

A New Approach to Urban Pedestrian Detection for Automatic Braking

Alberto Broggi, *Senior Member, IEEE*, Pietro Cerri, *Member, IEEE*, Stefano Ghidoni, *Member, IEEE*, Paolo Grisleri, *Member, IEEE*, and Ho Gi Jung, *Member, IEEE*

Abstract—This paper presents an application of a pedestrian-detection system aimed at localizing potentially dangerous situations under specific urban scenarios. The approach used in this paper differs from those implemented in traditional pedestrian-detection systems, which are designed to localize all pedestrians in the area in front of the vehicle. Conversely, this approach searches for pedestrians in critical areas only. The environment is reconstructed with a standard laser scanner, whereas the following check for the presence of pedestrians is performed due to the fusion with a vision system. The great advantages of such an approach are that pedestrian recognition is performed on limited image areas, therefore boosting its timewise performance, and no assessment on the danger level is finally required before providing the result to either the driver or an onboard computer for automatic maneuvers. A further advantage is the drastic reduction of false alarms, making this system robust enough to control nonreversible safety systems.

Index Terms—Artificial intelligence (AI), computer vision, fuzzy logic, image processing, pattern recognition, pedestrian detection.

I. INTRODUCTION

EXISTING pedestrian-detection systems are based on the search for pedestrians in the whole area in front of a vehicle. Candidates are located using pedestrian characteristics [1], such as shape, symmetry, texture, motion, and periodicity of human leg motion. When fusion between different sensing technologies is used, whether it is high level [2], [3] or low level [4], each sensor searches for pedestrian-specific features in the whole area in front of the vehicle.

In this paper, we describe a system for the detection of pedestrians based on a new approach. It is designed to work in a particularly challenging urban scenario, in which traditional pedestrian-detection approaches would yield nonoptimal results. Instead of searching for pedestrians in a large area in front

of the vehicle, making no assumption on the external environment, the system presented in this paper focuses on a specific urban scenario in which not only is the detection of a pedestrian of basic importance, but the danger of the situation can also clearly be assessed. In fact, in an advanced driving-assistance system, correct detection is just the first phase of a successful product: The localization of a traffic sign, a traffic light, an obstacle, or a pedestrian with no corresponding indication on their position with respect to the vehicle and the environment provides incomplete information. As an example, a pedestrian detection system that is able to correctly localize all pedestrians present in the scene provides a huge amount of information, which still needs to be filtered to be useful to either the driver or the onboard computer in charge of automatic maneuvers. A possible filter may be implemented by fusing information coming from other systems, such as lane detection or other situation analysis engines, and controller area network (CAN) data: A pedestrian exactly in front of the vehicle may or may not be considered dangerous, depending on the surrounding environment, as shown in Fig. 1.

In this paper, the Artificial Vision and Intelligent Systems Laboratory (VisLab) at the University of Parma approached the problem in the opposite way: Instead of detecting all possible candidates and filtering them after the analysis of the environment, we first analyze the scenario and then search for possible pedestrians in specific positions for that particular scenario. We call this approach *scenario-driven search* (SDS). This way, all detected pedestrians represent possible threats, and no further filtering is needed (apart from a validation and a possible final tracking step).

The scenarios that are considered here refer to the most common urban situations: When vehicles are moving on an urban road, the most common threat that a pedestrian may pose, therefore requiring successful detection, is road crossing. Stopped vehicles on the road or on the road edges may occlude visibility, thus making the detection of the pedestrian more complex.

The underlying idea of our SDS approach applied to this specific scenario is to localize stopped vehicles and then search for pedestrians in their close proximity or in the areas partly hidden by them. The stopped vehicles, whose edges will trigger the search for pedestrians, may be parked cars on the road edge, vehicles temporarily stopped on the road, vehicles queued in a line in front of a traffic light or zebra crossing, or simply jammed cars.

The first row of Fig. 2 shows some examples of situations in which the visibility of a crossing pedestrian is partly or completely occluded by stopped vehicles. The second row of Fig. 2 highlights, for each situation, the areas on which the system will perform a check for the presence of a possible pedestrian.

Manuscript received September 30, 2008; revised February 26, 2009, July 13, 2009, and August 31, 2009. First published October 20, 2009; current version published December 3, 2009. The Associate Editors for this paper was B. de Schutter and S. Shladover.

A. Broggi, P. Cerri, and P. Grisleri are with the Artificial Vision and Intelligent Systems Laboratory (VisLab), Dipartimento di Ingegneria dell'Informazione, Università di Parma, 43124 Parma, Italy (e-mail: broggi@vislab.it; cerri@vislab.it; grisleri@vislab.it).

S. Ghidoni was with the Artificial Vision and Intelligent Systems Laboratory (VisLab), Dipartimento di Ingegneria dell'Informazione, Università di Parma, 43124 Parma, Italy. He is now with the Intelligent Autonomous Systems Laboratory (IAS-Lab), Department of Information Engineering, University of Padua, 35131 Padova, Italy (e-mail: ghidoni@dei.unipd.it).

H. G. Jung was with MANDO Corporation Global R&D Headquarters, 777 Yongin, Korea. He is now with Yonsei University, Seoul 120-749, Korea (e-mail: hgjung@yonsei.ac.kr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TITS.2009.2032770

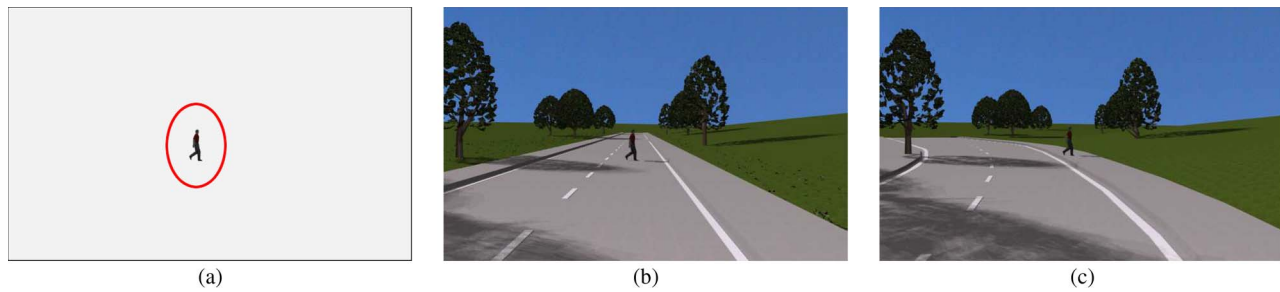


Fig. 1. When a pedestrian is localized but no reference with respect to the environment is provided like in (a), the detector is not able to assess the danger level. When environmental information are available, the very same pedestrian may (b) become a threat or (c) be in a safe position.

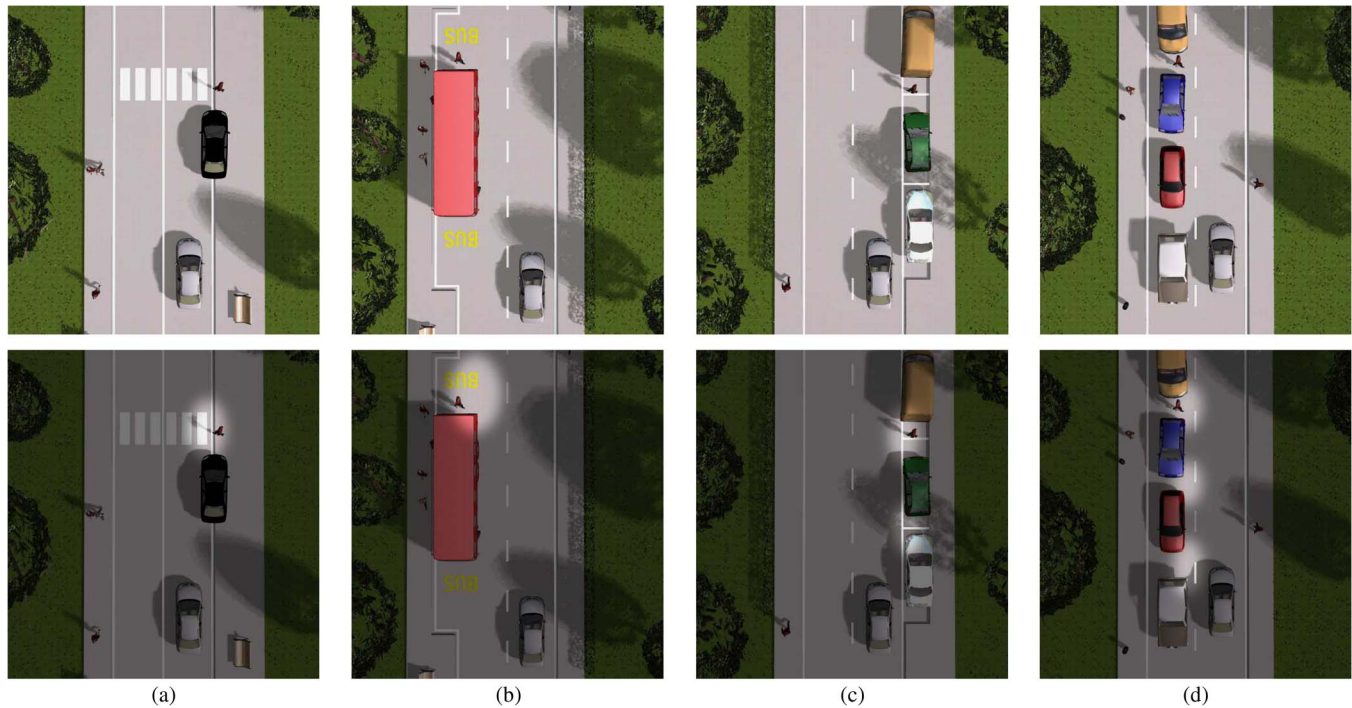


Fig. 2. (First row) Situations considered in this paper. (Second row) Areas of interest considered for the detection of a possible pedestrian. (a) A crossing pedestrian is hidden by a parked vehicle. (b) A pedestrian is crossing the road behind a stopped bus. (c) A pedestrian is appearing between two parked vehicles. (d) A pedestrian is crossing the road between two vehicles stopped on the other side of the road. It is important to note that situations (a) and (b) refer to specific and predetermined urban areas (zebra crossings and bus stops). Situations (c) and (d) may happen in any portion of the road network.

In other words, this paper focuses on the detection of pedestrians appearing just behind occluding obstacles; pedestrians that are clearly visible in the frontal area of the vehicle also need to be detected, but this function, which is also available on other systems [5], [6], is currently out of the scope of this paper.

The idea of focusing on a specific scene or scenario (referring to a dynamic or a static environment, respectively) is not new to pedestrian detection systems; in 2002, Franke and Heinrich [7] developed a module that is able to detect balls (which are usually a strong sign of the presence of a child). Another example of very specific systems is that developed by Curio *et al.* [8], which was based on the visual localization of the specific moving pattern of human legs.

It is known [9] that parked vehicles, blocking the visibility of pedestrians, are one of the main causes of accidents. Agran *et al.* [10] shows that the number of parked vehicles along a street is the strongest risk factor for pedestrian injury occurring on residential areas. Although, in these areas, parking spaces should diagonally be arranged, there are situations, as shown in Fig. 2, in which vehicles temporarily stop on the

road, and their position cannot strategically be organized and carefully controlled, as in the case of parking lots. Although some of the situations of Fig. 2 refer to specific urban areas (i.e., zebra crossings and bus stops) that could specifically be enhanced by intelligent infrastructures aimed at warning oncoming vehicles, other situations can happen in any part of the road network, making the installation of specific warning infrastructures impractical.

The main characteristic required by our system is the capability to perform the following:

- 1) quickly detect pedestrians, given the short working range and the particularly high danger of collision with a pedestrian suddenly appearing behind an obstacle;
- 2) detect pedestrians as soon as they appear, even when they are still partly occluded;
- 3) limit the search to specific areas, which are determined by a quick preprocessing.

Fig. 3 shows the coverage of different sensing technologies (laser and vision) in a specific scenario considered in this paper.

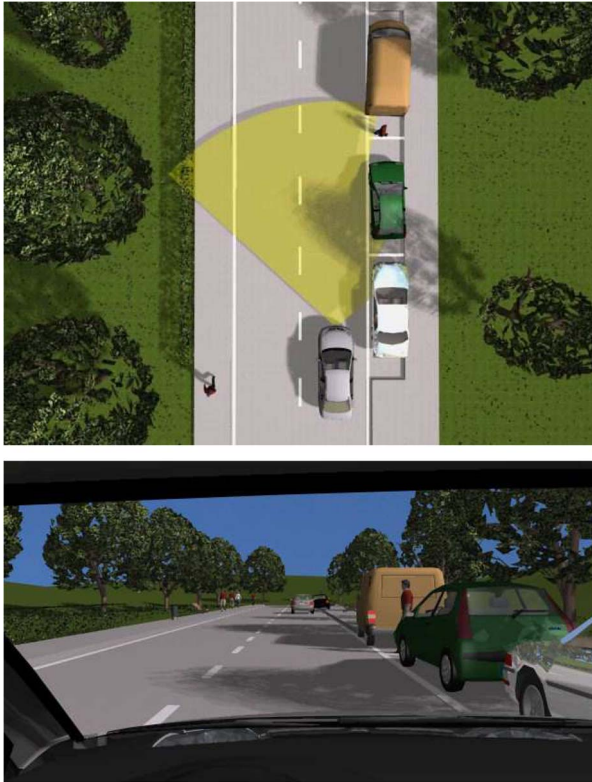


Fig. 3. Pedestrian partially hidden by parked cars may not be detected by a laser scanner positioned in the front bumper but can be detected using vision, even if partially occluded.

The fusion of laser scanner and vision can provide a quick and robust detection in case of suddenly appearing pedestrians: The laser scanner provides a list of areas in which a pedestrian may appear, whereas the camera is able to detect the pedestrian, even when he/she is not yet visible to the laser scanner.

A. Final Goal and Additional Requirements

The final goal of our pedestrian-detection system is to save lives and increase road safety through the use of both reversible and nonreversible driving-assistance systems. Once the pedestrian is detected with a sufficiently high confidence level, a warning is sent to the driver. In our system, we use audible warnings, but the final human machine interface is still to be defined. Should the driver not promptly react to the warning, the system would issue a second level of warning by blowing the vehicle's horn. This second warning is still considered a reversible system, although it is much more invasive than the former. The aim of this loud warning is to attract the attention of both the pedestrian itself and, once again, the driver. In case the danger level is not reduced due to a prompt action of the driver (or the pedestrian), the intelligent vehicle will trigger a nonreversible system, such as automatic braking.

Being a nonreversible and very invasive system, its triggering must be preceded by an extremely careful analysis of the danger level. Furthermore, the use of a nonreversible system requires the complete processing to be thoroughly tested with respect to false detections.

The false positive rate (i.e., the number of wrong detections per second) is a parameter that is gaining increasingly more

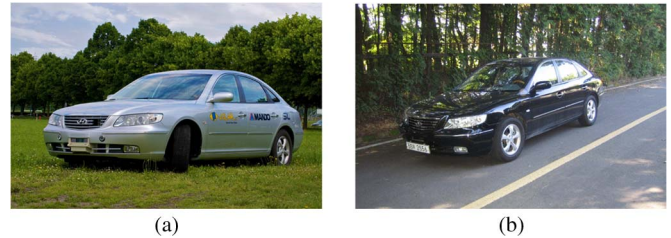


Fig. 4. (a) Hyundai Grandeur test vehicle equipped in Parma, Italy, and (b) the replica in Seoul, Korea.

interest as the research progresses toward the integration of higher levels of automation (i.e., nonreversible system). In the case of automatic braking, the number of false positives is even more important than the number of correct detections.

In the literature [1], many different pedestrian detection systems have been surveyed, and a large number of test methods for pedestrian detection systems have been proposed, which are based on single-frame analysis, temporal analysis of continuous video streams, or an event. Nevertheless, none of these methods provide sufficiently high performance to be considered safe enough to be fielded on a vehicle with a nonreversible safety system. Although the performance of each method in terms of false positive rate depends on a number of different parameters (such as the desired correct detection rate), the best systems are characterized by a false positive rate of about 0.01/frame. Considering a system running at 10 Hz, it would provide one false positive every 10 s. Clearly, this performance is very far from the ideal system.

The approach described in this paper is an alternative to conventional systems and aims to detect pedestrians in situations of clear danger by limiting the search to specific areas. In addition to being quicker than other systems, this method also aims to reduce the number of false detections to zero.

The system presented in this paper has been developed and tested by VisLab on a Hyundai Grandeur prototype vehicle; Mando is using an exact replica of this system in Korea (see Fig. 4) to double test time.

Section II presents the test vehicle setup. Section III describes the risky-area-detection subsystem.

In Section IV, considerations about vision fusion are discussed. In Section V, experimental results are presented, whereas Section VI concludes this paper with some final remarks.

II. TEST VEHICLE

This section describes the perception system installed on our test vehicle and the guidelines used to choose the sensors, the actuators, and the processing system.

A. Sensing Technologies

Being designed to address an urban scenario in which the prototype vehicle is running close to stopped vehicles, a limitation on vehicle speed and detection range can be accepted. Low-to-medium vehicle speeds of up to 50 km/h and a detection range of about 40 m can be considered as a safe choice.

Environmental sensing requires stopped vehicles and other standing obstacles to be detected; a laser-based solution is sufficiently strong to localize large obstacles, such as vehicles,

and classify them due to their shape. To overcome the lack of speed information associated with laser scans, we selected a particular laser scanner with the specific characteristic of providing interlaced data. It groups a number of interlaced scans together to form a single higher resolution scan. The analysis of this interlaced data coupled with vehicle inertial data allows the estimation of obstacle speed and, therefore, localizing static objects.

The best technology to check for the presence of even partly occluded pedestrians in given areas is vision. Monocular vision is sufficient since no 3-D reconstruction is needed, with distance measurements being already available. Although the system is now being tested under daylight conditions, the use of a near-infrared (NIR) camera and proper illumination allows the extension of its operational range to the night.

B. Sensor Selection

The camera, i.e., an AVT Guppy F-036B, has been chosen for several reasons. The sensor of 752×480 pixels has an aspect ratio of slightly less than 15/9, which is particularly suitable for automotive applications since it frames a large lateral area, which often contains relevant information. The sensitivity covers both the visible and NIR spectra. At night, a high response to the NIR radiation allows the detection of objects due to a specific illumination.

To achieve this, additional headlamps are mounted in front of the vehicle. An NIR light-emitting diode headlight with an aperture of about 25° is mounted in front of the radiator, whereas headlamp blocks are customized by LS and consist of two modules: 1) a visible-range bifunctional lamp acting as low beam and high beam and 2) a NIR spectrum-range-only lamp.

The laser scanner is a SICK LMS 211-30206. The detection capabilities (scanning angle of 100° , minimum angular resolution of 0.25° , up-to-80-m range, and fog correction) are suitable for our goal.

The laser scanner is capable of 1° resolution, but due to four subsequent rotations, the use of an interlacing system, and a phase gap of 0.25° , it is possible to decrease the final granularity to 0.25° . Every rotation takes 13.32 ms; therefore, in 53.28 ms, four rotations are performed. The time difference between measurements of the same scan is not negligible. Moreover, when the vehicle is moving, the shift between the laser scanner position when the first pulse is measured and its position when the last pulse is measured is appreciable: The laser scanner interlacing system makes this problem even more evident.

The camera and the laser scanner are not synchronized in hardware to relax the set of requirements. This means that a variable time shift exists between the data acquired by the two sensors; displacement due to the nonsynchronization is negligible, because the vehicle moves at a low speed, and the processing rate is sufficiently high.

C. Vehicle Setup

The laser scanner and NIR headlamps are located in the front bumper, as shown in Fig. 5. The NIR camera is placed inside the driving cabin near the rear-view mirror, as shown in Fig. 6.

A compact PC (an Intel Core 2 Duo-based Mini-ITX) is installed in the boot. The onboard FireWire A controller is



Fig. 5. Detail of the front bumper showing the laser scanner integration and the headlights.



Fig. 6. How the camera is installed inside the cabin.

used to connect the camera. An RS422-to-Universal-Serial-Bus (USB) adapter provides an easy connection with the laser. In addition, inertial data are gathered through the CAN bus using a USB adapter.

Mando modified the vehicle braking system, allowing control of the braking strength by sending appropriate CAN messages from the processing box to the actuators. The braking system installed on the test vehicle was replaced with Mando's *MGH-40 ESC plus* to control vehicle deceleration via CAN. The braking system incorporates a deceleration control interface (DCI) for high-level system functions, such as adaptive cruise control (ACC) and precrash safety. As DCI receives multiple deceleration commands, the deceleration coordinator selects the desired deceleration based on priorities and vehicle status. Once the desired deceleration is set, a feedforward controller with a feedback loop controls vehicle deceleration by controlling wheel brake pressures. The DCI braking range extends up to 1.0 g, with a resolution of 0.01 g, a stable error lower than 0.05 g, and a response time of 0.3 s. The DCI cooperatively controls deceleration with vehicle-stability control functions, such as electronic brake distribution, antilock braking system, and electronic stability control (ESC).

The horn was also modified so that it can also be controlled through a USB I/O board shown in Fig. 7(a). The processing system controls the horn by simply writing on a serial port.

An additional 100-Hz yaw rate sensor, which was manufactured by Siemens VDO and Mando, was also installed in the car's center of mass to get more accurate yaw measurements. The sensor is connected through the processing system via the CAN bus. The yaw rate sensor provides both lateral acceleration and yaw rate. Yaw rate values are provided with an accuracy of $0.0625^\circ/\text{s}$. Fig. 7(b) shows the sensor.

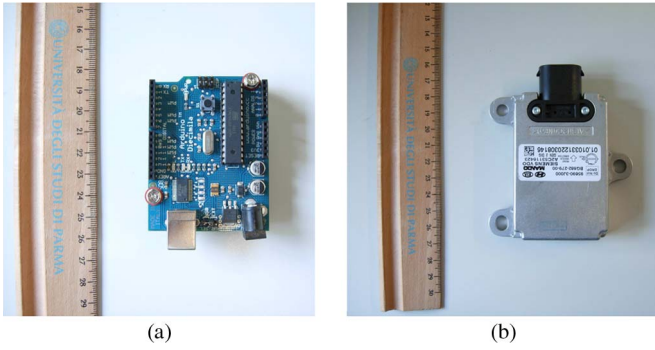


Fig. 7. (a) USB I/O board used to control the horn. (b) Special yaw rate sensor.

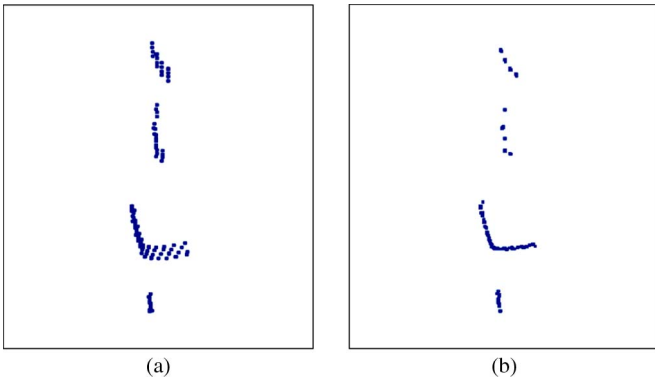


Fig. 8. Laser data (a) without correction and (b) with correction.

III. FOCUS OF ATTENTION

Two different laser data classifications are developed to make the algorithm robust. Both methods are based on clustering pulses into straight lines, but while the first method considers the last scan only to classify obstacles, the second method uses a number of subsequent scans to verify obstacle positions and hence estimate their speed.

A. Data Alignment

As already explained in Section II, shifts between subsequent laser scanner measurements can be appreciable and may cause clustering or classification problems since obstacle shapes may appear distorted.

By using ego-motion data provided by the ESP CAN box, it is possible to estimate vehicle rototranslation and therefore correct the position measured for each pulse. The vehicle speed is used to correct translation, whereas the steering wheel angle allows calculating the yaw rate and, therefore, the rotation correction matrix. The yaw rate can be computed as

$$\text{yaw rate} = \frac{V_{fr} + V_{fl}}{t_w \cos \delta}$$

where V_{fr} and V_{fl} are the front right and front left wheel speeds, respectively; δ indicates the wheel angle; and t_w is the vehicle front track.

As yaw rate measurements are affected by noise, an additional yaw rate sensor is installed on the vehicle to increase precision and perform a sharper data correction.

Fig. 8 shows raw and corrected laser data referring to non-moving obstacles: The four laser rotations that create a scan

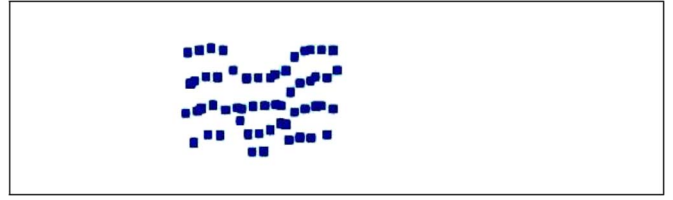


Fig. 9. Moving obstacle is represented by four different lines, even after data correction.

are clearly visible before the correction, whereas the obstacle appears as a single line after the correction.

Pulses echoed by moving obstacles cannot be compacted into a single line since the obstacle position changes during the multiscan: Moving obstacles are still characterized by four parallel lines after data correction, as shown in Fig. 9. This information will be used for obstacle classification.

B. Data Clustering

First, the pulses that belong to the same rotation are connected together: Moving obstacles are then identified by four different and parallel clusters, whereas standing obstacles are identified by four overlapping clusters. The pulses are clustered as chains of segments. The points that cannot be joined into any chain of the same rotation are then checked for merging with points of other rotations, considering proximity only. The points that cannot be connected to any other point or are close to the limit of the laser scanner range (about 80 m) are permanently discarded.

C. Segment Merging

Up to now, pulses are joined into chains of segments without any other information. Adjacent segments with approximately the same orientation can be merged into a longer segment, preserving the obstacle shape but reducing both the data-structure complexity and the details of the representation. Considering a chain of segments, it is possible to compute the straight line that connects the chain start and end points. Then, the distance between this line and all the internal points is computed, and if the maximum distance is larger than the threshold, the line is split into two lines. These steps are iterated while the maximum distance is larger than the threshold. The result is that each chain is therefore finally segmented into a polyline.

Fig. 10 shows a moving vehicle: The rear bumper is framed as four parallel lines, whereas its side, which is parallel to the vehicle movement direction, is marked by a single line.

D. Line Merging

Every obstacle in the laser scanner field of view is identified by four lines, i.e., one for each laser rotation that composes the whole high-resolution scan. In the case of static obstacles, the four lines are almost perfectly overlapping, due to the previous data correction, and can be merged. Conversely, in the case of moving obstacles, the lines are parallel but not overlapping; therefore, static and moving obstacles can be located and correctly classified. Fig. 11 shows the algorithm steps. This quick, simple, yet very effective process aimed at identifying static

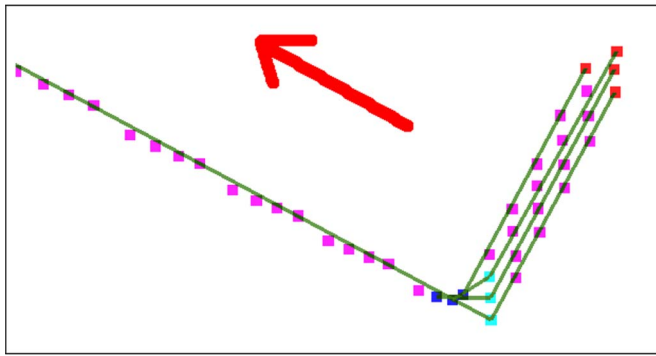
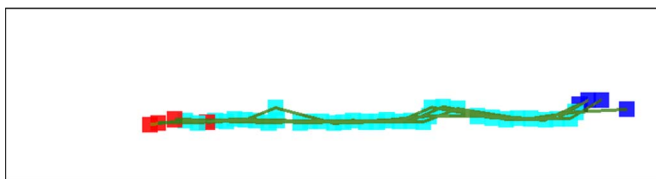
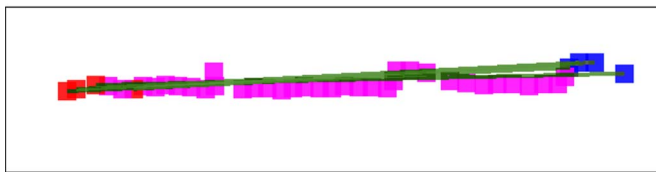


Fig. 10. Moving vehicle and its motion direction. The four parallel lines correspond to its rear bumper, whereas the single line corresponds to its left side. The blue points are polyline start points, the cyan points are corner points, the red points are polyline endpoints, and the violet points are generic line points.



(a)



(b)



(c)

Fig. 11. Steps of the algorithm. (a) Data clustering. (b) Approximation using polylines. (c) Line merging. All the points are merged into a single line. To simplify the following steps, small drifts are ignored.

obstacles may generate false negatives when the vehicle's pitch is not negligible. Section IV will discuss this issue.

E. Early Obstacle Classification

The polylines obtained so far can be classified according only to their size and *shape*. It is possible to divide obstacles into the following categories:

- 1) possible pedestrian;
- 2) road infrastructure;
- 3) L-shaped obstacle;
- 4) generic obstacle.

Obstacles that are eventually classified as pedestrians are supposed to have a limited size, whereas obstacles that have a large size and are almost parallel to the vehicle are assumed to be road edges (guardrails, buildings, road infrastructures, etc.). A simple and fast method based on line orientation is used to detect L-shaped obstacles. All obstacles that have yet to be classified are tagged as generic obstacles. The results obtained so far are satisfactory; Fig. 12 shows an example.

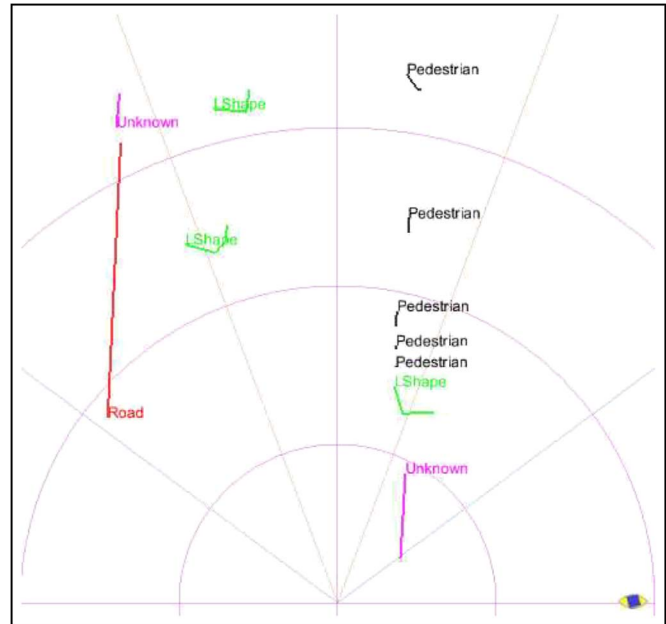


Fig. 12. Obstacle classification. All possible classifications based on *shape* are represented in the image.

Information about obstacle movement that was already estimated in the previous step is also stored for further usage.

F. Temporal and Shape-Based Correlation

The classification explained in the previous section classifies all small obstacles as possible pedestrians. Even if it is possible that all pedestrians are correctly classified, a number of false positives may indeed be present.

Fixed obstacles along the road (particularly parked vehicles) are used here to localize critical areas in front of the vehicle so that attention can be focussed on the immediate proximity of the edges of these risky areas, in which pedestrians can appear and become dangerous.

Polylines provided by the previous scans are aligned and rototranslated according to ego-motion; then, the overlap between the current and previous polylines is checked to provide a new classification in the following four classes: 1) *moving* obstacle; 2) *static* obstacle; 3) *changing-shape* obstacle; and 4) *new* obstacle.

Scan data referring to moving obstacles should have little or no spatial overlap when the time window is large, but unfortunately, the side of a moving vehicle appears as static (high spatial overlap). Anyway, due to the previous labeling of L-shaped objects, the bumper and the sides of the same vehicle belong to the same object, which therefore inherits the *moving* label.

Obstacles that are represented by overlapping scans in the given time window are marked as *static*, even when the newer scan lines are longer and more accurate than the older scan lines (due to the vehicle getting closer to the obstacle).

When there is no correspondence between the current and the old polyline, the object is classified as a *new* obstacle.

Static obstacles are important to locate the areas of interest in which vision will search for pedestrians; *changing-shape* obstacles are also of basic importance since they may contain a pedestrian in a very precise region of their shape.

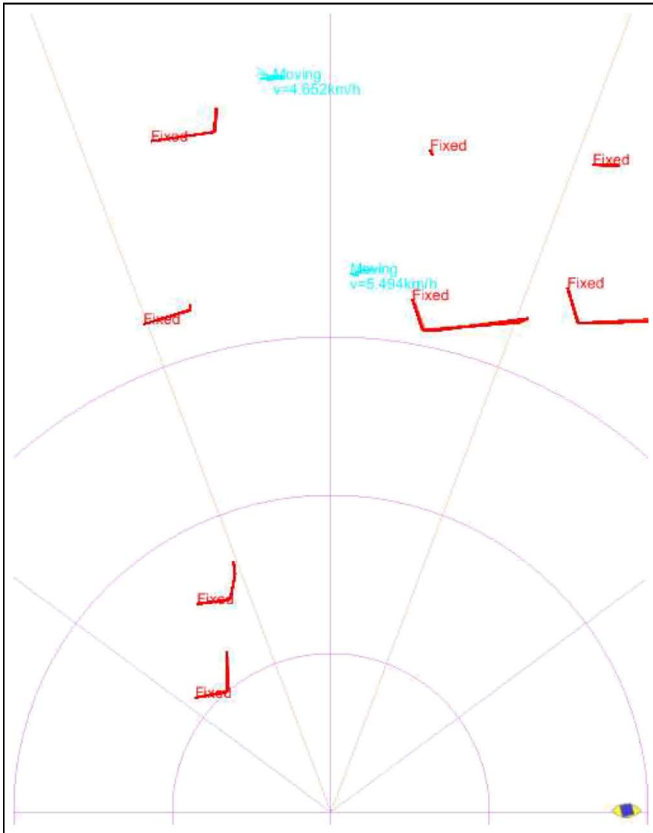


Fig. 13. Obstacle classification. All possible classifications based on *movement* are shown. Note that moving polylines change their position in world coordinates.

To trace obstacles, the centroid of each polyline is computed. Centroids are used to simplify the computation: more complex laser scanner data processing based on comparison between each point of the polylines are presented in [11]. The distances between the centroid of the last scan and the centroids of previous scans are computed and used for classification.

Polylines are labeled, considering all their points, as follows:

- 1) *fixed* obstacle;
- 2) *moving* obstacle;
- 3) *unknown* obstacle.

A polyline is labeled as *fixed* if all the distances between its current centroid and all the past centroids are below a threshold, whereas it is labeled as *moving* if the same distances increase with time. Another check is made to detect *fixed* obstacles: The threshold is increased according to the distance to the vehicle (to compensate for a decreasing accuracy) and the time gap (to partly compensate for errors in ego-motion reconstruction). A polyline is labeled as *unknown* if it can be labeled as neither *fixed* nor *moving*; this may happen the first time an obstacle appears or when the detection is not stable due to bad reflections of the laser beam. To compensate for inaccurate results in laser data processing, the five previous labels are also used; the polyline is labeled with the most common label.

A result obtained with this classification is shown in Fig. 13, which also shows an estimation of the speed of the object represented by the polyline.

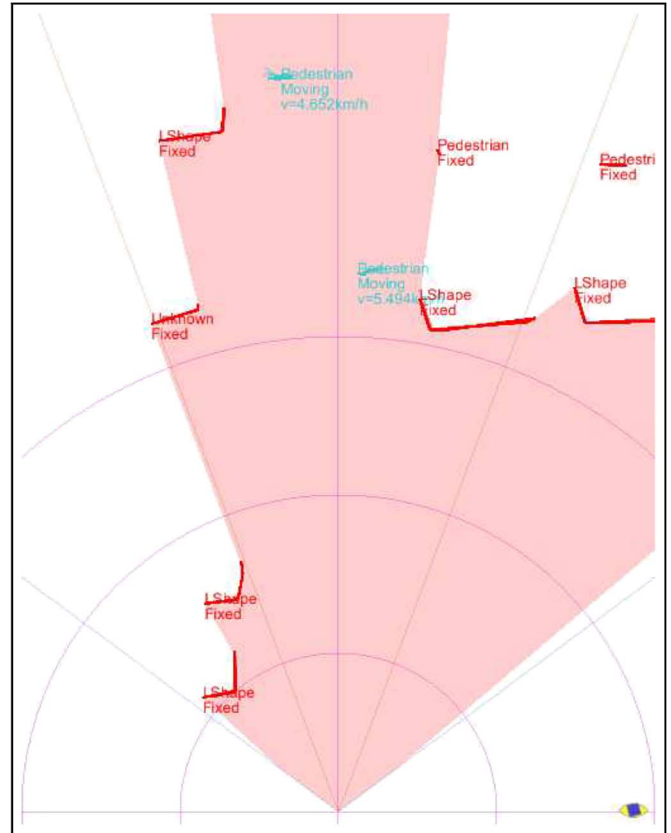


Fig. 14. Highlighted zone represents the driving corridor.

G. Driving Corridor

Polylines labeled according to their shape, size, and history represent a high-level and easy-to-use structure containing all the information needed for the following steps of the algorithm. This information can be used to identify the environment (static obstacles and vehicles) and the moving obstacles (pedestrians and vehicles). Obstacles identified as static by the second classification stage are used to define the environment structure. Fixed obstacles are used to build the driving corridor, i.e., the area that may be reached by the vehicle in the near future, as shown in Fig. 14.

The driving corridor can be composed of multiple forks, but the width of each subcorridor must be larger than the vehicle width; forks with a width narrower than this threshold are discarded. It is important to note that the corridor does not always overlap with the free space in front of the vehicle since it is built using only fixed obstacles and may also include moving obstacles.

H. Dangerous Areas and Possible Pedestrians

As already explained in the introduction, the algorithm is focused on the detection of suddenly appearing pedestrians and on areas hidden by a static obstacle or between two static obstacles. These areas are located along the corridor edges. Static obstacles that are used to build the corridor are also used to identify dangerous areas, i.e., areas in which pedestrians may appear. Dangerous areas are located behind the furthest point of each static polyline, as shown in Fig. 15.

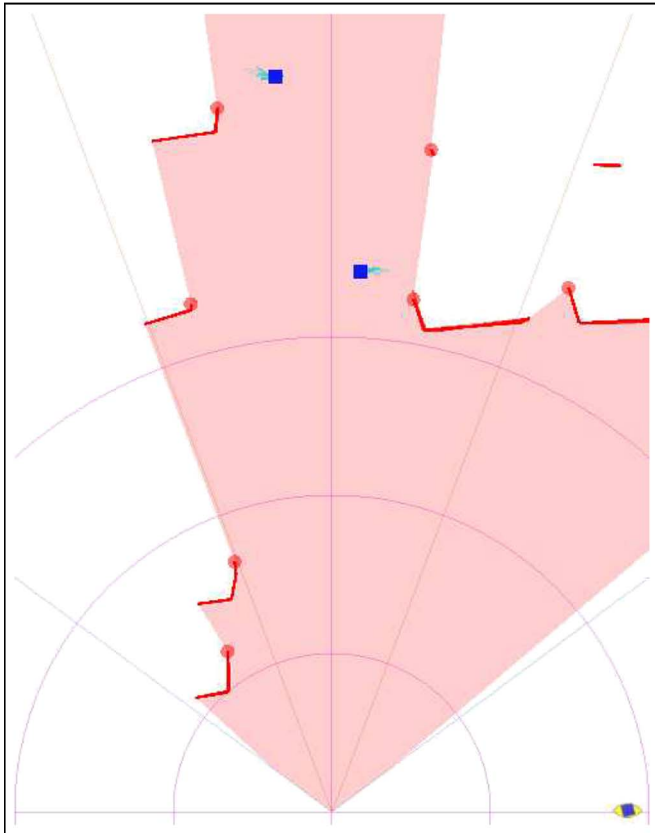


Fig. 15. Dangerous areas and possible pedestrians. (Circles) Dangerous areas. (Boxes) Possible pedestrians.

Obstacles moving inside the corridor, or near its edges, with a speed and a size typical of a pedestrian, are tagged as possible pedestrians.

IV. VISION FUSION

The regions of interest are 2-D areas in world coordinates; their corresponding areas in the image are then located by camera calibration and perspective mapping. Indeed, camera calibration and, more generally, *system* calibration play a basic role in a fusion system and need to be carefully tackled.

A large variation of vehicle pitch angles during motion may change camera orientation with respect to the world, but most importantly, it also causes the laser scanner to scan the scene using a different orientation. For on-road applications, this is generally considered a limited issue in the case of short distance sensing, and it is even less important when obstacles are vertical. Conversely, in the case of appreciable pitch angles, particularly negative angles that tilt the laser scanner toward the ground, the laser scanner's scanning plane may intersect with the ground or point to the sky, therefore yielding nonsignificant results.

An additional accelerometer is used to compute pitch. The instantaneous vehicle pitch is fed to the system (to inhibit the generation of wrong results during vehicle pitching).

The developed fusion is a laser-driven one: A laser is used to generate candidate, and vision is used to validate the candidate.

A. Classification

Once the areas of attention are located, the search for pedestrians is triggered in these areas. Specific image windows are defined using a perspective mapping transformation, considering 90 cm as the pedestrian width and 180 cm as the pedestrian height. These image areas are resampled to a fixed size (24×48 pixels).

AdaBoost was chosen to label each region of interest; AdaBoost is a technique that is widely used for the classification of pedestrians [12]. Haar features were chosen for the weak classifier [13]. Different Haar features are selected for each iteration, as suggested by Viola and Jones [14].

Instead of using two classes only (pedestrians and nonpedestrians), the following three classes are used here:

- 1) pedestrians;
- 2) nonpedestrians;
- 3) appearing pedestrians.

Appearing pedestrians are pedestrians that are initially not completely visible, i.e., partially occluded by obstacles, so that only a part of the pedestrian's shape can be framed, i.e., the upper or the side part only.

AdaBoost was trained using image windows determined by the previous steps of the algorithm. Roughly 100 000 images were dumped and manually divided into classes. Candidate selection is a complex and critical step for the AdaBoost training process [15]. Both normal and flipped samples are used in the training process. If two or more images are very similar (i.e., the pixelwise difference is less than a threshold), only one of them is used in the training process. In Fig. 16, some examples of images used in the training process are presented, i.e., images framed in Italy, the Netherlands (during tests before the IV 2008 demonstration), and Korea. Different pedestrians and pedestrians in different postures must be chosen. Different light conditions must be considered as well (see Fig. 16). Unfortunately, image sets freely available on the Internet [16] are not complete and cannot be used here since, in this case, we also need samples containing even partly occluded pedestrians (see the third row of Fig. 16).

As AdaBoost classification can be too specific on the images that are used during the training phase, a new sequence is used to assess the obtained classifier performance. To make the classification more precise, misclassified images are added to the training set for a new training.

B. Alerts and Warnings

Once a pedestrian partly hidden by a vehicle is detected by the vision system only, the system issues an internal *alert*; in this case, no warning is provided to the driver, because the danger level has yet to be determined.

When an *alert* is issued, the search continues in the same zone, and tracking is started. As soon as a tracked pedestrian also becomes visible to the laser scanner, the direction of its movement is considered; if a pedestrian moving from the corridor's edge to the corridor's center is detected, a *warning* is then issued to the driver. Fig. 17 shows an *alert* and the *warning* following it.



Fig. 16. Images used for training. (First row) Pedestrians. (Second row) Nonpedestrians. (Third row) Partly occluded pedestrians or, as defined in this paper, *appearing pedestrians*.



(a)



(b)

Fig. 17. Two subsequent frames of a sequence. (a) The system detects the partly occluded pedestrian and issues an internal *alert* but no warning to the driver. (b) When a detection of a fully visible pedestrian follows the internal *alert*, a *warning* is promptly sent to the driver.

It is of basic importance to note that the driver is warned only when the pedestrian is completely visible, like in other systems. However, the system presented in this paper is more reactive than others since the tracking starts when the pedestrian is only partially visible to one of the sensors; in an urban situation like this, promptness is an important key to the success of the system.

Fig. 18 shows some results obtained under complex conditions (bad weather condition, poor light condition, etc.) and a

misclassification, i.e., a suddenly appearing pedestrian detected as a normal pedestrian.

V. RESULTS

Usually, performance assessment of pedestrian-detection systems provides the percentage of correctly localized pedestrians and the false alarm rate. As this system is developed under a new perspective, performance figures are provided as the percentage of *appearing* pedestrians correctly identified and the false alarm rate. Indeed, it is not important to detect all the pedestrians present in the scene but only the dangerous pedestrians.

A set of sequences (different from that used for the training) was considered for extensive performance assessment, for a total driving time of about 10 h in complex urban scenarios. An entire week was dedicated to the final test; a total of 236 km was driven during the day and at night, under different weather conditions (sunny, cloudy, rain, and fog). Various scenarios were included: downtown, large and narrow roads, underground car parks, highways, and rural roads. During night tests, scenarios with and without external illumination were acquired. To complete the test sets, some specific situations were staged (such as very dangerous pedestrian crossings), but a number of dangerous scenarios were framed during normal driving anyway.

In the tests, the following performance indexes were considered:

- 1) number of pedestrians suddenly appearing in front of the vehicle (that must generate warnings to the driver);
- 2) number of pedestrians that appeared in front of the vehicle that have successfully been tracked and which can be hit by the vehicle (that must trigger the automatic brake);
- 3) number of fully visible pedestrians, which can be hit by the vehicle (that must trigger the automatic brake).

For each performance index, the number of correct detections, false positives, and false negatives were computed.

A total of 24 suddenly appearing pedestrians were correctly detected in the tests (case 1). Only one false positive is present (due to a misclassification of a parked scooter).

Considering the actual camera frame rate of 15 frame/s and the test duration, the false positive rate is about 2×10^{-6} false

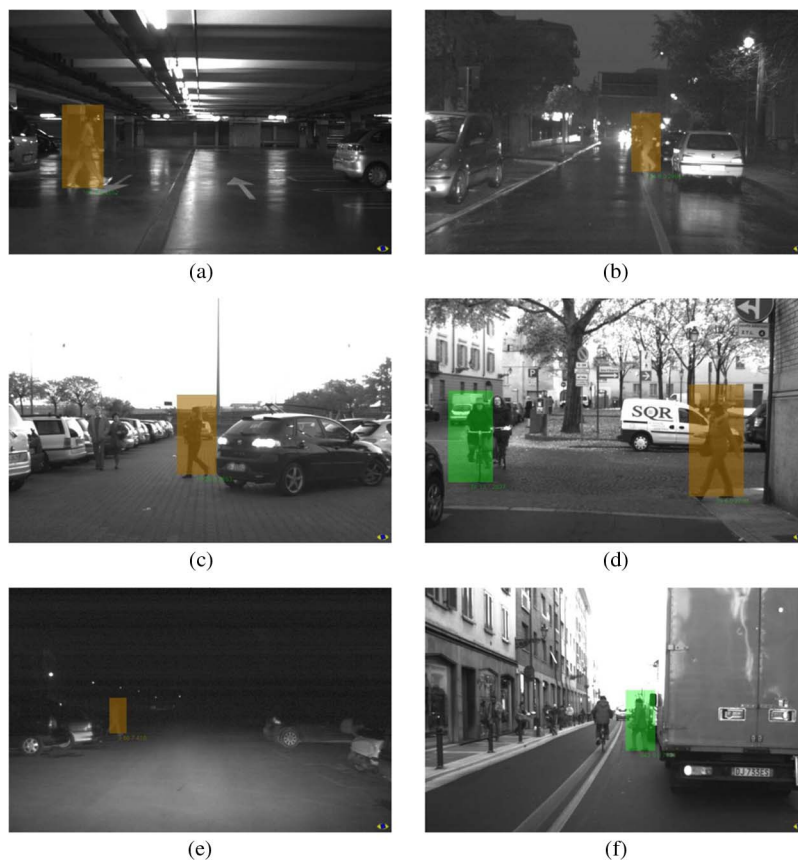


Fig. 18. Some suddenly appearing pedestrians correctly detected (a) in an underground parking, (b) in the rain, (c) behind a misaligned vehicle, (d) behind a wall, and (e) at night, and (f) a suddenly appearing pedestrian detected as a nondangerous pedestrian (false negative).

TABLE I
RESULTS OF A 10-H DRIVE (~236 km, ~540 000 IMAGES)

Perform. index	Correct Detections	False Negatives	False Positives
A	24	11	1
B	5	0	0
C	8	0	0

positive/frame. Only one missed warning out of 11 was due to a pedestrian misdetection; the others are caused by either alert misses or delayed detections. However, in these cases, even if the warning signal is not promptly issued to the driver, the automatic brake would have stopped the vehicle, avoiding the crash if the pedestrian would have been in a dangerous situation.

All pedestrians into the deceleration area are localized (no false negatives): Five fall under case 2, whereas eight fall under case 3. No false positives are present, due to the special attention paid to develop this SDS approach. Anyway, it is important to note that, for safety reasons, given that the tests were performed in real traffic, the size of the deceleration area was increased; nevertheless, the system behaved very satisfactorily.

Table I summarizes the results.

The limited number of dangerous events requiring the intervention of our safety system is not surprising: During normal driving, dangerous situations are not frequent; in addition, we staged some of them to challenge our system.

The analysis of the results obtained during the test also highlighted good results in the case of rain, when pedestrians

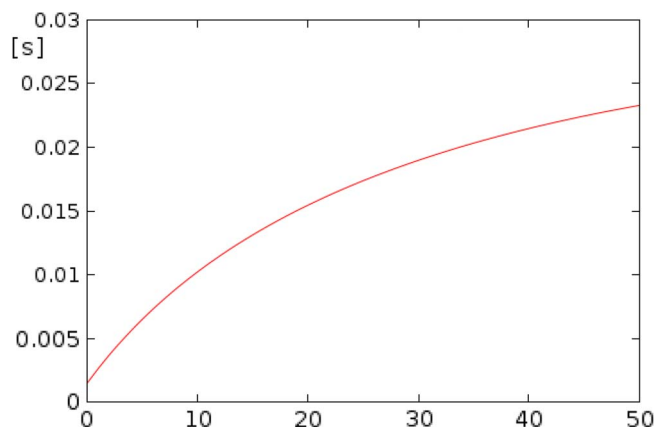


Fig. 19. Processing time versus the number of analyzed areas.

with umbrellas were detected as well. Anyway, the system is not able to discriminate between multiple pedestrians moving together or in situations in which the laser scanner is not able to obtain a clear picture, e.g., when pedestrians hold bags or other large objects. Moreover, suddenly appearing pedestrians walking very slowly or pedestrians under very critical lighting conditions can be missed.

Fig. 19 shows the processing time at the variation of the number of analyzed areas. (The values are computed using a Core-2-Duo PC running at 2.0 GHz.) Considering that the laser scanner works at about 20 Hz, full-speed processing is achieved.

TABLE II
MAIN STEPS OF THE ALGORITHM AND THE MAIN ADVANTAGES
OVER TRADITIONAL SYSTEMS

Action	Advantages over traditional systems
<ul style="list-style-type: none"> Look for specific spots in the 3D scene that might contain pedestrians (even partly occluded ones); in these areas visually search for pedestrians. 	Vision is activated only in selected areas providing a lower number of false detections and faster processing
<ul style="list-style-type: none"> Once a partly occluded pedestrian is detected, an internal alert is issued; if a completely visible pedestrian is detected after an initial alert, a warning is sent to the driver 	Very prompt, tracking already started when the pedestrian was partly occluded and visible only to one of the sensors (vision)

VI. CONCLUSION

This paper has presented a new scheme to increase safety and possibly avoid collisions with vulnerable road users. Instead of implementing a generic pedestrian-detection system, followed by both a validation step and the assessment of the danger level, this work has proposed an innovative approach: Whenever a specific environment structure is detected (i.e., stopped vehicles that may hide pedestrians), the possible presence of pedestrians is checked in specific areas only.

Not only does this solution approach the problem from a different perspective with respect to traditional implementations, but it also focuses on a particularly critical environment, which is typical of urban accidents. In addition to directly providing detections of dangerous situations, it also boosts timing performance since the computationally intensive part, i.e., vision-based pedestrian recognition, is performed only on limited portions of the image.

The system was developed to tackle a very specific yet very common scenario, whereas it does not cover the many other dangerous situations that may occur in an urban environment. The search for all the pedestrians present inside the so-called *driving corridor* can generalize the system yet maintain the scenario-driven approach, as the search is performed only in a corridor dynamically built using laser scanner data. As a conclusion, Table II summarizes the main steps of this system and highlights the specific advantages over traditional pedestrian-detection systems.

REFERENCES

- [1] T. Gandhi and M. M. Trivedi, "Pedestrian protection systems: Issues, survey, and challenges," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 3, pp. 413–430, Sep. 2007.
- [2] M. Szarvas, U. Sakai, and J. Ogata, "Real-time pedestrian detection using LIDAR and convolutional neural networks," in *Proc. IEEE Intell. Vehicles Symp.*, Tokyo, Japan, Jun. 2006, pp. 213–218.
- [3] S. Milch and M. Behrens, "Pedestrian Detection With Radar and Computer Vision." [Online]. Available: http://www.smart-microwave-sensors.de/Pedestrian_Detection.pdf
- [4] A. Broggi, A. Cappalunga, C. Caraffi, S. Cattani, S. Ghidoni, P. Grisleri, P. P. Porta, M. Posterli, P. Zani, and J. Beck, "The passive sensing suite of the TerraMax autonomous vehicle," in *Proc. IEEE Intell. Vehicles Symp.*, Eindhoven, The Netherlands, Jun. 2008, pp. 769–774.
- [5] C. Premebida, G. Monteiro, U. Nunes, and P. Peixoto, "A lidar and vision-based approach for pedestrian and vehicle detection and tracking," in

Proc. IEEE Int. Conf. Intell. Transp. Syst., Seattle, WA, Sep./Oct. 2007, pp. 1044–1049.

- [6] J. P. Hwang, S. E. Cho, K. J. Ryu, S. Park, and E. Kim, "Multi-classifier based LIDAR and camera fusion," in *Proc. IEEE Int. Conf. Intell. Transp. Syst.*, Seattle, WA, Sep./Oct. 2007, pp. 467–472.
- [7] U. Franke and S. Heinrich, "Fast obstacle detection for urban traffic situations," *IEEE Trans. Intell. Transp. Syst.*, vol. 3, no. 3, pp. 173–181, Sep. 2002.
- [8] C. Curio, J. Edelbrunner, T. Kalinke, C. Tzomakas, and W. von Seelen, "Walking pedestrian recognition," *IEEE Trans. Intell. Transp. Syst.*, vol. 1, no. 3, pp. 155–163, Sep. 2000.
- [9] R. A. Retting, S. A. Ferguson, and A. T. McCart, "A review of evidence-based traffic engineering measures designed to reduce pedestrian's motor vehicle crashes," *Amer. J. Public Health*, vol. 93, no. 9, pp. 1456–1463, Sep. 2003.
- [10] P. F. Agran, D. G. Winn, C. L. Anderson, C. Tran, and C. P. Del Valle, "The role of the physical and traffic environment in child pedestrian injuries," *Pediatrics*, vol. 98, no. 6, pp. 1096–1103, Dec. 1996.
- [11] A. Broggi, P. Cerri, S. Ghidoni, P. Grisleri, and H. G. Jung, "Localization and analysis of critical areas in urban scenarios," in *Proc. IEEE Intell. Vehicles Symp.*, Eindhoven, The Netherlands, Jun. 2008, pp. 1074–1079.
- [12] L. Leyrit, C. Chateau, C. Tournayre, and J.-T. Lapresté, "Association of AdaBoost and kernel based machine learning methods for visual pedestrian recognition," in *Proc. IEEE Intell. Vehicles Symp.*, Eindhoven, The Netherlands, Jun. 2008, pp. 67–72.
- [13] P. Geismann and G. Schneider, "A two-staged approach to vision-based pedestrian recognition using Haar and HOG features," in *Proc. IEEE Intell. Vehicles Symp.*, Eindhoven, The Netherlands, Jun. 2008, pp. 554–559.
- [14] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. Int. Conf. Comput. Vis. Pattern Recog.*, Dec. 2001, vol. 1, pp. 511–518.
- [15] G. M. Overett, L. Petersson, N. Brewer, L. Andersson, and N. Pettersson, "A new pedestrian dataset for supervised learning," in *Proc. IEEE Intell. Vehicles Symp.*, Eindhoven, The Netherlands, Jun. 2008, pp. 373–378.
- [16] INRIA, INRIA Person Dataset2005. [Online]. Available: <http://pascal.inrialpes.fr/data/human>



Alberto Broggi (SM'89–S'93–A'96–SM'06) received the Dr.Eng. degree in electronic engineering and the Ph.D. degree in information technology from the Università di Parma, Parma, Italy, in 1990 and 1994, respectively.

He is currently a Full Professor of computer science with the Università di Parma, where he was an Associate Researcher with the Dipartimento di Ingegneria dell'Informazione from 1994 to 1998 and an Associate Professor of artificial intelligence with the Dipartimento di Informatica e Sistemistica from 1998 to 2001. He is the President and Chief Executive Officer of the Artificial Vision and Intelligent Systems Laboratory (VisLab), which is a spin-off company of the Università di Parma, working in the field of signal processing for vehicular applications. He has authored more than 150 refereed publications in international journals, book chapters, and conference proceedings and has delivered invited talks at many international conferences. His research interests include real-time computer vision approaches for the navigation of unmanned vehicles and the development of low-cost computer systems to be used in autonomous agents.

Dr. Broggi was the Editor-in-Chief of the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS from 2004 to 2008. He is a member of the IEEE Intelligent Transportation Systems Society Executive Committee as President Elect, becoming President for the term 2010–2011.



Pietro Cerri (S'05–M'07) received the Dr.Eng. degree in computer engineering from the Università di Pavia, Pavia, Italy, in 2003 and the Ph.D. degree in information technologies from the Università di Parma, Parma, Italy, in 2007.

In 2003, he received a research grant from the Associazione Tecnica dell'Automobile, Italy. He is currently a Temporary Researcher with the Artificial Vision and Intelligent System Laboratory (VisLab), Dipartimento di Ingegneria dell'Informazione, Università di Parma. His research interests are computer vision and sensor fusion approaches for the development of advanced driver-assistance systems.



Stefano Ghidoni (S'04–M'09) received the M.S. degree in telecommunication engineering and the Ph.D. degree in information technology from the Università di Parma, Parma, Italy, in 2004 and 2008, respectively.

During 2004–2007, he worked on artificial vision applied to the automotive field as a Ph.D. student with the Artificial Vision and Intelligent Systems Laboratory (VisLab), Università di Parma. From 2008 to 2009, he worked in the artificial intelligence field at Henesis s.r.l., Parma, which is a spin-off company of the Scuola Superiore S. Anna di Pisa, Pisa, Italy. In March 2009, he joined the Intelligent Autonomous Systems Laboratory, University of Padua, Padua, Italy, as a Temporary Researcher, where he is currently working on intelligent video-surveillance systems and camera networks.



Paolo Grisleri (S'03–M'09) received the Dr.Eng. degree in electronic engineering and the Ph.D. degree in information technology from the Università di Parma, Parma, Italy, in 2002 and 2006, respectively.

He was a Temporary Researcher with the Artificial Vision and Intelligent System Laboratory (VisLab), Dipartimento di Ingegneria dell'Informazione, Università di Parma. He is responsible for software technologies of VisLab, which is a spin-off company of the Università di Parma, working in the field of signal processing for vehicular applications. His research interests include computer vision and software engineering applied to autonomous vehicles.



Ho Gi Jung (M'05) received the B.S., M.S., and Ph.D. degrees in electronic engineering from Yonsei University, Seoul, Korea, in 1995, 1997, and 2008, respectively.

From 1997 to April 2009, he was with the MANDO Corporation Global R&D Headquarters, Yongin, Korea. From 1997 to 2000, he developed environment-recognition algorithms for lane-departure warning systems and adaptive cruise control. From 2000 to 2004, he developed an electronic control unit and embedded software for an electrohydraulic braking system. Since 2003, he has developed environment-recognition algorithms for intelligent parking-assist systems, collision warning and avoidance, and an active pedestrian-protection system. Since May 2009, he has been with Yonsei University as a Senior Research Engineer. His research interests are automotive vision, driver-assistance systems, active safety vehicles, and intelligent surveillance.

Dr. Jung is a member of the International Society of Automotive Engineering; The International Society for Optical Engineers, which is an international society advancing an interdisciplinary approach to the science and application of light; the Institute of Electronics Engineers of Korea; and the Korean Society of Automotive Engineering.