Augmented Reality with Automatic Camera Calibration for Driver Assistance Systems


Graduate School of Applied Computing
Universidade do Vale do Rio dos Sinos - Unisinos

crjung@unisinos.br, jstreibel@gmail.com, lgonzaga@unisinos.br

Abstract
This paper presents a new approach for driver assistance systems based on automatic camera calibration and Augmented Reality. While an automatic camera calibration technique can obtain information about the real environment, an Augmented Reality component supplies the virtual content to produce an enriched scene. This resulting scene can assist drivers in their primary tasks and increase driving security. The virtual content can be alerts and warnings, enhanced obstacles and traffic signals, and any information from car sensors and GPS (global positioning system). Besides, the proposed solution decreases driver’s distraction by superimposing virtual content to real world through a head-up display. It gives the sensation that real and virtual contents are merged into the windshield, keeping the attention focus of drivers in their primary tasks.

1. Introduction
The increasing of navigation systems, mainly based on GPS (global positioning system) integrated with map providers, and cruise control have helped drivers to plan routes efficiently and turn driving tasks more comfortable, respectively. However, these instruments, jointly with other gadgets like cell phones, sound systems, television, voice command systems and so on, have increased the number of control buttons and displays in car cockpits, and therefore causing driver distractions and provoking accidents [4, 16]. We believe a semi-immersive solution, based on superimposing virtual content to the real environment, can conveniently supply drivers with essential and luxury information, decreasing the risks of distraction and accidents. But, it requires obtaining the real environment information, the virtual content itself, mechanisms to merge these information and conveniently display them.

This paper provides an integrated solution to make feasible driver assistance systems based mainly on the synergy between an automatic camera calibration method and Augmented Reality (AR) techniques [1, 12]. The system uses a high-quality digital camera fixed inside the car, which captures a video along the road. The camera calibration component processes the video for estimating relative position and orientation between the camera and the road. This issue has received lots of efforts in last years due to the request of industry to increase vehicle safety and to develop a autonomous navigation systems.

On the other hand, Augmented Reality (AR) provides the combination of real-world and virtual content, by superimposing graphics objects upon or composed with the real world in real-time. Virtual camera uses the real camera parameters to register virtual information to the real world. The main aim is to supply information to enhance and enrich user’s view about the world and transform the focus of interaction. Besides, AR has been exhaustively explored in computer applications, here we have adapted its functionalities to improve the
restrictive car environment.

The decreasing risks of driver’s distraction depends on keep their attention on the road. This approach fixes this problem changing the information viewing from the central display to a head-up display (HUD) projected directly in the car windshield. The HUD can beyond display textual information, it can take advantage of projective mapping of the road to vary the size of virtual 3-D objects according to real obstacles distance.

This paper is organized as follows: research papers and industrial applications about driver assistance systems and employed technologies are summarized in Section 2; Detailed description about system architecture is presented in Section 3, focusing an automatic camera calibration method and virtual content superimposing. Section 4 and Section 5 shows experimental results and conclusions, respectively.

2. Related work

Optical cameras are among the most advantageous sensors to equip a vehicle for obtaining information about the surrounding environment, since they provide rich information, are software configured, and can be acquired at low cost. For some applications, such as lane departure warning systems [11], the analysis of image coordinates only can be sufficient. However, camera calibration is required to determine the distance and relative speed to other vehicles or obstacles in world coordinates. Several calibration methods are found in literature, using different approaches and techniques. Some authors use more then one camera to have 3D information of the world or allow the camera to move to explore projective geometry constraints [17].

Other techniques rely on a single fixed camera. Guiducci [7] uses information of the road itself, like road width, length and spacing between longitudinal discontinuous stripes, allied with camera parameters. The method described in [3] makes use of geometrical structures (calibration pattern) with a motionless vehicle. This calibration also can be made in the production line, but it is very sensitive, any camera displacement or change in the vehicle, like tires change, jeopardizes the calibration and it must be remade. Other techniques include correction of geometrical lens distortion [8] increasing precision at cost of a complicated analysis.

The car technologies provide the growth of functionalities that enhance the safety, performance and level of comfort for drivers and passengers, through airbags, cruise control, navigation systems and so on. The majority of these functionalities is supplied through a central information display. Others sophisticated sensors can supply a lot of information, including a video camera. It can be mounted for capturing a video of the road. On the other hand, Augmented Reality has emerged as a new paradigm for immersive 3D user interaction in computer-based applications and under their requirements [2].

Nowadays, the main endeavour of researches are to make the synergy between original immersive computer-based interfaces and constrained automotive domain [13,15]. In this direction, the design of interfaces to drivers has been investigated [15] and experimental studies have compared different approaches to inform drivers about dangerous situations, employing Augmented Reality [16]. Car industry has provided active system to cruise control, integrated GPS, and many sensors providing simple information – speed, navigation directions and alerts – for visualization directly in the driver’s field of vision, for example BMW 5 Series [6] has adopted a similar solution. Indeed, BMW has supported the development of 3D interaction metaphors for in-vehicle information systems [5]. Beyond 3D interfaces, 2D widgets continue being developed to compound the complex car drivers interfaces, mainly to alert drivers and provide textual information through HUDs [10].

3. The proposed approach

We propose a new approach for driver assistance system through a synergy between automatic camera calibration and Augmented Reality. Figure 1 gives an overview about the entire system architecture and, illustrates the original road (left), virtual road boundary superimposed on the original (right) and crosses used to enhance the road and by the camera calibration process, respectively.

The camera mounted into the vehicle captures video of the environment in front of the car, while the camera calibration component
obtain, through a lane detection algorithm, the parameters for projection mapping. These parameters are so relevant for superimpose. For achieving correct image registration between real and virtual scenes, the relative position and orientation of real camera are employed by virtual one.

We have employed Augmented Reality techniques to perform coherent cameras positioning and orientation. Virtual camera can use the exactly or modified parameters of real one according with the system configuration. Unlike the conventional desktop or fully immersive systems where images are usually registered using the same parameter, in the current application, the driver (viewer) and the real camera are in slightly different positions. Then, small adjustments of virtual camera parameter can be performed to achieve best results. The real camera pose and equivalent adjustments of virtual one allow the correctly positioning and orientation of virtual objects into the scene, such as texts, signs, and mainly 3-d objects. Besides, 3-d objects can vary their dimensions at a varying distance. Because this kind of approach can help drivers to detect the proximity of obstacles and other vehicles, and system can emit warning alerts to avoid collision and accidents.

3.1. Visualization Subsystem

Let us assume that the camera is mounted in the interior of the vehicle, with its optical axis aligned with the central axis of the vehicle (only yaw, no pitch or roll), as shown in Figure 2. In normal driving conditions on a straight portion of the road, it is also reasonable to assume that the central axis of vehicle is parallel to both lanes. If \( x_w = (x_w; y_w; z_w) \) denote the coordinates of a 3d point in world coordinates, the corresponding coordinates \( x_c = (x_c; y_c; z_c) \) in the camera 3d coordinate system is:

\[
x_c = R(x_w - x_0)
\]

where \( x_0 = (h; 0; 0) \) is the position of the camera in world coordinates, as shown in Figure 2.

Using homogeneous coordinates \( X_w = (x_w; y_w; z_w) \) and \( X_c = (x_c; y_c; z_c) \), such transformation can be written as

\[
X_c = G X_w, \quad G = \begin{bmatrix} R & -Rx_0 \\ 0 & 1 \end{bmatrix},
\]

where

\[
R = \begin{bmatrix}
\cos \alpha & 0 & -\sin \alpha \\
0 & 1 & 0 \\
\sin \alpha & 0 & \cos \alpha
\end{bmatrix}
\]

is the rotation matrix around the y axis. Finally, the image coordinates \( (u; v) \) using perspective projection (pinhole camera) are given by

\[
u = \frac{fx_c}{z_c}, \quad v = \frac{fy_c}{z_c},
\]

and \( f \) is the focal length of the camera. For sake of notation, it is important to mention that \( u \) is the vertical coordinate (grows upwards) and \( v \) is the horizontal component (grows rightwards), and the origin \( (c; 0) \) is the optical axis of the camera.
Assuming that the road is flat, points on the road will lie on the $y_w x z_w$ plane, as illustrated in Figure 3. If the vehicle is moving parallel to lane boundaries in a straight of the road (which usually happens in normal driving conditions), with a lateral offset $y_0$ w.r.t. central axis of the lane, the equations of the left and right lane boundaries are:

Left: $y_w = y_0, z_w = 0$. Right: $y_w = y_0 + W, z_w = 0$, \hspace{1cm} (5)

where $W$ is the width of the lane. Applying the perspective projection given by Equation (4), these lines relate to two other lines $v = f(u)$ and $v = f'(u)$ in image coordinates, given by:

$$f(u) = \frac{(\cos(\alpha)u + \sin(\alpha)f) y_0}{h}, \hspace{1cm} (6)$$

$$f'(u) = \frac{(y_0 + W) \cos(\alpha)u + f \sin(\alpha)(y_0 + W)}{h} \hspace{1cm} (7)$$

The intersection of the two lane boundaries in image coordinates happens at the vanishing point $v = (v_u, v_v)$, given by:

$$v_u = -f \tan(\alpha), \hspace{0.5cm} v_v = 0. \hspace{1cm} (8)$$

Also, after some algebraic manipulation, it can be shown that

$$y_0 = \frac{W m_l}{m_r - m_l}, \hspace{1cm} (9)$$

where

$$m_l = f(u) = \frac{-y_0 \cos(\alpha)}{h}, \hspace{0.5cm} m_r = f'(u) = \frac{-(y_0 + W) \cos(\alpha)}{h} \hspace{1cm} (10)$$

are the angular coefficients of the left and right lane boundaries in image coordinates, respectively. Also, let us consider a rectangular portion of the road, delimited by the bottom-left and top-right points $W_1 = (x_1; y_1; z_1)$ and $W_2 = (x_2; y_2; z_2)$, given respectively by (clockwise, from bottom-left point):

$$W_1 = (0, y_0, z_0), W_2 = (0, y_0 + W, z_0 + R), \hspace{1cm} (11)$$

which are mapped using Equations (2) and (4) to the corresponding image points $I_1 = (u_1; v_1)$ and $I_2 = (u_2; v_2)$:

$$I_1 = \left( \frac{f((y_0 + W) \sin(\alpha) + x_1 \cos(\alpha) - f y_0)}{h \sin(\alpha) - f y_0 \cos(\alpha)} \right), \hspace{1cm} (12)$$

$$I_2 = \left( \frac{f((\cos(\alpha)u + \sin(\alpha)f) y_0)}{h \sin(\alpha) - f y_0 \cos(\alpha)} \right). \hspace{1cm} (13)$$

It is not difficult to show that

$$\frac{v_y(v_2 - v_1)}{v_1 v_2} = \frac{R \sin(\alpha)}{y_0}, \hspace{1cm} (14)$$

and then the yaw angle $\alpha$ can be computed through:

$$\alpha = \arcsin \left( \frac{v_0 v_y (v_2 - v_1)}{R v_1 v_2} \right). \hspace{1cm} (15)$$

After computing $\alpha$, $h$ can be isolated in Equation (10), leading to

$$h = \frac{-y_0 \cos(\alpha)}{m_l}. \hspace{1cm} (16)$$

and the focal length $f$ can be isolated in Equation (8), yielding

$$f = \frac{-v_u}{\tan(\alpha)}. \hspace{1cm} (17)$$

It is important to notice that all parameters needed to compute $\alpha$, $f$ and $h$ can be obtained automatically using a lane detection scheme (in this work, we used the approach described in [9], except for the lane width $W$ (which can be known a priori, since the width is determined when road is built) and the longitudinal parameter $R$. In fact, $R$ can be also obtained

![Figure 3. Left: road in world coordinates. Right: road in image coordinates](image-url)
automatically if the road contains dashed markings, following the approach presented in [14]. Basically, the distance between two adjacent markings can be computed by counting the number of markings within a time interval, and also the distance traveled by the vehicle within such interval. Then, given a frame of the video sequence, we can define the control points \( W_1 \) and \( W_2 \), described in Equation (11) to match the beginning of two adjacent lane markings, as illustrated in Figure 4. More details on estimating \( R \) can be found in [14].

![Figure 4. Placement of control points \( I_1 \) and \( I_2 \) in the beginning of adjacent lane markings](image)

### 3.2 Adding The Virtual Content

The Augmented Reality techniques require two steps to superimpose the virtual content to real scene: the virtual camera adjustments and the content itself. The former depends on real camera pose obtained automatically by the calibration, while the last one depends on application requirements. This architecture system supports pure virtual information display and visualization of measurements obtained directly by car sensors and external equipments (ex. GPS), such as speed, oil pressure, global position, distances, directions, alerts and warnings messages. The system capabilities are huge, but a large number of information can cause distraction, too. Essential information is required and luxury can be activated on demand and under excellent drive conditions.

The parameters \( \alpha \), \( f \) and \( h \) are then used to place a virtual camera with focal length \( f \) and yaw angle \( \alpha \) at the position \( (h; 0; 0) \), as shown in Figure 5. To visually check if the camera parameters found were coherent with the real camera parameters, we have drawn virtual lanes on the road image the lanes on the ground plane after applying the camera transformations as calculated before. The results, although very coherent, still don’t quite match the lane’s borders. This can be understood, since the algorithm used to find the camera’s parameters only consider the vanishing point’s height, and assumes the optic center’s coordinate to match the center of the image (which is not always true). As mentioned early, this position would be corrected if the driver’s head was aligned to video camera mounted in the middle of car.

![Figure 5. Virtual lanes drawn in 3D space without image center correction](image)

Usually, it is false. The driver’s head is translated to side and few above or below. Few empirically adjustments at the camera position to \( (h + \epsilon, w, 0) \) were performed to correctly register images for the driver viewpoint and video camera. Figure 6 shows the virtual lanes drawn with correction.

![Figure 6. Virtual lanes drawn in 3D space with image center correction](image)

The \( w < 0 \), for steering wheel at left side of the car. The \( \epsilon \) depends on driver height, it represents the difference between driver’s eyes...
and the video camera. Then, virtual objects can be placed and overlaid to the real video sequence. For our purpose, and considering the camera is in the xy-plane, we have admitted keeping original values. But, for more precisely applications, $\alpha$ and $f$ values can be changed.

The effectiveness of this approach depends on supplying information without causing driver distraction and avoid increasing the number of control buttons. Conventional interfaces, such as desktop solution and central information display, do not fulfill these requirements. Then, our approach is to superpose virtual content in a head-up display projected on the windshield. The main restriction is that these new interfaces are concerned with real environment occlusion and driver focus of attention. This limitation are fixed through two simple artifacts, transparency and activation control. Transparency allows display virtual information and real one, without occlusion. The activation control is employed by drivers to turn on/off assistance system according with their needs.

4 Experimental Results

The main experiments were realized using videos, notebook and desktop computer with usual configuration: 2.2 / 2.4 GHz, 2GB / 4GB RAM, GeForce 9600GT / GeForce 9800GTX video adapter. All video sequences were captured with a resolution of 240 x 320 pixels at 15 fps. OpenGL\(^1\) has been employed to supply graphics rendering and OpenCV is responsible for video processing. The system setup consists of video input containing an image of a road, the road’s width and the depth distance between the points on the road that are chosen for camera calibration. Then, the system find lane borders, using computer vision algorithms (OpenCV\(^2\) library). Once these borders are found, five points are required in order to compute the camera parameters. Such points are two points on the left lane, two points on the right lane and one point in the lane intersection, the vanishing point. Points on the right and left lane must be at the same height, Figure 7. The Figure 8 shows only the road, without any virtual information. It can be enriched through virtual information, main textual information about global positioning, car speed, and the identification of another vehicle in opposite direction (red square), Figure 9.

To illustrate the enrichment of the scene, The Figure 10 shows a virtual car (3-d model) in the same direction of test vehicle, keeping the same speed. The 3-d object has been re-scaled and projected to the road, giving the impression that is travelling in the front of the real car.

\(1 \)http://www.opengl.org/

\(2 \)http://opencv.willowgarage.com/wiki/

![Figure 7. Roads enriched with virtual lanes, plus points on the borders used by camera calibration](image)

![Figure 8. Snapshot of roads used in the experiments, without any virtual content.](image)

![Figure 9. Snapshot of roads used in the experiments, with a hud displaying virtual information and the vehicle identification](image)
5 Conclusion and Future Work

This paper has presented a new approach to driver assistance system heavily based on an automatic camera calibration method and its synergy with Augmented Reality techniques. The camera calibration can obtain real environment information, supplying required parameters to adjust virtual camera and, correct and robust registering of virtual and real scenes, AR offers all techniques to superimpose virtual objects to real world. Besides, a head-up display decreases drivers distraction compared to central information display, assuring more driving security.

The camera calibration process has been shown robust and efficient in finding correct camera parameters. It requires a minimum amount of adjustment for the camera, quietly plug the camera and provide road measurements. Besides, the synergy between camera calibration and Augmented Reality has been appropriated, because all parameters required to setup virtual camera are provided by camera calibration.

Unfortunately, current approach presents few limitations. Input parameters of the system could be more controllable by users. For example, instead of finding the camera height based on the lane width, we are exploring the opposite way: to find the lane width based on the camera height. This is also of interest since the system could be employed in precarious roads.

This paper has explored the viability of improving driver assistance systems, using videos captured through camera into vehicles. For future investigations, it’s necessary to create a car cockpit to explore and evaluate the use of interaction metaphors with real drivers, such as alerts, enhance existing traffic signals and add 3-D objects. Besides, this cockpit can provide a more realistic environment to accomplish experiments with drivers. After, real test using real vehicles under usual traffic condition would be tested, including subjective evaluation about HUD adoption. The cockpit will provide a convenient environment for investigating and testing some projection technologies to superimpose virtual content to real scene on the windshield.

6. References


