44 Vision-Based Blind Spot Monitoring

Elena Cardarelli
Dip. Ing. Informazione, Università di Parma, Parma, Italy

1 Image Acquisition .......................................................... 1072

2 Vehicle Detection .......................................................... 1074
2.1 Pattern Analysis ............................................................. 1074
2.2 Optical Flow Estimation .................................................... 1076
2.3 Features Extraction ......................................................... 1077
2.4 Features Tracking ......................................................... 1079

3 Algorithm Overview ....................................................... 1083

4 Conclusions ............................................................... 1087
Abstract: These sections introduce a vision-based system designed for monitoring the area that a driver cannot see from exterior mirrors, usually referred to as blind spot. This is a challenging task that requires to discriminate from vehicles and background when both are not static and also to cope with usual automotive problems like camera vibrations and oscillations.

The development of ADAS has involved an improvement of the safety on the road, ensuring vehicle reliability and supporting driver for accidents preventing, as seen in the previous chapters. Particularly, some studies have been focused on the assistance during lane change, monitoring the area covered by the blind spot, which drivers are unable to see in the exterior mirrors (Fig. 44.1). If the system detects an overtaking vehicle in the dangerous zone, visual and acoustic signals warn the driver about the risk of collision.

Radar technology and laser scanner are sometimes employed for the lane changing assistance, although they do not guarantee always a complete representation of the scene: When the vehicle is pitching down or up, the laser beam hits the ground or points to the sky, making data acquisition worthless. Moreover, in complex scenario, like urban road traffic, the radar has limited functionalities because its reliability depends on the radar cross section of the object to identify. For this reason, vision data fusion is necessary to provide high-resolution performance.

A common choice for the blind spot monitoring is the using of a camera mounted on the wing mirror (Fig. 44.2) in order to extend the driver’s view.

1 Image Acquisition

The use of a camera mounted under the side-view mirror increases the system complexity: The operability scenario is not static and different elements such as camera angle, perspective deformation, and camera vibration have been considered because they may affect the system performance. Moreover, when the camera roll angle is not null, the acquired images are rotated (Fig. 44.3).
To cope with this problem, images have to be transformed: With a rotation, this effect may be removed, but the relation between camera parameters and images becomes incongruous. For this reason, image rectification is performed: In this case, camera parameters are acquired and modified in order to obtain a null roll angle; then through the homographic transformation, the new parameters are assigned to the camera. The results of images rectification are shown in Fig. 44.4.
In order to remove the region occupied by the vehicle where the camera is mounted on, a gray level mask is applied to the image; an example is shown in Fig. 44.5: Only the images pixels corresponding to the black zone are considered useful information.

2 Vehicle Detection

The first step for the vision-based lane changing assistance is the detection of vehicles in the critical area; to perform this, task pattern analysis (VOLVO Technologies 2007) or optical flow estimation (Batavia et al. 1997) could be employed.

2.1 Pattern Analysis

In the blind spot application, pattern analysis is based on the detection of vehicle discriminative parts: Signature, lights, wheels, bumper, and plate could be searched in the acquired images in order to verify, with the support of a classifier, the presence of
a vehicle. Also gradient information could be useful for this task. Moreover, tracking algorithms are employed to improve the reliability of the results.

When the overtaking vehicle is closed to the camera and its frontal part is occluded by the presence of the vehicle where the camera is mounted on, template matching could be used (Fig. 44.6).

The drawback of these approaches concerns the generalization of the detection for all type of vehicles providing, at the same time, low computational costs: since cars, motorcycles, and trucks have different features, the recognition step could be complex, affecting real-time performance.

Moreover, in the pattern analysis, the shadows under the vehicles could be considered in order to delimit the region of interest where vehicle are detected using distinctive information, like symmetries, edges, or shape (Fig. 44.7).

In this case, the processing involves low computational costs, but the system performance is limited because the extraction from a single image of information about object position is not accurate.
2.2 Optical Flow Estimation

Many approaches (Batavia et al. 1997; MVT Ltd 2004) are based on the evaluation of the object moving with respect to the camera’s one: Since the camera is mounted on the side-mirror, the vehicles’ detection have to be performed in a dynamic scenario (Fig. 44.8).

To recognize an overtaking vehicle all directions of image elements are processed: If an object is moving closer to the camera, it is considered as an overtaking vehicle; otherwise, it is a background element.

The optical flow estimation is a common choice to represent the apparent speeds distribution generated by the objects motions, and usually, the estimation is supported by a tracking algorithm that allows to search in the current frame the position of previously detected objects in order to estimate their speed and direction. Moreover, to increase the accuracy, the optical flow estimation could be combined with edges detection (Sun et al. Mar. 2006; Mae et al. 1996), color information or Kalman snakes.

Since the optical flow of the background elements is similar to the moving objects’ one, the studying of the FOE (Focus Of Expansion) could be a possible solution to increase the reliability of the results and limit, at the same time, the computational costs. In this case, the road plane is estimated in the image in order to delimit the region of interest for the vehicle searching.

Either dense or spread optical flow analysis may be performed: In the first approach, all image pixels are considered as contribution for the motion estimation; in the

![Fig. 44.8](image)

*Description of the scenario and the motion of the objects in the image*
second case, only some distinctive features are processed. In Fig. 44.9, an example for overtaking vehicle detection based on sparse optical flow analysis and eigenvalues processing is shown: In the image, dynamic and static areas are separated, then sparse optical flow is used to subtract dynamic region from the background in order to make the detection robust to camera shocks and vibrations.

An important aspect of the optical flow is its versatility: it allows to discriminate among static objects, elements with the same direction of the vehicle where the camera is mounted on and vehicles that follow an opposite direction with respect to the camera’s one.

For the optical flow estimation, it is firstly important to extract from the images a set of discriminative features: These points are tracked among difference frames, in order to determine their motion with respect to camera’s one. Then the features are classified according to their direction, in order to select and pack together all points belonging to a possible overtaking vehicle.

### 2.3 Features Extraction

Every system based on features tracking strictly depends on the feature’s extraction algorithm: If a feature detected in a frame does not appear in the next one, the input data for the tracking are unreliable. An important aspect to consider in the feature extraction is that in the image, some points could not be tracked: In shape and color homogeneous areas, where there is not texture, it is impossible to determine exactly where a previously detected point appears in the current frame; thus for some group of pixels, it is impossible to track their motion. Moreover, an object edge allows to determine the element motion only along the direction perpendicular to the edge. Instead in the presence of a corner, there is a significant brightness change that allows motion tracking along both axes.

Feature extraction can be performed directly processing gray level images or using edge-based algorithms; in this case, all points with maximum curvatures or edges crossing points are considered. In Figs. 44.10 and 44.11 are shown the images resulting by the application of two common methods for edges enhancement: gradient operator and Sobel filter.
In order to guarantee the reliability of the results, in automotive application, the principal aspects to be considered for the choice of the best features extractor are: temporal stability, accuracy on feature detection, and computational costs. Some examples of feature extractor are shown in Figs. 44.12 and 44.13.
2.4 Features Tracking

Feature tracking is based on the generation of a correlation among objects in different frames: it focuses on the detection in the current frame of previously detected features, in order to determine their motion, and estimate their future position.

The information to track consists of color distribution or edges, even if the use of color images is not a common choice because illumination and weather condition could affect the system performance. The choice of the suitable feature tracker depends on different aspects: the tracking model to use (statistics, volume, etc.), the dimension of the features space (2D or 3D), the type of acquisition system (monocular, stereo, trinocular), and the camera mobility (static or not).

Tracking could be focused on the extraction of specific features, useful to identify vehicles’ parts such as wheels: In this case, a model of the object to track is defined using a priori knowledge. In practice, model-based tracking exploits distinctive information about the elements to track or their motions to limit the set of region of interest to process.

Otherwise tracking may be performed after the detection, using as input all the region of interest, without any information about high level knowledge: In this case, the system is based on motion detection and tracking of the detected areas in order to determine the trajectories of the moving objects. This choice allows to provide generalization about the application scenario.

An example of non-model-based tracking approaches is the Hidden Markov Model, a probabilistic technique based on the analysis of time-discrete series: A basing structure is shown in Fig. 44.14. A generic state is connected with a certain probability to other states; the interconnection parameters may be estimated using feed forward techniques such as the Baum–Welch algorithm. The features to track could be points, lines, or two-dimensional blobs.

Another method for feature tracking could be based on optical flow estimation (Fig. 44.15): In this case, the information about objects motion is important to detect
scene dynamics and correlate spatial information with time variation. Starting from image points, the optical flow estimation provides a two-dimensional approximation of the 3D world points motions. To perform the tracking, all features with a uniform motion are packed together. The processing is repeated in different frames in order to determine in the images all elements with a similar motion and speed. Finally, elements’ trajectories are determined exploiting borders and center information of the detected region of interest. A priori knowledge about objects to track may improve the performance.

Differences between optical estimation and real motion may affect the performance: Since only the apparent motion may be extracted from the processed images, to obtain a quantitative evaluation of the scene, some additional assumptions have to be done, for example, brightness changes, objects specific features, and the relation between world points and 2D projections motion. Template matching is another possible tracking

\[ y(t) \]

**Fig. 44.14**
An example of three states HMM, \( y(t) \) is the observation obtained by each state

**Fig. 44.15**
Optical flow estimation for features tracking
approach: It is based on the localization of a specific region tracked using correlation evaluation. Usually, the area of interest is selected in order to make its detection easier: It involves the processing of specific image points called corner points (i.e., where the brightness gradient is high) and their neighborhood. The advantage of this technique concerns its flexibility. The template is frequently updated in order to make the tracking robust to scene variation: If the template in the current frame is significantly different from the previously detected one, a new template is defined according to the new region of interest; in the presence of slight template changes, a mean template version is processed to cope with small noise deviation. In the template matching, the correlation techniques could be region based or feature based according to the information to track.

In Fig. 44.16, an example of features’ tracking based on template matching is shown: The gray level intensity in a neighborhood around each detected feature is tracked between different frames; moreover, the position with respect to the vanishing point is considered in order to verify that the features’ motion is compatible with an overtaking vehicle motion. In particular, in a typical dynamite scene, the direction of the overtaking vehicle’s features is opposite to the vanishing point motion.

To evaluate the features’ motion with respect to the vanishing point, it is possible to calculate the angle between the line that crosses the previous detected feature and the vanishing point, and the line defined by the current feature and the previous detected one. This angle is equal to:

$$\cos \alpha = \frac{v \cdot w}{\|v\| \|w\|}$$  \hspace{1cm} (44.1)

Figure 44.17 shows a possible approach for the determination of the background and overtaking vehicles features according to their direction: $V_p$ is the vanishing point, $prec$ is the previous detected point, and $curr$ is the current feature. $\alpha$ is the angle
between the lines defined by the vectors $v$ and $w$. The matched corners are classified according to their motion as:

- **Overtaking corners**: They represent the overtaking vehicles, with a positive speed greater than the speed of the car where the camera is mounted on. The features pairs associated to this category have the $\cos$ positive and greater than a certain threshold.
- **Background corners**: They are associated to the background elements and to the objects with a negative speed; in this case, the features matches have $\cos$ negative and lower than a specific threshold.
- **Other corners**: They represent all other image elements with a motion not compatible with the previously described categories.

In **Fig. 44.18**, the tracked features are plotted with different colors, according to the included angle; in particular, the right image shows the classification results: Overtaking
vehicles are represented by red segments, background elements are enhanced with green lines, and all other objects are plotted with blue segments.

To track the pairs of features, best matches between the current points and the previously detected ones could be considered, associating a specific vote to each couple of features. Moreover, if the vote is not greater than a default threshold, the relative features pair is discarded in order to remove noisy contribution. The results obtained with this noise filtering are shown in Fig. 44.19.

To improve the results, it is possible to maintain the history information for each tracked features pair: For example, the position of the first instance, the current, and the previous features coordinates may be considered to count the number of frames on which that feature’s pair has been correctly detected. In this way, it is possible to calculate the “age” of each instance and use this value to further filter the results. In Fig. 44.20 are shown the results of the described features history tracking.

3 Algorithm Overview

A general scheme for blind spot monitoring is shown in Fig. 44.21: The top module defines a high level algorithm that interacts with three independent blocks, that
implement low level routines. These blocks operate simultaneously and communicate their outputs to the top level module. All the results are fuse together by the top level module in order to determine the appropriate behavior.

For each image acquired by the camera, Harris corners are extracted and tracked in order to detect all moving objects in the scene. According to pattern matching approaches, overtaking vehicles are recognized. Simultaneously lane detection is performed, allowing the high level application to warn the driver when an overtaking vehicle will occupy the lane closer to his car; moreover, with multi lanes detection, it is possible to evaluate their intersection in order to improve the FOE estimation and obtain a value that does not strictly depend on the calibration parameters.

To identify all dangerous overtaking vehicles, the high level module selects among all possible candidates only the cars within the limits defined by the second lane, if it is detected. To reduce the number of false positives, the selected elements have to be tracked in different frames and they have to be described by a significant number of Harris corners. If the distance of the overtaking vehicle from the camera is lower than a certain threshold, an acoustic and visual signal is activated to warn the driver about the dangerous situation.

An example of functioning is shown in Fig. 44.22: The left images do not represent dangerous situations because the vehicles detected in the second lane are far from the camera, thus the warning signal is not emitted. Vice-versa the scenarios in the right images described a critical situation: The vehicles detected are close to the camera; therefore, the red light is on and the car is delimited by a red bounding box to enhance its dangerousness.

When the acoustic/visual signal is activated, the blind spot area is monitored to warn the driver about the presence of vehicles in that region. Therefore, the signal remains enabled if the number of Harris corners relative to overtaking vehicles is greater than a specific threshold; otherwise, it is automatically switched off. In Fig. 44.23, the presence of a vehicle in the spot area involves the enabling of the warning signal (red light on) in the right image; when the vehicle leaves the dangerous zone (left image), the signal is deactivated.
Fig. 44.22
Example of lane change assistance system

Fig. 44.23
Example of vehicle detection in the blind spot area
In the system, different thresholds are defined in order to provide the reliability of the results:

- **Dangerous Distance**: It determines if the spatial limits within a lane change maneuver are considered dangerous.
- **Min Corner Vehicle**: It sets the minimal number of features needed to consider a candidate as an overtaking vehicle.
- **Min Age Vehicle**: It is the minimal tracking history for valid vehicles.
- **Min Dangerous Corner Blind Spot**: It defines the minimal number of corners in the blind spot area to consider the region occupied by a vehicle.

In Fig. 44.24 are shown the Harris corners tracked with the previously described approach, varying the dimension of the window used to detect the maximum values in the Harris image. By increasing this dimension, the detected corners are more stable, thus the tracking algorithm provides better performance.

![Harris corner tracking with different windows dimensions: (a)–(b) 5 pixel, (c)–d) 7 pixel (e)–(f) 9 pixel](image-url)
4 Conclusions

The implementation of a blind spot monitoring system, based on computer vision algorithms, for lane change assistance is an important issue for providing safety on the road. If the lane change maneuver is performed while an overtaking vehicle occupies the dangerous zone (i.e., delimited by a critical distance), an acoustic/visual signal is activated to warn the driver. A particular attention has to be reserved for the camera calibration phase to guarantee the reliability in the measurements of the overtaking vehicles distance. Moreover, with the image pre-processing, it is possible to provide image rectification and improve the vehicles detection performance. A blind spot monitoring system may be based on a layered architecture where different modules perform three principal tasks: vehicle identification, dynamic objects extraction, and lane detection. Several methods may be used for the detection of the overtaking vehicles: pattern analysis, optical flow estimation, features extraction, and tracking. A possible approach for overtaking vehicles detection could be based on their shadows analysis; the Harris corners may be used to discriminate between static and dynamic objects.

References
