Calibration of an embedded camera for driver-assistant systems

Mario Bellino, Frédéric Holzmann, Sascha Kolski, Yuri Lopez de Meneses, Jacques Jacot

***Abstract*— One of the main technical goals in the actual automotive industry is to increase vehicle safety. The European project SPARC (Secure Propulsion using Advanced Redundant Control) is developing the next generation of trucks towards this aim. The SPARC consortium intends to do so by providing the truck with active security systems. Specifically, by equipping the vehicle with different sensors, it can be made aware of its environment, such as other vehicles, pedestrians, etc. By combining all sensor data and processing it with internal proprioceptive information[1], the truck can advice, warn or even override the driver in case of non-response.**

Camera systems are particularly advantageous for sensing purposes, because they are passive sensors and provide very rich information. Moreover, they can easily be software-reconfigured to extract new or additional data from the input-image. Typical information that SPARC aims to extract is the position of the vehicle within the lane, the presence and distance of other vehicles or obstacles, and the identification of roadsigns. In this paper, a lane-detection algorithm will be presented and discussed.

Some of the resulting information needs to be given in world coordinates, as opposed to image coordinates. To carry out the necessary conversion, a previous calibration is needed. The challenge is to determine a procedure to calibrate a camera mounted on a truck to precisely determine the position of obstacles situated in a 100 meter range. The two-step calibration procedure presented here has been designed to simplify the calibration of the mounted cameras in the truck production line.

***Index Terms*— Vehicle safety, calibration procedure, SPARC, lane detection, assistant system, production line.**

I. INTRODUCTION

PASSENGER SAFETY is one of the most important axes of research in the automotive industry. This goes beyond increasing vehicle reliability or equipping cars with passive security systems. Indeed, a statistical investigation, presented in [1], pointed out that 95% of accidents are due to human behavior and only 5% to defective vehicles. Moreover, 80% of these accidents involve improper driving reaction, speed and U-turn manoeuvre [2]. Analyses of these accident scenarios show that more than 40% of the accidents might have been avoided if the vehicles had been equipped with a warning system. This level of safety could rise to 95% if the vehicle could autonomously engage a safety driving response in critical situations.

Therefore it is necessary to develop active security systems capable of sensing the environment where the vehicle is evolving and analyzing the situation in real-time. This driver-assistance system should further be capable of interacting with the driver, in order to inform or warn him of a potentially dangerous situation, and to act on the powertrain and steering to avoid the accident, should the driver not respond on time.

The authors are with Faculty of Engineering (STI "Sciences et Techniques de l'Ingénieur"), Institute of Production and Robotics, École Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland. Contact: mario.bellino@a3.epfl.ch

[1]defining the internal state of the vehicle as mainly yaw, pitch and roll value, speed, and acceleration but also powertrain status, etc...

A. European project SPARC

The European project "Secure Propulsion using Advanced Redundant Control" (SPARC) aims to build the new generation of driver-assistant systems for heavy goods vehicles. Trucks are particularly interesting for this project compared to cars, because accidents with heavy vehicles cause more than two times more heavy damage, nearly two times more injuries and more than three times lethal casualties. Furthermore, because of their driving distance, driving time and professional use, trucks stand out as pioneers for car technologies. To prove this principle, SPARC intends to demonstrate the scalability of the security systems by porting the developments made on trucks to a small passenger-car.

To achieve these aims, SPARC is based on two concurrent developments. The first is the X-by-wire technology that enables the steering wheel of the driver to be mechanically disconnected from the wheels, as in the aircraft industry, the steering column is replaced by servomotors and the driver commands are sent through wires. This equipment has been successfully tested in the Powertrain Equipped with Intelligent Technologies project (PEIT, contract no.: IST-2000-28722), and allows some automatic controls to maintain driving stability and reduce braking distance in critical situations. The second development is the creation of a safety assistant, or Co-Pilot, to compute and decide truck behavior. This module is composed by the Human-Machine Interface (HMI) where the driver chooses the direction and velocity to apply to the truck. This information (or stimuli) can thus be characterized by a 2-Dimensional vector. In parallel, the Co-Pilot technology fuses all sensor information, and provides a redundant vector which expresses the safest vehicle behavior. Finally, the safety decision controller will generate a secure motion vector based on both previous vectors, and send this third vector to the powertrain in order to avoid accidents in case of driver failure (see Fig. 1).

The Co-Pilot builds an internal representation, or map, of the world surrounding the vehicle in order to select the best action to be taken to follow a safe trajectory. This action is then encoded as a *redundant vector*. It is redundant because usually the driver has taken that same action. In order to build this internal representation of the environment, the Co-Pilot relies on a set of sensors, such as radars, GPS or camera systems. Indeed, no single sensor technology is capable of providing accurate, robust information in all weather or traffic conditions. Therefore, the Co-Pilot exploits the complementarity of sensors by fusing their information.

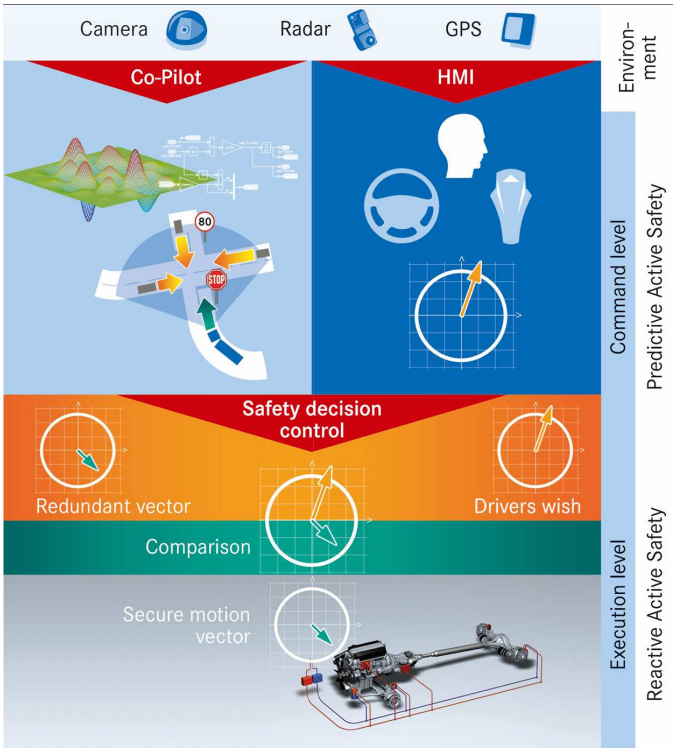

Fig. 1. SPARC system diagram: The safety decision controller receives the driver command through the Human-Machine Interface (HMI), and from the Co-Pilot a sensor-based redundant command vector. The Co-Pilot establishes the safety envelop of the environment by fusing the result of various sensors. Finally, the safety controller decides which one is more reliable and acts accordingly on the powertrain using X-by-wire technology.

B. SPARC camera system

Camera systems are particularly advantageous for sensing purposes, because they are passive sensors and they provide a very large volume of information. However, the amount of information provided by the camera is too much to be processed with traditional automotive electronic control units. To this end, an embedded platform for image analysis is being developed within the SPARC project. The goal of the vision platform is to extract from the image the relevant information in each situation and provide it to the Co-Pilot.

Information that has been identified as relevant for most traffic scenarios is

- Position of the car relative to the current lane
- Lane's width
- Curvature radius of the lane
- Position and size of obstacles and objects on lane and immediate neighboring lanes
- Time to contact, or time that will elapse before the impact with the objects
- Content and position of roadsigns

Some of the resulting information, such as lane width or obstacle distance, needs to be given in world coordinates, as opposed to image coordinates. This is necessary, either because it only has sense in physical coordinates or because this information has to be correlated with other sensors and thus a common metric is needed. To carry out the necessary conversion from image to world coordinates, a previous calibration should be performed.

Thus, we will initially present the lane detection principle and then, in section III, we will introduce a method that can be used to calibrate the camera directly at the end of the production line. This approach tries to minimize the effort needed to perform the calibration process while minimizing the measurement error. Then, the section IV will highlight the experimental results that are obtained with such a calibration method. Finally, a conclusion based on these results will be presented.

II. Lane detection algorithm

The task of detecting and tracking road limits or lane marking is particularly difficult; mainly because the evolving scene is a complex blend of elements, with a high level of changes and variability, on which the system has no control. To be able to avoid building a system that works only in specific situations, which will not fulfill the SPARC objectives, the algorithm will implement several approaches to detect the desired lane. Thus, as described in [3], the algorithm will use multiple hypotheses of detection which will track multiple models of lanes. This method will then fuse the results of all the tested models, and provide a solution with a higher level of confidence.

Actually, several models are tested and implemented in experimental vehicles. These models rely on the following specific hypothesis:

- In most conventional situations, the road-side has more color, or gray, variability than the vehicle's lane. Thus, analyzing the gradient variation helps to separate road area from its side.
- The road or lane width is a quite stable value, thus the distance between left and right lane cannot vary too much during two consecutive images. Moreover, this value must be contained in a given interval of possible road widths.
- If the acquisition time between two successive images is short, then the distance between two successive road limits detection has to be small.
- The line width can also be found by analyzing two similar gradient peaks at a given distance. If a given width is found, we are quite sure to have detected a road line.
- Line color can also be a reliable data. However, the user must be aware that illumination changes can have a severe influence on this measurement. Thus, a solution can be found by using a relative value, between center of line color and road color.
- As the field of view of the camera has to deal with road perspective, the algorithm will have difficulties to find reliable data objects that are far away. Thus, as the density, in image plane, of objects situated far away is big, it increases the possibility of finding a transition that has no direct relation with the lane. Thus, the points that are close to the top of the image must have a bigger uncertainty than the bottom ones.
- Finally, after having estimated the road-side, it is possible to compute the gradient of the image along this estimator.

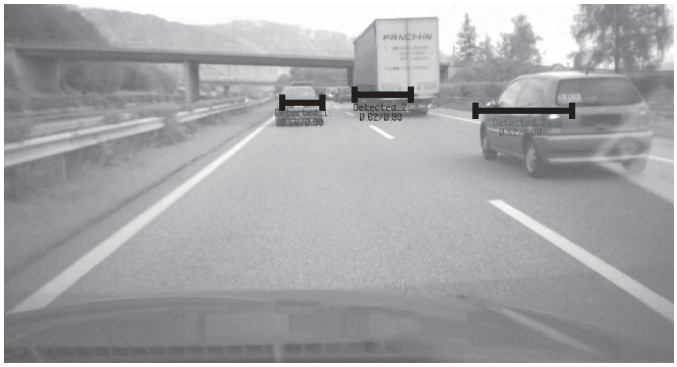

Fig. 3. Typical vehicles detection result. The information of surrounding vehicle is found in the camera coordinate frame, but should be converted in the real 3D space in order to compare the results of the camera module within the fusion step.

The relative value of this gradient along the lane estimation can help to qualify and improve the road boundary.

This last hypothesis has a severe influence on lane detection performance. Indeed, all the other assumptions are local and could not use the globality and continuity of the lane limits (the improvements of such method are shown in Fig. 2).

III. Calibration process

It is not only necessary to detect the lane where the vehicle is evolving, but also to provide quantitative information about it. For instance, the radius of curvature is necessary to decide if the vehicle is entering a curve too fast. Other information, such as the distance between the truck and the closest obstacle (Fig. 3), require the vision platform to provide metric measurements.

This article will thus focus on the calibration of monocular camera, which is a cheaper initial solution, and will therefor not cover the calibration of stereo-vision systems. However, the calibration procedure can be repeated on every isolated camera composing a vision system.

Through calibration, we intend to relate pixel-based information read by the camera, to scene information. The dimensional analysis of a 3D scene with a single, fixed camera forces the user to constrain the supplementary degree of freedom of the system. Indeed, although a point in the scene has three degrees of freedom, the same point in the camera has only two. In order to constrain the system, we suppose that the road is completely flat. It is clear that this supposition is a theoretical model of route, and that real roads match only partially this hypothesis. Moreover, the calibration procedure supposes the camera orientation and position to be fixed, which is particularly false when a camera is embedded in a truck. Indeed, the pitching of its cabin will largely modify the position and orientation of the camera, which will increase the error of the dimensional estimation. Several solutions can be proposed to eliminate, or decrease, this behavior. The first one is to attach accelerometers to the cabin to estimate its actual position. Thus, it is possible to correct the calibration of the camera by taking into account the modification of its position and orientation. A second approach could consist in estimating the pitching of the cabin based on a model containing mainly the steering angle, load, speed and acceleration of the vehicle. Other solutions consist in analyzing the evolution of the vanishing line. These methods can be used to increase the behavior of the calibration method, but will not be further detailed because they are not the aim of this article.

If we have to define an imaging system [4], we can say that it collects radiation emitted or reflected by objects for future processing. These emitted rays (particle flow, magnetic, or acoustic wave) are projected on a sensor that is constituted by small sensitive surfaces. Thus, the basic concept of calibration is to link the world coordinates of three-dimensional emitting points with their corresponding ones in the image defined by the sensor. Classical calibration methods take at least one picture of a calibration pattern. This calibration pattern can be a chessboard, a pattern of full circles, or whatever known shapes, but the calibration method will rely on these specific points, as the chessboard corners or the center of the circles. The geometry of these points, expressed in the 3D world coordinate system, has to be perfectly known, because they will then be linked to their corresponding points in the image frame.

A. Calibration technique overview

There are several calibration algorithms described in the computer vision literature. They differ mostly by their scope, precision and requirements for their application. Some algorithms [5] [6] require a good initial guess of some parameters, typically the focal length and camera geometry (as yaw, pitch and roll angles). This implies an a-priori knowledge that has to be given by the user. A possible solution to avoid giving initial values, is to use a two-step method [7] where the first run supposes a distortion-free camera model, and the second step uses the prior result as an initial guess in order to find the optimal solution including distortion parameters.
Other algorithms require a non-coplanar calibration pattern [8] or two different calibration patterns [9], implying a longer calibration procedure. [10] and [11] use at least two images where the movement of the camera, or the calibration plane, does not need to be known. Concerning the attainable precision, algorithms based on linear systems do not consider lens radial distortion [12] [13]. This introduces significant errors when working with short focal length optics. Only algorithms with nonlinear optimization [14] can handle radial distortion.

Finally, several methods are being investigated in order to find an optimal calibration procedure using Genetic Algorithms [15], but do not require necessarily a physical model of the optical system.

B. Calibration model

To calibrate a camera mounted on a truck the logical choice would be to place the calibration pattern on the surface where most measurements will be carried out: the road. However, this is not a practical situation because road flatness is difficult to be guaranteed, and thus, the precision of the three-dimensional points of calibration could lead to an inaccurate model. Moreover, if the camera is also used to detect traffic signs, or other long-range vehicles, the area covered by the calibration pattern

Fig. 2. Lane detection without (left image) and with (right image) cumulated gradient information along the lane. Using this global hypothesis, the algorithm succeeds in situations where the contrast of one street side (left one) is not constant and the other side (right one) has no particular marking.

in the image would be very small. Indeed, although the size of calibration pattern laid on the road could be of significant dimension, its projection in the image plane will hardly cover the whole sensor size. The original contribution of this project is to use a vertical calibration plane, placed in front of the camera and covering most of the image, and then to virtually rotate this plane to fit the road surface (see Figs. 4 and 5).

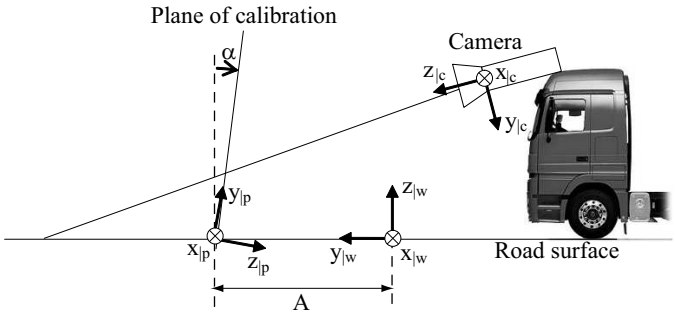

Fig. 4. Sketch of camera geometry with the different coordinate systems (scene projection on the left-side of the truck).

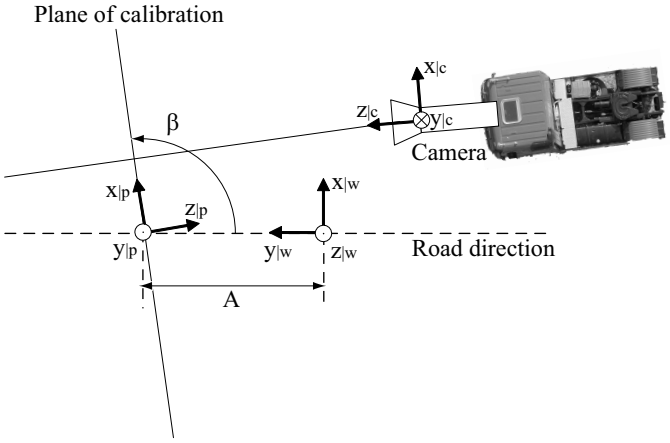

Fig. 5. Sketch of camera geometry with the different coordinate systems (scene projection on the floor of the road).

The aim of this research is not to develop a new calibration algorithm but to use an existing one to calibrate the camera against this vertical plane, and then use simple geometry to recover the position of objects on the road placed in front of the vehicle. This approach is particularly advantageous for calibrating trucks at the end of the production line, because it only requires a vertical plane aligned to the cabin frame.

The first step of our calibration technique will start by calibrating the vertical plane with the vision system. As described in sub-section III-A different procedures exist and could be used according to their performances and user's problem knowledge, as coplanar or non-coplanar procedure. However, several benefits could be highlighted by using the Tsai's calibration [16]. Indeed, this last quoted method computes the image distortion if the calibration patterns covers a large size of the image and it needs only 7 points to compute the calibration. Moreover, this algorithm can deal with coplanar calibration planes and has one of the best calibration accuracies (see [16]). The calibration procedure extracts the following unknowns:

- Extrinsic parameters: the rotation matrix (3 by 3 matrix with 3 unknown angles) and translation vector (3 unknowns), that link the camera position and orientation to the calibration plane.
- Intrinsic parameters: focal length, radial lens distortion (2 unknowns), and scale factor of the camera's optical system.

The second step of our calibration technique is to retrieve the points located on the road by using the extrinsic parameters of the previous calibration (see Figs. 4 and 5).

- $c = O(x_{|c}\,;\, y_{|c}\,;\, z_{|c})$ describes the coordinate system of the camera, where $x_{|c}$ increases from the left side to the right side of the image, $y_{|c}$ is the axis that starts from top side to bottom, and finally $z_{|z}$ is aligned with the optical axis of the imaging system.
- $p = O(x_{|p}\,;\, y_{|p}\,;\, z_{|p})$ describes the coordinate system of the calibration plane. To perform Tsai's [16] calibration, $z_{|p} = 0$.
- $w = O(x_{|w}\,;\, y_{|w}\,;\, z_{|w})$ describes the coordinate system of the road in the 3D world, indeed the straight road is expressed with $x_{|w}$ being a constant. Moreover, a critical hypothesis of algorithm is to define that road is perfectly flat, i.e., $z_{|w} = 0$. Without any loss of generality, the direction defined by the origin of the coordinate systems

w and p should be parallel to $y_{|w}$, which allow to simplify the calibration equations.
- α defines the inclination of the calibration plane with respect to a perpendicular to the road surface.
- β defines the angle between the direction of road and the calibration plane (see Fig. 5).
- the length A is the offset in the straight forward road direction between 3D world coordinate w and calibration plane system p.

Thus, the coordinate system of the 3D world is assigned by the knowledge of the calibration plane system, the angles α and β, and the distance A.

C. 3D world reconstruction after plane calibration

After having calibrated the camera against the calibration plane using one of the method in section III-A, we can recover the extrinsic parameters of the system. Then, it is possible to construct the different transformations that link the various coordinate systems. The second step is to compute the projection of the point situated on the calibration plane to the road surface.

1) Coordinate transformations: Using the extrinsic parameters, we know explicitly

$$\vec{x}_{|c} = R_{pc} \cdot \vec{x}_{|p} + \vec{T}_{cp} \quad (1)$$

$$\Rightarrow \vec{x}_{|p} = R_{pc}^{-1} \cdot \vec{x}_{|c} - R_{pc}^{-1} \cdot \vec{T}_{cp} \quad (2)$$

which describes the relationship between the camera coordinate model c and the calibration plane system p. This transformation can be summarized by a rotation matrix R_{pc} from coordinate system p to c, and a translation vector $\vec{T}_{cp}$.

Following a similar approach, the relationship between the systems w to p can be written

$$\vec{x}_{|p} = R_{wp} \cdot \vec{x}_{|w} + \vec{T}_{pw} \quad (3)$$

with $R_{wp} = \begin{bmatrix} s\beta & -c\alpha \cdot c\beta & -s\alpha \cdot c\beta \\ 0 & -s\alpha & c\alpha \\ -c\beta & -c\alpha \cdot s\beta & -s\alpha \cdot s\beta \end{bmatrix}$

and $\vec{T}_{pw} = A \cdot (c\alpha \cdot c\beta\,;\, s\alpha\,;\, c\alpha \cdot s\beta)^{\mathrm{T}}$ where $s\alpha$ stands for $\sin(\alpha)$, and $c\alpha$ for $\cos(\alpha)$. Thus, the position of points situated on the calibration plane and expressed in w coordinates is defined by

$$\vec{x}_{|w} = R_{wp}^{-1} \cdot \left(\vec{x}_{|p} - \vec{T}_{pw}\right) \quad (4)$$

where $R_{wp}^{-1} = R_{wp}$.

2) Reconstruction of road model: This problem is solved by stating that the object position R expressed in w, its corresponding point B onto the calibration plane and the optical center C of the camera are aligned (see Fig. 6). Thus, using simple geometry manipulation, it is possible to extract the estimated position of the object onto the modeled road by the knowledge of the two other points. The projection onto the calibration plane $\vec{x}_{B|w}$ is directly given by applying successively (2) and (4). The position of the focal of the camera had to be extracted from (2), which gives $\vec{x}_{C|p} = -R_{cp}^{-1} \cdot \vec{T}_{pc}$ when expressed in the p system. Thus, with the help of (4), it is possible to extract the position of the optical center of the camera in w, denoted $\vec{x}_{C|w}$. Finally, if the estimated position into the road model is described by $\vec{x}_{R|w} = \left(x_{R|w}\,;\, y_{R|w}\,;\, z_{R|w}\right)^{\mathrm{T}}$, then

$$x_{R|w} = \frac{x_{B|w} - x_{C|w}}{z_{B|w} - z_{C|w}} \cdot \left(z_{R|w} - z_{C|w}\right) + x_{C|w} \quad (5)$$

$$y_{R|w} = \frac{y_{B|w} - y_{C|w}}{z_{B|w} - z_{C|w}} \cdot \left(z_{R|w} - z_{C|w}\right) + y_{C|w} \quad (6)$$

D. Camera coordinate based on 3D world measurements

This section describes the other way around. Indeed, it is not only useful to get 3D measurements from the image plane of the camera, but in some applications, it is needed to get the pixel coordinates based on a 3D measurement. This is particularly useful when the system has a representation of the world in the natural coordinates, and tries to know the representation of an object in the camera frame. Typical examples are sensor fusion problems. Indeed, such technologies allow to exchange information between different sensors. Thus, just like human drivers, the fusion module can be suddenly interested in a specific region of the environment. This ROI can then be deeper analyzed by several sensors which, in order to achieve such a goal, should be able to translate the information coming from the environment to the pixel coordinates of the camera frame.

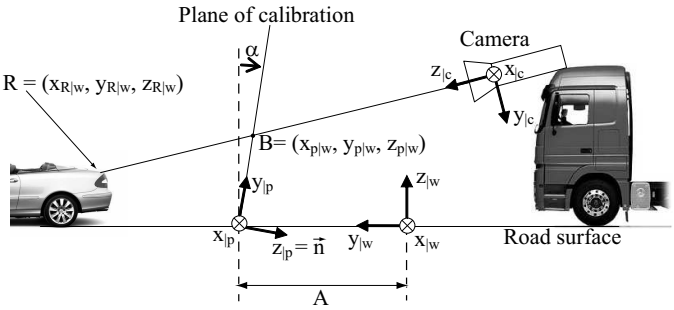

Fig. 6. Sketch of geometrical representation of object position R, and its ray intersection B with the calibration plane.

Analyzing figure 6, we are trying to find a methodology to compute $\vec{x}_{|c}$ when we know $R = \left(x_{R|w}, y_{R|w}, z_{R|w}\right)$. Referring to appendix I, we use equations (8), (9) and (10), in order to get the coordinates of $B = \left(x_{p|w}, y_{p|w}, z_{p|w}\right)$. Then, it is obvious that B should be expressed in the coordinate frame of the calibration plane, which can be easily done with (3). Finally, the information in the coordinate system of the camera ($O(x_{|c}\,;\, y_{|c}\,;\, z_{|c})$) can be found by using the previous calibration technique (see section III).

E. Calibration procedure

The calibration procedure can be separated in two steps: the initialization, and the iterative part which has to be performed on each step. Thus, the initialization phase begins by placing a calibration plane in front of the camera, and pre-compute R_{wp}^{-1} and $-\ R_{wp}^{-1} \cdot \vec{T}_{pw}$ described in equations (3) and (4). Then,

using one of the existing calibration procedure (see section III-A), extract the extrinsic parameter of the calibration and perform the computation of $\vec{x}_{C|w}$ using equation (4).
In order to recover the 3D position $\vec{x}_{R|w}$ corresponding to a specific pixel position, apply the calibration transformation to recover the $\vec{x}_{B|p}$ position on the calibration plane, which can be converted to $B = \vec{x}_{B|w}$ with (4). Finally, using equations (5) and (6), lead to recover $R = \vec{x}_{R|w}$.
Similarly, in order to recover the pixel position corresponding to a given point $R = \vec{x}_{R|w}$, then apply equations (8), (9) and (10) in order to recover $B = \vec{x}_{B|w}$. Note that several elements of (10) can also be pre-computed. Then, $\vec{x}_{B|p}$ is found by using (3) and finally, the reverse computation of the chosen calibration will lead to the pixel position into the camera image.

IV. Experimental results

This section presents the initial tests that have been performed in order to characterize the robustness of the approach. The different experiments depicted here use a camera placed in a car with a vertical resolution of 400 pixels, positioned at a height of $1.15\ m$ and where the calibration plane has been placed with $\beta \cong 90^{\text{o}}$ and $A = 1.148\ m$.

The calibration process has been done using the coplanar algorithm of [16], on a target of three by five patterns. These 15 points are used to calibrate the camera, and hence to find its intrinsic and extrinsic parameters. In order to correct the radial distortion, the calibration target covers the whole image taken by the camera. The focal length was of $6.49\ mm$ and the orientation of c has an angle of 8^{o} around $x_{|c}$.

After having performed the calibration, we moved a vertical test target of $40\ cm$ per $40\ cm$ between 2.8 and $11.5\ m$ along the road direction $y_{|w}$. The distance has been measured with a laser that has a precision of $\pm 2\ mm$, and a repeatability of 95%. Then, the computational procedure described in section III has been applied.

The results are summarized in figure 7, which shows that the factor α has a severe influence on results. Indeed, if the calibration plane is supposed to have $\alpha = 0^{\text{o}}$, we can see that the relative estimation error seems to rapidly explode (more than 55% of error for an object placed at $11.5\ m$). A negative error stands when the calibration result underestimate the correct measurment; similarly, a positive error on figure 7, describes a calibration result that is greater or equal to the real measurment. Replacing α by -1.3^{o} results in a mean error of -2% when computed on the whole range of measured distance. Moreover, by setting $\alpha = -3.0^{\text{o}}$, the relative error tends to be quite constant.

As shown in figure 8, by subtracting the mean of measurements error (with $\alpha = -3^{\text{o}}$), the precision reaches less than $\pm 1\%$ in the distance range of $[2.8\ m; 11.5\ m]$. This method requires several measurements to be able to compute the mean of experiments. An alternative solution is to average the errors at the range limits. In this case the maximum relative error was found to be 1.4%.

Similarly, another experiment has been conducted in order to increase the measurement range. In this second case, the result are presented in figure 9. It shows that this experiment exhibits less than 1.4% of error in the range of $[4.3\ m; 49.7\ m]$. This last peformance should be taken really carefully. Indeed, as previously described, the target object used to measure the distance has a square dimension of about $40\ cm$ which could be considered as big when seen at a distance of $4.3\ m$. However, the same object seen by the camera at a distance of $49.7\ m$ is represented by only about 4 pixels! The problem is then not the precision of the calibration technique, but how precisely it is possible to determine the target with so few pixels. However, the goal of this experiment is not to caracterize the calibration procedure, because it is an exact mathematical development, and thus have the same characteristics of the calibration method used (in this experiment it is

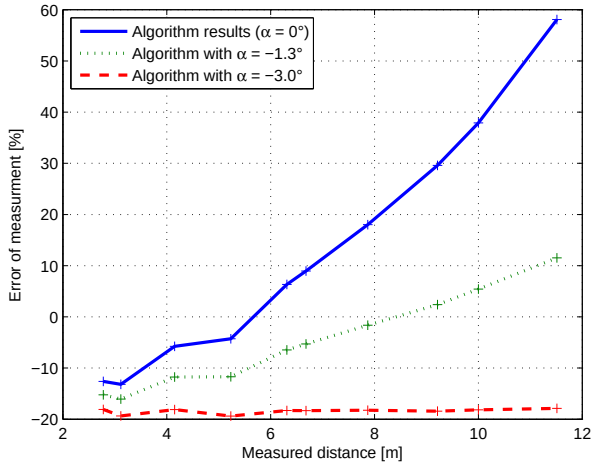

Fig. 7. Result of calibration method error. The graphics highlight that the influence of parameter α is critic to have a perfect reconstruction of road model. By supposing a perfect verticality of calibration plane ($\alpha = 0^{\text{o}}$), the reader can see that an error of more than 55% at $11.5\ m$ could take place in this particular example. However, if $\alpha = -3^{\text{o}}$, the relative error of measurement is quite constant on the whole range of measured distance.

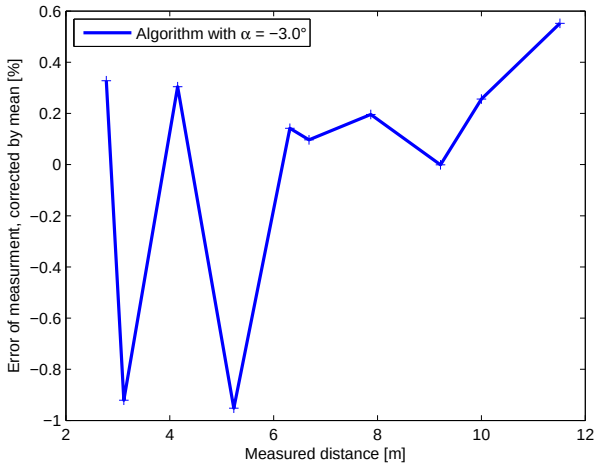

Fig. 8. Relative error results of calibration method for $\alpha = -3^{\text{o}}$ with suppression of relative error offset.

Tsai's algorithm [16]). But rather, to extract the result of a practical calibration technique which could be easily applied at the end of a production line.

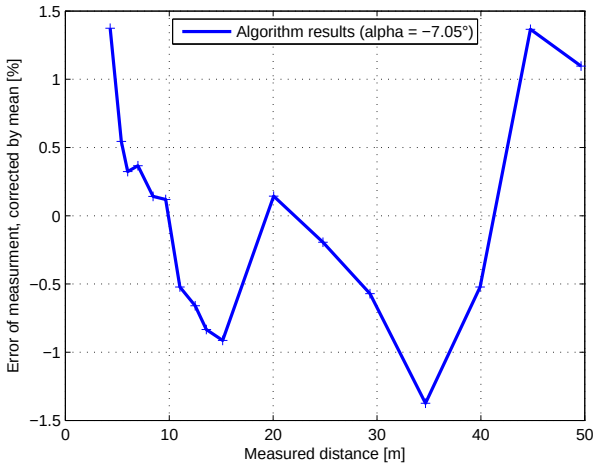

Fig. 9. Result of calibration method for the second experimental measurement. This data set has been calibrated with an angle $\alpha = -7.05^{\circ}$, and the results are displayed as explained in chapter IV with suppression of the relative error offset.

A. Calibration improvement

As it was exposed by experimental analysis, the measurement of the angle α is a critical step for calibration result. It was observed in experimental results (see Fig. 7), that the relative error of measurement when $\alpha = -3.0^{\circ}$ has an important offset. After the initial investigations, it could be pointed out that this bias was partially due to an inaccurate estimation in the camera position. Indeed, the procedure used to calibrate p to the camera coordinate system c, presents an average measurement error of about 13%. Moreover, further experiments show that the initial position of the calibration plane has also a severe influence on this bias. However, we will now present two procedures to find and suppress this offset:

- In the production process, the position of truck and calibration plane is perfectly mastered. Thus, the knowledge of angle the α can be measured with a precision of less than 0.05°. In this case, the environment can be changed to fit the requirement of the calibration method.
- For experimental tests, the method consists in a two-step calibration process. The first step has been described in the previous section and will use the most precise measure, of the critical angle α. This measure will largely depend on the tool that will be used to determine the inclination of calibration plane. However, this measurement does not need to be extremely precise, because the second step of the calibration will be used to determine it much more accurately. Indeed, after the first calibration, we measure the distance to two objects, one situated in a close range, for example $3\ m$, and a second one in a far range $40\ m$. Having these two measurements, it is now possible to determine experimentally the angle α to fit precisely these distances.
  Otherwise, if it is possible to guarantee that the deviation of the angles α and β from their default value are small, then it is possible to compute their estimated value by using equations (13) and (14) of appendix II.

Both methods, not only allow to find the optimal α, but they can determine the correction factor that could be applied to correct the shift in the estimation error.

B. Parameters analysis

The calibration procedure of this article has three main parameters, which are α, β and A (see Figs. 4 and 5 and section III-B). However, the results only describe the influence of α on the calibration. This can be explained because this angle is certainly difficult to measure, but also because it has the bigger impact on the calibration results for a longitudinal measurement (along $y_{|w}$). Indeed, looking at equation (6) and (4), it can be shown that β has a smaller influence that α on $y_{|w}$. In order to validate this statement, we can observe the error measurements when varying independently these angles.

As seen in Fig. 10, the variation of relative error of $y_{|w}$ is much less important when β varies from 80° to 100°, than when α varies from -8° to -4° (see Fig. 11).

V. Conclusion

The SPARC project develops new technologies to improve drivers' security in the next generation of commercial vehicles. It states that there is no sensor from which it is possible to extract a sufficient amount of information to protect the driver in every situation that can occur on roads. Thus, different sensors based on different physical principles are fused to reconstruct the 3D scene of the truck's environment. This fusion has two advantages, the first one is to be quite robust if a single sensor become defective. Moreover, by fusing different data types, it is possible to obtain more robust and precise information.

One of these sensors is a camera that is connected to a vision platform that will extract the more useful information to achieve the security goal of SPARC. In this paper, we rapidly describe an algorithm to perform lane detection. The approach is based on a multiple lane models and multiple hypotheses of detection.

Based on these results, it is now necessary to determine the position of objects and obstacles on the road. For that, it is needed to relate the pixel-based information of the camera to the 3D scene. This calibration process will also be useful for other algorithms, such as computation of lane curvature and width. This article presents a calibration method that can be used at the end of production line. Instead of performing the calibration directly on the road surface, which is particularly difficult to realize for space and cost reasons, this method performs the calibration on a vertical plane set in front of the truck. We present the transformation that can be applied to compute the calibration, and present the results of a car calibration. These results show that the verticality (α) of the calibration plane is a critical parameter. Two solutions have been proposed, the first one is to control perfectly this

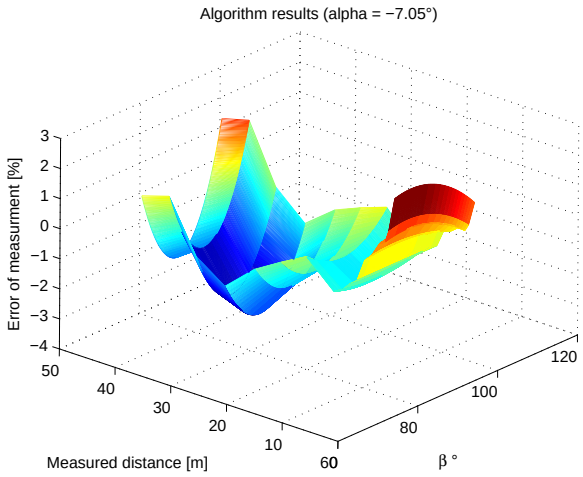

Fig. 10. Relative error measurement of $y_{|w}$ when $\beta \in [80^{\text{o}}; 100^{\text{o}}]$.

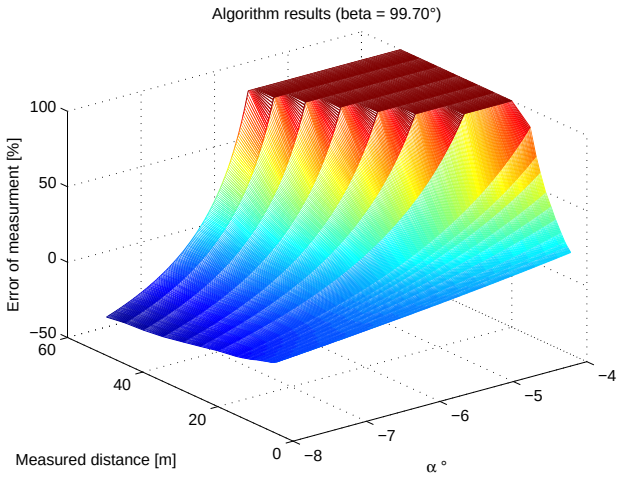

Fig. 11. Relative error measurement of $y_{|w}$ when $\alpha \in [-8^{\text{o}}; -4^{\text{o}}]$. The results have been truncated to a maximal error of 100% in order to improve lisibility.

angle with a precision of less than 0.05^{o}. However, should the measurement of α be difficult, a second method was proposed which does not require to know the exact inclination of calibration plane.

Finally, the results show that a precision of less than $\pm 1\%$ in the range of $[2.8\ m; 11.5\ m]$ can be obtained with an extremely simple setup, and less than $\pm 1.4\%$ in the range $[4.3\ m; 49.7\ m]$.

Appendix I
Proof of the reverse computation

It could be essential to be able to reconstruct the image position based on a known point in the world coordinate frame. To achieve this goal, we have to first find the intersection of the light ray linking the desired point R to the camera C and the projection plane (see Fig. 6). Then, using the reverse formula of the chosen calibration technique, it is possible to get the image coordinates of the given point. Initially we have to characterize the calibration plane in the world coordinate. By definition, a plane can be modeled with its normal vector $\vec{n}$, and a point placed on it. By using equation (4) and setting $\vec{x}_{|p} = (0\,;\,0\,;\,1)^{\text{T}}$, which is expressed in the plane coordinate and is by definition orthogonal to the calibration plane, we get $\vec{n}_{|w} = (-c\beta\,;\,-c\alpha \cdot s\beta\,;\,-s\alpha \cdot s\beta)^{\text{T}}$. Thus, we can extract a possible representation of the calibration plane in the world coordinate which passes through the specific point $(0\,;\,A\,;\,0)_{|w}$:

$$c\beta \cdot x_{|w} + c\alpha \cdot s\beta \cdot y_{|w} + s\alpha \cdot s\beta \cdot z_{|w} - c\alpha \cdot s\beta \cdot A = 0 \quad (7)$$

If the coordinate $z_{R|w}$ is different from $z_{C|w}$, which implies that the points R and C have different heights in the world coordinate w, then the intersection between the object seen by the camera and the calibration plane, can be found by isolating $x_{B|w}$ and $y_{B|w}$ from equation (5) and (6):

$$x_{B|w} = \frac{x_{R|w} - x_{C|w}}{z_{R|w} - z_{C|w}} \cdot \left(z_{B|w} - z_{C|w}\right) + x_{C|w} \quad (8)$$

$$y_{B|w} = \frac{y_{R|w} - y_{C|w}}{z_{R|w} - z_{C|w}} \cdot \left(z_{B|w} - z_{C|w}\right) + y_{C|w} \quad (9)$$

and replace them in the equation (7) of the calibration plane, which gives the last unknow parameter $z_{B|w}$.

Thus, the light ray starting from the observed object to the camera, intersects the calibration plane in the point $B = \left(x_{B|w}, y_{B|w}, z_{B|w}\right)$ when expressed in world coordinate, where $x_{B|w}$ is given by equation(8), $y_{B|w}$ by (9) and finally, $z_{B|w}$ is obtained with equation (10).

Appendix II
Calibration plane inclination computation

In order to test the inclination of calibration plane, we will provide a control formula. Thus, knowing the exact position of a point $R = \left(x_{R|w}, y_{R|w}, z_{R|w}\right)$ in the world coordinate system, and its corresponding point in the image coordinate, it is possible to determine the angles α and β.

Starting from the image coordinate system, and using the selected calibration technique (for example Tsai's), it is possible to get the corresponding point $B = \left(x_{B|w}, y_{B|w}, z_{B|w}\right)$. Moreover, it is easy to have the position of the camera $C = \left(x_{C|w}, y_{C|w}, z_{C|w}\right)$, which is an extrinsic parameter of the calibration procedure. As the points B and C are naturally expressed in the coordinate system of the calibration plane, we will transform $R_{|w}$ in $R_{|p}$ by using the equation (3). Applying this transformation, the point $R_{|p}$ will let appear the angles α and β which are the unknown. Stating that the three points (R, B and C) are on a straight line, we get the following

$$
\begin{aligned}
z_{B|w} &= \frac{\left(c\alpha \cdot s\beta \cdot (A - y_{C|w}) - c\beta \cdot x_{C|w}\right) \cdot \left[z_{R|w} - z_{C|w}\right] + c\beta \cdot z_{C|w} \cdot \left[x_{R|w} - x_{C|w}\right] + c\alpha \cdot s\beta \cdot z_{C|w} \cdot \left[y_{R|w} - y_{C|w}\right]}{c\beta \cdot \left[x_{R|w} - x_{C|w}\right] + c\alpha \cdot s\beta \cdot \left[y_{R|w} - y_{C|w}\right] + s\alpha \cdot s\beta \cdot \left[z_{R|w} - z_{C|w}\right]} \\
&= \frac{\begin{pmatrix} z_{C|w} \cdot c\beta \\ z_{C|w} \cdot c\alpha \cdot s\beta \\ \left(A - y_{C|w}\right) \cdot c\alpha \cdot s\beta - x_{C|w} \cdot c\beta \end{pmatrix} \cdot \left[\vec{x}_{R|w} - \vec{x}_{C|w}\right]}{\begin{pmatrix} c\beta \\ c\alpha \cdot s\beta \\ s\alpha \cdot s\beta \end{pmatrix} \cdot \left[\vec{x}_{R|w} - \vec{x}_{C|w}\right]}
\end{aligned} \tag{10}
$$

conditions:

$$\frac{\left(x_{B|p} - x_{R|p}\right)}{\left(z_{B|p} - z_{R|p}\right)} = \frac{\left(x_{C|p} - x_{B|p}\right)}{\left(z_{C|p} - z_{B|p}\right)} \tag{11}$$

$$\frac{\left(y_{B|p} - y_{R|p}\right)}{\left(z_{B|p} - z_{R|p}\right)} = \frac{\left(y_{C|p} - y_{B|p}\right)}{\left(z_{C|p} - z_{B|p}\right)} \tag{12}$$

In order to simplify the resolution of the above equations, the angle α is approximated with a limited development of first order around $0\ rad$ which gives $c\alpha \to 1$, $s\alpha \to \alpha$ and similarly, a small angle approximation on β around $\frac{\pi}{2}\ rad$ leads to replace $c\beta \to \frac{\pi}{2} - \beta$ and $s\beta \to 1$ in the last set of conditions. With some simple, but fastidious, computations, one can find that the angle α is approximated by equation (13), and β is obtained from equation (14). This approximation is the general solution of the equation system, but several particular solutions exist and could lead to simpler solutions, as example for $z_{R|w} = 0$.

APPENDIX III
ACKNOWLEDGMENTS

The authors are funded by the European project SPARC, No. IST-507859.

REFERENCES

[1] M. Shell. (2002) Final report of the european eSafety working group on Road Safety. [Online]. Available: http://www.eu.int/information_society/programmes/esafety/index_en.htm

[2] A. Maisch. (2004) An european funded project : Powertrain Equipped with Intelligent Technologies (PEIT). [Online]. Available: http://www.eu-peit.net/

[3] M. Bellino and Y. L. de Meneses, "Lane detection algorithm for an onboard camera," in *Proc. EPIC/SPIE first Workshop on Photonics in the Automobile*, Geneva, Switzerland, Dec. 2004.

[4] B. Jähne, Ed., *Digital Image Processing*. Springer, 2002, vol. 5th ed., pp. 143–280.

[5] Y. I. Abdel-Aziz and H. M. Karara, "Direct linear transformation into object space coordinates in close-range photogrammetry," in *Proc. Symp. Close-Range Photogrammetry*, Urbana, Ill., 1971, pp. 1–18.

[6] M. Fischler and R. Bolles, "Random sample consensus: A paradigm for model fitting applications to image analysis and automated cartography," in *in Proc. Image Understanding Workshop*, 1980, pp. 71–88.

[7] J. Weng, P. Cohen, and M. Herniou, "Camera calibration with distortion models and accuracy evaluation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 10, pp. 965–980, October 1992.

[8] Q.-T. Luong and O. Faugeras, "Self-calibration of a moving camera from point correspondences and fundamental matrices," *International Journal of Computer Vision*, vol. 22, no. 3, pp. 261–289, 1997.

[9] H. A. Martins, J. R. Birk, and R. B. Kelly, "Camera models based on data from two calibration plane," in *IEEE Computer Graphics and Applications*, vol. 17, 1981, pp. 173–180.

[10] Z. Zhang, "Flexible camera calibration by viewing a plane from unknown orientations," in *International Conference on Computer Vision (ICCV'99)*, One Microsoft Way, Redmond, WA 98052, September 1999, pp. 666–673, corfu, Greece. [Online]. Available: http://research.microsoft.com/~zhang

[11] ——, "A flexible new technique for camera calibration," Microsoft Research, Microsoft Corporation, One Microsoft Way, Redmond, WA 98052, Tech. Rep. MSR-TR-98-71, 1998 (revisited in 2002). [Online]. Available: http://research.microsoft.com/~zhang

[12] W. Faig, "Calibration of close-range photogrammetry systems: Mathematical formulation," in *Photogrammetric Eng. Remote Sensing*, vol. 41, 1975, pp. 1479–1486.

[13] I. Sutherland, "Three-dimensional data input by tablet," in *Special Issue of the IEEE Computer Graphics and Applications*, vol. 62, 1974, pp. 453–456.

[14] R. Hartley and A. Zisserman, *Multiple View Geometry in computer vision*. Cambridge university press, 2000, vol. 5th ed.

[15] Q. Ji and Y. Zhang, "Camera calibration with genetic algorithms," *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, vol. 31, no. 2, pp. 120–130, March 2001.

[16] R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE Journal of Robotics and Automation*, vol. RA-3, no. 4, pp. 323–344, 1987.

[17] (2003) Basics of projective geometry. [Online]. Available: http://cmp.felk.cvut.cz/~hlavac/Public/Pu/33PVRleto2003/p1Geom1cameraFeb2003.pdf

$$\alpha \cong \frac{\gamma+\delta+2\cdot z_{R|w}\cdot\big((y_{B|p}-y_{C|p})\cdot y_{R|w}+A\cdot(y_{C|p}-y_{B|p})\big)\mp\sqrt{\eta}}{2\cdot z_{R|w}\cdot\big((A-y_{R|w})\cdot(z_{C|p}-z_{B|p})+(y_{C|p}-y_{B|p})\cdot z_{R|w}\big)} \tag{13}$$

$$\beta \cong \frac{\gamma-\delta+\pi\cdot x_{R|w}\cdot z_{R|w}\cdot(y_{C|p}-y_{B|p})\mp\sqrt{\eta}}{2\cdot x_{R|w}\cdot(y_{C|p}-y_{B|p})\cdot z_{R|w}} \tag{14}$$

$$\gamma = (x_{B|p}-x_{C|p})\cdot x_{R|w}\cdot y_{R|w}+(y_{R|w}^2+A^2)\cdot(z_{C|p}-z_{B|p})+A\cdot\big((x_{C|p}-x_{B|p})\cdot x_{R|w}+(z_{B|p}-z_{C|p})\cdot 2\cdot y_{R|w}\big) \tag{15}$$

$$\delta = z_{R|w}\cdot\big(y_{B|p}\cdot z_{C|p}-y_{C|p}\cdot z_{B|p}-z_{R|w}\cdot(z_{C|p}-z_{B|p})\big) \tag{16}$$

$$\begin{aligned}\eta = {} & \Big((A-y_{R|w})\cdot\big((x_{C|p}-x_{B|p})\cdot x_{R|w}+(A-y_{R|w})\cdot(z_{C|p}-z_{B|p})\big)\cdots\\ & +\big(2\cdot A\cdot(y_{C|p}-y_{B|p})+y_{B|p}\cdot(2\cdot y_{R|w}+z_{C|p})+y_{C|p}\cdot(2\cdot y_{R|w}+z_{B|p})\big)\cdot z_{R|w}-(z_{C|p}-z_{B|p})\cdot z_{R|w}^2\Big)^2\cdots\\ & -4\cdot z_{R|w}\cdot\big((A-y_{R|w})\cdot(z_{C|p}-z_{B|p})+(y_{C|p}-y_{B|p})\cdot z_{R|w}\big)\cdot\Big((x_{R|w}^2+A^2+y_{R|w}^2)\cdot(y_{C|p}-y_{B|p})\cdots\\ & +x_{R|w}\cdot\big(x_{C|p}\cdot(y_{B|p}-z_{R|w})+x_{B|p}\cdot(z_{R|w}-y_{C|p})\big)+y_{R|w}\cdot\big(y_{C|p}\cdot z_{B|p}-y_{B|p}\cdot z_{C|p}+z_{R|w}\cdot(z_{C|p}-z_{B|p})\big)\cdots\\ & +A\cdot\big(y_{B|p}\cdot(2\cdot y_{R|w}+z_{C|p})-y_{C|p}\cdot(2\cdot y_{R|w}+z_{B|p})+z_{R|w}\cdot(z_{B|p}-z_{C|p})\big)\Big)\end{aligned} \tag{17}$$

[18] I. Shimizu, Z. Zhang, S. Akamatsu, and K. Deguchi, "Head pose determination from one image using a generic model," *IEEE 3rd. International Conference on Face & Gesture Recognition*, pp. 100–105, 1998.

[19] F. Maxim, "Real-time infrared tracking system for virtual environments: from ideas to a real prototype," *Virtual Environment on a PC Cluster Workshop 2002*, 2002.

[20] P. F. Sturm and S. J. Maybank, "On plane-based camera calibration: A general algorithm, singularities, applications," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, p. 437, 1999.