ | **motion classes** |
|---|---|---|---|
| 1 | 0 | 0 | constant positive speed |
| -1 | 0 | 0 | constant negative speed |
| 0 | 1 | 0 | constant acceleration |
| 0 | -1 | 0 | constant deceleration |
| 0 | 0 | 1 | turning to the left hand side |
| 0 | 0 | -1 | turning to the right hand side |

Following Thompson and Kearney's definition of qualitative representation [15], we define the qualitative motion representation as a set of labels of the kinematic behaviors of vehicles, including constant velocity mode (CV), constant acceleration mode (CA) and turning mode (T). The first step is to assign to each kinematic behavior CV, CA, T a qualitative variable

$V_{cv}$, $V_{ca}$, and $V_t$, respectively. Each variable has three values: -1, 0, 1. So there are $3^3$ possible combinations or labels in the ($V_{cv}$, $V_{ca}$, $V_t$) base. However, the actual qualitative description includes only 6 motion classes corresponding to a subset of all these combinations, as shown in Table 1.

### 3.2 Quantitative Motion Representation

The quantitative motion representation is straightforward. It is assumed that the car moves in the ground plane; thus, the state of the tracked car is described by a six dimensional vector $(x, y, \theta, s, \dot{\theta}, \dot{s})^T$, where $(x, y)^T$ is the position of the car, $\theta$ is the direction of travel, $s$ is the speed, $\dot{\theta}$ is the turning rate and $\dot{s}$ is the acceleration. The reason that we choose to represent the direction of travel ($\theta$) instead of the direction of heading ($\theta_h$) is that they are not equivalent when the car is moving at high speed. In fact $\theta = \theta_h + \delta$, where $\delta$ is the slip angle [16]. Since the direction of travel is more relevant to our autonomous driving task, accordingly we choose to represent the turning rate of the travel direction instead of that of heading. All variables are defined with respect to the world coordinate frame attached to the initial position of the observer car, so that the tracking is carried out in the world coordinates. This has a clear advantage for any predictive or gaze control scheme, but requires prior knowledge of the observer's egomotion which, in this case, is obtained from a GPS receiver.

## 4 Motion Analysis Based on the EIMM Algorithm

In order to get the qualitative motion interpretation of the tracked car and to detect its maneuvers, which is critical to later motion planning and control, we construct three models corresponding to the three kinematic behaviors of vehicles as shown in Section 3.1. Each model is incorporated into an Extended Kalman Filter (EKF), and the set competes in the framework of the Extended Interacting Multiple Model (EIMM), which selects the most appropriate model to represent the dynamics of the tracked car at any time. The following two subsections describe the EKFs for three models and the EIMM algorithm, respectively.

### 4.1 The Filters for Three Models

For the purpose of computation uniformity in the EIMM algorithm, we selected a unique state vector as described in Section 3.2 for all filters: $X = (x, y, \theta, s, \dot{\theta}, \dot{s})^T$. The measurement vector is $Z = (x, y, \theta)^T$, and so the measurement equation is the same for all filters: $Z = HX + w$, where $H$ is the measurement matrix, and $w$ the measurement error vector. The state transition functions are obtained by integrating the following stochastic differential equations over time $t \geq 0$.

**Model 1.** The Constant Velocity Model (CV)

$$\dot{X} = \begin{bmatrix} s\cos\theta \\ s\sin\theta \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ e_{\dot{\theta}} \\ e_{\dot{s}} \\ 0 \\ 0 \end{bmatrix} \tag{1}$$

Here $e_{\dot{\theta}}$ and $e_{\dot{s}}$ are gaussian white noise terms used to absorb the error made by using the constant velocity assumption.

**Model 2.** The Constant Acceleration Model (CA)

$$\dot{X} = \begin{bmatrix} (s+\dot{s}t)\cos\theta \\ (s+\dot{s}t)\sin\theta \\ 0 \\ \dot{s} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ e_{\dot{\theta}} \\ 0 \\ 0 \\ e_{\ddot{s}} \end{bmatrix} \tag{2}$$

Here $e_{\dot{\theta}}$ and $e_{\ddot{s}}$ are noise terms used to absorb the error made by using the constant acceleration assumption.

**Model 3.** The Turning Model (T)

$$\dot{X} = \begin{bmatrix} (s+\dot{s}t)\cos(\theta+\dot{\theta}t) \\ (s+\dot{s}t)\sin(\theta+\dot{\theta}t) \\ \dot{\theta} \\ \dot{s} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ e_{\ddot{\theta}} \\ e_{\ddot{s}} \end{bmatrix} \tag{3}$$

Similarly, $e_{\ddot{s}}$ and $e_{\ddot{\theta}}$ are noise terms used to account for jerk and angular acceleration, respectively.

### 4.2 Motion Estimation Using the EIMM Algorithm

After motion models have been developed, the next step is to use these models to analyze the dynamic behaviors of the tracked car. The IMM algorithm is a superior and an attractive technique for managing multiple models and for tracking maneuvering targets. It provides good state estimates and it is real time. Good state estimates result in good model probability estimates and vice versa. Since its model probability computation [10] is based on the model likelihoods and the model switching matrix governed by the Markov chain, which provides a good statistical criterion for motion classification or qualitative motion interpretation, it is therefore straightforward to extend the IMM algorithm to provide qualitative motion analysis. We call the extended algorithm the EIMM algorithm.

Let $\hat{X}_k^j$ be the state estimate at time step $k$ based on model $j$, $\Lambda_k$ be the vector of model likelihoods at time $k$, $\hat{X}_k$ be the state combined from all model estimates and $\mu_k$ be the vector of the model probabilities at time $k$ when all the likelihoods have been considered.

One cycle of the EIMM algorithm for tracking with N models is outlined in the following 5 steps:
**Step 1:** Mixing of State Estimates

Starting with *a priori* state estimates $\hat{X}^i_{k-1}$, one computes the mixed state estimate and the mixed state error covariance as follows:

$$\hat{X}^{0j}_{k-1} = \sum_{i=1}^{N} \hat{X}^i_{k-1} J_{k-1}(ij) \tag{4}$$

where

$$J_{k-1}(ij) = p_{ij}\mu^i_{k-1}/\bar{c}_j, \qquad \bar{c}_j = \sum_{i=1}^{N} p_{ij}\mu^i_{k-1} \tag{5}$$

and $p_{ij}$ is the assumed transition probability for switching from model $i$ to model $j$, i.e., the $(i,j)$ element of the model switching matrix $\Pi_{k-1}$.

$$P^{0j}_{k-1} = \sum_{i=1}^{N}[P^i_{k-1} + (\hat{X}^i_{k-1} - \hat{X}^{0j}_{k-1})$$

$$(\hat{X}^i_{k-1} - \hat{X}^{0j}_{k-1})^T]J_{k-1}(ij) \tag{6}$$

**Step 2:** Model State Updates

The EKFs provide the model updates. The values of $\hat{X}^{0j}_{k-1}$ and $P^{0j}_{k-1}$ are used as input to the EKF matched to model $j$ to yield $\hat{X}^j_k$ and $P^j_k$.

**Step 3:** Model Likelihood Computations

Model Likelihood Computations are based on the filter residuals, the covariance of the filter residuals, and the assumption of Gaussian statistics.

$$\Lambda^j_k = \frac{1}{\sqrt{|2\pi T^j_k|}} \exp[-0.5(\tilde{Z}^j_k)^T (T^j_k)^{-1} \tilde{Z}^j_k] \tag{7}$$

where $\tilde{Z}^j_k$ is the filter residuals, and $T^j_k$ the covariance of the filter residuals.

**Step 4:** Model Probabilities Update

The model probabilities are updated according to model likelihoods and the model switching matrix governed by the Markov chain.

$$\mu_k(j) = \frac{\Lambda^j_k \bar{c}_j}{\sum_{i=1}^{N} \Lambda^i_k \bar{c}_i} \tag{8}$$

**Step 5:** Quantitative and Qualitative Motion Estimates

The quantitative motion estimates are obtained from a probabilistic sum of the individual filter inputs:

$$\hat{X}_k = \sum_{i=1}^{N} \hat{X}^i_k \mu^i_k \tag{9}$$

$$P_k = \sum_{i=1}^{N} \mu^i_k [P^i_k + (\hat{X}^i_k - \hat{X}_k)(\hat{X}^i_k - \hat{X}_k)^T] \tag{10}$$

The qualitative motion estimate or the motion labeling is based on the model probabilities:

$$Label_k = \begin{cases} (sign(s), 0, 0) & \text{if } m = 1 \\ (0, sign(\dot{s}), 0) & \text{if } m = 2 \\ (0, 0, sign(\dot{\theta})) & \text{if } m = 3 \end{cases} \tag{11}$$

where

$$m = \underset{j}{argmax}\, \mu^j_k \tag{12}$$

## 5 Experimental Results

Experiments on simulated and real range data were conducted to investigate the performance of our tracking system. The real range images were obtained from the laser scanner mounted on Navlab5 travelling on Rte. 376(Figure 3) ( The Navlabs are the experimental platforms for the Automated Highway System (AHS) research being conducted at Carnegie Mellon University.).

**Figure 3.** the Navlab5 testbed

The frame rate of the laser scanner is 0.16 seconds/frame, while the EIMM algorithm runs parallel to the range image grabbing process at 0.05 seconds/frame, so the whole tracking system runs in real time on a Pentium Pro PC. Before conducting experiments, we need to choose the parameters for the prior model probabilities $\mu_0$ and the model switching probability matrix $\mu$. In this case, we empirically chose $\mu_0 = (0.8, 0.2, 0.0)$ and

$$\mu = \begin{bmatrix} 0.85 & 0.14 & 0.01 \\ 0.20 & 0.70 & 0.10 \\ 0.10 & 0.20 & 0.70 \end{bmatrix}.$$

One of our key aims in this work is to detect car maneuvers, especially lane changing, at the very beginning of the maneuver execution. Figure 4 illustrates the recovered trajectory of a tracked car changing lanes once, and the corresponding model probabilities. It can be seen that our tracking system can detect the model switching reliably, and select the correct model at each time step. This also reflects the accuracy of the quantitative estimates on which motion classification or qualitative interpretation is based.

The bouncing of the observer vehicle can result in missing data of the tracked car. In this case, the tracker relies completely on the prediction (Figure 5: dotted box) from the EIMM algorithm for motion estimation. Figure 5 shows that the tracker can still keep on tracking when the tracked car was missed in Frames 513 and 514. The EIMM algorithm predicted the position of the car accurately enough that the car was re-acquired in Frame 515. It demonstrates the robustness of our tracking system.

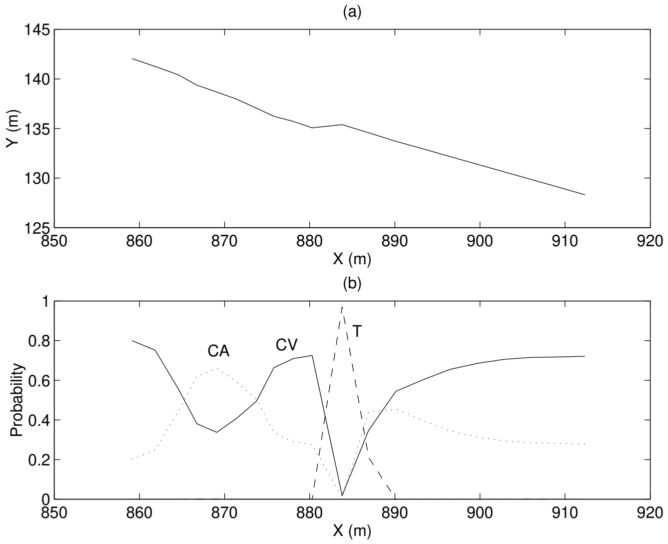


**Figure 4.** (a) the trajectory of the tracked car (in the world coordinate system) (b) model probabilities

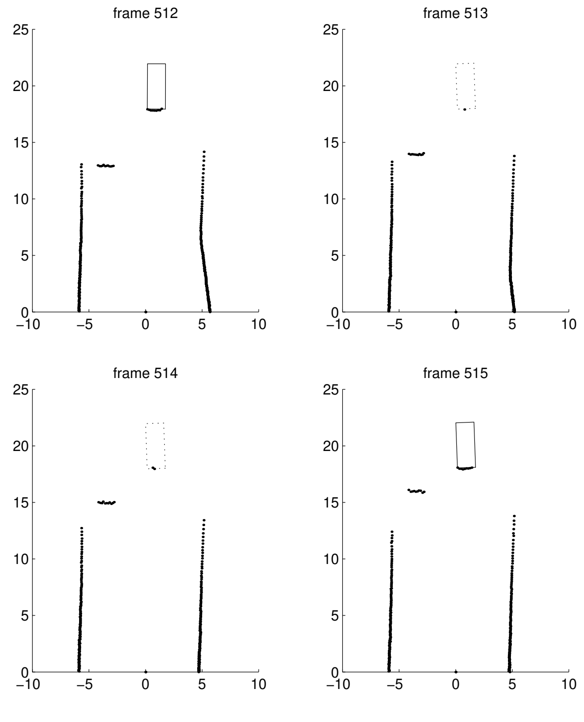


**Figure 5.** Car tracking over missing data of the tracked car

The tracking system was tested on simulated data to evaluate its accuracy. The simulations consisted of 100 Monte Carlo runs using the same filtering parameters and the model switching probabilites as those used in real experiments. Since the state estimate errors are relatively very small to the true state values, we plot only the state Root Mean Square (RMS) errors in Figure 6. The curves presented in Figure 6 lead to the following conclusion: the errors of the estimated states including position, direction of travel, speed, turning rate,and acceleration are quite low.

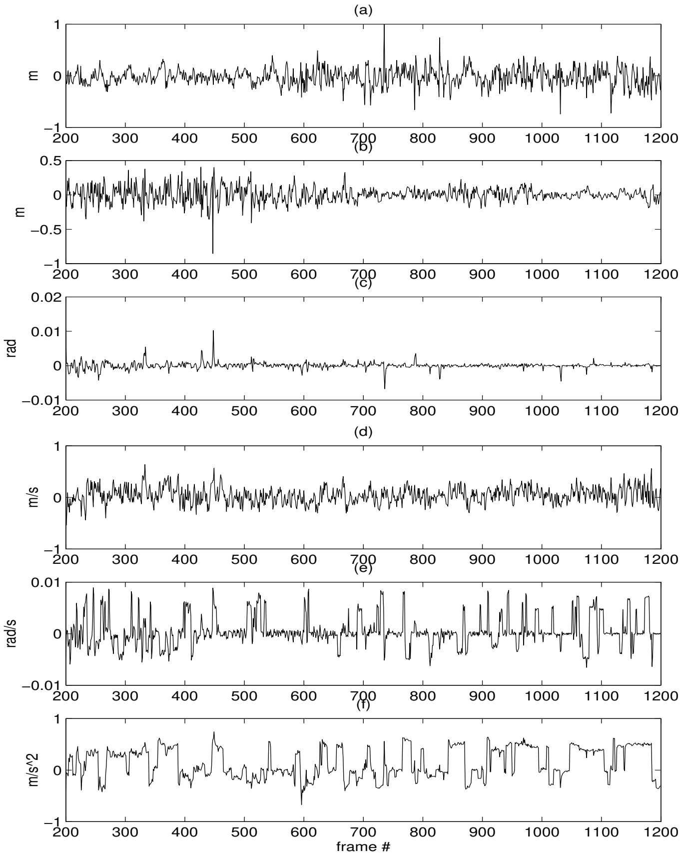


**Figure 6.** RMS errors of estimation on: (a) $x$ (b) $y$ (c) $\theta$ (d) $s$ (e) $\dot{\theta}$ (f) $\dot{s}$

## 6 Conclusions and Future Work

We have presented a range image-based real-time car tracking system. In contrast to previous tracking systems which used a single motion representation, our tracking system interprets the kinematic behaviors of the tracked car both quantitatively and qualitatively. This has been achieved by classifying the car behaviors into three significant motion modes distinguished by three motion models competing in the framework of the EIMM algorithm. Thanks to the accurate motion modeling and efficient model management, the resulting system provides accurate motion estimates, motion classification and reliable maneuver detection.

Its multiple motion representation makes the tracking system useful in various autonomous driving tasks involving low level control or high level reasoning.

In the current EIMM algorithm, the probability $p_{ij}$ of switching from model $i$ to model $j$ is assumed to be uniform between each measurement update. However, for a tracked car moving under changing traffic situations, for a sensor with missed data points or for multiple sensors operating asynchronously, the probability of maneuvers may differ between measurements. Therefore, we are constructing time-dependent model switching probabilities to overcome this limitation.

Although the results presented in this paper used simulated or 2D range data, the EIMM algorithm is independent of the image nature of the underlying data. It is readily applicable to the tracking systems based on intensity images or 3D range images. In fact, we are investigating car tracking from sequences of 3D range images, so that the ground plane assumption can be removed. We also plan to incorporate the information from the intensity images into our tracking system in order to distinguish the lane changing maneuver from curved road following behavior by taking road curvature into consideration. Our tracking system can also track multiple vehicles simultaneously, and we will addess this problem specifically in the future.

Considering the various levels of tasks in an autonomous driving system, multiple motion representation is a very promising research topic. We hope that our current work and future work will result in a general car tracking system, enabling other autonomous driving tasks to be successfully accomplished and finally brings fully autonomous vehicles into reality.

### Acknowledgements

We would like to thank Bala Kumar, Michelle Bayouth, Jhon Kozar, Todd Jochem, Dave Duggins, Dean Pomerleau, Jhon Hancock and Parag Batavia for their help with the experiments on highways. Support for this research came from USDOT under Cooperative Agreement Number DTFH61-94-X-00001 as part of the National Automated Highway System Consortium.

## References

[1] C. Thorpe, "Mixed Traffic and Automated Highways," *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, Vol. 2, No. 8, Sep. 1997.

[2] W.F. Gardner, D.T. Lawton, "Interactive Model-Based Vehicle Tracking," *IEEE Trans. Pattern Recognition and Machine Intelligence*, Vol. 18, No. 11, pp. 1151-1121, 1996.

[3] M. Betke, E. Haritaoglu, L.S. Davis, "Multiple Vehicle Detection and Tracking in Hard Real-Time," *Technical Report CS-TR-3667*, University of Maryland, College Park, 1996.

[4] D.Beymer, P. McLauchlan, B. Coifman, and J. Malik, "A Real-Time Computer Vision System for Measuring Traffic Parameters," *Proc. Conf. Computer Vision and Pattern Recognition*, pp. 495-501, 1997.

[5] R.C. Nelson, "Qualitative Detection of Motion by a Moving Observer," *Int. J. of Computer Vision*, Vol. 7, No. 1, pp. 33-46, 1991.

[6] E. Framcois, P. Bouthemy, "Derivation of Qualitative Information in Motion Analysis," *Image and Vision Computing*, Vol. 8, No. 4, pp. 279-287, 1990.

[7] J. Gonzalez, R. Gutierrez, "Mobile Robot Motion Estimation from a Range Scan Sequence," *Proc. Int. Conf. on Robotics and Automation*, pp. 1034-1039, 1997.

[8] P.W. Smith, N. Nandhakumar, C-H Chien, "Object Motion and Structure Recovery for Robotic Vision Using Scanning Laser Range Sensors," *IEEE trans. on Robotics and Automation*, Vol. 13, No. 1, pp. 74-80, 1997.

[9] B.Sabata, J.K. Aggarwal, "Surface Correspondence and Motion Computation from a Sequence of Range Images," *Computer Vision and Image Understanding*, Vol. 63, No. 2, pp. 232-250, 1996.

[10] P. Maybeck, *Stochastic Models, Estimation and Control*, Vol. 1 Academic Press, New York, 1979.

[11] H.A.K. Blom, "An Efficient Filter for Abruptly Changing Systems," *Proc. 23rd IEEE Conf. on Decision and Control*, pp. 656-658, 1984.

[12] K.J. Bradshaw, I.D. Reid, D.W. Murray, "The Active Recovery of 3D Motion Trajectories and Their Use in Prediction," *IEEE Trans. Pattern Recognition and Machine Intelligence*, Vol. 19, No. 3, pp. 219-233, 1997.

[13] G.A. Watson, W.D. Blair, "IMM Algorithm for Tracking Targets that Maneuver through Coordinated Turns," *Proc. SPIE Signal and Data Processing of Small Targets*, Vol. 1698, pp. 236-247, 1992.

[14] P.M. Sharkey and D.W. Murray, "Coping with Delays for Real-Time Gaze Control," *Proc. SPIE Sensor Fusion VI*, Boston Mass., 1993.

[15] W.B. Thompson, J.K. Kearney, "Inexact Vision," *Proc. IEEE Workshop on Motion: Representation and Analysis*, pp. 15-21, 1986.

[16] T.D. Gillespie, *Fundamentals of Vehicle Dynamics,* Society of Automative Engineerings, Inc. 1992.