

# Online Extrinsic Parameters Calibration for Stereovision Systems Used in Far-Range Detection Vehicle Applications

Sergiu Nedevschi, *Member, IEEE*, Cristian Vancea, Tiberiu Marita, and Thorsten Graf

**Abstract**—This paper presents a high-accuracy online calibration method for the absolute extrinsic parameters of a stereovision system that is suited for far-distance, vision-based vehicle applications. The method uses as prior knowledge the intrinsic parameters and the relative extrinsic parameters (relative position and orientation) of the two cameras, which are calibrated using offline procedures. These parameters remain unchanged if the two cameras are mounted on a rigid frame (stereo rig). The absolute extrinsic parameters define the position and orientation of the stereo system relative to a world coordinate system. They must be calibrated every time after mounting the stereo rig in the vehicle and are subject to changes due to static load factors for the used car setup. The proposed method is able to perform online the estimation of the absolute extrinsic parameters by driving the car on a flat and straight road, parallel with the longitudinal lane markers. The edge points of the longitudinal lane markers are extracted after a 2-D image classification process and reconstructed by stereovision in the stereo-rig coordinate system. After filtering out the noisy 3-D points, the normal vectors of the world coordinate system axes are estimated in the stereo-rig coordinate system by 3-D data fitting. The output of the method is the height and the orientation of the stereo cameras that are relative to the world coordinate system.

**Index Terms**—Computational geometry, online camera calibration, real-time systems, stereovision, vanishing point.

## I. INTRODUCTION

VISION-BASED sensors are gaining greater importance over active sensors (radars, laser scanners, etc.) in the field of driving-assistance systems due to their decreasing costs and increasing hardware processing capabilities. Moreover, a task such as lane detection is one of the exclusive privileges of computer vision.

The measuring capabilities of vision sensors with monocular cameras are reduced because corresponding methods are relying on some prior knowledge of the scene (such as flat road assumption and constant pitch angle), which cannot be always fulfilled in real scenarios. However, a full 3-D reconstruction of

Manuscript received December 27, 2006; revised April 17, 2007, July 13, 2007, and July 19, 2007. This work was supported by Volkswagen A.G., Germany. The Associate Editor for this paper was U. Nunes.

S. Nedevschi, C. Vancea, and T. Marita are with the Computer Science Department, Technical University of Cluj-Napoca, 400020 Cluj-Napoca, Romania (e-mail: Sergiu.Nedevschi@cs.utcluj.ro; Cristian.Vancea@cs.utcluj.ro; Tiberiu.Marita@cs.utcluj.ro).

T. Graf is with the Electronic Research, Volkswagen AG, 38436 Wolfsburg, Germany (e-mail: Thorsten.Graf@volkswagen.de).

Digital Object Identifier 10.1109/TITS.2007.908576

the driving environment is possible when using a stereovision system.

In the second case, the quality of the reconstruction is strictly dependent on the estimation (calibration) accuracy of the following parameters:

- *internal parameters of the stereo system*: intrinsic parameters (focal length and the principal point positions) and relative position and orientation of the two cameras (relative extrinsic parameters);
- *absolute extrinsic parameters of the stereo system*: position and orientation of the stereo rig (usually the left camera coordinate system) relative to a world coordinate system in which the measurements are reported.

The quality of the internal parameters influences the accuracy of the essential and fundamental matrix and, consequently, the position of the epipolar lines [1], which is essential in the stereo correlation process. Wrong estimation of these parameters can lead to a lack of correlated points or to false correlations, which is disastrous to the whole stereo reconstruction process. The quality of the absolute extrinsic parameters is linked only to the accuracy of the 3-D measurements relative to the world coordinate system.

Section II discusses the various calibration strategies that were published in the literature and makes a brief introduction of the proposed method—the stereo setup—and its importance for other stereovision techniques. In Section III, we present the importance of precise estimation of the extrinsic parameters and the propagation of errors in the process of 3-D reconstruction. In this context, we identify the most critical parameters and set them as the primary goal of our algorithm. In Section IV, we introduce a short description of the algorithm along with some prior constraints, which make the subject of a suitable environment. Our strategy is based on detecting longitudinal lane-marking points, which are further used in a robust 3-D data fitting for extraction of high-precision extrinsic parameters. Section V describes the algorithm that is used to classify lane-marking points from 2-D images, whereas Section VI deals with parameter estimation based on 3-D information. Section VII presents the experimental results, and the conclusion is presented in Section VIII.

## II. RELATED WORK

Regarding the intrinsic parameter calibration, there are many general-purpose methods found in the literature [2]–[5], which

can be applied offline. Most of them estimate the parameters by minimizing the projection errors of a set of control points from a calibration object/pattern, with known structure, against the detected 2-D images of the same control points. Multiple views of the calibration object are taken, and the accuracy of the results usually increases with the total number of considered control points. The Bouguet method implemented in the Caltech Camera Calibration Toolbox [5] also has the possibility to estimate the relative extrinsic parameters of a stereovision system. Moreover, other integrated stereovision systems [6], [7] are delivered with offline calibration tools that are able to estimate the internal parameters of the stereo system.

Regarding the absolute extrinsic parameters estimation, the existing methods can be grouped in two categories—offline methods and online methods.

For the offline methods, the principle is the same: minimizing the projection error of some 3-D control points with known positions in the world coordinate system (measurement coordinate system). Therefore, all the absolute extrinsic parameters can be estimated (the rotation matrix and the translation vector [1]). Some general-purpose methods [5], [8] use the same calibration object as for the intrinsic parameters estimation. Such an approach is suited only for near-range applications (e.g., indoor visual navigation of robots). Dedicated methods for vision-based driving assistance systems use painted markers or calibration objects in known positions on a flat road surface, as in [9], [10], and [17]. Because the markers are placed in a large depth range, the obtained accuracy for far-range detection is higher compared to the general-purpose methods.

Regarding the online methods, a dedicated calibration object with known geometry is no longer available. Nevertheless, some prior knowledge about the calibration scene structure must be known in advance. The most facile approach is the usage of a planar road surface with some painted markers. The absence of a calibration object limits the number of detected parameters. In [11]–[13], Weber *et al.* had proposed two ground plane-fitting methods based on the residual disparity of road points to estimate a subset of parameters of a stereo camera that is mounted on a vehicle: the height, the pitch angle, and a sort of roll angle but no yaw angle. In [10], a method for online adjustment of the extrinsic parameters based on painted markers on the car's hood was introduced; therefore, the accuracy for far-range 3-D reconstruction is reduced, and the extrinsic parameters are relative to the ego-car coordinate system and not to a global road-related one. Moreover, each adjustment of the extrinsic parameters requires that all feature points should be reprojected from the image planes into the ground plane, based on inverse perspective mapping. As a consequence, due to several such tuning iterations, which are required for each pair of captured images, this approach suffers from increased processing time. Another approach based on inverse perspective mapping is presented in [18], only this time, the adjustment is performed using a genetic algorithm. A fitness function is used, which measures the parallelism of the reprojected longitudinal road markings, detected in both left and right images. The method works only for canonical configurations (the images have to be rectified), and it is able to estimate the height, pitch, and roll angles but no yaw angle. The

calibration strategy introduced in [19] mixes correlation based on epipolar constraint with minimization of the reprojection error for tracked object points to obtain a convergence of the estimated camera parameters. Only an initial rough estimate is needed before starting the calibration process. The cameras can independently rotate around their yaw axes. All extrinsic parameters are estimated relative to the baseline center, except for the pitch angle of the right camera, which is considered  $0^\circ$ . On the other hand, the method suffers from increased processing time due to recursive parameter estimation based on a robust iterated extended Kalman filter. Even the accuracy of the estimated yaw, pitch, and roll angles (which are critical for far-range 3-D reconstruction) is in question, as the experimental results in their figures are spanning in a range greater than  $15^\circ$ .

In this paper, we present a fast online calibration method that is suitable for far-range 3-D reconstruction, which is able to estimate the orientation and height of the stereo cameras, related to the ego-vehicle world coordinate system, with a precision of 10 min of arc for orientation and a few centimeters for height. The world coordinate system has the **OX** and **OZ** axes in the road plane, with **OZ** coinciding with the longitudinal axis of the ego vehicle (Fig. 6). Its origin is the projection of the ego vehicle's front wheel axis central point on the road plane and is moving with the car along the road. The height of the stereo system and its orientation relative to this world coordinate system are influenced by the variable load of the car. Therefore, we propose an online calibration of the absolute extrinsic parameters for the initial conditions of a specific vehicle setup, which is essential to perform a complete reconstruction of the driving environment (implying 3-D stereo reconstruction, road-obstacle separation, and obstacle detection, as in [14] and [16]). Furthermore, dynamic variations of the ego vehicle's orientation and position relative to a lane-related world coordinate system can be estimated by the use of specialized lane detection modules, as in [15], [16], and [26].

### III. PROBLEM STATEMENT

Considering the intrinsic parameters and the relative extrinsic parameters of both cameras already known, because they can be estimated offline using dedicated calibration methods [17], it is possible to perform 3-D reconstruction [14] in the left camera coordinate system, which is also referred as the camera reference frame (CRF).

The purpose of a vision-based vehicle application is to perform the reconstruction in a world coordinate system [referred as the world reference frame (WRF)]. To do that, the rotation matrix and the translation vector of the left camera ( $\mathbf{R}_C^{\text{left}}$ ,  $\mathbf{T}_C^{\text{left}}$ ), with respect to the WRF, have to be determined. Furthermore, using the relative extrinsic parameters ( $\mathbf{T}_{\text{rel}}$ ,  $\mathbf{R}_{\text{rel}}$ ), right camera position and orientation can be calculated from the left camera, thus obtaining the complete calibration of the stereovision system's absolute extrinsic parameters, i.e.,

$$\begin{cases} \mathbf{T}_C^{\text{right}} = \mathbf{T}_C^{\text{left}} + \mathbf{R}_C^{\text{left}} \cdot \mathbf{T}_{\text{rel}} \\ \mathbf{R}_C^{\text{right}} = \mathbf{R}_C^{\text{left}} \cdot \mathbf{R}_{\text{rel}} \end{cases} \quad (1)$$

TABLE I  
ERRORS IN 3-D RECONSTRUCTION DUE TO CHANGES IN ORIENTATION

Orientation [deg]		X <sub>err</sub> [mm]	Y <sub>err</sub> [mm]	Z <sub>err</sub> [mm]
Pitch	-8	≈ 0	-2003	-1194
	-7		-910	-413
	-6.5		-470	-63
	-6		247	291
	-5		1337	1022
Yaw	-1	2036	≈ 0	802
	0	1009		488
	0.5	502		311
	1	-13		181
	2	-1030		-102
Roll	-1	72	2.5	≈ 0
	0	33	5.02	
	0.5	16	7.9	
	1	-9	13.1	
	2	-30	25.1	

When the rig is mounted into the car, it has a dedicated mounting position and orientation, which will be the same, with very small changes, even if the rig is repeatedly mounted or unmounted. We also consider that the cameras are tightly fixed on the rig, and their relative position and orientation will not change, if not specifically wanted. The position of the left camera (CRF) in the WRF is defined as the translation vector  $\mathbf{T}_C^{\text{left}} = [T_{CX} \ T_{CY} \ T_{CZ}]^T$ . Its estimation errors have a proportional error in the reconstructed 3-D coordinates, being independent of the position in space for the reconstructed points.

The lateral offset  $T_{CX}$  has a relative meaning to the car longitudinal axis; the longitudinal offset  $T_{CZ}$  has a relative meaning to the car front wheel axis. Both parameters can be measured (with errors of a few centimeters) or estimated (using the offline calibration procedure), and there is no need to estimate them online because they insignificantly suffer over time due to possible changes in the orientation of the ego car related to the ground plane.

The only remaining problem to be solved is the estimation of the rotation matrix  $\mathbf{R}_C^{\text{left}}$  of the CRF as well as its height  $T_{CY}$ , which is highly dependent on the carload.

In the case of the three angles (pitch, yaw, and roll) characterizing the camera orientation, the estimation errors are propagating in the reconstruction of the 3-D coordinates in a linear high slope or exponential manner. Table I briefly presents several simulation results that we have obtained on this matter.

The reconstruction error for a 3-D point at 60-m depth is insignificant for lateral offset, critical (about 1 m for a 1° pitch error) for height, and smaller for depth (about 0.5 m for a 1° pitch error).

The influence of the estimated yaw angle error on the reconstructed coordinates for the same 3-D point at 60-m depth is critical for lateral offset (about 1 m for a 1° yaw error), insignificant for height, and smaller for depth (about 0.5 m for a 1° yaw error).

Finally, roll angle errors have a small effect on the lateral offset (about 4 cm for a 1° roll error) and height (about 1.5 cm

TABLE II  
EXTRINSIC PARAMETERS ERRORS' EFFECT ON 3-D RECONSTRUCTION

Extrinsic parameter error	Error of the reconstructed 3D coordinate		
	X (lateral offset)	Y (height)	Z (depth)
$T_{CX}$	≈ $T_{CX}$ error	≈ 0	≈ 0
$T_{CY}$	≈ 0	≈ $T_{CY}$ error	≈ 0
$T_{CZ}$	≈ 0	≈ 0	≈ $T_{CZ}$ error
Pitch	≈ 0	<b>critical</b>	small error
Yaw	<b>critical</b>	≈ 0	small error
Roll	small error	small error	≈ 0

for a 1° roll error) of a point situated at 60-m depth, and an insignificant effect on depth.

The influence of the absolute extrinsic parameters error on the camera position and orientation, relative to the road-related world coordinate system, could be synthesized in Table II.

If only the detection of the free space in front of the ego car (depth) is required, the online pitch angle estimation would be enough. However, we are interested in the complete 3-D reconstruction of the driving environment with high accuracy. Therefore, we must take into account the influence of the other two car angles—the yaw angle and the roll angle.

#### IV. OVERVIEW OF THE METHOD

The proposed solution is based on the 3-D coordinates of reconstructed points (in CRF) from lane markings' edges that are painted on the road surface. This method requires neither calibration objects with a known structure nor calibration patterns, but it does require a suitable environment respecting the following constraints.

- There must be a flat road surface with no curvature or inclination.
- There must be the presence of clear longitudinal lane markings on each side of the car and no other objects on the road with edges parallel to those of lane markings.
- The longitudinal axis of the car should be parallel with the road longitudinal axis as well as with the lane markings all the time (no change in car's moving direction is allowed).
- For better results, it is recommended to have continuous lane markings, slow speed, and (as much as possible) a canonical placement of the cameras on the rig and small values for yaw and roll angles of the left camera related to road plane.

The purpose of our method is only for the initial estimation of the extrinsic parameters (considering that the intrinsic and relative extrinsic parameters do not change) whenever these suffer from changes of the carload or repeated mounts and unmounts of the stereo rig. Therefore, the driver is released

from the burden of performing an offline calibration, which requires special calibration fields and objects.

The environmental constraints required by the currently proposed method are easily found on a highway or uncrowded urban roads. As the algorithm requires only 4 ms for each frame, it can continuously run without affecting too much the overall processing time of other stereovision processes (lane/object detection, pedestrian recognition). Whenever a patch of road satisfying the above constraints is met, the extrinsic parameters are automatically refined. Further dynamic variations of the orientation and position expressed in a lane-related world coordinate system can be estimated by using specialized lane detection algorithms [15].

The online calibration process is performed along a fixed number of consecutive pairs of images, captured while driving the car. For each pair, two major steps are applied.

- The first step consists of extracting the current lane markings' edge points. The only important edges are those that are parallel with the longitudinal axis of the road; therefore, each lane marking has only two useful edges.
- The second step takes as input the 3-D coordinates of the previously classified edge points, reconstructed in the CRF using a fast and high-accuracy stereo reconstruction algorithm [14]. The output will be the rotation matrix  $\mathbf{R}_C^{\text{left}}$  and the height  $T_{CY}$  of the CRF that is related to the WRF.

Even if both steps independently work, each of them imposes specific constraints combined with tracking of the selected lane-marking points. Although the first step is based only on 2-D information, its output is further exploited using stereo information in the second step. Such a combination allows us to obtain an increased accuracy in results.

## V. CLASSIFICATION OF THE ROAD MARKING POINTS

Many algorithms have as purpose the extraction of road features or tracking of the lane boundaries.

A robust method for detecting road features is introduced in [20]. It is based on the presence of planar horizontal edges situated on the road surface and some prior knowledge about the vanishing line of the road plane. In our case, we are interested in extracting only longitudinal markings, which appear as oblique edges when projected in a 2-D image.

Extraction of lane markers based on steerable filters [23] is also a reliable approach, particularly when combined with template matching and Kalman filtering techniques, to increase the detection range and solve the problem of occlusions.

Other methods that are able to detect a road/navigable area from an image, based on probabilistic reasoning, are introduced in [21] and [24]. Probabilistic reasoning is also used in [22] for tracking lane markers. These methods have a satisfactory behavior even for difficult scenarios (e.g., merging/splitting lanes, lane curvature). On the other hand, they suffer from increased processing time, which does not make them suitable for the purposes of this paper.

Moreover, a robust behavior in difficult scenarios is not subject to our needs, as we are interested in a high-accuracy calibration, which is possible only in the presence of planar road surface with parallel longitudinal lane markings and no

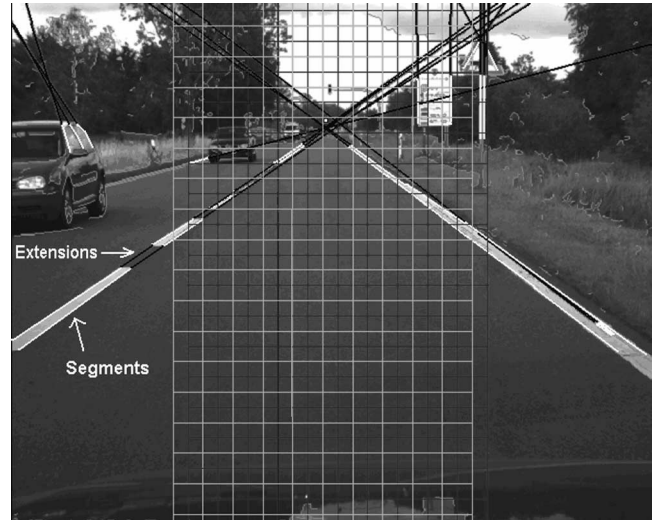


Fig. 1. Initial vanishing point estimation: the two interlaced grids; the filtered segments (white) before applying vanishing point discrimination and their extensions (black).

curvature. For such scenarios, we propose a fast classification algorithm, in which road features are searched only through a 2-D analysis on the left image. Parallel longitudinal lane-marking edges are searched as pairs of segments respecting the following constraints.

- 1) There is a limited distance among them (which increases as the vertical position drops in the image).
- 2) A similar intensity difference is encountered between the inner-bright and outer-dark sides.
- 3) The edges are situated below the vanishing point.
- 4) Their extensions intersect the vanishing point.

For the last two constraints, the vanishing point position is determined using an initial vanishing point estimation method followed by a vanishing point position tracking throughout the pairs of captured images. The initial estimation is also used as a means to reinitialize the tracking procedure whenever the vanishing point tracking fails.

### A. Initial Vanishing Point Estimation

The lines containing segments that are filtered with respect to the first two constraints are plotted on two interlaced grids. Each grid is formed by similar squared cells of size equal to the permitted error in estimating the vanishing point position (Fig. 1). Each cell has an associated value (initialized with 0), which is incremented with the length of any segment (without extension part) that passes through it. Finally, the cells that have the maximum values in the two grids, as well as the cells that have values equal or between the two maximum values, are selected, and their center of weight is returned as an approximation of the vanishing point position. The two maximum values must have neighboring positions; otherwise, the vanishing point is not validated, and the entire calibration process is aborted in the current frame.

The reason for using two interlaced grids is only to eliminate the uncertainty that might appear in case two cells of the same grid encounter similar values.



Fig. 2. Result after filtering out segments passing too far away from the currently estimated vanishing point. The remaining segments are used to estimate the new vanishing point position for the next frame. (White) Lane-marking segments. (Black) Their extensions.

The area covered by the grids should represent the searching area for the vanishing point.

The relative position between the grids was chosen such that the center of a cell in one grid should intersect a node in the other grid.

### B. Vanishing Point Tracking

Once we know the vanishing point position, it can be used as an extra constraint for the returned lane-marking segments, as their extensions must be passing close to it (Fig. 2).

In the end of the whole feature selection algorithm, the segments that are classified as lane markings' edges are known, and their equations in the image plane are used to determine the vanishing point position for the next frame as their intersection (Fig. 2), i.e.,

$$\begin{cases} x_{vp} = \text{slope}_1 \cdot y_{vp} + \text{offset}_1 \\ \vdots \\ x_{vp} = \text{slope}_n \cdot y_{vp} + \text{offset}_n \end{cases} \quad (2)$$

where  $(x_{vp}, y_{vp})$  is the predicted vanishing point position in the next frame, and  $\text{slope}_n$  and  $\text{offset}_n$  are parameters describing the  $n$ th segment equation.

Applying least squares, we obtain  $(x_{vp}, y_{vp})$  as the point having the minimum sum of the squared distances to every line containing lane markings' segments.

In case the distance between the new and current vanishing points is greater than the cell dimension, the tracking procedure is invalidated, and the initial vanishing point estimation is applied for the next frame.

### C. Selection of the Proper Lane Markings

Normally, all lane markings can be used in the calibration process; however, for accuracy reasons, we selected only those belonging to the current lane. Considering that the stereo

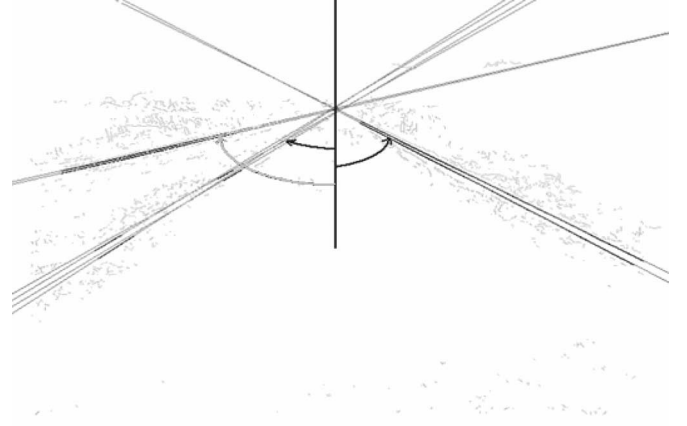


Fig. 3. Vertical line used as reference. Current lane markings form with this reference line minimum angles by each side.

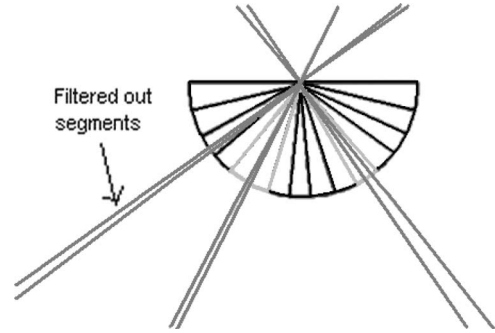


Fig. 4. Polar histogram used to track current lane markings' segments. Segments not passing the gray sectors are filtered out.

system is forward oriented, current lane markings will appear by the two sides of the vertical line that passes through the vanishing point. They are selected by their orientations, considering the vertical line as reference and the positive angles counterclockwise (Fig. 3).

- The right current lane markings are found having the minimum positive angles.
- The left current lane markings are found having the maximum negative angles.

The angles found for the current lane are tracked from frame to frame using a polar histogram (Fig. 4) containing the results from a certain number of previous frames. In case the markings from one side of the current lane are lost, nothing is returned for that frame.

Finally, we perform a supplementary filtering to eliminate the eventually false mark elements that respected all the conditions imposed until now. We assume that there is a single pair of markings on both lane sides; therefore, we select only the segments that respect this condition according to their strength and length (Fig. 5).

## VI. ESTIMATION OF THE ABSOLUTE EXTRINSIC PARAMETERS

All the points of each segment that are classified as belonging to the current lane markings are further reconstructed in the

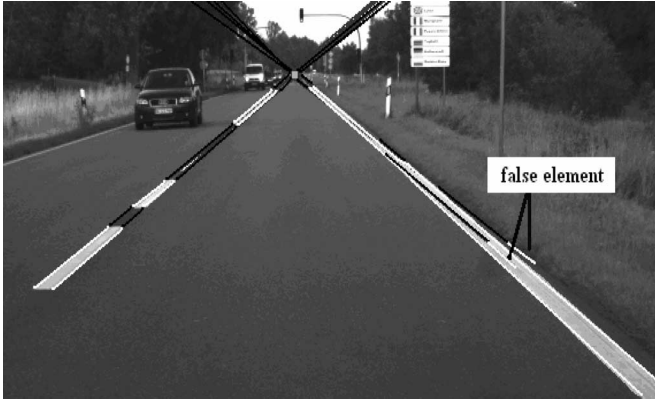


Fig. 5. False element that is removed by the final filtering of the marking elements based on the assumption that the current lane is bordered by simple markings. (White) Lane-marking segments. (Black) Their extensions.

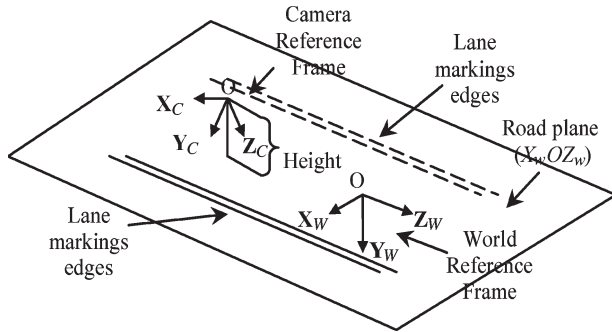


Fig. 6. WRF ( $O X_W Y_W Z_W$ ) and CRF ( $O X_C Y_C Z_C$ ). Scene overview.

CRF, and their 3-D coordinates are used in the process of estimating the rotation matrix  $\mathbf{R}_C^{\text{left}}$  and height  $T_{CY}$ .

We will further consider that the axes of the WRF are  $\mathbf{O}X_W$ ,  $\mathbf{O}Y_W$ , and  $\mathbf{O}Z_W$ , and the axes of the CRF are  $\mathbf{O}X_C$ ,  $\mathbf{O}Y_C$ , and  $\mathbf{O}Z_C$  (Fig. 6).

Given a point position in the CRF, described by vector  $\mathbf{P}_C$ , its position in the WRF will be given by vector  $\mathbf{P}_W$  according to the following formula:

$$\mathbf{P}_W = \mathbf{R}_C^{\text{left}} \cdot \mathbf{P}_C + \mathbf{T}_C^{\text{left}}. \quad (3)$$

The  $T_{CX}$  and  $T_{CZ}$  components of the  $\mathbf{T}_C^{\text{left}}$  vector are considered to be known (they can be measured or estimated); therefore, it only remains to estimate  $\mathbf{R}_C^{\text{left}}$  and  $T_{CY}$ , where

$$\mathbf{R}_C^{\text{left}} = [\mathbf{O}X_W^T \quad \mathbf{O}Y_W^T \quad \mathbf{O}Z_W^T]^T \quad (4)$$

and  $\mathbf{O}X_W$ ,  $\mathbf{O}Y_W$ , and  $\mathbf{O}Z_W$  are the normalized vectors corresponding to the WRF axes in the CRF (Fig. 7). To have a better view, the normalized vectors of the WRF were translated

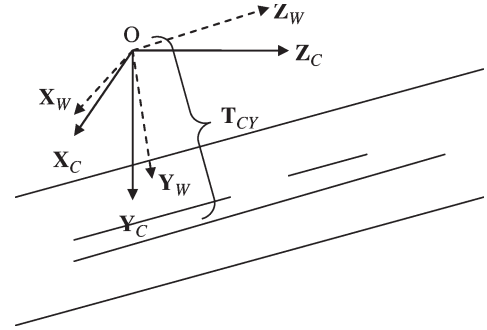


Fig. 7. WRF ( $O X_W Y_W Z_W$ ) orientation in CRF ( $O X_C Y_C Z_C$ ).

in Fig. 7 in the center of the CRF. This process does not affect the rotation matrix.

The problem of computing  $\mathbf{R}_C^{\text{left}}$  (4) is reduced to estimating only  $\mathbf{O}Y_W$  and  $\mathbf{O}Z_W$  because  $\mathbf{O}X_W$  is the cross product of the first two, i.e.,

$$\mathbf{O}X_W = \mathbf{O}Y_W \times \mathbf{O}Z_W. \quad (5)$$

As the car is driven parallel with the lane markings, their 3-D reconstructed edges in the CRF should be parallel with the  $\mathbf{O}Z_W$  axis; therefore, its orientation can be obtained simply by estimating the normalized direction of any of the edges. To obtain a robust result, we calculated the average orientation of all reconstructed edges.

The 3-D orientation of an edge can be calculated from its projections on planes  $X_C O Z_C$  and  $Y_C O Z_C$  (Fig. 8). Each edge is represented by a set of  $n$  reconstructed points in the CRF. Their projections on  $X_C O Z_C$  and  $Y_C O Z_C$  are constrained by the following set of equations:

$$\begin{cases} a_{XOZ} \cdot Z_1 + b_{XOZ} = X_1 \\ \vdots \\ a_{XOZ} \cdot Z_n + b_{XOZ} = X_n \end{cases} \quad \begin{cases} a_{YOZ} \cdot Z_1 + b_{YOZ} = Y_1 \\ \vdots \\ a_{YOZ} \cdot Z_n + b_{YOZ} = Y_n \end{cases} \quad (6)$$

where  $\mathbf{P}_C^k = [X_k \quad Y_k \quad Z_k]^T$ . Initially, we filter out points having depth and lateral offsets outside of a given range [outside our space of interest (SOI) [16]]. Afterward, we interpolate the remaining points through a weighted least squares process that is applied on (6), where  $a_{XOZ}$ ,  $b_{XOZ}$ ,  $a_{YOZ}$ , and  $b_{YOZ}$  are the unknowns. Finally, the 3-D orientation of the edge is obtained as the vector in (7), shown at the bottom of the page.

The weight associated to each point is inversely proportional with its reconstruction error and respects the following formula:

$$\text{Weight}_k = \max W - \frac{\max W - \min W}{\max E} \cdot \text{Error}_k \quad (8)$$

where weights are in the range  $\min W$  to  $\max W$ ,  $\max E$  is the maximum reconstruction error allowed (points with greater errors are removed), and  $\text{Error}_k$  is the estimated reconstruction

$$\left[ \frac{a_{XOZ}}{\sqrt{a_{XOZ}^2 + a_{YOZ}^2 + 1}} \quad \frac{a_{YOZ}}{\sqrt{a_{XOZ}^2 + a_{YOZ}^2 + 1}} \quad \frac{1}{\sqrt{a_{XOZ}^2 + a_{YOZ}^2 + 1}} \right]^T \quad (7)$$

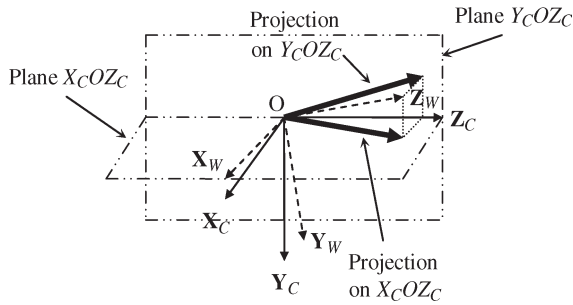


Fig. 8. Projection of lane-marking segments on planes  $X_C O Z_C$  and  $Y_C O Z_C$ . Each reconstructed segment should have identical orientation as the  $OZ_W$  axis.

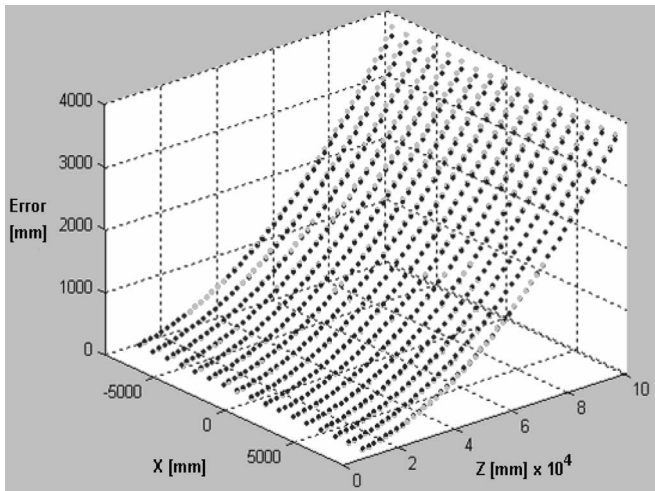


Fig. 9. Similarity between reconstruction errors for a horizontal plane of points situated at 10-m height, when using (black) canonical camera configuration and (gray) general camera configuration.

error for the current point. Considering the case of a canonical stereo system, it is possible to derive a formula that estimates the reconstruction error in the CRF by simply knowing the correlation error  $\Delta c$ , the focal distance  $f$ , and the baseline  $b$ , i.e.,

$$\text{Error}_k = \left| \frac{Z_k \cdot \Delta c \cdot \sqrt{X_k^2 + Y_k^2 + Z_k^2}}{f \cdot b + \Delta c \cdot Z_k^2} \right|. \quad (9)$$

We were able to also use this approximation in a general camera configuration with small relative angles because the differences between estimated errors, for a canonical configuration compared to our general configuration, were up to 150 mm (Fig. 9) in the case of points situated at 10-m height, 100 m ahead, and 8-m lateral offset (the extreme limit of our SOI). These are maximum errors, and only a very small percentage of the points will encounter such coordinates/errors.

The weighted least squares process is repeated several times; each time, the points are filtered out if their fitting error is greater than a fixed percentage of the average fitting error. We imposed several conditions for stopping the iterations.

- The number of iterations equals a certain value—concerning the speed.



Fig. 10. Typical calibration scenario.

TABLE III  
CONSTANTS USED FOR THE EXPERIMENTAL SETUP

Symbol	Value	Description
$\Delta c$	0.2 pixels	correlation error in (9); it was chosen as the most frequent correlation error obtained after performing several experiments
$[minW, maxW]$	[0, 10]	allowed weight range in (8)
$maxE$	0.5 m	maximum reconstruction error in (8)
$maxAllowedError$	550 mm	maximum allowed error in (10)
$FixedDist$	50 m	distance for maximum allowed error in (10)

- The number of remaining points becomes lower than a certain percentage of their initial number.
- The average fitting error drops under a certain value (based on experimental results).
- The average fitting error does not converge.

After the orientation of an edge is obtained using (7), it is compared with the average direction of the  $OZ_W$  axis over a fixed number of previous frames. A filtering threshold is defined in terms of a fixed distance and a fixed maximum error that is allowed at that distance, i.e.,

$$\text{Threshold} = \tan^{-1} \frac{\text{max Allowed Error}}{\text{Fixed Dist}}. \quad (10)$$

This approach is based on the fact that lane markings' edges are parallel with the  $OZ_W$  axis, and their orientation remains constant from frame to frame, assuming that the car does not change its heading direction.

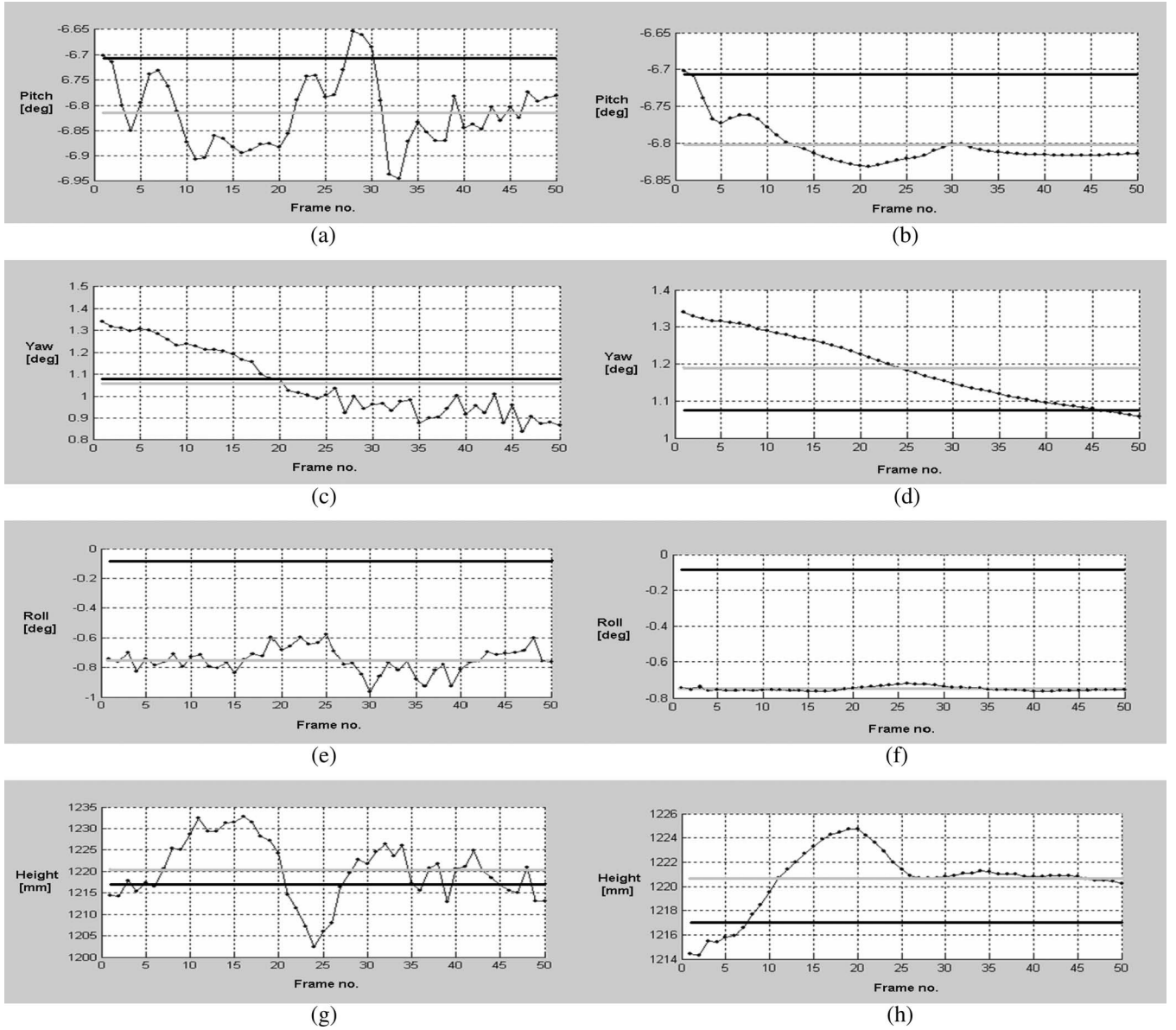


Fig. 11. Experimental results on 50 consecutive frames. (a) Estimated pitch angle on each frame. (b) Pitch angle by averaging rotation matrix on previous frames. (c) Estimated yaw angle on each frame. (d) Yaw angle by averaging rotation matrix on previous frames. (e) Estimated roll angle on each frame. (f) Roll angle by averaging rotation matrix on previous frames. (g) Estimated height on each frame. (h) Height by averaging results on previous frames. (Dots) Estimated values. (Thick gray line) Mean value. (Thick black line) Values of the high-precision offline calibration that are suitable for far-range detection.

Once the aforementioned filtering techniques are applied, the remaining lane-marking points are known to be points from the road plane and respect the following constraints:

$$\begin{cases} A \cdot X_1 + B \cdot Y_1 + C \cdot Z_1 = H \\ \vdots \\ A \cdot X_n + B \cdot Y_n + C \cdot Z_n = H \\ A^2 + B^2 + C^2 = 1 \end{cases} \quad (11)$$

where  $H$  is the height of the plane, and  $\mathbf{n} = [A \ B \ C]^T$  is the normal to the plane. The above system of equations can be solved using a weighted least squares process, in which the weights are calculated according to (8). The result is a plane fitting of the lane-marking points, thus obtaining the normal vector  $\mathbf{n} = \mathbf{OY}_W$  and the height  $H = T_{CY}$ . This process

is performed several times until one of the above stopping conditions becomes true.

Once  $\mathbf{OY}_W$  and  $\mathbf{OZ}_W$  axes are estimated, we calculate  $\mathbf{OX}_W$  using (5). Therefore, we obtain the rotation matrix  $\mathbf{R}_C^{\text{left}}$  according to (4). Finally, we apply a cleaning process on rotation matrix elements based on singular value decomposition to eliminate possible axis distortions that might appear due to precision errors in computations or imperfect 3-D reconstruction of points in the CRF. This final stage insures an orthogonal property for the rotation matrix [25] with a considerable improvement on further 3-D reconstruction.

As the online calibration is performed on several consecutive frames, the results are averaged over frames to avoid unstable results due to small car oscillations or small changes in heading direction.



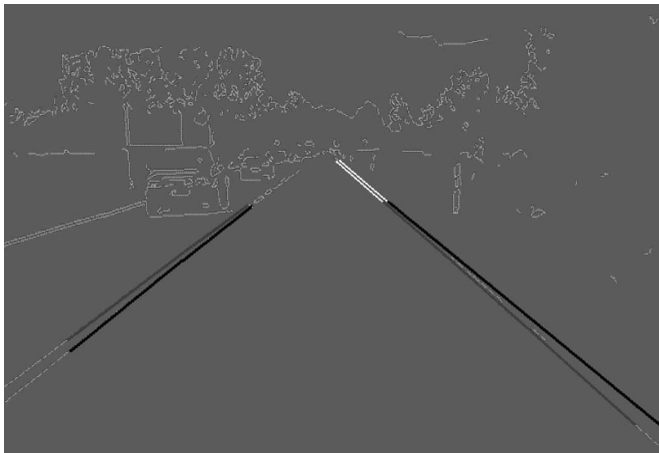


Fig. 12. Segments classified as lane markings are returned as an array of 3-D reconstructed points. They are represented with one of the following colors: white, black, or dark gray. White areas represent points classified as lane markings, which were filtered out in the parameters' estimation phase.

## VII. EXPERIMENTS

We performed several experiments in different environmental and illumination conditions while driving the car on an approximately flat road with clear lane markings (Fig. 10). The intrinsic parameters of the cameras and relative extrinsic parameters were considered known, as they were obtained using an offline calibration process. Table III contains several initialization parameters that we used while performing the online calibration of the absolute extrinsic parameters.

The tests were performed on 50 consecutive frames. The estimated rotation matrix and height were compared with those obtained using a high-precision offline calibration process that is suitable for far-range detection. Results concerning height and derived rotation angles (pitch, yaw, and roll) corresponding to the computed rotation matrix are presented in Fig. 11.

As it may be noticed, all four parameters are close to the offline estimated values, with offsets induced by different vehicle load, as compared to the offline setup, proving method stability and accuracy. The differences encountered in the case of the average values were no more than 10 min of arc for orientation and a few millimeters for height. As a consequence, we were able to perform an accurate 3-D reconstruction (with errors below 3% of the detection range) up to 100 m ahead based on the estimated values. Tests performed on other sequences revealed the same small error range. The yaw angle estimation is highly dependent on a driver's ability to drive the car parallel with the lane markings. The obtained precision allows us to use this method for far-range detection applications.

Regarding the accuracy of the lane-marking detection phase, the results were quite satisfactory, even for bad weather and illumination conditions. Each lane-marking segment is delivered as an array of reconstructed points (Fig. 12). The accuracy of the results is very much improved, due to further 3-D filtering in the parameters' estimation phase.

On the other hand, the set of thresholds used for filtering constraints can be set to less-permissive values. As a consequence, the number of total points actually involved in the parameter estimation might decrease to a considerable value (Fig. 13),

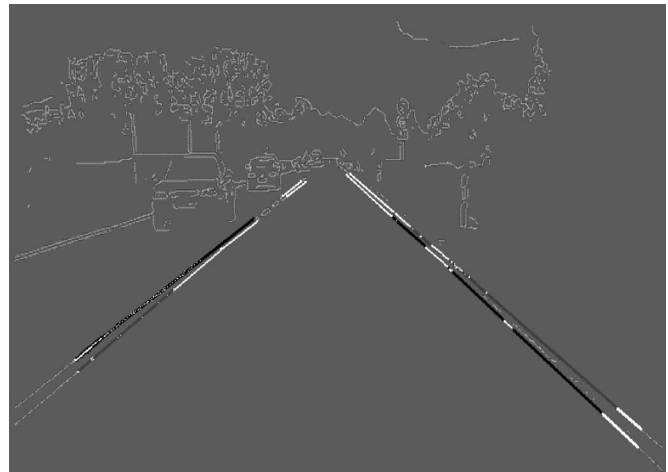


Fig. 13. Effect of using a less permissive set of threshold values. Segments are represented with one of the following colors: white, black, or dark gray. White areas represent points classified as lane markings, which were filtered out in the parameters' estimation phase.

which can have a negative effect on the final results (we need as many points as possible that are spread in our reconstruction range to obtain a trustful set of extrinsic parameters). Therefore, we deal with a tradeoff in what concerns the establishment of an optimal set of threshold values.

Regarding the implementation of the extrinsic parameters' estimation phase, its complexity respects the following formula:

$$\text{No computations} = 2 \cdot \sum_{i=1}^n I_i \cdot P_i + 2 \cdot I \cdot P \quad (12)$$

where  $n$  is the number of lane-classified marking edges,  $I_i$  is the number of weighted least squares iterations performed when estimating edge $_i$ 's direction,  $P_i$  is the number of points from edge $_i$ ,  $I$  is the number of weighted least squares iterations performed for road plane fitting, and  $P$  is the number of points involved in road plane fitting.

The average processing time needed for extracting the lane markings' edges was about 3.9 ms ( $644 \times 512$  image), and the extrinsic parameters' estimation was about 0.2 ms, resulting in a total time of 4.1 ms. Tests were deployed on a 2.6-GHz Intel Pentium IV Processor.

## VIII. CONCLUSION

In this paper, we proposed a fast high-precision online calibration method for the absolute extrinsic parameters characterizing a stereovision system, which is suited for far-range detection vehicle applications. The intrinsic and relative extrinsic parameters of the cameras are considered to be known, as they can be estimated using a high-precision offline calibration procedure.

The process requires no any calibration objects with a known structure, nor calibration patterns, but only a suitable environment consisting of a flat and straight road with clearly painted

lane markings. Such an environment may be found on highways or marked roads; therefore, it does not require the construction of special calibration fields. The results were very close to those obtained using a high-precision offline calibration method that is suitable for far-range detection, with offsets determined by static load factors for the used car setup.

## REFERENCES

- [1] E. Trucco, *Introductory Techniques for 3D Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall, 1998.
- [2] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE J. Robot. Autom.*, vol. RA-3, no. 4, pp. 323–344, Aug. 1987.
- [3] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *Proc. ICCV*, Corfu, Greece, Sep. 1999, pp. 666–673.
- [4] J. Heikkilä and O. Silven, "A four-step camera calibration procedure with implicit image correction," in *Proc. Conf. CVPR*, 1997, pp. 1106–1112.
- [5] J. Y. Bouguet, *Camera Calibration Toolbox for Matlab*. [Online]. Available: [http://www.vision.caltech.edu/bouguetj/calib\\_doc](http://www.vision.caltech.edu/bouguetj/calib_doc)
- [6] Videre Design. [Online]. Available: <http://www.videredesign.com>
- [7] Tyzx. [Online]. Available: <http://www.tyzz.com/>
- [8] N. Kämpchen, U. Franke, and R. Ott, "Stereo vision based pose estimation of parking lots using 3D vehicle models," in *Proc. IEEE Intell. Veh. Symp.*, 2002, pp. 459–464.
- [9] S. Ernst, C. Stiller, J. Goldbeck, and C. Roessig, "Camera calibration for lane and obstacle detection," in *Proc. Int. Conf. IEEE Intell. Transp. Syst.*, 1999, pp. 356–361.
- [10] A. Broggi, M. Bertozzi, and A. Fascioli, "Self-calibration of a stereo vision system for automotive applications," in *Proc. IEEE Conf. Robot. Autom.*, Seoul, Korea, May 2001, vol. 4, pp. 3698–3703.
- [11] J. Weber and M. Atkin, "Further results on the use of binocular vision for highway driving," in *Proc. SPIE—Transportation Sensors Controls: Collision Avoidance, Traffic Management ITS*, 1997, vol. 2902, pp. 52–61.
- [12] J. Weber, D. Koller, Q.-T. Luong, and J. Malik, "An integrated stereo-based approach to automatic vehicle guidance," in *Proc. 5th Int. Conf. Comput. Vis.*, Boston, MA, 1995, pp. 52–57.
- [13] J. Weber, D. Koller, Q.-T. Luong, and J. Malik, "An integrated stereo-based approach to automatic vehicle guidance," in *Proc. SPIE—Collision Avoidance Automated Traffic Management Sensors*, 1995, vol. 2592, pp. 116–127.
- [14] S. Nedevschi, R. Schmidt, T. Graf, R. Danescu, D. Frentiu, T. Marita, F. Oniga, and C. Pocol, "High accuracy stereo vision system for far distance obstacle detection," in *Proc. IEEE Intell. Veh. Symp.*, Parma, Italy, Jun. 14–17, 2004, pp. 161–166.
- [15] S. Nedevschi, R. Schmidt, T. Graf, R. Danescu, D. Frentiu, T. Marita, F. Oniga, and C. Pocol, "3D lane detection system based on stereovision," in *Proc. IEEE ITSC*, Washington DC, Oct. 4–6, 2004, pp. 292–297.
- [16] S. Nedevschi, R. Danescu, T. Marita, F. Oniga, C. Pocol, S. Sobol, T. Graf, and R. Schmidt, "Driving environment perception using stereovision," in *Proc. IEEE IV Symp.*, Las Vegas, NV, Jun. 2005, pp. 331–336.
- [17] T. Marita, F. Oniga, S. Nedevschi, T. Graf, and R. Schmidt, "Camera calibration method for far range stereovision sensors used in vehicles," in *Proc. IEEE IV Symp.*, Tokyo, Japan, Jun. 2006, pp. 356–363.
- [18] J. M. Collado, C. Hilarío, A. Escalera, and J. M. Armingol, "Self-calibration of an on-board stereo-vision system for driver assistance systems," in *Proc. IEEE IV Symp.*, Tokyo, Japan, Jun. 2006, pp. 156–162.
- [19] T. Dang, C. Hoffman, and C. Stiller, "Self-calibration for active automotive stereo vision," in *Proc. IEEE IV Symp.*, Tokyo, Japan, Jun. 2006, pp. 364–369.
- [20] R. Okada, Y. Taniguchi, K. Furukawa, and K. Onoguchi, "Obstacle detection using projective invariant and vanishing lines," in *Proc. 9th Int. Conf. Comput. Vis.*, Oct. 2003, vol. 1, pp. 330–337.
- [21] P. Jeong and S. Nedevschi, "Local difference probability (LDP)-based environment adaptive algorithm for unmanned ground vehicle," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 3, pp. 282–292, Sep. 2006.
- [22] Z. Kim, "Realtime lane tracking of curved local road," in *Proc. IEEE ITSC*, Toronto, ON, Canada, Sep. 2006, pp. 1149–1155.
- [23] J. C. McCall and M. M. Trivedi, "Video-based lane estimation and tracking for driver assistance: Survey, system, and evaluation," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 20–37, Mar. 2006.
- [24] Y. He, H. Wang, and B. Zhang, "Color-based road detection in urban traffic scenes," *IEEE Trans. Intell. Transp. Syst.*, vol. 5, no. 4, pp. 309–318, Dec. 2004.
- [25] T. Gandhi and M. M. Trivedi, "Vehicle surround capture: Survey of techniques and a novel omni video-based approach for dynamic panoramic surround maps," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 3, pp. 293–308, Sep. 2006.
- [26] H.-Y. Cheng, B.-S. Jeng, P.-T. Tseng, and K.-C. Fan, "Lane detection with moving vehicles in the traffic scenes," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 4, pp. 571–582, Dec. 2006.



**Sergiu Nedevschi** (M'99) received the M.S. and Ph.D. degrees in electrical engineering from the Technical University of Cluj-Napoca (TUCN), Cluj-Napoca, Romania, in 1975 and 1993, respectively.

From 1976 to 1983, he was with the Research Institute for Computer Technologies, Cluj-Napoca, as a Researcher. In 1998, he was appointed Professor in computer science and founded the Image Processing and Pattern Recognition Research Laboratory at the TUCN. From 2000 to 2004, he was the Head of the Computer Science Department, TUCN, and is

currently the Dean of the Faculty of Automation and Computer Science. He has published nearly 150 technical papers and has edited over ten volumes, including books and conference proceedings. His research interests include image processing, pattern recognition, computer vision, intelligent vehicles, signal processing, and computer architecture.



**Cristian Vancea** received the M.S. degree in computer science in 2004 from the Technical University of Cluj-Napoca, Cluj-Napoca, Romania, where he is currently working toward the Ph.D. degree, also in computer science, specializing in stereovision systems for intelligent vehicles and robot navigation.

His research interests include image processing, pattern recognition, camera calibration, stereovision, and digital design with reconfigurable hardware.



**Tiberiu Marita** received the M.S. degree in computer science in 1995 from the Technical University of Cluj-Napoca, Cluj-Napoca, Romania, where he is currently working toward the Ph.D. degree, also in computer science, specializing in computer vision.

He is a Senior Lecturer with the Computer Science Department, Technical University of Cluj-Napoca, where he teaches image processing, pattern recognition, and design with microprocessors. His research interests include camera calibration, stereovision, vision-based automotive applications, and medical imaging.



**Thorsten Graf** received the M.S. degree in computer science and the Ph.D. degree from the University of Bielefeld, Bielefeld, Germany, in 1997 and 2000, respectively.

In 1997, he became a Member of the "Task Oriented Communication" graduate program, University of Bielefeld, which was funded by the German Research Foundation DFG. In June 2001, he joined Volkswagen AG Group Research, Wolfsburg, Germany. Since then, he has worked on different projects in the area of driver assistance systems as

a Researcher and Project Leader. He is the author or coauthor of more than 40 publications and is the holder of about ten patents. His research interests include image processing and analysis dedicated to advanced comfort/safety automotive applications.