Title

Development Of Binocular Stereopsis For Vehicle Lateral Control, Longitudinal Control And Obstacle Detection

Authors

Malik, Jitendra
Taylor, Camillo J.
Mclauchlan, Philip
et al.

Publication Date

1997

CALIFORNIA PATH PROGRAM
INSTITUTE OF TRANSPORTATION STUDIES
UNIVERSITY OF CALIFORNIA, BERKELEY

Development of Binocular Stereopsis for Vehicle Lateral Control, Longitudinal Control and Obstable Detection

Jitendra Malik, Camillo J. Taylor, Philip McLauchlan, Jana Kosecka

California PATH Research Report
UCB-ITS-PRR-97-41

This work was performed as part of the California PATH Program of the University of California, in cooperation with the State of California Business, Transportation, and Housing Agency, Department of Transportation; and the United States Department of Transportation, Federal Highway Administration.

Report for MOU 257

November 1997

ISSN 1055-1425

Development of Binocular Stereopsis for Vehicle Lateral Control, Longitudinal Control and Obstacle Detection

PATH MOU-257 Final Report

Department of Electrical Engineering and Computer Sciences
University of California at Berkeley
Berkeley, CA 94720

Principal Investigator: Prof. Jitendra Malik
Postdoctoral Researches: Camillo J. Taylor, Philip McLauchlan, Jana Košecká

Contents

List of Figures

Abstract

This report describes the progress that we have made in the application of computer vision techniques to the lateral and longitudinal control of an autonomous highway vehicle.

From the lateral control standpoint, we focus on an analysis of the vehicle's lateral dynamics and the design of an appropriate controller for lateral control. We investigate various static feedback strategies where the measurements obtained from vision, namely offset from the centerline and angle between the road tangent and the orientation of the vehicle at some look-ahead distance, are directly used for control. We explore the role of look-ahead, its relation to the vision processing delay, longitudinal velocity and road geometry and present experiments with our autonomous vehicle system along with simulation results.

For the longitudinal control problem, we investigated the possibility of using stereo vision to provide the range information, in conjunction with a scanning laser radar sensor. The vision based tracking system utilizes a layered architecture wherein the bottom layer computes motion in both images using a simple correlation algorithm, and the upper level performs stereo fixation and reconstruction using an algorithm designed for active vision systems. We present some initial results comparing the quality of range measurements provided by a vision system with the laser radar system.

We report the results from the experimental demonstration of the system as part of the National Automated Highway Systems Consortium (NAHSC) Demonstration which took place in August 1997 in San Diego. The overall system was demonstrated as a part of the main highway scenario as well as part of a small public demonstration of the vision based lateral control on a highly curved test track.

1 Introduction

This report describes the work and research results carried out under MOU-257. Within this proposal we explored the feasibility of the use of visual sensing as a part of the Advanced Vehicle Control System (AVCS). The basic theoretical foundation on which this work was based had been developed under a previous program [24]. In this phase of the research program we demonstrated improved versions of the proposed algorithms, real-time implementations and a novel stereo tracking algorithm for longitudinal control. The integration of the vision subsystem with the vehicle control subsystem enabled us to perform real-world experiments with vision as an integral component of the vehicles control system.

The experimental part of the work has been carried out in collaboration with Honda R&D North America Inc. and Honda R&D Company Limited who provided us with three Honda Accords and a team of Honda engineers responsible for the low-level control of steering and throttle actuators as well as maintenance and overall organization leading towards the final demonstration which was part of NAHSC DEMO 1997 in San Diego.

Several sensing technologies have been proposed for use in an Advanced Vehicle Control System, including vision, magnetic sensors and active range sensors. Some of the most influential work on visually guided control of autonomous vehicles has been done by E.D.Dickmanns and his colleagues [26]. In their system vision was used to provide input for both lateral and longitudinal control of the vehicle on free roads as well as in the presence of other vehicles. The basic approach was to recursively estimate a set of road and vehicle state parameters which included the horizontal and vertical curvature of the road, the vehicle heading angle, the slip angle and the lateral offset of the vehicle with respect to the road. A dynamical model of the vehicle captured their knowledge about the motion of the vehicle and served as a tool for both fusion of sensor data, design of control strategies and prediction of the effects of control input on the evolution of the measured data.

In the United States one of the leading efforts has been the CMU NavLab Project [35] which has employed a variety of different sensing and control strategies including a neural network based lane following algorithm called ALVINN. A number of other groups throughout the country explored the possibility of using visual sensing for vehicle guidance both in outdoor and indoor environments [36, 33] concentrating primarily on the lane keeping or path following modality on the free roads.

For vision-based lateral control we undertook to explore a strategy which uses direcly the information from the vision at some look-ahead distance. Taking into account the vehicle dynamic model, we formulated the vision dynamics as well as the control objective at the look-ahead distance.

This resulted in the development of simpler control strategies Within this setting we evaluated a variety of feedback strategies including lead-lag control, state feedback via pole-placement and I/O linearization. We also implemented observer schemes which provided estimates for the curvature of the roadway which allowed us to incorporate a feedforward component into our control laws. This feedforward component was particularly crucial on highly curved roads.

In case of vision-based algorithms for longitudinal control, in spite of the fact that the context of highways and vehicles is clearly very structured, we avoided using direct scene models in the low-level tracking algorithms, and this distinguishes our work from that of Dickmanns's group [5], for instance. We draw on the large amount of work on scene reconstruction from multiple images in unstructured scenes, in particular the work on robust motion segmentation [22], and affine reconstruction [20]. These approaches are able to take advantage of the redundant information in images, because they latch onto whatever features are available, whereas model-based methods are restricted to the features associated with the chosen model. Redundancy is a vital issue here, because vision is a massively redundant sensor, and approaches which negate this aspect are likely to be discarded in the long term. Because we reconstruct the geometry of the lead vehicle, the algorithms generalize naturally to other vehicle types.

Section 1 outlines the model of the lateral dynamics of the vehicle, the analysis of the vision based lateral control problem and the controller design. In Section 2 we describe the algorithm for tracking the lead vehicle together and provide the results from off-line range estimation as well as from real-time implementation of the tracking algorithm. The last section outlines the overall scenario demonstrated in NAHSC DEMO in San Diego in August 1997.

2 Vision-Based lateral Control

This section addresses the problem of designing control systems for steering a motor vehicle along a highway using the output from a video camera mounted inside the vehicle. Several aspects of this problem have been examined extensively in the past, both in the psychophysics literature [31] as well as in control theoretic studies. In the kinematic setting there have been several attempts to formulate the vision-based steering task in the image plane [36, 27]. A stability analysis was provided for an omnidirectional mobile base trying to align itself with a straight road [27] or non-holonomic mobile base following an arbitrary ground analytic curve [32]. The controllers designed based on kinematic models were either tested in simulation or in experiments at speeds below 20 m/s. However at higher speeds dynamic effects are quite pertinent and the need for a dynamic model becomes apparent.

The control problem in a dynamic setting, using measurements ahead of the vehicle, has been explored by [33] who proposed a constant control law proportional to the offset from the centerline at a look-ahead distance. Their analysis showed that closed loop stability for this controller can always be obtained by increasing the look-ahead distance to an appropriate value. Dickmanns, et al [26] developed a Kalman-filter based observer which estimated the state of the vehicle with respect to the road along with the road geometry and used the estimate for full state feedback using a pole-placement method. Further studies typically use a small and fixed look-ahead distance and the control objective is formulated either at the look-ahead distance [29] or at the center of gravity of the vehicle [34]. An analysis of the tradeoffs between the performance requirements and robustness of the system can be found in [29].

We discuss the problem of automated steering using computer vision, focusing on the analysis of the problem and controller design choice. We propose a static feedback strategy where the measurements obtained from vision, namely offset from the centerline and angle between the road tangent and the orientation of the vehicle at some look-ahead distance, are directly used for control. Within this setting we explore the role of lookahead, its relation to the vision processing delay, longitudinal velocity and road geometry.

2.1 Modeling

The dynamics of the vehicle can be described by a detailed 6-DOF nonlinear model [34]. Since it is possible to decouple the longitudinal and lateral dynamics, a linearized model of the lateral vehicle dynamics is used for controller design. The linearized model of the vehicle retains only lateral and yaw dynamics, assumes small steering angles and a linear tire model, and is parameterized by the

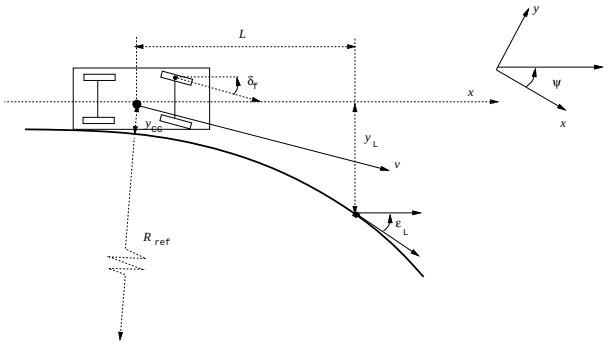


Figure 1: The vision system estimates the offset from the centerline y_L and the angle between the road tangent and heading of the vehicle ε_L at some look-ahead distance L.

current longitudinal velocity. Coupling the two front wheels and two rear wheels together, the resulting bicycle model is described by the following variables and parameters: $\boldsymbol{v}$ linear velocity vector $(v_x,\, v_y)$, v_x denotes speed, α_f, α_r side slip angles of the front and rear tires, ψ vehicle yaw angle within a fixed inertial frame, δ_f front wheel steering angle, δ commanded steering angle, m total mass of the vehicle, I_ψ total inertia vehicle around center of gravity (CG), l_f, l_r distance of the front and rear axles from the CG, l distance between the front and the rear axle $l_f + l_r$, c_f, c_r cornering stiffness of the front and rear tires.

The lateral dynamics equations are obtained by computing the net lateral force and torque acting on the vehicle following Newton-Euler equations [30] and choosing $\dot{\psi}$ and v_y, as state variables. The state equations have the following form:

$$\begin{bmatrix} \dot{v}_y \\ \\ \ddot{\psi} \end{bmatrix} = \begin{bmatrix} -\frac{a_1}{m v_x} & \frac{-m v_x^2 + a_2}{m v_x} \\ \\ \frac{a_3}{I_\psi v_x} & -\frac{a_4}{I_\psi v_x} \end{bmatrix} \begin{bmatrix} v_y \\ \\ \dot{\psi} \end{bmatrix} + \begin{bmatrix} b_1 \\ \\ b_2 \end{bmatrix} \delta_f \tag{1}$$

where $a_1 = c_f + c_r$, $a_2 = c_r l_r - c_f l_f$, $a_3 = -l_f c_f + l_r c_r$, $a_4 = l_f^2 c_f + l_r^2 c_r$, $b_1 = \frac{c_f}{m}$ and $b_2 = \frac{l_f c_f}{I_\psi}$.

Vision Dynamics. The additional measurements provided by the vision system (see Figure 1) are: y_L the offset from the centerline at the look-ahead, ε_L the angle between the tangent to the road and the vehicle orientation K_L the curvature of the road at the look-ahead. L denotes the look-ahead distance. The equations capturing the evolution of these measurements due to the motion of the car and changes in the road geometry are:

$$\dot{y}_L = v_x\, \varepsilon_L - v_y - \dot{\psi} L \tag{2}$$

$$\dot{\varepsilon}_L = v_x\, K_L - \dot{\psi} \tag{3}$$

We can combine the vehicle lateral dynamics and the vision dynamics into a single dynamical system of the form:

$$\begin{aligned} \dot{\boldsymbol{x}} &= A\,\boldsymbol{x} + B\,\boldsymbol{u} + E\,\boldsymbol{w} \\ \boldsymbol{y} &= C\,\boldsymbol{x} + D\,\boldsymbol{u} + F\,\boldsymbol{w} \end{aligned}$$

with the state vector $\boldsymbol{x} = [v_y, \dot{\psi}, y_L, \varepsilon_L]^T$, the output $y = [\dot{\psi}, y_L, \varepsilon_L]^T$ and control input $\boldsymbol{u} = \delta_f$, disturbance $\boldsymbol{w} = K_L$ and matrices:

$$A = \begin{bmatrix} -\frac{a_1}{m v_x} & \frac{-m v_x^2 + a_2}{m v_x} & 0 & 0 \\ \frac{a_3}{I_\psi v_x} & -\frac{a_4}{I_\psi v_x} & 0 & 0 \\ -1 & -L & 0 & v_x \\ 0 & -1 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ 0 \\ 0 \end{bmatrix} \quad E = \begin{bmatrix} 0 \\ 0 \\ 0 \\ v_x \end{bmatrix}$$

The output equations have following form:

$$y = C\,\boldsymbol{x} \text{ where } C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

The road curvature K_L enters the model as an exogenous disturbance signal.

2.2 Lane Tracking

The vision-based lane tracking system used in our experiments takes its input from a single forward-looking CCD video camera. It extracts potential lane markers from the input using a simple template-based scheme. It then finds the best linear fits to the left and right lane markers over a certain lookahead range through a variant of the Hough transform. From these measurements we can compute an estimate for the lateral position and orientation of the vehicle with respect to the roadway at a particular lookahead distance, L.

The vision system is implemented on an array of TMS320C40 digital signal processors which are hosted on the bus of an Intel-based industrial computer. The system processes images from the video camera at a rate of 30 frames per second. The total delay between the time the shutter on the CCD video camera is closed and the time the measurements for that image are available to the control computer is 57 milliseconds. Since the delay is quite substantial we will explicitly consider it in the controller design.

Figure 2: Camera's view of the Honda Accord used in experiments

2.3 Analysis

The block diagram of the overall system following the state equations is shown in Figure 3. The transfer function $V_1(\boldsymbol{s})$ between the steering angle δ_f and offset at the look-ahead y_L can be obtained by taking a Laplace transform of the state equations and has the following form:

$$V_1(\boldsymbol{s}) = \frac{1}{\boldsymbol{s}^2} \frac{a\boldsymbol{s}^2 + b\boldsymbol{s} + c}{d\boldsymbol{s}^2 + e\boldsymbol{s} + f} \tag{5}$$

where the numerator is a function of both speed and lookahead distance and the denominator is parameterized by the speed of the car. $V_1(\boldsymbol{s})$ can be rewritten according to Figure 3 by singling out the vehicle dynamics in terms of $\ddot{y}_{CG}$ and $\ddot{\psi}$ followed by the integrating action $1/s^2$:

$$V_1(\boldsymbol{s}) = \frac{1}{\boldsymbol{s}^2}(G(\boldsymbol{s}) + L\, G_2(\boldsymbol{s})) \tag{6}$$

where $G(\boldsymbol{s})$ and $G_2(\boldsymbol{s})$ are transfer functions between steering angle and lateral acceleration and yaw acceleration respectively. There are two additional components which appear in the block diagram. The actuator $A(\boldsymbol{s})$ is modeled as a low pass filter of the commanded steering angle δ and a pure time delay element $D(\boldsymbol{s}) = e^{-T_d s}$ representing the latency T_d of the vision subsystem. In our system $T_d = 0.057$ s. The transfer function $C(\boldsymbol{s})$ corresponds to the controller to be designed.

Control objective. The vehicle control objective is to follow the reference path specified by radius R_{ref} (curvature $K_{ref} = \frac{1}{R_{ref}}$). Perfect tracking of the road in steady state corresponds to the zero offset $y_{CG} = 0$ of the vehicle's center of the gravity from the centerline, with the orientation of the vehicle aligned with the tangent to the road. The speed is chosen such as not to exceed lateral acceleration of 0.3-0.4g, where g = 9.81 m/s^2, which has been shown to be comfortably accepted by humans. In addition to limits on the steady-state lateral acceleration an important design criterion

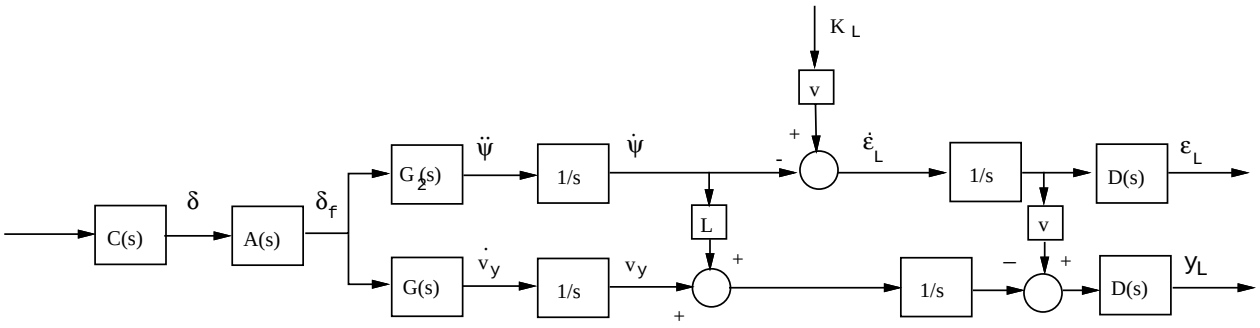


Figure 3: The block diagram of the overall system with the two outputs provided by the vision system.

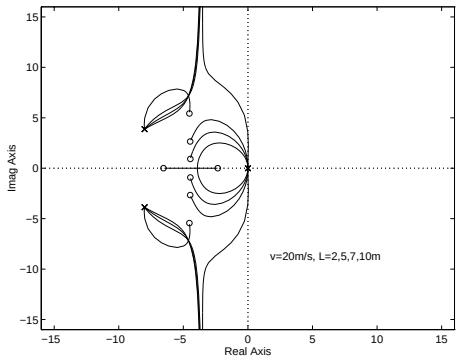


Figure 4: Increasing the lookahead distance L moves the zeros of the transfer function $V_1(s)$ closer to the real axis, which improves their damping.

is that of passenger comfort. This is typically expressed in terms of jerk, corresponding to the rate of change of acceleration. For a comfortable ride no frequency above 0.1-0.5 Hz should be amplified in the path to lateral acceleration [29]. Additional road following criteria can be specified in terms of maximal allowable offset y_{Lmax} as a response to the step change in curvature as well as bandwidth requirements on the transfer function $F(s) = \frac{y_L(s)}{K_{ref}(s)}$. Since the primary advantage of the vision system is the availability of measurements at a point ahead of the vehicle, we will analyze how the choice of the look-ahead distance L affects the transfer function $V_1(s)$ between steering angle and the offset at the look-ahead. The analysis will also take into account the processing delay T_d inherent in the vision system which substantially affects the stability of the system.

Lookahead. A root locus of the transfer function $V_1(s)$ is in Figure 4. The transfer function $V_1(s)$ has four poles and two zeros, where the damping of the zero pair affects the location of closed loop poles and subsequently the transient response of the system more profoundly. As the lookahead increases the zeros move closer to the real axis, improving the damping of the closed loop poles of $V_1(s)$. Increasing the velocity moves both poles and zeros of $V_1(s)$ towards the imaginary

axis, resulting in a poor damping of the poles. The choice of proper lookahead distance is therefore important for stability and performance of the system.

Delay. Another parameter which affects the behavior of the overall system is the delay associated with the vision system. The delay element adds an additional phase lag over the whole range of frequencies having a clear destabilizing effect on the overall system and limiting the system's bandwidth. More detailed analysis can be found in [30, 25].

Controller Design. Analysis reveals that up to 15 m/s the lookahead one can guarantee satisfactory damping of the closed loop poles of $V_1(\boldsymbol{s})$ and compensate for the delay using simple unity feedback control with proportional gain in the forward loop. As the velocity increases the transient response is affected more by the poor damping of the poles of $V_1(\boldsymbol{s})$ introducing additional phase lag around the 0.1-2 Hz. Since further increasing the lookahead does not improve the damping, gain compensation only cannot achieve satisfactory performance. The natural choice for obtaining an additional phase lead in the frequency range 0.1-2 Hz would be to introduce some derivative action. In order to keep the bandwidth low an additional lag term is necessary. One satisfactory lead-lag controller has the following form:

$$C(\boldsymbol{s}) = \frac{0.09\boldsymbol{s} + 0.18}{0.025\boldsymbol{s}^2 + 1.5\boldsymbol{s} + 20} \tag{7}$$

where $C(\boldsymbol{s})$ is a lead network in series with a single pole. The above controller was designed for a velocity of 30 m/s (108 km/h, 65 mph), a lookahead of 15 m and 60 ms delay. The resulting closed loop system has a bandwidth of 0.45 Hz with a phase lead of 45° at the crossover frequency. A discretized version of the above controller taking into account the 30 ms sampling time of the vision system have been used in our experiments.

Since increasing the speed has a destabilizing effect on $V_1(\boldsymbol{s})$, designing the controller for the highest intended speed guarantees stability at lower speeds and achieves satisfactory ride quality. In order to tighten the tracking performance at lower speeds individual controllers can be designed for various speed ranges and gain scheduling techniques used to interpolate between them.

The steady state behavior of the system during perfect tracking of a curve with radius R_{ref}, is characterized by particular values of $\dot{\psi}_{ref}, v_{yref}$ and δ_{ref}. By setting the $[\dot{v}_y, \ddot{\psi}, \dot{y}_L, \dot{\varepsilon}_L]^T$ to 0, the steering angle δ_{ref} can be obtained from state equations and becomes:

$$\delta_{ref} = K_{ref}\,(l - \frac{(l_f c_f - l_r c_r) v_x^2 m}{c_r c_f l})\,. \tag{8}$$

The feedforward control law essentially provides information about the disturbance ahead of the car and improves the transient behavior of the system when encountering changes in curvature.

The effectiveness of the feedforward term depends on the quality of the curvature estimates. We discuss the curvature estimation process as part of the observer design in the next section 2.4.

2.4 Observer Issues and Design

In order to apply modern state space control techniques we require access to the states of the system. This is usually accomplished by constructing an observer. Our first step is to rewrite state equations in the following form:

$$\dot{\boldsymbol{x}} = A(v_x)\boldsymbol{x} + B\delta_f \tag{9}$$

where $\boldsymbol{x} = [v_y, \dot{\psi}, y_L, \varepsilon_L, K_L]^T$ and

$$A(v_x) = \begin{bmatrix} -\frac{a_1}{mv_x} & \frac{-mv_x^2+a_2}{mv_x} & 0 & 0 & 0 \\ \frac{a_3}{I_\psi v_x} & -\frac{a_4}{I_\psi v_x} & 0 & 0 & 0 \\ -1 & -L & 0 & v_x & 0 \\ 0 & -1 & 0 & 0 & v_x \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Note that the state vector $\boldsymbol{x}$ has been augmented with the road curvature K_L since we are interested in estimating this parameter as well. This differential equation can be converted to discrete time in the usual manner by assuming that the control input, δ_f, is constant over the sampling interval T.

$$\boldsymbol{x}(k+1) = \Phi(v_x)\boldsymbol{x}(k) + \beta\boldsymbol{u}(k) \tag{10}$$

Equation (10) allows us to predict how the state of the system will evolve between sampling intervals.

Measurements of the system state can be obtained from two sources: the vision system provides us with measurements of y_L and ε_L, while the on-board fiber optic gyro provides us with measurements of the yaw rate of the vehicle, $\dot{\psi}$. Our use of the yaw rate sensor measurements is analogous to the way in which information from the proprioceptive system is used in animate vision. Considering $\boldsymbol{y} = \left[\dot{\psi}, y_L, \varepsilon_L\right]^T$ the measurement equation for our system can be written as in Equation 4:

$$\boldsymbol{y} = C\boldsymbol{x} \tag{11}$$

The measurement vector $\boldsymbol{y}$ is used to update an estimate for the state of the system $\hat{\boldsymbol{x}}$ as shown in the following equation:

$$\hat{\boldsymbol{x}}^+(k) = \hat{\boldsymbol{x}}^-(k) + L(\boldsymbol{y}(k) - C\hat{\boldsymbol{x}}^-(k)) \tag{12}$$

where $\hat{\boldsymbol{x}}^-(k)$ and $\hat{\boldsymbol{x}}^+(k)$ denote the state estimate before and after the sensor update respectively.

The gain matrix L can be chosen in a number of ways [28], depending on the assumptions one makes about the availability of noise statistics and the criterion one chooses to optimize.

3 Vision for longitudinal vehicle control

Another modality of the automated vehicle is a longitudinal control system which combines both laser radar and vision sensors, enabling throttle and brake control to maintain a fixed distance to a lead car. Currently, steering control can be driven by a combination of lane tracking described in the previous section and magnetometers detecting magnetic "nails" in the road, and longitudinal control achieved using the laser radar sensor. In our work we explored possibility of using vision which can potentially provide higher bandwidth (30/60Hz) output than is available from the laser radar system [9].

Our approach combines shape reconstruction (2D planar reconstruction rather than the usual 3D) from stereo/motion with motion estimation, using recently developed robust and efficient feature matching methods. The resulting algorithm runs at 3-5Hz, and frame-rate performance (30Hz) is achieved by combining this with a normalized cross-correlation algorithm for stereo and motion computation, which runs in parallel. Employing these quite sophisticated tools we are capable to effectively track a vehicle over an extended time, i.e. a period measured in minutes rather than frames.

We shall assume that the viewed vehicle is an affine projection of a planar object. We shall not assume any knowledge of the camera calibration for the purpose of tracking the vehicle. Thus we have possibly the simplest available imaging model. The lead vehicle presents its rear end to the cameras, with little change in orientation, which might otherwise induce the 3D cues that we are ignoring. The depth relief is also small, since only the rear of the car is visible. The smallest range we are considering in our experiments is 10m, so that cars of size 2m will subtend an angle of at most 10^{o}, justifying the assumption of parallel projection from scene to image. The results at the end of this paper will demonstrate that these minimalist assumptions are appropriate.

3.1 Visual tracking

In order to track an object over an extended time, it is necessary to compute the *position* of the object, which in this context may only be computed by integrating the velocity over time. Since there are inevitable errors in the computed velocities, these errors will tend to accumulate over time. Thus we can expect the computed position to drift. The drift problem can be eliminated in correlation trackers by fixing the template used for correlation, converting it into a position-based tracker; however then the tracker will work only for a short time, because relatively small motions, especially rotations, will break the tracker, although Shi & Tomasi [18] suggest a partial solution to that problem through an affinely deformable template. Correlation with a fixed template forms

the lower level of our tracker, and we use it in such a way that a single template is used for only a short period of time.

Feature tracking algorithms have the capability to allow the position of an object to be accurately estimated over an extended time. This aspect is of vital importance in the context of sensing for control, where the sensor is required to return accurate error feedback during the whole period of the control task. Moreover we have demonstrated in previous work [1] that vehicle tracking using features can be made robust both to partial occlusion of the vehicles and to lighting changes in the environment. The major problems that have to be overcome when tracking features are the fragmentary nature of the data (features appear, disappear and change shape) and the integration of feature data from multiple images in a statistically valid manner. We now describe briefly some of the problems. More detailed description can be found in [13].

3.2 Fixation/Scene Reconstruction

In [17] a fixation technique was described that allows a single "fixation point" to be chosen from a cluster of tracked features, in a way that is robust to losing track of individual features, while allowing the same *object* point to be fixated over time.

While we have a simpler problem, with a fixed camera and a 2D scene representation, similar principles apply. The "fixation point" here refers to the point chosen as representative of the vehicle for the purpose of estimating its position, and hence its range. Maintaining scene structure explicitly within a reconstruction algorithm stabilizes the computation of motion over time, so a logical extension of the fixation point transfer algorithm is to recursively update the structure of the tracked object, and employ the improved motion estimates to perform fixation point transfer. This is the method we have implemented. The reconstruction technique detailed below is a 2D affine version of the Variable State Dimension Filter (VSDF) algorithm [14]. The VSDF is a general algorithm for visual reconstruction that deals naturally with fragmentary data and combines data from multiple images in a near-optimal manner.

3.3 2D Affine Reconstruction

The 2D affine projection from scene to image can then be written in its most general form as

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} M_{1X} & M_{1Y} \\ M_{2X} & M_{2Y} \end{pmatrix} \begin{pmatrix} X \\ Y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad \text{or} \quad \mathbf{z} = M\mathbf{X} + \mathbf{t}. \tag{13}$$

Here $\mathbf{z}$ is the image point, $\mathbf{X}$ is a 2D scene point, and M, $\mathbf{t}$ constitute the camera matrix parameters, which we will term the *motion* parameters because they represent the camera/scene

motion over time. The 2D affine reconstruction problem for point features is: given $\mathbf{z}_i^{(j)}$ for multiple features i in multiple images j, determine the motion $M^{(j)}$, $\mathbf{t}^{(j)}$ and structure $\mathbf{X}_i$. We employ the variable state dimension filter algorithm which achieves virtually the same accuracy as previously used batch algorithms [20], but has the advantages of being recursive, not requiring complete data, and allowing new features to be added to the reconstruction as they appear and discarded features to be removed.

Given the recursively computed scene reconstruction, the transfer of the fixation point into the new stereo pair with computed motion parameters $M_l, \mathbf{t}_l$ and $M_r, \mathbf{t}_r$ is simply $\mathbf{z}_{fl} = M_l\mathbf{X}_f + \mathbf{t}_l, \quad \mathbf{z}_{fr} = M_r\mathbf{X}_f + \mathbf{t}_r$. The transferred positions $\mathbf{z}_{fl}$, $\mathbf{z}_{fr}$ are now converted via triangulation to 3D world coordinates to be interpreted as a range measurement. We choose $\mathbf{X}_f$ to be the centroid of the structure vectors computed from the initial batch computation of the 2D reconstruction.

3.4 Stereo/Temporal Matching

A major issue with all reconstruction techniques is their reliance on high quality, essentially outlier-free input data. The method that has recently been proposed to achieve this is to select a set of feature matches that are globally consistent, in the sense of satisfying the rigidity constraint.

We follow [21] and apply the RANSAC algorithm of Fischler & Bolles [7] to compute a large subset of feature matches consistent with a single set of 2D transformation parameters. Part of the stereo matching algorithm is also an algorithm for enforcing uniqueness of individual matches. Both algorithms are described in more detail in [13].

3.5 Layered Tracking

The frame-rate (30Hz) performance of the tracker results from the coordination of two separate tracking algorithms, a frame-rate correlator which computes the motion in both images, and the corner feature-based fixation algorithm, which runs at 3-5Hz depending on the size of the region used for the corner detection (maximum 140×100 pixels). The two processes run on separate C40s, and are coordinated in such a way that the correlator is always using an image region centered around the latest fixation point for the correlation. The two processes communicate whenever the fixation algorithm has any results to pass on. In addition, the laser radar provides the fixation algorithm with the initial bounding boxes around the vehicle, so the radar may be considered a third layer of the tracker. This provides the system with a great deal of robustness. If the correlator fails for any reason (usually due to not finding a motion with a high enough correlation score) it simply waits for the fixation algorithm to provide it with a new template/position pair. If the

fixation algorithm fails, then it also must wait for the laser to provide it with a new bounding box pair.

4 Experimental Results

Both modalities were tested in the simulation and in the real experiments. In the simulation of the lane keeping modality the full nonlinear model of the vehicle has been used and the lateral controler design has been tested for various road scenarios (see Figure 5). In simulation the maximum offset did not exceed 10 cm and the lateral acceleration was within passenger comfort standards.

With the introduction of the lead-lag controller and an observer to filter noise from our measurements, we were able to take our experimental vehicle to speeds of 100 mph for extended periods of time without any degradation in passenger comfort. The testing was performed on the oval track at HPCC (Honda Proving Center of California) near Mojave which consists of two straight line segments connected by curved segments with the 1200 m radius of curvature. Further increase of the speed up to 110 mph resulted in an uncomfortable ride for passengers, due in large part to the increased lateral acceleration. We also evaluated various control strategies including state-feedback using pole placement method and I/O linearization via state feedback. The tracking performance in the curved segments improved in all cases with the addition of the feedforward control component.

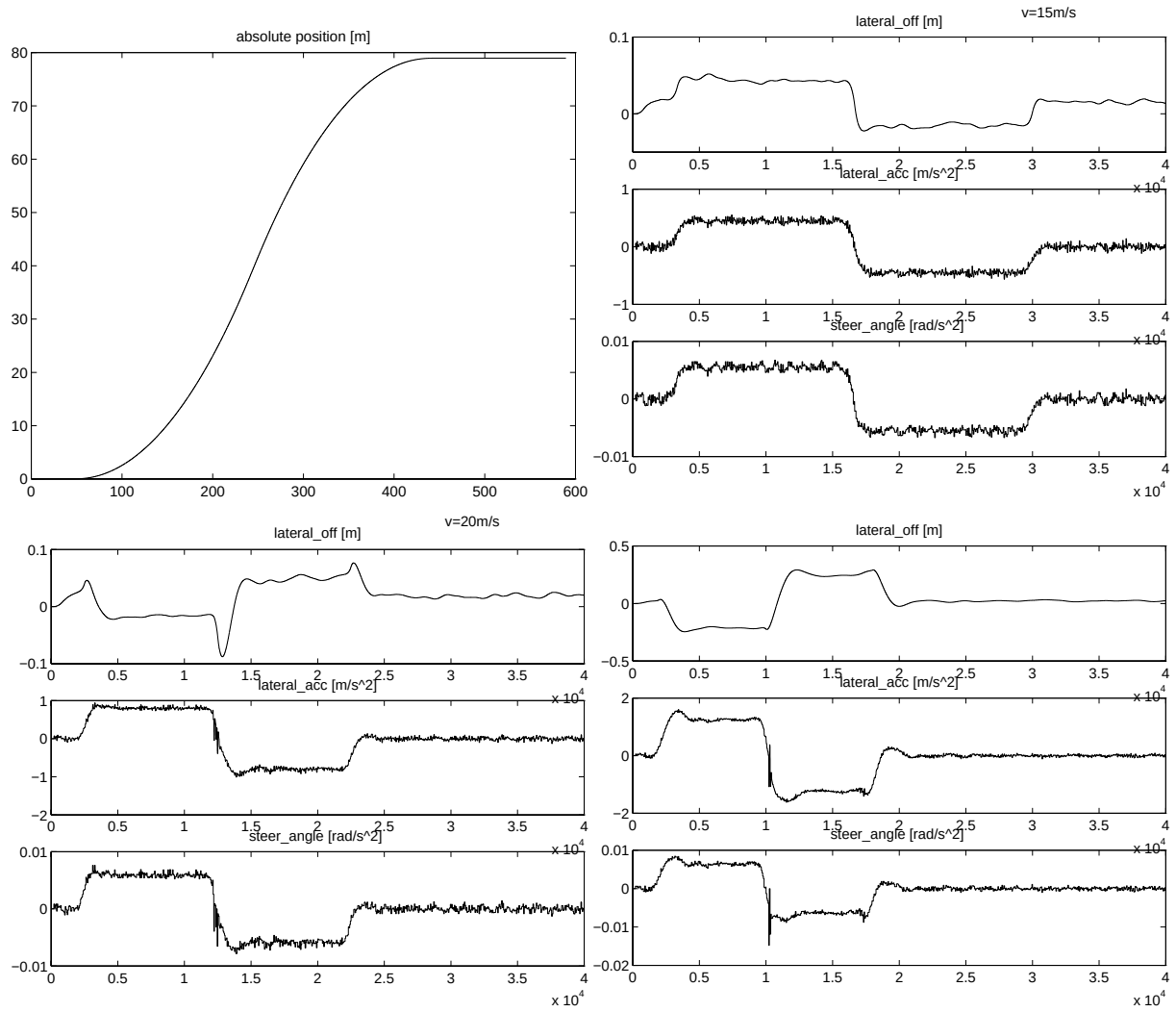


Figure 5: Tracking changes in curvature without intermediate straight line segments for various velocities. (a) reference path with straight line segment, followed by two curved segments with K_{1ref} = 0.002 m^{-1} and K_{2ref} = -0.002m^{-1}. (b) v = 15m/s (c) v = 20m/s (d) v = 25m/s. The look-ahead distance used in all experiments was L = v 0.9s.

Figure 6: Honda Accord used in experiments.

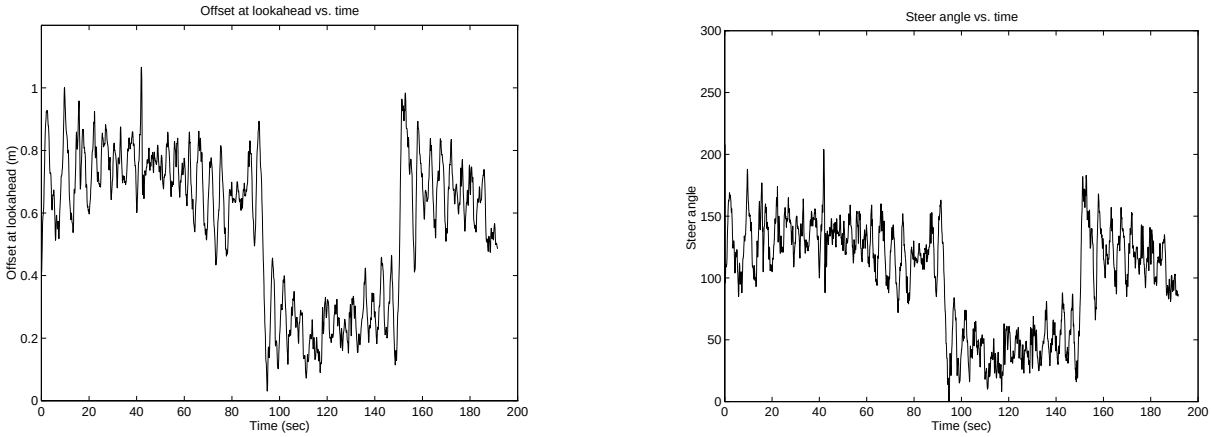


Figure 7: (a) The offset at the look-ahead y_L used for control purposes. (b) Commanded steering angle.

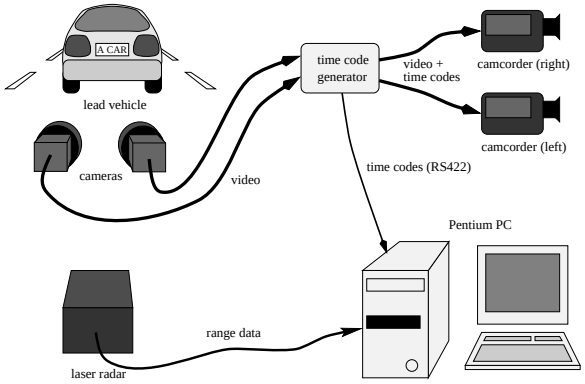
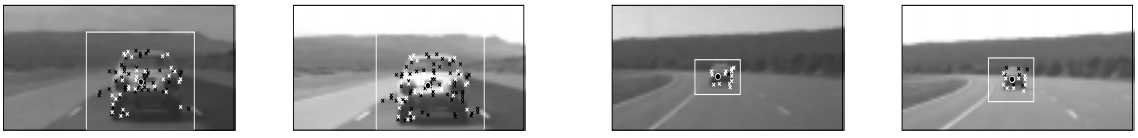


Figure 8: The experimental setup.

The experimental setup for the vision-based tracker for longitudinal control is illustrated in Figure 8. The off-line version of the tracking algorithm was tested on approximately 20 minutes of synchronized video and laser radar [9]. We selected an initial window surrounding the lead vehicle, although subsequent processing was completely automatic. Figure 9 show some example images, with the tracking results superimposed. The corner features are shown as small crosses, white for those matched over time or in stereo, and black for unmatched features. The black and white circle indicates the position of the fixation point, which ideally should remain at the same point on the lead car throughout the sequence. The white rectangle describes the latest estimate of the bounding box for the vehicle.

We have attempted here to summarize significant aspects of our data. Images 1 and 2 show the first stereo pairs in the sequence, where the vehicle is close (17m) to the camera and range estimates from stereo disparity may be expected to be accurate. By contrast images 421 and 422 were taken when the vehicle was 60m from the camera (the greatest distance achieved during the sequence). Here we can predict that depth estimates from stereo will be unreliable, since the disparity relative to infinity is only a few pixels and so difficult to measure. It will still be feasible to use the change in apparent size measured by the motion processing to obtain reasonable range estimates. We computed the range and bearing estimated from the laser radar range finder and



1 **421...**

Figure 9: Example stereo-pairs from the tracking sequence.

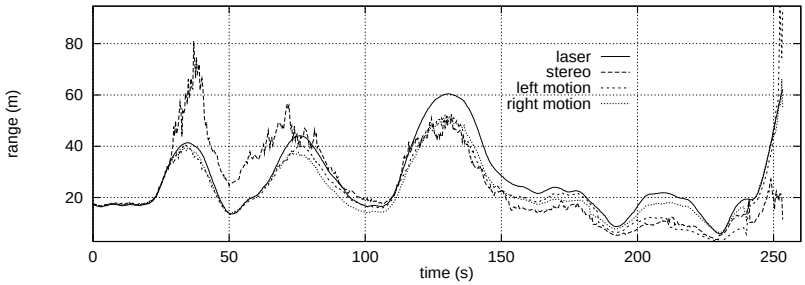


Figure 10: Comparison of range estimates from laser radar and vision.

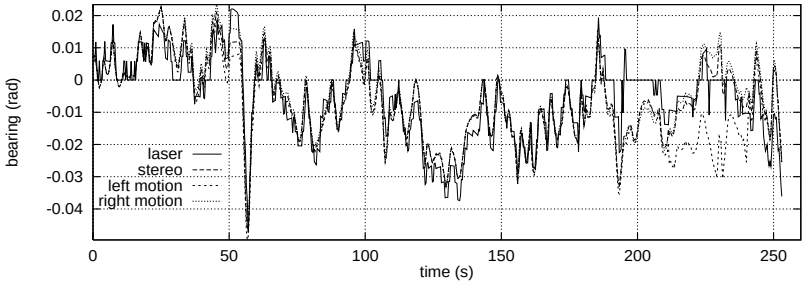


Figure 11: Comparison of bearing estimates from laser radar and vision.

plot them together with the corresponding data collected from the vision algorithms in figures 10 and 11. Depth from stereo is computed by inverting the projection of the fixation point at each image pair and finding the closest point of intersection of the two resulting space rays. The cameras are very roughly calibrated.

The results for the real-time version of the algorithm, in a simple scenario wherein the lead car is stationary and the following car moves backwards and forwards, are shown in figure 12. Depth from stereo is computed by inverting the projection of the fixation point at each image pair and finding the closest point of intersection of the two resulting space rays. Figure 12 shows the comparison of the vision and laser radar data in a 200 second segment of data. There is clearly an offset between the graphs, due to a calibration error, which we could correct fairly easily. The vision data is quite noisy; this is partly due to the small baseline (26cm) used for triangulation to obtain the range measurements, and again this could be improved by increasing the baseline. A comparison of the heading angle is shown in figure 13.

The software was written using the C library "Horatio" which has been developed at Oxford and Berkeley for the purpose of supporting efficient computer vision applications. HTML documentation for Horatio may viewed from the first author's WWW home page at `http://www.cs.berkeley.edu/~pm/`, and the complete library may also be downloaded from that site.

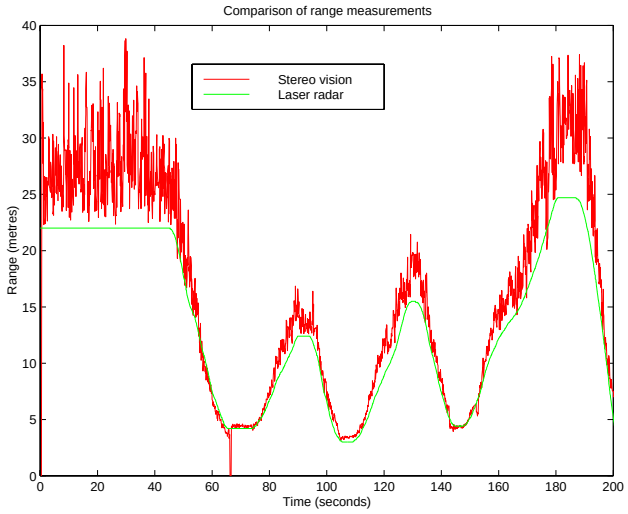


Figure 12: Comparison of range estimates between stereo vision and laser radar from a real-time run.

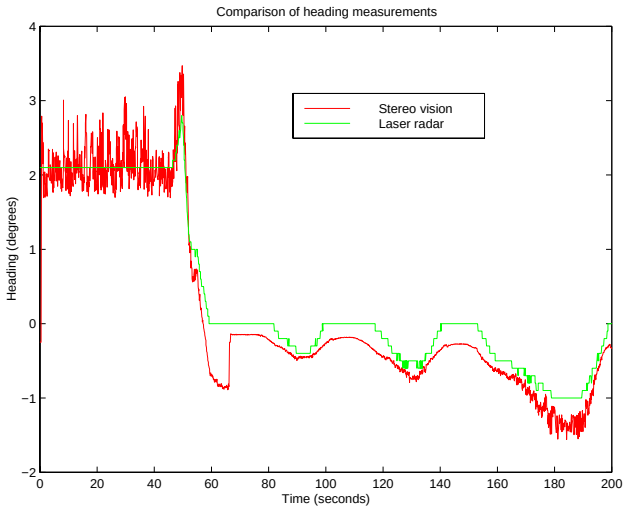


Figure 13: Comparison of heading estimates between stereo vision and laser radar from a real-time run..

4.1 NAHSC Demonstration

The lateral controller has been subsequently integrated into a system with a velocity controller, obstacle detection and avoidance system, and an intra-vehicle communication system. The resulting system was successfully demonstrated in the National Automated Highway Demonstration in August 1997 in San Diego on a segment of freeway **I15**. The highway scenario involved demonstration of the vision based lateral control at 60 mph operating in parallel with obstacle detection module. The detection of an obstacle by laser range sensor triggered the vision based lane change maneuver, followed by another lane change back to the original lane. The vision based lane change algorithm managed to keep the lane in the field of view during the entire duration of the maneuver providing the necessary feedback signal for the lateral controller. Lane changes were followed by a platoon formation with the lead vehicle. The information about the lead vehicle for longitudinal control of the follower vehicle was obtained from laser radar data. The synchronization between vehicles for the platoon formation was achieved by vehicle intra-communication system.

The vision based lateral controller has been also demonstrated at the parking lot of Miramar College in San Diego. The vehicle was following a single line track with curvature segments as small as 8 m radius of curvature. The speed of the vehicle was limited to 18 mph in order to satisfy the passenger comfort requirements. The system was in operation during the entire duration of the DEMO'97 (4 days) and we gave rides to 1200 people.

5 Conclusions

We have demonstrated an approach for vision based lateral and longitudinal vehicle control. The vision based lane tracking system used a robust fitting strategy which allowed us to overcome spurious lane markings and other distracting features that are common on CALIFORNIA highways. Closed-loop simulations and careful analysis revealed the importance of the look-ahead information provided by the vision system, both in the presence of the delay caused by processing of the visual information and changes in the vehicle dynamics with increasing speed. The delay plays an important role in the system and should be taken into account explicitly in case of output feedback strategies, such as the ones we presented. We showed that sufficiently large look-ahead and appropriate choice of gain can compensate for the additional phase lag introduced by delay and vehicle dynamics at lower velocities. At higher velocities additional lead action was introduced in order to achieve desired phase margin. Introducing a real-time observer process into the system not only reduced the noise inherent in the system's sensor measurements, but also provided an accurate estimate of the current vehicle state, circumventing the delay in the vision system and permitting the implementation of more advanced state-space based controllers. The curvature of the road was incorporated into the observer process and the estimates were used for feedforward control strategies. The presence of the feedforward term improved tracking performance in curved road segments.

The information about the vehicle ahead for longitudinal control was provided by a stereo vision algorithm in conjunction with a scanning laser radar sensor. The vision algorithm is built on fixation and reconstruction algorithms designed for active vision systems, which combine stereo and motion cues. The current version of the algorithm has been implemented on a network of C40 DSP's and preliminary results indicate that the range estimates are comparable with the laser radar system. The layered approach to vehicle tracking that involves the use of a simple correlation tracker and a more sophisticated temporal/stereo reconstruction algorithm was designed to take advantage of the merits of the two approaches. Correlation tracking is easy to implement in real time, but is not suitable for extended tracking. Temporal/stereo reconstruction is more suited to extended tracking, but it involves a large and variable amount of computation. By utilizing both approaches in parallel, we have achieved a real-time tracker that can maintain tracking and compute the distance to the vehicle for a long period of time.

The resulting system has been tested extensively during the development time and during the NAHSC Demonstration in August 1997 in San Diego when the system was operational for 4 days.

We are currently doing a thorough analysis and comparison of the tested vision based control stategies for lateral control.

References

[1] D. Beymer, P.F. McLauchlan, B. Coifman, and J. Malik. A real-time computer vision system for measuring traffic parameters. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 1997.

[2] A. Blake, R. Curwen, and A. Zisserman. A framework for spatiotemporal control in the tracking of visual contours. *International Journal of Computer Vision*, 11(2):127–146, October 1993.

[3] P.J. Burt, J.R. Bergen, R. Hingorani, R. Kolczynski, W. A. Lee, A. Leung, J. Lubin, and H. Shvaytser. Object tracking with a moving camera. In *Proc. IEEE Workshop on Visual Motion*, pages 2–12, 1989.

[4] T.H. Cormen, C.E. Leiserson, and R.L. Rives. *Introduction to Algorithms*. MIT Press, 1994.

[5] E.D. Dickmanns and B.D. Mysliwetz. Recursive 3-d road and relative ego-state recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):199–213, February 1992.

[6] S.M. Fairley, I.D. Reid, and D.W. Murray. Transfer of fixation for an active stereo platform via affine structure recovery. In *Proc. 5th Int'l Conf. on Computer Vision, Boston*, pages 1100–1105. IEEE Computer Society Press, 1995.

[7] M.A. Fischler and R.C. Bolles. Random sample concensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm. ACM*, 24(6):381–395, 1981.

[8] C. Harris. Tracking with rigid models. In A. Blake and A. Yuille, editors, *Active Vision*. MIT Press, Cambridge, MA, 1992.

[9] H. Kikuchi, M. Ishiyama, and T Nakajima. Development of laser radar for radar brake system. In *Proceedings of the International Symposium on Advanced Vehicle Control*, pages 385–389, Tsukuba Research Centre, AIST, MITI, Japan, 1994.

[10] D. Koller, J. Weber, and J. Malik. Robust multiple car tracking with occlusion reasoning. In *Proc. 3rd European Conf. on Computer Vision, Stockholm*, volume 1, pages 189–196, May 1994.

[11] J. Košecká, R. Blasi, C.J. Taylor, and J. Malik. Vision-based lateral control of vehicles. In *Proc. Intelligent Transportation Systems Conference, Boston*, 1997.

[12] P. F. McLauchlan and J. Malik. Vision for longitudinal control. In A. Clark, editor, *Proc. 8th British Machine Vision Conf., Essex*. BMVA Press, 1997.

[13] P. F. McLauchlan and J. Malik. Vision for longitudinal vehicle control. In *Proc. Intelligent Transportation Systems Conference, Boston*, 1997.

[14] P.F. McLauchlan and D.W. Murray. A unifying framework for structure and motion recovery from image sequences. In *Proc. 5th Int'l Conf. on Computer Vision, Boston*, June 1995.

[15] P.A.Beardsley, A.Zisserman, and D.W.Murray. Sequential updating of projective and affine structure from motion. *International Journal of Computer Vision*, 23(3), 1997.

[16] K. Pahlavan, T. Uhlin, and J.O. Eklundh. Dynamic fixation and active perception. *IJCV*, 17(2):113–135, February 1996.

[17] I. D. Reid and D. W. Murray. Active tracking of foveated feature clusters using affine structure. *International Journal of Computer Vision*, 18(1):1–20, April 1996.

[18] J. Shi and C. Tomasi. Good features to track. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pages 593–600, 1994.

[19] S. M. Smith. ASSET-2: Real-Time Motion Segmentation and Shape Tracking. In *Proc. 5th Int'l Conf. on Computer Vision, Boston*, pages 237–244, 1995.

[20] C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: A factorization approach. *International Journal of Computer Vision*, 9(2):137–154, 1992.

[21] P.H.S. Torr, P.A. Beardsley, and D.W. Murray. Robust vision. In E. Hancock, editor, *Proc. 5th British Machine Vision Conf., York*, pages 145–154. BMVA Press, 1994.

[22] P.H.S. Torr, A. Zisserman, and S.J. Maybank. Robust detection of degenerate configurations for the fundamental matrix. In *Proc. 5th Int'l Conf. on Computer Vision, Boston*, pages 1037–1042, 1995.

[23] Z. Zhang and O. Faugeras. *3D Dynamic Scene Analysis*. Springer-Verlag, 1992.

[24] Jitendra Malik, Camillo J. Taylor, Joseph Weber, Dieter Koller and Quang-Tuan Luong. A combined approach to stereopsis and lane finding. UCB-ITS-PRR-97-27, California PATH, final report, 1997.

[25] Robert. S. Blasi. A study of lateral controllers for the stereo drive project. Master's thesis, Department of Computer Science, University of California at Berkeley, 1997.

[26] E. D. Dickmanns and B. D. Mysliwetz. Recursive 3-D road and relative ego-state estimation. *IEEE Transactions on PAMI*, 14(2):199–213, February 1992.

[27] B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Transactions on Robotics and Automation*, 8(3):313 – 326, June 1992.

[28] Arthur Gelb *et al. Applied optimal estimation.* MIT Press, 1994.

[29] J. Guldner, H.-S. Tan, and S. Patwarddhan. Analysis of automated steering control for highway vehicles with look-down lateral reference systems. *Vehicle System Dynamics (to appear)*, 1996.

[30] Jana Košecká. Vision-based lateral control of vehicles:look-ahead and delay issues. Internal Memo, Department of EECS, University of California Berkeley, 1997.

[31] M. F. Land and D. N. Lee. Where we look when we steer? *Nature*, 369(30), June 1994.

[32] Yi Ma, Jana Košecká, and Shankar Sastry. Vision guided navigation for a nonholonomic mobile robot. In *submitted to CDC'98*, 1997.

[33] Ü. Özgüner, K. A. Ünyelioglu, and C. Hatipoğlu. An analytical study of vehicle steering control. In *Proceedings of the 4th IEEE Conference on Control Applications*, pages 125 – 130, 1995.

[34] H. Peng. *Vehicle Lateral Control for Highway Automation.* PhD thesis, Department of Mechanical Engineering, University of California, Berkeley, 1992.

[35] Chuck E. Thorpe, Martial Herbert, Takeo Kanade and Steve Shafer. Vision and navigation for the Carnegie-Mellon Navlab. In *IEEE Trans. Pattern Analysis and Machine Intelligence*, 10(3):342-361, 1988.

[36] D. Raviv and M. Herman. A 'non-reconstruction' approach for road following. In Proceedings of the SPIE, editor, *Intelligent Robots and Computer Vision*, pages 2–12, 1991.

[37] Camillo J. Taylor, Jitendra Malik, and Joseph Weber. A real-time approach to stereopsis and lane-finding. In *Proceedings of the 1996 IEEE Intelligent Vehicles Symposium*, pages 207–213, Seikei University, Tokyo, Japan, September 19-20 1996.