HAL
open science

Object Tracking with a pan-tilt-zoom camera : application to car driving assistance

Xavier Clady, François Collange, Frédéric Jurie, Philippe Martinet

To cite this version:

Xavier Clady, François Collange, Frédéric Jurie, Philippe Martinet. Object Tracking with a pan-tilt-zoom camera : application to car driving assistance. International Conference on Robotics & Automation (ICRA '01), May 2001, Seoul, South Korea. pp.1653–1658, 10.1109/ROBOT.2001.932848 . inria-00548285

HAL Id: inria-00548285

https://inria.hal.science/inria-00548285v1

Submitted on 21 Dec 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Object Tracking with a Pan-Tilt-Zoom Camera: *application to car driving assistance.*

Xavier Clady, François Collange, Frédéric Jurie, Philippe Martinet.
LASMEA - UMR 6602 du CNRS
24 avenue des Landais
63177 Aubiere Cedex
e-mail: {*Firstname.Name@lasmea.univ-bpclermont.fr*}

Abstract

In this paper, visual perception in car driving assistance is considered. The work concerns the development of a system combining a Pan-Tilt-Zoom (PTZ) camera and a standard camera, in order to track front vehicles. The standard camera has a small focal length, and is devoted to the analyze of the whole frontal scene.

Here, the PTZ camera is used to track the closest vehicle. An original approach to control it is integrated. Camera rotations and zoom are controlled by visual servoing and by an efficient real time target tracking algorithm. The aim of this work is to keep the rear view image of target vehicle stable in scale and position. The methods presented here were tested on real road sequences within the VELAC[1] demonstration vehicle. Experimental results show the effectiveness of such an approach.

1 Introduction

The development of driving assistance systems [11] has retained many attention since the last decade. They face to the problems due to the increase of the world road traffic (particularly in industrialized countries): pollution, traffic congestion, safety... They are to help drivers to prevent and avoid collisions and to alleviate traffic congestion.

Within this framework, several devices for intelligent vehicles are developed by research laboratories and car manufacturers:

- "Stop and Go" devices, stop and start vehicles in congestions.
- "Platooning", allows to create trains of vehicles (several vehicles follow the trajectory of a car-leader).
- "Adaptive Cruise Control" [1, 12], controls automatically the speed of a vehicle in order to maintain safety distances relative to front vehicles.
- "Collision Warning or Avoidance", informs or replaces the driver when risks of accident are detected.
- "Intelligent Highways", use communication between the road infrastructure and vehicles to optimize traffic conditions.

These various devices need devoted perceptual systems, in particular to capture the vehicle environment (road, obstacles,...). Therefore several sensors are studied and developed at the moment (GPS, radar, lidar, telemeter, camera,...).

The vision based equipment described in this article is well adapted for A.C.C. applications because it is devoted to the capture of frontal views. However, it can be integrated in other applications when combined with other sensors. This system should make it possible to answer the problems existing in ACC application, by providing informations both on the obstacles and the driving lanes and by increasing the sensor range (to detect far vehicles).

It is made of an hardware part and a software part. The first uses a PTZ (Pan-Tilt-Zoom) camera. This camera can be controlled in zoom and in pan tilt angles. It provides image sequences of an object to a tracking algorithm. This algorithm is the software part of the sensor system. It computes the position, the orientation and the size of the target object in the image. These data are used to control the PTZ camera, in order to keep the target centered in the image with a constant size.

The originality of the algorithm is its ability to track any kind of patterns in real time video, on low cost hardware. These characteristics make the sensor sys-

[1] VELAC, Véhicule Expérimental du LASMEA pour l'Aide à la Conduite (see fig. 8).

tem very efficient for ACC applications, when the target object is the rear view of a front vehicle.

This article is made of three sections. The first one describes the context of the application. It explains how the PTZ camera is combined with a standard camera to realize the sensor equipment. The second section presents the tracking algorithm and the control law. The experimentations were realized within the demonstration VELAC vehicle and are presented in the third section.

2 Context of the application

Three functions of perception are essential in ACC application: road and lanes detection (in order to locate the experimental vehicle), front vehicles detection and localization and finally, tracking of the most dangerous vehicle[2] [1]. The kinematic characteristics (i.e. speed and relative position) of dangerous vehicles can be computed by combining these three functions.

These perception tasks can efficiently be achieved using video cameras. Despite the fact that cameras can not directly perceive the 3D environment, they have numerous advantages. They can provide very usefull information about the structure of the environment like the position of the lanes or the position of obstacles and surrounding vehicles. These equipments are generally not very heavy and not very expensive, and therefore can be easily embedded inside a vehicle. In addition, future development in microsystem technology should decrease the cost of such equipment.

PTZ cameras have a variable focal length. This is a great advantage for A.C.C. applications. Indeed, the average braking distance for a vehicle circulating at the speed of 100 km/h is 100 meters and it can reach 200 meters at 140 km/h. Thus, for applications on motor-ways, it is necessary to detect vehicles as far as possible. An acceptable limit would be around 300 meters (in straight line). PTZ camera is the only equipment satisfying these constraints, at the moment.

But an increase of the camera focal length limits the visual field. Target can get out from the field of view, especially in turns or declivities of the road.

The originality of the presented approach is to combine a short fixed focal length camera with the PTZ camera to overcome these limitations. The PTZ camera offers a local but directional vision while the standard camera offers a global vision of the frontal scene.

The first camera will be used to detect lanes, to localize the vehicle and to detect vehicles stated as potentially dangerous (by taking into account of the

[2]Generally, the most dangerous vehicle is the closest vehicle in the same lane as the host vehicle is driving. But this notion can change according the traffic situation.

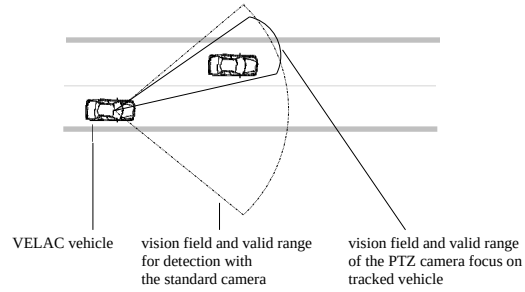

Figure 1: PTZ camera used for tracking.

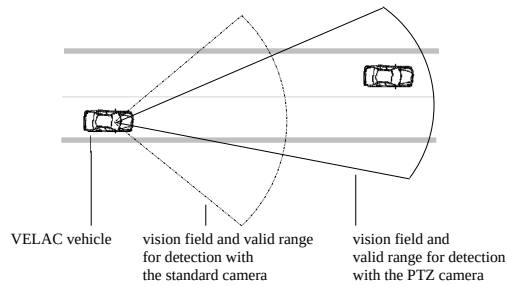

Figure 2: PTZ camera used for detection.

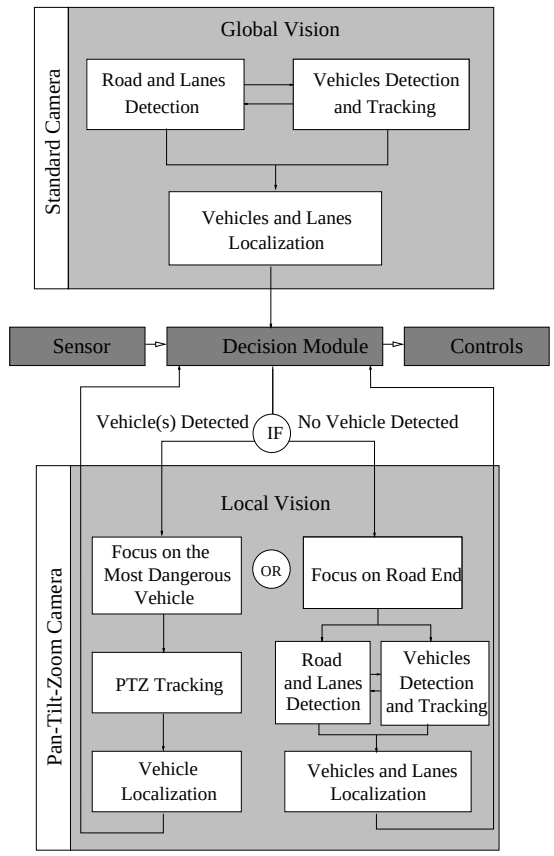

Figure 3: Application General Synopsis

distances of the frontal vehicles and their positions on the road lanes) within the limits of its visual field.

The PTZ camera is focused on the detected vehicle (see fig. 1). However, when no vehicle is detected with the first camera, the PTZ camera is used to detect far vehicles (see fig. 2). The switch between these 2 functions of the PTZ camera is decided according to traffic evolution.

This work requires an intelligent module, called *Decision Module* (see fig. 3). It analyzes the information given by both cameras (and proprioceptive sensors) and compute the speed of the vehicle, in order to respect the safety distances [12]. This speed will be reached by acting on the vehicle actuators (brakes, accelerators).

3 Tracking an Object with a PTZ camera

In this application, tracking an object means tracking a 2D pattern in image sequences, in order to keep it along the camera optical axis with a constant scale. In this section, the tracking algorithm is first presented. Second, the control of the camera is explained.

3.1 The tracking algorithm[8]

The algorithm related in this section allows to estimate the displacement (location, scale and orientation) of a 2D image pattern. It is based on the general framework proposed by Hager and Belhumeur [5].

It consists in relating the displacement of a *reference pattern* to the variations of intensity in an region of interest (ROI) centered on the pattern. This relation, supposed to be linear, is included in a matrix called *interaction matrix*. This matrix is computed during a learning stage. The following paragraphs show how this matrix is defined and used to track the pattern.

3.2 Definitions

To represent the position of the ROI (including pattern) and its content, two vectors are defined: a *parameters vector* $\mathbf{E}$ and a *pattern vector* $\mathbf{I}$. In this article, ROI are elliptic and their position vectors contain five parameters: the position of the center (X_c, Y_c), the orientation (θ) and the largest/smallest axes (R_1,R_2) lengths (see Fig. 4). The vector $\mathbf{I}$ is obtained by sampling the pattern included in the ROI. The samples are distributed on concentric ellipses deduced by scaling the ROI edges, in order to keep constant their density all over the ROI (in Fig. 5, dots represent the points where samples are taken). This representation of patterns in a local reference is independent of their position, orientation and scale in the image.

With these definitions, a reference pattern vector corresponding to the samples of the reference pattern is computed. By writing $\mathbf{\Delta I} = \mathbf{I_{ref}} - \mathbf{I_{cur}}$, $\mathbf{\Delta I}$ is

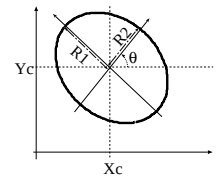

Figure 4: The five parameters

Figure 5: Patterns vectors

null when the current pattern is equal to the reference pattern. An optimal parameter for $\mathbf{E_{opt}}$ leading to this configuration has to be estimated.

The knowledge of a function f defined by:

$$\begin{aligned} \mathbf{\Delta E} = \mathbf{E_{opt}} - \mathbf{E_{cur}} = f(\mathbf{\Delta I}) \\ \Leftrightarrow \mathbf{E_{opt}} = \mathbf{E_{cur}} + f(\mathbf{\Delta I}) \end{aligned} \tag{1}$$

is a powerful tool to track patterns. This function can be approximated by the first order term of its Taylor development, and can be written under the following matrix form:

$$\mathbf{\Delta E^T} = \mathbf{A}.\mathbf{\Delta I^T} \tag{2}$$

where $\mathbf{E^T}$ denotes the transposition of $\mathbf{E}$.

This matrix $\mathbf{A}$ is called *interaction matrix*.

3.3 Estimation and use of Interaction Matrix

The interaction matrix is learnt in a learning stage. At the beginning of this stage, the ROI including the reference pattern is (manually or automatically) selected. Then, it is randomly modified: small perturbations are applied to each of its five parameters. At each time, the pattern vectors corresponding to this new ROI are sampled and kept for further treatments.

When a number of perturbations superior to the number of samples (giving a system with more equations than unknowns), a least square approximation of $\mathbf{A}$ can be estimated by using a Singular Values Decomposition.

Then, the use of the interaction matrix is very straightforward. Multiplied by the difference between the current and the reference pattern vectors, it gives an estimation of the transformation aligning the current ROI with the location of reference pattern in the current image (see eq.2).

3.4 Conclusions about the tracking algorithm

This tracking algorithm presents the following advantages:

- it is a generic approach: any kind of pattern can be tracked after being learnt (in a relatively short time - see section 3.3),

- it runs in real time: the computation consists in a single matrix multiplication,
- it tracks small displacements in location, scale and orientation with great precision,
- it could be enlarged to large displacements by adding a prediction step.

This algorithm is well suited for the presented application where, any kind of vehicle have to be tracked in real time. Current vehicles, except motor-cycle, often present a 2D surface on its rear view. A precise computation of small displacements is needed to control the PTZ camera.

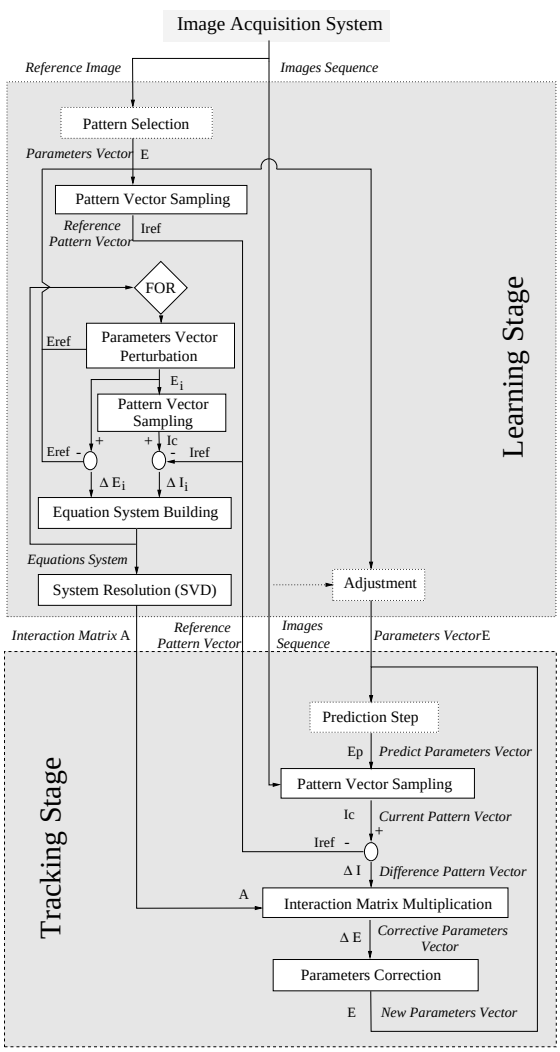

Figure 6: Tracking Process Scheme

3.5 Control law

The control of the PTZ camera is done by a classic visual servoing approach. The information provided by the tracking algorithm are the parameters of the ellipse. One way is to use the work done in [9] where the parameters of ellipse are used to control 5 degrees of freedom (d.o.f). As the sensor equipment has only 3 d.o.f, a more simple way has been chosen.

Here, the sensor signals are fixed to:

$$\mathbf{s} = \begin{pmatrix} u \\ v \\ h \end{pmatrix} \tag{3}$$

where $\begin{pmatrix} u \\ v \end{pmatrix}$ represents the center of the ellipse in image space, and h the size of the ellipse in x direction (which is modelized by the length of a particular segment). Let ${}^e\mathbf{v}$ be the kinematic screw applied to the sensor equipment. The relationship between the time derivative of the feature vector and this screw is:

$$\dot{\mathbf{s}} = \mathbf{L}_s {}^e\mathbf{v} + \frac{\partial \mathbf{s}}{\partial t} \tag{4}$$

where

- $\mathbf{L}_s$ is the interaction matrix relative to the camera. At the equilibrium, $\mathbf{s}^* = (0, 0, h^*)^T$, then $\mathbf{L}_{s^*}$ is defined by:

$$\begin{pmatrix} -\frac{Fu}{z^*} & 0 & 0 & 0 & -Fu & 0 \\ 0 & -\frac{Fv}{z^*} & 0 & Fv & 0 & 0 \\ 0 & 0 & \frac{h^*}{z^*} & 0 & 0 & 0 \end{pmatrix}$$

  Fu and Fv represent the focal lengths in x and y direction respectively. The third line is obtained when considering that the segment is centered in image space at a given distance z^*.

- $\frac{\partial \mathbf{s}}{\partial t}$ represents the contribution of the target motion. In this first approach, as the main goal is to maintain the "virtual distance" constant with the front vehicle, this term has been considered as null. This is a realistic assumption near the equilibrium.

Building the *Task function* $\mathbf{e} = C.(\mathbf{s} - \mathbf{s}^*)$, a proportional control law is defined by the relation [3]:

$${}^e\mathbf{v} = -\lambda.\widehat{\mathbf{L}_{s^*}}^{+}(\mathbf{s} - \mathbf{s}^*) \tag{5}$$

where $\widehat{\mathbf{L}_{s^*}}^{+}$ is the pseudo-inverse of an estimation (at the equilibrium) of the image jacobian. Then, the sensor signal vector $\mathbf{s}$ has an exponential decrease.

The sensor equipment has only 3 d.o.f. So, the reduced control vector ${}^e\mathbf{v}_r$ is defined by:

$${}^e\mathbf{v}_r = (T_z, \Omega_x, \Omega_y)^T \tag{6}$$

and the interaction matrix becomes:

$$\widehat{\mathbf{L}_{s^*}} = \begin{pmatrix} 0 & 0 & -Fu \\ 0 & Fv & 0 \\ \frac{h^*}{z^*} & 0 & 0 \end{pmatrix} \quad (7)$$

Then, each d.o.f can be controlled separately as it has been done in [7]. The pan and tilt angles of the camera are controlled with the center of the ellipse, and the focal length with the size of the segment.

The control of the pan-tilt camera is done by:

$$\begin{pmatrix} \Omega_x \\ \Omega_y \end{pmatrix} = -\lambda \begin{pmatrix} 0 & \frac{1}{F_v} \\ -\frac{1}{F_u} & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \quad (8)$$

The control of the zoom is more rudimentary. An ideal length for x's axe h^* of the ellipse is fixed. The zoom is increased (respectively decreased) if the axe length is under (above) a given threshold. In fact, the control of the zoom is done through a small range of speed (hardware limits).

Thus the command proposed here is a basic solution for the problem of target tracking using a PTZ camera. In future development, improvements of the control law can be done using the modelling presented in [2, 4, 6]. Therefore, as it will be shown in the next section the first results obtained in real conditions are satisfying.

4 Application to visual sensor

In this section, hardware of the sensor equipment is described. Then results obtained on real motor-way with the VELAC vehicle (see fig. 8) are presented. These experiments consist in tracking a front vehicle in real condition.

4.1 PTZ Camera

The PTZ camera is a Sony camera EVI-G21 (see fig. 7). The following table presents its principal characteristics:

focal length	4.5 - 13.5mm (×3 zoom)
pan angle	±30°
tilt angle	±15°

Figure 7: SONY camera

Figure 8: VELAC

The tracking process needs a calibration stage in order to evaluate the position of the principal point (u_0,v_0) of image and the focal lengths (in pixels). The focal lengths (F_u,F_v) are functions of the *zoom* and *focus* camera values through the relation: $a_* \cdot focus + b_* \cdot zoom + c_* \cdot F_* + d_* = 0$, $(* = u, v)$. By using the calibration method given in [10], the results give:

u_0	379.97 ± 0.05 pixels
v_0	295.53 ± 0.05 pixels

-	$F_u(pixels)$	$F_v(pixels)$
a_*	-0.05049	-0.07326
b_*	0.1208	0.1207
c_*	-0.9914	-0.9900
d_*	1028	1055

4.2 Results

Figure 9: Tracking a truck

Figure 10: Tracking a truck under a bridge

The algorithm has been tested on a SGI Indy Station embedded in the experimental vehicle VELAC. During these tests in real conditions on motor-ways, the reference patterns are manually selected on rear view of vehicles.

Figure 11: Tracking a camping-car during an overtaking

Results obtained show that the training phase of the tracking algorithm is rather constraining because of its excessive computation time (approximately 7 seconds). So, the camera position has to be readjusted before to start the tracking.

Furthermore, the current algorithm cannot track specular surfaces. Fortunately, they are rare on standard cars and have small size (the license plate, for example). It is necessary to have a camera equipped

with a powerful zoom. This is not the case of the sensor equipment presented in this article. Thus, the tests in real conditions were only efficient with vehicles of significant size and less prone to local variations of illumination, like trucks or camping-car.

On the other hand, the tracking procedure is completely suited to real time (20 ms for the tracking algorithm and less than 25 ms for the acquisition and control stage) and to real conditions. The first results obtained are presented in Fig.9, 11 and 10 (left: standard camera view, right: PTZ camera view). They show the abilities of the proposed new sensor:

- to keep a constant image-size of the tracked vehicle (Fig.9),
- to track a front vehicle in different situations. For example, in Fig.11, the PTZ camera tracks the camping-car during an overtaking procedure whereas the camping-car disappears of the standard camera view.

The algorithm process is relatively robust because:

- the tracking resists to global illumination changes (passage under a bridge, see Fig.10),
- the tracking time duration are about several minutes (until the target vehicles quit the motorway).

5 Conclusion and future work

This work represents a first step in the development of a new visual sensor for car driving assistance systems. A PTZ camera is combined with a standard camera having a small focal length. This sensor equipment allows to track front vehicles. It is based on a 2D pattern tracking algorithm and a PTZ camera controlled by visual servoing.

The results obtained in real condition on motorways are very encouraging. The visual servoing process allows to focus the PTZ camera on the front vehicle along image sequences. The tracking algorithm runs at the video rate (20ms) and can track any kind of patterns. However, a comparaison with other tracking algorithms will be necessary to demonstrate the effectivness of such algorithm.

Recent works in zoom control [4] consider complex models of PTZ cameras. These models should allow to improve the visual servoing process. Another improvement is to introduce a speed estimator of the front vehicle in the control law. Further developments of the tracking algorithm will make it robust to occultations. The sensor equipment will be complete when the *Decision Module* will be fully implemented.

References

[1] R. Chapuis, F. Marmoiton, and R. Aufrere. Road detection and vehicles tracking by vision for an on-board acc system in the velac vehicle. In *Proc. of the 3rd Int. Conf. on Information Fusion*, volume 2, pages WeB5:11–18, Paris, France, July 2000. FUSION'2000.

[2] B. Espiau. Visual servoing with zoom control. Technical Report 2613, INRIA, July 1995.

[3] B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Trans. on Robotics and Automation*, 8(3), 1992.

[4] J.A. Fayman, O. Sudarsky, and E. Rivlin. Zoom tracking. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, volume 4, pages 2783–2788, Leuven, Belgium, May 1998. ICRA'98.

[5] G.D. Hager and P.N. Belhumeur. Efficient region tracking with parametric models of geometry and illumination. *PAMI*, 20(10):1025–1039, October 1998.

[6] K. Hosoda, H. Moriyama, and M. Asada. Visual servoing utilizing zoom mechanism. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, volume 1, pages 178–183, Nagoya, Japan, May 1995. ICRA'95.

[7] F. Huynh. *Vision-based manipulation for supervision and grasping tasks on a static or mobile object.* PhD thesis, Paul Sabatier University, Toulouse, France, 1998.

[8] F. Jurie and M. Dhome. Tracking objects in image sequences : an efficient approach. In *Reconnaissance des Formes et Intelligence Artificielle*, volume 1, pages 537–546, Paris, France, February 2000. RFIA'2000.

[9] D. Khadraoui, G. Motyl, P. Martinet, J. Gallice, and F. Chaumette. Visual servoing in robotics scheme using a camera/laser-stripe sensor. *IEEE Trans. on Robotics and Automation*, 12(5):743–750, October 1996.

[10] M. Li and J.M. Lavest. Some aspects of zoom lens camera calibration. *Trans. on Pattern Analysis and Machine Intelligence*, 18:1105–1110, November 1996.

[11] H. Pfannschmidt. Safety measures in transportation systems. In *Proc. of the World Engineers' Convention : Professional Congress Mobility*, pages 153–176, Hannover, Germany, June 2000.

[12] S. Schultz, J.W. Rozenblit, and K. Buchenrieder. Towards an application of model-based codesign: An intelligent autonomous cruise controller. In *Proc. of the IEEE Conf. and Workshop on Engineering of Computer Based Systems*, volume 1, Monterey, Canada, May 1997.