Metadata of the chapter that will be visualized online

ChapterTitle	Camera-Based Automotive Systems	
Chapter Sub-Title		
Chapter CopyRight - Year	Springer Science+Business Media, LLC 2009 (This will be the copyright line in the final PDF)	
Book Name	Smart Cameras	

Corresponding Author	Family Name	**Broggi**
	Particle	
	Given Name	**Alberto**
	Suffix	
	Division	
	Organization	VisLab
	Address	Query, Italy
	Email	broggi@CE.UniPR.IT
Author	Family Name	**Bertozzi**
	Particle	
	Given Name	**Massimo**
	Suffix	
	Division	
	Organization	VisLab
	Address	Query, Italy
	Email	
Author	Family Name	**Bombini**
	Particle	
	Given Name	**Luca**
	Suffix	
	Division	
	Organization	VisLab
	Address	Query, Italy
	Email	
Author	Family Name	**Grisleri**
	Particle	
	Given Name	**Paolo**
	Suffix	
	Division	
	Organization	VisLab
	Address	Query, Italy
	Email	
Author	Family Name	**Porta**
	Particle	
	Given Name	**Pier Paolo**
	Suffix	
	Division	
	Organization	VisLab
	Address	Query, Italy

Email

Abstract

This chapter addresses the most important issues involved in the selection, installation, and calibration of a camera system onboard a vehicle, taking into consideration all the specific characteristics of the automotive environment and the requirements of the various applications.

Chapter 18
Camera-Based Automotive Systems

Massimo Bertozzi, Luca Bombini, Alberto Broggi, Paolo Grisleri, and Pier Paolo Porta

Abstract This chapter addresses the most important issues involved in the selection, installation, and calibration of a camera system onboard a vehicle, taking into consideration all the specific characteristics of the automotive environment and the requirements of the various applications.

18.1 Introduction

An extremely challenging application of artificial vision and smart cameras is visual perception onboard vehicles. The ability to perceive and understand the surrounding environment is of paramount importance for advanced driver assistance systems (ADAS), be they just warning systems to alert the human driver or autonomous systems directly controlling vehicle motion.

Environmental perception can be carried out, thanks to a great variety of different sensors and technologies (including cameras, laser, radar, sonar) but the processing of a picture can deliver an extremely rich quantity of information, much more articulated than with other sensors.

The integration of cameras on vehicles is indeed a topic that has been addressed for a long time; the first vehicle prototypes with cameras onboard were demonstrated in the late 1980s. At that time the main impeding factor was the limited processing power available for real-time image analysis. Other problems were indeed present, but computational resources kept researchers focused on the processing architecture rather than the sensor itself. In the last few years, on the other hand, computational constraints have been eased by the availability of sufficiently powerful and low-cost processing engines and a considerable effort has been put into the design of smart cameras that fit on vehicles.

A. Broggi (✉)
VisLab, Italy
e-mail: broggi@ce.unipr.it

A.N. Belbachir (ed.), *Smart Cameras*, DOI 10.1007/978-1-4419-0953-4_18,

An ever increasing number of vehicles have been equipped with cameras for environmental sensing within different research projects; on the industrial side, however, car manufacturers started with the integration of other technologies first, such as radars or sonars. The use of a camera on series vehicles is still currently limited to very basic applications, such as parking monitoring or night vision enhancement, in which no processing is performed: no automatic recognition is done, focusing only on image display.

As mentioned, on the contrary, research projects are actively pursuing the benefit of imaging, although the processing is definitely more complex and additional issues must be considered.

Functions such as detecting the driving lane or obstacles on the path, recognizing traffic signs, or localizing pedestrians need cameras installed on the frontal part of the vehicle looking forward, but each function has specific requirements on orientation, field of view, and sensitivity that must be carefully addressed and that are part of the camera selection and design process. Other applications such as parking assistance, blind stop monitoring, or junction management require cameras oriented differently, and – again – also specific considerations regarding the previously mentioned parameters.

Nonetheless, a primary constraint that is generally overlooked is the appearance and integration of the sensor in the vehicle. The sensor position must be carefully localized according to both functionality and style: Besides its obvious need to perform a given function, the integration of a new sensor must not be invasive, must not occlude the driver visibility, and must not alter the vehicle aesthetic aspect, yet providing enough evidence for the sensor to be perceived as a real value.

This chapter outlines all the technological issues, the setup constraints, and the calibration problems that must be considered when using applications involving the use of a camera onboard a moving vehicle. Additional problems, specific to the automotive environment such as vehicle dynamics, system temperature, environmental illumination, and camera vibrations, are also discussed and possible solutions highlighted.

18.2 Technology

This section gives an overview of the main technologies used in automotive applications to capture road images and feed the processing stage with relevant information. Camera selection – the first important degree of freedom when designing a vision system – usually requires to slightly change the parameters obtained from theoretical considerations to match the available models on the market. It is hard to evaluate theoretically how the performance of the final system is impacted by this choice. Mechanical constraints and price are other important constraints that have to be carefully considered when choosing a camera to develop an application for the mass market.

Common advanced driving assistant systems with typical camera features are summarized in Table 18.1.

Table 18.1 Common automotive applications with typical camera features

Application	Pedestrian detection	Lane detection	Obstacle detection	Traffic-signs recognition	Parking	Blind spot	Automatic headlight control
Sensor/camera (global shutter)	Thermal	Visible/NIR	Visible/NIR	Visible/NIR	Visible/NIR	Visible/NIR	NIR
Pixel resolution	320×240	VGA or greater	VGA or greater	VGA or greater	VGA or greater	VGA or greater	VGA or greater
Dynamic range	Thermal	Highest	Highest	Highest	Highest	Highest	Highest
Temporal resolution	At least 10 Hz	At least 10 Hz	At least 10 Hz	At least 10 Hz	At least 10 Hz	At least 10 Hz	At least 10 Hz
Technology	Micro-bolometer	CCD	CCD/CMOS	CCD	CMOS	CCD/CMOS	CMOS
Advantages	Day/night, no texture on targets	High S/N ratio	Low cost	Low cost	Low cost	Low cost	No smear
Limitations	Expensive, hot backgrounds	Night: smear	Smear or noise	Dynamic range	Ego-motion detection	Dynamic range	Dynamic range
Heads	Mono or stereo	Mono or stereo	Mono or stereo	Mono	Mono	Mono	Mono or stereo

18.2.1 NIR and FIR Sensors

Automotive cameras suitable for ADAS are designed to make processing system recognition tasks easier and more robust. This includes the use of special sensors able to capture radiations with wavelengths outside the perception range of the human eye. There are two main interesting ranges of frequencies used in automotive applications: near infrared (NIR) and far infrared (FIR).

NIR is the acronym used to indicate the electromagnetic waves with wavelength between 700 and 3000 nm. This radiation cannot be perceived by the human eye but is full of important information, especially during the night. Moreover, when the scene is illuminated with high beams emitting in the NIR range, humans in other cars do not perceive the dazzling[1] by the lighting system while vision systems can perform its detection on a fully illuminated scene. The NIR light is valuable, thanks to the different reflectivity of objects in this range. Scenes looking poorly illuminated in the visible domain have a good amount of information still available in the NIR domain. Thus, capturing and performing detections on these kinds of images usually lead to better results than using visible images.

Thanks to the physical silicon characteristics, most of the commercial/industrial devices already capture the NIR radiation and a filtering glass is inserted between the lens and the sensor to filter out the intensity content in this range in order to obtain images with more realistic colors.

Special fabrication processes may improve the sensitivity in the NIR domain. In these cases the visible contribution can be filtered out obtaining pure NIR images. Due to the visible cutoff filter, the amount of light reaching the sensor is strongly reduced, thus longer shutters and/or higher gains must be used to obtain a correct exposure. Images taken in the NIR domain are suitable for lane detection and pedestrian detection. Not all materials have a good NIR reflectivity. Some PVC (polyvinyl chloride) clothes adsorb this kind of light which makes the recognition task more difficult. This is a strong limiting factor to the use of this kind of images for pedestrian detection.

Far infrared is conventionally intended to be the spectrum of the electromagnetic waves with wavelength between 3 and 25–40 μm. Measuring the intensity of this type of radiation emitted by a body allows to measure its temperature. Range from 3 to 5 μm is sometimes also referred to as MWIR (medium wavelength infrared) while from 5 to 14 μm is named LWIR (long wavelength infrared).

Sensors able to produce a thermal image of the framed scene are used in aftermarket products for night vision-driving assistants [327]. In a thermal image each pixel value is related to the temperature of the surface covered by the pixel projection through the lenses via a typical transfer function.

One of the interesting points about FIR in the automotive industry is related to the high air transparency of these wavelengths. Figure 18.1 describes how the air transmittance changes with the radiation wavelength. This function looks like a

[1] This solution can be unsafe for the human eye in proximity (0–10 m) of the car.

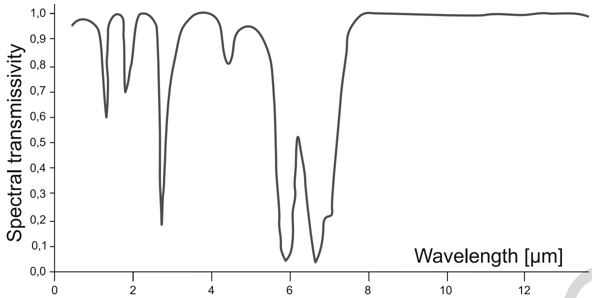

Fig. 18.1 The transmittance of the air has a window in the far-infrared range

sequence of high-attenuation and high-transmittance ranges. These ranges are called *atmospheric windows*. The transmittance in the LWIR range is high thus the radiation is not attenuated over a long distance. On the other hand, in the MWIR range a sensible attenuation occurs and makes air opaque for this radiation at distances of some tens of meters.

FIR images are used in automotive systems to detect pedestrians and animals, thanks to the different temperature of their body with respect to the background (see Fig. 18.2(a). However, limitations of this approach can be due to heavy clothes worn in winter – masking the human emissivity – and hot backgrounds during summer – where the background may be hotter than the subjects – as shown in Fig. 18.2(b).

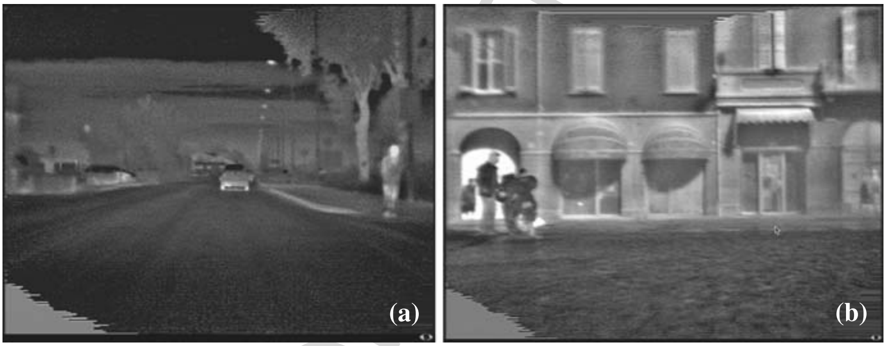

Fig. 18.2 FIR images allow an easy detection for pedestrians and vehicles in winter (**a**). In the summer scene (**b**) the pedestrian at left in the *right* image is cooler than the background gate

Although these systems are really effective for human detection at medium-low temperatures only, they are not widespread due to the sensor cost.

18.2.2 Color Sensors

The color information is important for the detection of structured elements such as vehicles, lanes, road markings, pedestrian crossings, traffic lights, and traffic signs.

A cheap way to obtain color using a single imager is to install on the top of each pixel a color filter. This way each pixel is exposed to a specific wavelength range depending on the color filter.

The most common solution is to use a Bayer pattern. Referring to a 2×2 pixel cell, four possible combinations of R, G, and B wavelengths are commonly used. Usually two pixels are dedicated to the green radiation since the human eye is more sensitive to this color; the other two are used to filter the red and blue components.

The color image can be reconstructed using different algorithms. These can be fast or accurate, depending on the needed accuracy and on the available computational power. The best reconstruction can be obtained by sub-sampling: each RGB pixel of the final color image is obtained directly from a 2×2 pixel in the raw image. Other techniques may show artifacts due to the spread of the color info over several adjacent pixels. A Bayer image uses one-third of a RGB image size; this is beneficial to reduce bandwidth during transmission. Software libraries can be deployed to perform transformations and detection directly on the Bayer image.

18.2.3 Global and Rolling Shutter

A sensor with a global shutter exposes all pixels at the same time. Each pixel integrates an amount of light coming from different portions of the framed scene at the same time. If the shutter is too slow, when the sensor is mounted on a fast moving vehicle, motion blur can occur. While in some situations like lane detection this can lead to a more uniform and easier image to work with, in other situations like obstacle detection it can lead to blurred and unusable images. A global shutter can be expensive since it requires high-speed electronics. A global shutter is typical for CCD technology.

On the other hand, a simpler alternative for CMOS sensors is a rolling shutter. Different areas of the image (often each line) are electronically exposed at different times. This can lead to image artifacts like skew for still images or wobbling for image sequences. These artifacts occur when the camera is moving or the framed scene contains moving objects. The results are as annoying for camcorder and consumer camera users, as they are unacceptable for image analysis, especially in automotive systems, since the constraint that the source image is taken at the same time is violated.

18.2.4 Multi-imager Sensors

Vision systems may use one or more sensors. A system with two imagers connected or integrated onto the same board and specifically oriented to frame the same scene but from two different angles is called stereoscopic. Such a system has an additional feature to be considered in the design phase: the baseline, i.e., the distance between the two imagers. Depending on the baseline width, the stereo head will be tailored to detect objects at different distances: with a large baseline far objects are better detected while with a short baseline close 3D reconstruction is possible. If the sensors are mounted on the same, rigid mechanical bar a factory calibration can be performed to adjust the relative orientation. If the sensors are separated (such as in case of two separate cameras) inter-calibration problems may occur if one of the sensors loses its orientation.

Simple detection or recognition systems use one (monocular) single camera to capture the image stream that will be analyzed. This kind of system relies on the knowledge of camera orientation with respect to the world reference system to provide an accurate distance estimation of the detected objects. If the camera is mounted on a moving vehicle, its orientation changes continuously, depending on many variables like bumpers response, vehicle speed, or road roughness. An accurate stabilization system, which may be obtained by analyzing the acquired images or using other sensors such as inertial measurement units (IMU), can help make the system more robust to many of these changes. However, this is true only on flat grounds. The estimation of road slope is possible for monocular systems only if additional assumptions are made, such as constant road width or specific lane markings structure. Stereo systems, on the other hand, can provide distance estimation even on non-flat terrains. Other features like road slope and instantaneous pitch can be detected from the processing of a stereo pair. The drawback of this technique is the high cost involved in duplicating the sensor, keeping it calibrated, and providing additional computing power. Systems with more than two heads have also been designed for very specific applications [82].

18.2.5 High Dynamic Range

The brightness of an automotive scene can range from 0.1 to 50000 $\mathrm{cd/m^2}$. With current sensors it is not possible to capture this dynamic range[2] in a single shot. An alternative, and cheaper, way to work with high dynamic range (HDR) images is to take different shots, two for instance, with different shutter values, one underexposed and one overexposed, and then use an algorithm to merge them together in a unique image with an extended dynamic range.

The shots can be taken using different sensors or the same sensor. When using different sensors, the images must be taken at the same time, so that moving objects are in the same position in both images. Unfortunately, this technique is expensive since it requires two sensors with exactly the same orientation. Moreover since the two sensors are not in the same position but skewed at least along one axis, some artifacts will appear in the final image. For these reasons the other technique is preferred, i.e., taking the images from the same sensor but at different times; this is commonly done to take HDR pictures of still scenes. In the case of automotive

[2] Dynamic range is intended to be the ratio between the greatest amount of light to produce the brightest pixel and the least amount of light to produce the darkest pixel within the same frame.

applications the movement of the camera is an even worse source of artifacts, since two consecutive images will indeed frame the world from two different points of view regardless of moving or still objects. In the next few years the market will see the introduction of new CMOS sensors where the pixel response can be different in different areas of the image, which is useful especially in challenging illumination conditions such as entering or exiting tunnels and with strong shadows.

18.2.6 Frame Rate and Processing Rate

Automotive systems need to respond in the shortest amount of time possible to the situations they are designed to recognize giving the driver or to the actuation system the longest time possible to actuate the proper countermeasures. This requires to sample the world at a sufficiently high frequency. Obviously there is not a unique rule to compute this frequency, and the solution is a trade-off between different constraints, one of the most important being the cost. Vehicle speed and motion type – such as rectilinear, curve, or abrupt maneuver – can strongly influence the requirement of more samples per second. On the other hand, the processing system should have enough computational power to run the algorithm fast enough to utilize the amount of data produced by the sensors. This, especially for images and depending on the resolution, creates the need for a new generation of processing units able to supply the proper processing rate. A specific generation of embedded processing chips is needed to match the difficult constraints of the ADAS market: high computational power, minimum need of external components, low power consumption, large temperature range, large shock resistance, and finally small size to be directly integrated into the smart camera.

18.2.7 Optics

Besides the sensor, optics are primarily responsible for image quality. Depending on the focal length, the optic will introduce some geometric distortion in the image. This distortion can be removed using a lookup table obtained from accurate theoretical models or from experimental measurements. The experimental solution is preferable since it delivers more accurate results for the whole system.

The amount of light captured by the lens is important in night applications. Also NIR-only systems should use bright lenses to compensate for the visible filter attenuation. When choosing lenses, size, cost, and the amount of light (f number) should be carefully considered.

High-definition sensors with a large size (1/2" or more) require finding an appropriate lens which is able to cover the entire sensor area, as otherwise vignetting[3] phenomena might occur.

AQ1

[3] Vignetting is a noticeable gradient within the image brightness. The image has a high brightness near its center which decreases in the peripheral areas.

Depending on the glass quality, lenses have an optical resolution. Resolution is the capability of an optical system to distinguish between two adjacent points. The lens quality must be selected appropriately according to the sensor. High-definition lenses help to resolve details which are needed for the detection of far features (such as traffic signs or pedestrians).

Optics usually have a few regulation gears for focus or iris aperture. In automotive environments, locking screws for these gears are useful to contrast unwanted movements due to vibrations. Any change in focus or iris causes a variation of the aperture angle and likely a movement of the optical center over the sensor. To avoid calibration problems, in extremely harsh environments or for series production, fixed focus and fixed iris are preferable. Adjustable optics are useful in prototyping phases.

18.3 Setup

This section gives an overview of the main problems and constraints in the system setup for automotive applications.

Typical problems of artificial vision systems, such as background noise, camera movements, illumination conditions, and characteristics of the target to detect, are amplified in the automotive field. Unlike industrial applications or video surveillance systems, the scenario changes continuously, the camera is on the move, the vehicle is subject to vibrations and oscillations due to road coarseness, some targets like pedestrians can only be defined in a statistical and non-exhaustive way. Moreover the area for system wiring and sensors positioning is very limited and usually a camera-based system must be connected to other vehicle sensors or devices.

For these reasons, the setup design is one of the most complex challenges in the implementation of a complete system. System designers have to consider the constraints discussed in the following section.

18.3.1 Functionality

A wide range of ADAS are currently available on the market and others will come out shortly [410]: Adaptive Cruise Control, All-Round-View, Collision Warning and Auto Break, Pre-crash Safety, Lane Departure Warning, Lane Keeping Assistant, Stop-and-Go Assistant, Blind Spot Detection, Lane Change Assistant, and Night Vision. AQ2

The hardware setup strongly depends on the specific functionality. Some of these systems like Lane Departure Warning or Blind Spot Detection require a simple hardware setup: one smart camera connected to an integrated display on the vehicle. Other systems like Stop-and-Go or Collision Warning and Auto Break require a more complex setup: a stereo system or a sensor fusion with other devices.

ADAS providing complex pre-crash features, such as pedestrian detectors, require a more complex design since they need to process data from several sensors which might already be used for other purposes – such as a single wheel speed detector for ESP (electronic stability program) – to perform their task.

For multi-sensor systems synchronization must be ensured to avoid artificial data re-alignment inside the ECU (electronic control unit). Synchronization must be supported by sensors and is usually distributed as a square wave triggering the sampling instant. If sensors only supply a strobe signal, a robust time stamping inside the ECU is required in order to allow real-time data alignment. Data alignment can be problematic from some sensors like forward-looking cameras.

18.3.2 Technical Feasibility of Device Positioning

In the prototyping phase, the sensors installation must follow a feasibility analysis. During this phase, constraints like installation cost and system performance must be considered together with aesthetics or ergonomics. The perception system components can be placed all around the vehicle depending on the application without limiting the visibility for the driver and can be placed both inside or outside the cabin.

These choices are driven by both the target application and technological issues. Inside the cabin the camera is protected from rain, snow, and dust but has to follow some aesthetic and ergonomic constraints. Moreover, modern heat-treated windshields filter the near-infrared wavelength causing loss of information if the system use an infrared camera sensor. This problem can be solved in different ways such as replacing the windscreen or changing the camera position outside the cabin.

Far-infrared cameras cannot be placed inside the cabin since glass is opaque to these wavelengths. Figure 18.3(a) shows an example of FIR camera integration. However, an outdoor installation has to cope with environment-related problems such as cleaning the device, waterproof resistance, and in some cases shock resistance. Devices mounted in peripheral positions – such as behind the bumper – need a protection system from shocks. Figure 18.3(b) shows a possible solution for the camera setup in a Start-Inhibit system on a truck [84].

Fig. 18.3 Example of integrating of a FIR vision system: an infrared camera is mounted in a central position in the front of the vehicle (**a**). Integration of a stereo vision system on a truck (**b**)

18.3.3 Wiring and Positioning

Device positioning and wiring have to be carefully considered if needed. As discussed in previous sections, ADAS systems require extremely small cameras that are suitable for integration with low impact inside the car and reasonable processing power to perform recognition task. For these reasons, industrial smart cameras including sensor and processing unit in one enclosure are still not sufficient for ADAS applications. A good solution is to keep the vision sensor and a compact processing unit separated and connected by some robust interface, such as Ethernet, firewire, or USB cables. Is this way the embedded processing unit of appropriate power can be placed more freely on the vehicle where there is space available. Some systems may have the ECU placed in proximity of the sensor and produce the results – such as driver warnings – directly there. However, if sensors are placed outside the cabin, connection cables between the sensor and the ECU must be placed taking into account problems such as temperature range, electromagnetic interferences generated by the engine, and thermal noise, all of which cause signal degradation. These problems are critical if the signal has a high frequency, such as for high-resolution or high frame rate cameras. Differential buses such as CAN (controller area network), firewire, or LVDS (low-voltage differential signal) provide the necessary robustness for communication.

18.3.4 Lighting Control

During the day scene, illumination is determined by weather conditions. When the camera is placed inside the cabin, internal illumination can cause reflections on the glass (see Fig. 18.4(a)); to avoid this effect a small black chamber can be installed around the camera (like the one in Fig. 18.4(b)).

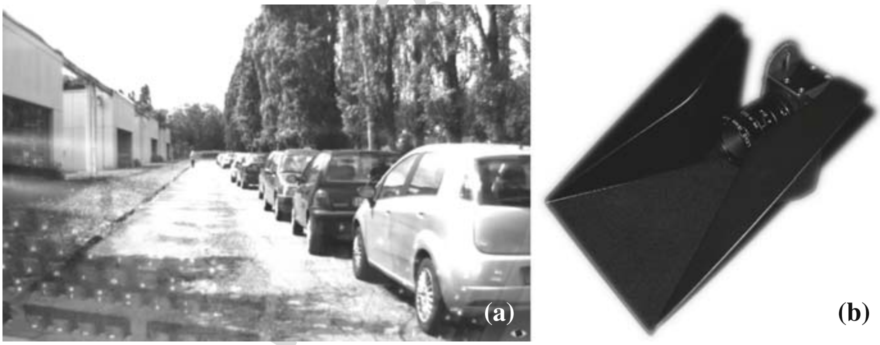

Fig. 18.4 Color image acquired by an onboard camera with reflections in the bottom (**a**). A possible solution to protect the camera sensor from reflections (**b**)

At night, on the other hand, illumination is poor even with a NIR camera and the system requires a proper illuminating hardware (with respect to the camera

Fig. 18.5 Two different types of near-infrared lamps installed on an experimental vehicle (**a**). Night vision lighting system composed by low/high beam light to the *left*, NIR lamp in the *middle*, and parking light to the *right* (**b**)

sensitivity spectrum). In Fig. 18.5(a), the setup with two different NIR lamps is shown. In Fig. 18.5(b) the NIR illuminator has been integrated within the head lamp assembly.

18.4 Calibration

Camera calibration is one of the main issues in machine vision applications. In general calibration provides a correspondence between the results of an algorithm and the real world. It is a hard link between the software and its application.

The calibration process consists in finding the extrinsic and intrinsic parameters of a camera; the former being the camera position and orientation, the latter being the internal parameters, i.e., focal length, optical center, etc.

In the literature many algorithms for camera calibration have been proposed for monocular systems [544, 566, 600], stereo systems [606, 311], etc., but many of them are linked to some particular hypotheses that ease the calibration step; usually these hypotheses are not verified in automotive environments, for example, short distance perception, still scenarios, or still camera.

Calibration assumes a particular importance in stereo systems, or in general in systems that involve more sensors; in fact, when a monocular system is affected by a miscalibration, its results are projected in wrong real-world coordinates; therefore, the last step of the processing is degraded. On the other hand, when a multi-sensor system (such as a stereo system for example) is affected by miscalibration, the correctness of the whole processing is compromised since it gets deeply degraded by wrong information matching between the sensors.

The most common way to perform calibration in an automotive system is to use a large calibration grid as, for example, the one shown in Fig. 18.6. Through a calibration tool it is possible to pinpoint all the known 3D world points in the image thus performing an association; then the calibration parameters can be extracted. Figure 18.7 shows one of these tools: the acquired image is shown on the left-hand side while its corresponding view of the world coordinates is shown on the right.

SPB-176214 ChapterID 18 August 6, 2009 Time: 01:43pm t1-v1.4 Proof1

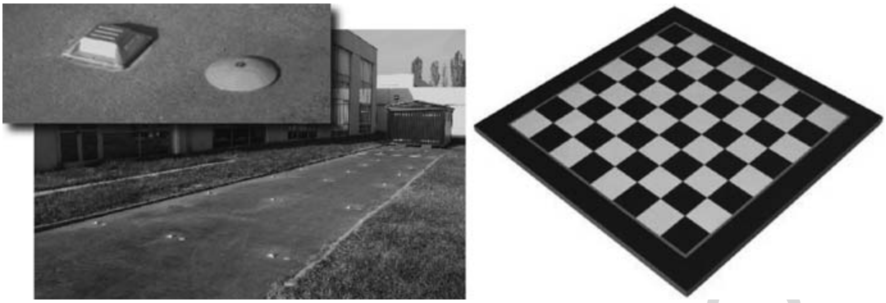

Fig. 18.6 On the *left*, example of an automotive calibration grid available at VisLab (http://www.vislab.it). Each grid point is indicated by two nails: a *yellow* plastic one, used for visible cameras, and a metal, light reflecting one, to calibrate FIR cameras and visible cameras at night. On the *right*, chessboard used for calibration for indoor applications

Fig. 18.7 A handy camera calibration tool developed by VisLab

18.4.1 Mechanical Issues

During the prototyping phase cameras have to be placed with many degrees of control. It is important to choose a good trade-off between comfort during camera adjustment and robustness of the system itself. To perform camera orientation in the best way, the choice of an appropriate camera mount is fundamental. In many applications it is useful to have all three axes available. On the other hand, the more degrees of freedom are available, the more the camera can be moved by vibrations or other mechanical causes. However in the final configuration, cameras will be fixed into a rigid cameramount to improve robustness and stillness. The first step with adjustable cameramounts is useful because for some applications a particular orientation relates the system to under specific hypotheses that can be used to simplify the algorithm and speed up the computation process [83].

Once calibration is achieved, it must be kept throughout the whole product lifetime. Otherwise, either an automatic or manual procedure mustbe made to compensate for drifts. In case of an automatic procedure, the system periodically checks the system calibration and runs a recalibration algorithm if needed. In case of a manual procedure, the system, once a miscalibration has been detected, can either offer a simple recalibration procedure to the end customer (e.g., drive on a straight road) or suggest to approach an authorized garage for a recalibration by means, for example, of a grid.

There are some mechanical issues affecting calibration that have to be considered: vibrations, for example, are very critical in automotive applications, especially with trucks. The choice of the optics is critical as well, since the optical center is different from the image center. The optical center is one of the intrinsic parameters that has to be extracted with the calibration procedure. In particular, if a focus adjustment is needed after recalibration, this will result in a change of the intrinsic parameters: focal length and also optical center position can change. The former is a direct consequence of focus adjustment while the latter is a result of imprecisely mounted optics on the sensor. In fact, if the lens axis is not perfectly perpendicular to the sensor plane, then the rotation movement can cause a movement of the optic center.

18.5 Specific Automotive Issues

Camera-based automotive systems have to face issues specific to the automotive domain. The main issue is due to the fact that cameras are installed on moving vehicles and therefore the vision system and its related processing steps must be robust with respect to vehicle movements. In most cases, the vehicle's ego-motion must be taken into account. Besides ego-motion, also other kinds of movements like vibrations or oscillations represent a source of noise for vision-based systems.

Other issues are related to specific environmental conditions in outdoor environments. In fact, temperature and illumination conditions can vary and can be barely controlled. Especially for illumination, extreme situations like direct sunlight or strong reflections must be taken into account. In addition, other light sources, such as car headlights or reflectors, might be present in a typical automotive scene.

Specific camera issues related to the automotive environment are summarized in Table 18.2.

Table 18.2 Common automotive applications with typical camera features

Issue	Ego-motion	Oscillations and vibrations	Illumination conditions
Properties	Moving background, perspective changes	Noise overlapped to ego-motion	Object texture changes, bad reflections
Impact	Motion blur, object changes	Tracking problems	Camera dazzling, bad recognition
Workaround	Faster shutters, better processing	Better ego-motion detection	Better processing, higher dynamic range

18.5.1 Vehicle Ego-Motion

When the vision system is installed onboard a vehicle, it has to be robust with respect to vehicle movements. This design issue can be examined at two different levels: vision devices (i.e., cameras configuration) and processing (algorithms).

Concerning the cameras, some technologies are not robust to motion artifacts and moving objects are blurred in acquired images. This effect is particularly evident when the vehicle makes a sharp turn and the whole background begins to move. Figure 18.8 shows the effect for a FIR camera.

Fig. 18.8 An example of motion blur in FIR images acquired by an onboard vision system: the *left* image was captured when the vehicle was still, while the *right* photograph was taken only a few seconds later – when the vehicle turned *left*– and shows a heavy horizontal motion blur effect

While blurring can even help in some scenarios and for specific vehicle movements by hiding unnecessary details, generally it has to be avoided. Therefore, a careful camera selection is a mandatory step in designing the setup. Specifically, old CMOS-based cameras more likely feature a slow sensor and, thus, can be affected by this problem. Conversely, the effect is not appreciable for CCD-based cameras and, generally, in recent CMOS models.

Vehicle movements, namely ego-motion, must be considered as input for many image processing algorithms. The computation of ego-motion for a vision system can be performed using machine vision techniques like the analysis of background movements or visual odometry [147]; however, these techniques require additional computing and are not always applicable, in such cases added (and often expensive) sensors like gyroscopes, odometers, or inertial devices are generally used.

18.5.2 Oscillations and Vibrations

Oscillation and vibrations have been already discussed from a mechanical point of view for calibration; in this section, specific issues for processing in automotive applications are covered.

Besides tracking, other vision-based applications are affected by vehicle movements as well. In fact, many systems rely on calibration to recover 3D information or to detect objects. Unfortunately, vibrations or oscillations, induced by normal vehicle functioning, affect calibration and may lead to incorrect results.

Therefore, image stabilization techniques are widely used to cope with this problem. In some cases, this can be done during the acquisition step, since some cameras feature image stabilization at sensor level. Another hardware-based solution is the use of electromechanical stabilization platforms [475] or lenses-based mechanisms [93]. These approaches are generally effective for suppressing really abrupt movements but are less suited for removing the specific range of movements due to oscillations or vibrations typical of the automotive field [68].

In most situations, a specific processing phase devoted to removing this source of noise has to be developed. This is a difficult task, since only unwanted vision system motions have to be removed, while the motion components due to vehicle ego-motion have to be preserved.

Vibrations and oscillations are considered the *high-frequency* component of global motion and therefore image stabilization can be applied in an attempt to smooth inter-frame motions. In specific situations, this task can be simplified in order to remove critical noise components only; in fact, the definition of *unwanted motions* can depend on the specific application; as an example, pitch variations can highly affect distance estimation for monocular systems which often relies on the vertical features positioning in the acquired images to estimate distance; in such a specific case only pitch deviations should be removed to avoid wrong distance estimation [68]. Conversely, in a stereo system, distance can be computed exploiting 3D triangulations but, at the same time, a large number of stereo vision-based systems are based on the assumption of a null roll. In such cases, pitch oscillations barely affect the processing while roll variations have to be compensated.

An image stabilization process is generally divided into two different steps: inter-frame motion detection and motion compensation.

In the first step, most systems exploit feature detection and tracking to recover motion. Again, the nature of the features to extract highly depends on the stabilization requirements: for simple stabilization techniques or when real-time constraints apply, simple features are generally extracted like edges [68]. More complex features, like lane markings, are used when a more precise stabilization process is required [320].

A different approach for motion detection is based on the use of dense matching techniques like image disparity or optical flow computation.

The motion compensation stage is used to compute the roto-translation which is applied to consecutive frames to minimize noise introduced by vibrations and oscillations. In simple cases, it is based on a low-pass filter to remove high-frequency components of movements, but also more complex approaches that exploit supplementary information on the scenario, like object or background position, are widely used.

18.5.3 Illumination Conditions

In the automotive environment, illumination can be barely controlled and therefore represents a major issue.

In fact, weather conditions, the different sun positions, and artificial light sources such as headlights or street lamps highly affect scene illumination. This is true for daylight or near-infrared cameras, while in the case of far-infrared cameras the problem arises only in extreme situations like direct sun framing or when light sources also produce thermal effects. Shadows represent a critical issue for image acquisition and processing; in fact, the simultaneous presence of shady and fully illuminated areas in the scene may lead to acquiring images in which shady areas are too dark or illuminated objects are too bright. Moreover, shadows represent a pattern that can interfere with the image processing systems based on pattern matching techniques. One case, in which shadow's presence indirectly impacts on FIR domain as well, is due to thermal effect that light can have. In fact, sun or even artificial lights increase the temperature of objects exposed to lights creating *thermal shadows*; Figure 18.2(b) showed this effect on the wall below the tents, which is colder than the other portions of the wall that are heated by the sun.

In addition, vehicle movements can lead to abrupt changes in illumination conditions. The worst situation happens when the sun is suddenly framed or exiting/entering a tunnel, making the entire image completely dark or white.

In such cases, cameras that have a fast automatic exposure control (AEC) are recommended. AEC acts on both camera gain and control to compensate global illumination changes. Since a large gain value introduces noise at camera sensor level, it would be better to have a system that primarily acts on the shutter trying to maintain a low gain value; in addition, such a system can avoid to monitor the whole image reducing the area used for exposure control to the one actually processed. Figure 18.9 shows the result of an evolute exposure control algorithm that has been conceived to compute the most suitable exposure for the lower portion of the image, since the area of interest is the road and not the sky. In this case, the pedestrian can be recognized, while the computation of the exposure using also the upper portion of the image would have left the road completely dark. This requires a camera that features inputs for controlling gain and shutter like most IEEE1394 or IP-based cameras or a smart camera with some processing inside.

18.5.3.1 Smear Effect

The *smear effect* is another artifact degrading image quality for cameras in the visible domain: a strong light that directly hits the sensor in low illumination conditions produces bright artifacts, especially vertical bright lines (see Fig. 18.10(a)). This effect is typical for visible cameras and, in the automotive environment, can be easily caused by reflectors or other vehicles' headlights at night or inside tunnels. This effect represents a source of noise for image processing leading to wrong results, i.e., a lane markings detection system, that is typically based on the detection of

Fig. 18.9 Example of automatic exposure control obtained by defining a specific area (matched with the application) in which the contrast and brightness should assume optimal values

Fig. 18.10 Smear effect in (**a**) visible cameras and (**b**) NIR devices

bright lines on the road surface, can be fooled to interpret smear effects as lane markings.

Smear effect is caused by internal reflections inside the camera and lens system and is lower at the infrared wavelength. Therefore, near infrared cameras are less affected by this problem (see Fig. 18.10(b)) and can be evaluated as a replacement for standard daylight cameras in many situations.

18.5.3.2 Reflections and Glares

Reflection represents another source of problems for onboard systems.

The worst case is due to strong light reflections that dazzle the camera and lead to saturated images, but also weak reflections can create artifacts in acquired images. As an example, Fig. 18.11 shows how a wet asphalt road behaves as a mirror in

Fig. 18.11 Reflection of a wet surface for far-infrared radiations

the FIR domain and produces ghost pedestrians in the acquired image. In order to reduce reflections, a polarized lens can be used for cameras.

18.6 Concluding Remarks

Indeed the use of cameras onboard vehicles opens a great deal of opportunities to provide the vehicle itself with full awareness about the surrounding environment. Cameras have the capability to record many details – even small ones – of the environment and are based on the same technology used by humans when driving. Unfortunately, although cameras may achieve a resolution and sharpness higher than the human eyes, there are still issues to be solved: mimicking the driver requires not only the ability to process iconic data like images at a high frame rate but also the capability to select the important parts of the scene, which – in turn – may require moving/rotating the head as well. In the case of the electronic driver, this might lead to either move the camera or fuse the information coming from different cameras around the vehicle, pointing at different directions.

The issues discussed in this chapter also clearly show that developing a smart camera for the ADAS market is a difficult endeavor. In fact, it is mandatory to match specific constraints and, in particular size, processing power, and reliability that the current technology can provide have to be still improved for an effective deployment for smart cameras in the automotive scenario.

Cameras have the great advantage over other sensors, such as laserscanners, that nothing is moving inside the box, and adding mechanical parts to move the box would translate into degrading their applicability for harsh environments. Besides requiring additional hardware, power, and connections, changes in camera orientation may result in losing the sensor calibration in case the mechanics are not precise enough; therefore, the cost of a robust gazing system might be comparable – if not higher – than a solution based on multiple cameras.

The last solution, based on the integration of multiple cameras, is indeed preferred in off-road applications where vehicles are subject to heavy and strong vibrations, therefore, requiring a much stronger and ruggedized – and hence expensive – gazing mechanics.

Another great advantage of cameras is that they are based on a passive sensing technology and are therefore specifically preferred for military purposes.

Much like the other sensors, cameras require a calibration before being able to deliver information that can be registered with the environmental reference system. Other sensors like laserscanners require proper positioning and orientation since all their measurements are used for environmental reconstruction (due to the limited number of samples of the 3D scene). On the other hand, thanks to the high resolution of imaging devices, even if the camera is not perfectly aligned, data can be manipulated (the image is rototranslated), compensating for small orientation errors. In other words, a laserscanner oriented toward the ground will always yield unusable data, while a camera might still be able to provide images that – after a specific preprocessing – can contain meaningful data.

This preprocessing step, aimed at compensating for small errors in camera orientation, is performed on each image acquired by the camera, with parameters defined after installation. However, some vision systems are able to recalibrate themselves on the fly, recomputing the preprocessing parameters in order to compensate also for drifts in the orientation due – for example – to strong vibrations or accidental camera movements.

Finally, camera installation has another great advantage over other sensors: radars or laserscanners need to be positioned in front of the vehicle, typically in or near the bumper; this forefront position allows to acquire data without any occlusion caused by vehicle parts. However, unintentional small bumps against obstacles or other vehicles during parking maneuvers or dirt and rocks thrown by the preceding vehicle while driving at high speeds may damage the sensor. Cameras, on the other hand, are generally installed inside the cabin, behind the windshield, thus, besides being automatically protected against bumps, rocks, or dirt, they are also kept at an ideal operative temperature. Moreover, in some installations the wiper keeps the glass in front of the cameras clean when it rains.

Unfortunately cameras suffer from the main problem that affects human drivers and which is usually one of the causes of accidents: bad visibility. During fog or heavy rain – and also in particularly bad illumination conditions such as the sun low on the horizon and in front of the sensor – the cameras cannot deliver meaningful data. Some wavelengths, such as far infrared, are able to penetrate fog and light rain, but generic daylight or near-infrared sensors cannot.

As concluding remark it is important to note that due to their potentialities, low cost, and high range of applications cameras offer very promising possibilities; nevertheless cameras alone may not be able to disambiguate and correctly perceive every situation. For example, a textureless wall in front of the vehicle is barely perceivable, just like a gray obstacle on a gray background. To be sure to successfully handle every single situation, data fusion with other sensors based on different perceiving technologies is mandatory.

Chapter 18

Query No.	Page No.	Query
AQ1	326	Please confirm the edit made to the sentence "The image has a . . . areas"
AQ2	327	Please suggest whether "Adaptive cruise control . . .. Night Vision" in the sentence "A wide range . . ." can be set in lower case.