An Experimental Study on Pitch Compensation in Pedestrian-Protection Systems for Collision Avoidance and Mitigation

D. F. Llorca,   M. A. Sotelo,   I. Parra,
J. E. Naranjo, M. Gavilán, and S. Álvarez

Abstract—This paper describes an improved stereovision system for the anticipated detection of car-to-pedestrian accidents. An improvement of the previous versions of the pedestrian-detection system is achieved by compensation of the camera's pitch angle, since it results in higher accuracy in the location of the ground plane and more accurate depth measurements. The system has been mounted on two different prototype cars, and several real collision-avoidance and collision-mitigation experiments have been carried out in private circuits using actors and dummies, which represents one of the main contributions of this paper. Collision avoidance is carried out by means of deceleration strategies whenever the accident is avoidable. Likewise, collision mitigation is accomplished by triggering an active hood system.

Index Terms—Collision avoidance and mitigation, pedestrian protection, pitch compensation, stereovision, virtual disparity image.

I. INTRODUCTION

The range of active safety measures for pedestrian protection is quite wide [1], including ideas like active hood systems, outside airbags, active bumpers, or automatic deceleration [2]. Since these actuators have to be activated just before the crash occurs, sensors such as radar and cameras have to compulsorily be used to provide a measure of the time to collision well in advance. The study of the cumulative frequency of crashes between vehicles and pedestrians [3] shows that a crash speed of up to 40 km/h can cover more than 75% of the total pedestrian injuries. Thus, if a speed of up to 40 km/h is considered, then the levels of injury suffered by pedestrians involved in frontal impacts with motor vehicles will significantly be reduced. Furthermore, some accidents are likely to be avoided, e.g., for velocities well below 40 km/h if pedestrians are detected by the sensors onboard the car with enough anticipation. In such cases, the deployment of deceleration strategies makes sense not only for collision mitigation but for collision avoidance as well.

One of the most popular measures for collision mitigation is the use of the so-called active hood system [4], [5]. This type of system raises the hood of the vehicle in case of unavoidable crash. This way, a more elastic deformation of the hood achieves a reduced force over the pedestrian, particularly over the head. Most of the pedestrians involved in a car-to-pedestrian accident have first contact with the car's frontal region. This usually means that the legs make contact with the front bumper, and after 50–150 ms, the body and particularly the head hit the bonnet or the windscreen of the car, as stated in [6]. For adult leg injuries, the major source is the front bumper of vehicles. When an adult pedestrian is struck by a vehicle, the first impact is generally between the pedestrian knee region and the vehicle's front bumper. Because this initial contact is below the pedestrian's center of gravity, the upper body begins to rotate toward the vehicle. The pedestrian's body linearly accelerates relative to the ground because the pedestrian is being carried along by the vehicle. The second contact is between the upper part of the grille or front edge of the bonnet and the pedestrian's pelvic area. The final phase of the collision involves the head and thorax striking the vehicle with a linear velocity approaching that of the initial striking velocity of the vehicle. Research has shown that the linear head impact velocity is about 90% of the initial contact velocity [3]. Child and adult heads and adults legs are the body regions to be most affected by contact with the front end of vehicles. On vehicles, the bonnet top and the windscreen are the vehicle regions mostly identified with a high potential for contact in a car-to-pedestrian accident. These areas can cover more than 65% of all fatal and serious injuries.

Sensor systems onboard the car are mandatorily required for predicting the car-to-pedestrian distance and the time to collision for both collision avoidance and mitigation. Cameras are the most commonly used sensors for that purpose. A number of remarkable pedestrian-detection systems have been developed by scientists around the world using cameras in the visible spectrum [7]–[11] and infrared cameras [12]–[14]. When using cameras, an accurate estimation of the ground plane inclination, by making use of appropriate pitch angle estimation, is a key aspect in pedestrian-detection systems, particularly in urban environments, for three principal reasons. First, to avoid objects belonging to the ground plane, such as zebra crossing lines, from being mistakenly interpreted as 3-D obstacles due to an excessive (wrongly estimated) negative road inclination. Fake obstacles cause upsetting false alarms to drivers. Second, to assure that no pedestrians are misdetected due to an excessive (wrongly estimated) positive road inclination. Third, to increase the accuracy of car-to-pedestrian distance measurements. This is particularly useful in collision-mitigation applications to precisely calculate the active hood-triggering time.

This paper aims to improve pedestrian detection by compensation of the camera's pitch angle for both collision-avoidance and collision-mitigation applications. To that effect, two pitch compensation methods have been developed and compared. Real experiments have been carried out for collision avoidance and mitigation. Collision avoidance is performed by means of deceleration strategies whenever the accident is avoidable. Likewise, collision mitigation is accomplished by triggering an active hood system. The collision avoidance module has been tested on a Citroen C3 Pluriel car equipped with a stereovision system. Tests were carried out on private circuits with actors. The collision-mitigation module was mounted on a Seat Cordoba car equipped with an active hood system that is triggered by the stereovision system. Real experiments were performed using dummies. Practical results and discussion are provided in this paper.

The rest of this paper is organized as follows. Section II provides a brief summary of the pedestrian-detection system developed in the previous works of the authors. Section III presents current improvements by means of pitch angle compensation. Implementation and experimental results are described in Section IV. Finally, Section V summarizes the conclusions and future work.

II. PEDESTRIAN DETECTION

Pedestrian detection is carried out by using an improved version of the system described in [7]. Nondense 3-D maps are generated using an improved correlation process, which reduces the number of matching errors up to 37% [15]. This is particularly important since most
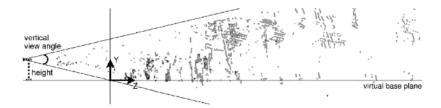
Fig. 1. Three-dimensional projected points on the YOZ plane up to 30 m with respect to the left camera center.
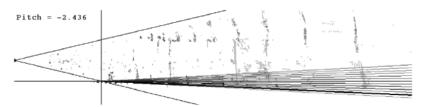

Fig. 2. Pitch angle estimation example. The darker the ray, the higher the number of accumulated points. The bold line represents the estimated pitch angle.

matching errors correspond to points located under the road. Regions of interest are located using a subtractive clustering attention mechanism based on stereovision. This method yields generic obstacles with a 3-D shape that is similar to that of pedestrians. The 2-D candidates are then produced by projecting the 3-D points over the left image and computing their bounding box. A by-component Support Vector Machine-based classifier is then applied using a combination of feature-extraction methods.

Nonetheless, the 2-D bounding box corresponding to a 3-D candidate might not perfectly match the candidate appearance in the image plane, which decreases the performance of the classifier (as demonstrated in [7]) and yields inaccurate pedestrian depth measurements (i.e., inaccurate time-to-collision estimations). To reduce the impact of the "badly bounded" pedestrians, several candidates are generated for each candidate cluster by slightly shifting the original candidate bounding box in the u and v axes in the image plane. The so-called multicandidate (MC) generation strategy increases the detection rate and the accuracy of depth measurements. In addition, pedestrians are detected at larger distances, which increases the time available for anticipated actions [7].

III. PITCH ESTIMATION

The most common way to estimate pitch variations has been carried out by using the so-called v-disparity map [16], where a single dominant line feature is computed by using the Hough transform. This method has successfully been proposed when a dense representation of the scene is available [11], [17]. However, with a nondense representation of the scene, this approach has only been applied over synthetic data [18] or in nonurban environments [16]. To deal with robust pitch estimation from nondense 3-D data points in urban environments, a more complex processing is required. Two different methods have been tested in this paper and are explained in the following sections.

A. Pitch Estimation Based on YOZ Projection Map

Considering that a virtual base plane without pitch variations can be computed beforehand, as well as the camera height relative to the ground and the camera vertical field of view, the origin of the world coordinate system is placed at the intersection point between the virtual base plane and the lower boundary of the vertical field of view. Fig. 1 depicts the lateral projection of 3-D points on the Y- and Z-axes (YOZ) plane. Note that the actual representation of the ground plane will be different with the virtual base plane, particularly when there are pitch variations different from that of the virtual ground plane.

The YOZ map can be observed similar to that of the v-disparity map. If a dense representation of the YOZ map is available, then the use of the Hough transform will be enough to estimate the ground, as in [11] and [17]. However, the use of nondense 3-D maps along with the heavy disparity clutter typical of urban environments suggests the use of a more robust method for pitch estimation. The number of 3-D projected points over the same 2-D point in the lateral view is coded in a gray-scale image, which reduces the weight of matching errors. As in [19], we consider the vertical displacement due to roll negligible in comparison with the displacement due to pitch. From the origin of the world coordinate system, and varying the slope to cover all possible pitch values, uniformly spaced rays are cast. Gray-level values (number of points) along each ray i are counted in a histogram $H(i)$. The histogram is normalized, and the mean value $\bar{h}$ is computed. A stable *jump* over $2/3\bar{h}$ in the histogram is looked for from the bottom of the road upward. Being that $i = 0$ is the lowest ray and $i = N$ the highest ray, the pitch angle is selected as follows:

$$
\begin{aligned}
&\text{for } i = 0 \text{ to } N \\
&\quad \text{if } (H(i), H(i+1), H(i+2) > \frac{2}{3}\bar{h} \\
&\quad \text{and } H(i), H(i+1), H(i+2) > H_{\min}) \\
&\quad \text{then } \alpha = \alpha_i; \text{ break;} \\
&\quad \text{else } \alpha = 0.
\end{aligned}
\tag{1}
$$

Parameter $H_{\min}$ is used to avoid pitch estimation errors when the number of detected road points is not enough. Fig. 2 depicts an example of the pitch estimation. The darker the ray, the higher the number of accumulated points. The bold line represents the estimated pitch angle.

B. Pitch Estimation Using Virtual Disparity Map

An improvement in pitch estimation accuracy is proposed based on the idea suggested in [20]. A rigid 3-D transformation is carried out from the left camera with regard to the ground plane. For this purpose, a rotation around the X-axis and a translation along the Y-axis are performed, as depicted in Fig. 3, where pitch variations are represented with regard to the actual ground plane. The parameters h (camera height) and α (camera pitch angle) obtained in the calibration stage are used in the transformation process.

After changing the reference frame, 3-D points are backprojected on the image plane using the camera intrinsic parameters. This way, the so-called virtual disparity image I_Δ is obtained. The gray level of pixels in virtual disparity images is inversely proportional to their
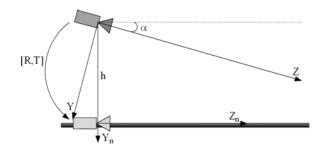
Fig. 3. Rigid transformation. Rotation around the X-axis and translation along the Y-axis to virtually place the camera on the ground plane. Pitch variations are represented with regard to the ground plane.

distance Z with regard to the camera, as depicted in Fig. 4(b). As can be observed, the shorter the distance between the point and the camera, the higher the corresponding pixel gray level. After applying the rigid transformation and backprojecting pixels on the image plane, the resulting virtual disparity image is shown in Fig. 4(c).

The backprojection of all points along the ground plane in the virtual camera must be located on the v coordinate corresponding to the camera central point v_{ol}, assuming that no variation with respect to the calibrated camera pitch angle takes place, as depicted in Fig. 5(a). Points in the virtual camera image plane lying on coordinate v, which are different from v_{0l}, imply a variation in the pitch angle with regard to the calibrated angle. Fig. 5(b) and (c) depicts a change in the ground plane projection due to the positive and negative variations of the pitch angle, respectively. To obtain the projection over the ground plane, a vertical histogram is computed by accounting for the number of points lying on each row in the virtual disparity image, as illustrated in Fig. 4(d). The vertical histogram is smoothed using a 3×3 window. After that, the average value $\bar{h}$ is computed. The proposed algorithm seeks a stable significant *jump* in the vertical histogram starting from the bottom line $i = $ height up to the top line $i = 0$, implementing the following process:

$$\text{for } i = \text{height to } 0$$

$$\text{if } (H_s(i), H_s(i-1), H_s(i-2) > \bar{h}$$

$$\text{and } H_s(i), H_s(i-1), H_s(i-2) > H_{\min})$$

$$\text{then } v_a = i; \text{break}$$

$$\text{else } v_a = v_{0l}. \tag{2}$$

Parameter $H_{\min}$ is again used to set the minimum number of points requested for achieving robust estimation. The computation of pitch angle with regard to the ground plane is carried out in the image plane using the camera intrinsic parameters f, f_{yl}, and d_{yl} (in millimeters per pixel in the Y-axis) and the camera optical center (u_{0l}, v_{0l}), which yields

$$v_a = \frac{y}{d_{yl}} + v_{0l} \Rightarrow \Delta v = v_a - v_{0l}$$

$$= \frac{y}{d_{yl}} \Rightarrow y = \Delta v d_{yl} \tag{3}$$

$$\tan \alpha = \frac{y}{f} = \frac{\Delta v d_{yl}}{f_{yl} d_{yl}} \Rightarrow \alpha = \tan^{-1}\left(\frac{\Delta v}{f_{yl}}\right). \tag{4}$$

This method can easily be extended to estimate the camera roll angle [20].

C. Pitch Correction

To have a steady estimation of the pitch angle, a linear Kalman filter is applied. The state vector contains the pitch angle and its velocity $x_k = \{\alpha_k, \dot{\alpha}_k\}$, whereas the measurement vector contains the pitch angle $z_k = \{\alpha_k\}$. Accordingly, a smoother pitch angle value is obtained. A rotation matrix R_α around the X-axis is then applied to perform 3-D point correction. Once all the 3-D point coordinates are corrected, separation between objects/ground plane is carried out based on a simple height criteria. Given that $P_l' = (X_l', Y_l', Z_l')^T = R_\alpha P_l$, the separation criteria is defined as

$$\begin{cases} \text{if,} & Y_l' \leq H_{\min} \Rightarrow \text{noise} \\ \text{if,} & Y_l' > H_{\min} \text{ AND } Y_l' \leq H_{\text{ground}} \Rightarrow \text{ground} \\ \text{if,} & Y_l' > H_{\text{ground}} \text{ AND } Y_l' \leq H_{\max} \Rightarrow \text{object} \\ \text{if,} & Y_l' > H_{\max} \Rightarrow \text{high object}. \end{cases} \tag{5}$$

Parameter $H_{\min}$ determines the minimum distance under which the detected points are considered to be noise. H_{ground} determines the maximum height value for a 3-D point to be considered as belonging to the ground plane. In practice, this value is set in the range between 10 and 20 cm to reduce noise in the 3-D reconstruction process. Parameter $H_{\max}$ establishes a maximum height value for a 3-D point to be considered as part of a pedestrian. In practice, this parameter has been set to 2.5 m, taking into account that pedestrians may appear over raised crosswalks, speed bumps, sidewalks, etc. By doing so, road points and very high points do not perturb the clustering step.

IV. IMPLEMENTATION AND RESULTS

The system was implemented on two different platforms for testing collision avoidance and collision mitigation, respectively. On one hand, a Pentium IV at 2.4 GHz was mounted on a SEAT Cordoba prototype vehicle equipped with an active hood system. This vehicle was tested on a private circuit using light dummies so that the vehicle could run them down as many times as required without being damaged. On the other hand, a second prototype based on a G4 Power PC was installed on a Citroen C3 Pluriel equipped with automatic steering wheel, brake, and accelerator pedals. This vehicle was tested on a different test circuit, which emulates an urban quarter, where velocity deceleration techniques and avoidance maneuvers could properly be executed. In both cases, the stereovision system used 320×240 pixel images. The complete algorithm runs at an average rate of 15–20 frames/s, depending on the number of pedestrians being tracked and their positions.

A. Pitch Correction Results

To compare the performance between both methods, a ground truth of the pitch variations in a sequence of 1000 frames was built up. A software tool was developed in our lab to allow the users to manually select the most suitable pitch value for each frame. Road and object points are displayed in different colors as a function of the pitch. Then, the user can modify the pitch and visualize the changes until a reasonable solution is achieved. Six different users have run the software three times; therefore, the ground truth is finally defined as the average of 18 different data sets.

Fig. 6 depicts the pitch angle estimation results in the same sequence using the YOZ projection map and virtual disparity image, along with the manually generated ground truth. As can be observed, the results obtained using virtual disparity image are closer to the ground truth in comparison with the results obtained using the YOZ projection map. This is also proved by means of the root mean square error analysis since $\text{RMSE}_{YOZ} = 0.6472$, and $\text{RMSE}_{VD} = 0.3601$.
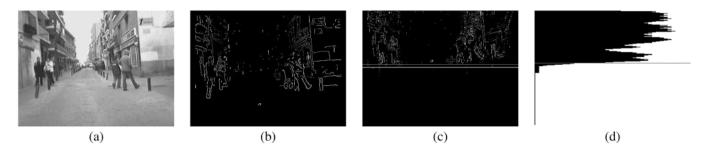
Fig. 4. (a) Original left image. (b) Nondense disparity image I_Δ. (c) Virtual disparity image. (d) Vertical histogram of virtual disparity image.
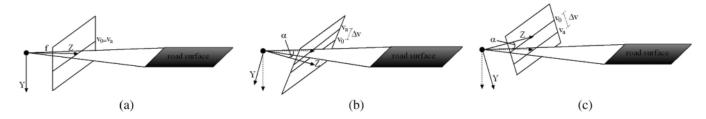

Fig. 5. Ground plane projection on the virtual camera image plane. (a) Without pitch variation. (b) With positive pitch variation. (c) With negative pitch variation.
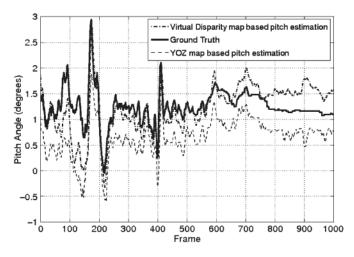

Fig. 6. Pitch angle estimation comparison between the YOZ projection map and virtual disparity image approaches and the ground truth.
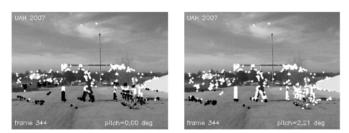

Fig. 7. Separation between ground-plane (black)/objects (white). (Left) Without pitch estimation. (Right) With pitch angle correction. When no pitch compensation is applied, many points belonging to the dummy legs are considered as part of the ground plane. These points will be removed and not considered for selecting candidates.
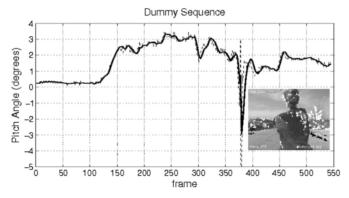

Fig. 8. Pitch estimation in a collision-mitigation experiment. An abrupt change in the camera pitch angle can be appreciated when the car collides with the dummy.

There are other improvements achieved by using the virtual disparity image. The computational cost of pitch estimation processing is reduced ($T_{YOZ} \approx 5$ ms and $T_{VD} \approx 1$ ms). The results are more stable due to the fact that, whereas the method based on the YOZ projection map needs ground points at all distances, the method based on the virtual disparity image works well in situations where only a few points of the road are available. This statement can also be applied when comparing v-disparity map approaches to the virtual disparity image approach. In general, a lower amount of road points is needed to compensate the pitch variations by the virtual disparity image approach in comparison with the YOZ map approach (or even with v-disparity map ones). In spite of that, the system will fail if a minimum number of road points is not available.

Correct estimations of pitch angle have been used for compensating 3-D measurements. As an example, Fig. 7 depicts a sequence in which the prototype car runs over a dummy. Points below $Y < 10$ cm are depicted in black, and points for which $Y \geq 10$ cm are depicted in white. When no pitch compensation is applied, many points belonging to the dummy legs are considered as part of the ground plane. Accordingly, these points will be removed and will not be considered for selecting candidates. In the case of excessive positive road inclinations, the difficulty of selecting candidates and classifying them increases considerably.

The estimated pitch in a collision-mitigation experiment is depicted in Fig. 8, where an abrupt change in the camera pitch angle can be appreciated when the car collides with the dummy.

Correct pitch estimation allows for correcting depth measurements, which implies that a more accurate car-to-pedestrian distance and car-to-collision estimated times can be achieved. Fig. 9(a) shows the
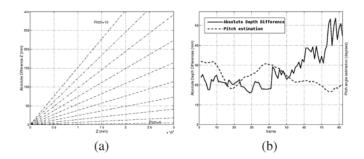
Fig. 9. (a) Theoretical absolute depth difference with/without pitch estimation. (b) Absolute depth difference with/without pitch estimation and pitch estimation in the sequence running over a dummy.

absolute difference between depth measurements performed with and without pitch correction. The figure provides the absolute difference for pitch angles between 1° and 10°. For angles below ±5°, the depth differences are not high (less than 10 cm for depths of up to 25 m). However, for larger pitch values (up to 10°), the errors in depth estimation can be between 15 and 40 cm for depths ranging between 10 and 25 m. In the dummy sequence, the maximum depth correction is below 6 cm, as can be observed in Fig. 9(b).

B. Collision Avoidance

As previously stated, collision-avoidance experiments consist of emergency stop maneuvers by brusquely decelerating the vehicle. Several experiments were conducted to achieve an estimation of the distance that is needed to prevent an accident as a function of the current vehicle speed. Although this computation can easily be done using theoretical equations, we decided to execute real emergency stop maneuvers and derive a realistic estimation of the stopping time at different velocities. After conducting several experiments, where an emergency stop maneuver was executed, an approximate equation for computing the stopping time was estimated using linear regression techniques. The final expression of the estimated stopping time t_{stop}, including actuator latencies, is provided in

$$t_{\text{stop}} = 0.0003 \cdot x^2 + 0.006 \cdot x + 0.5757 \qquad (6)$$

where x stands for the vehicle velocity at the time of starting the emergency stop maneuver. Similarly, the estimated distance required by the vehicle to come to a full stop D_{stop} is provided as

$$D_{\text{stop}} = 0.5 \cdot (0.0003 \cdot x^3 + 0.006 \cdot x^2 + 0.5757 \cdot x). \qquad (7)$$

The system was implemented on a G4 Power PC that was installed on a Citroen C3 Pluriel equipped with automatic steering wheel, brake, and accelerator pedals. The Citroen C3 was tested on several private circuits for live demonstration purposes, including the live demo carried out at the EUROCAST 2007 International Conference at the port of Las Palmas de Gran Canaria, Spain, in February 2007 and the live demo performed in the framework of the PREVENT European Project in Versailles, France, in September 2007. Video files showing the system performance in urban environments, test circuits, and in the aforementioned demos can be retrieved from ftp://www.depeca.uah.es/pub/vision/pedestrian.

C. Collision Mitigation

Collision-mitigation experiments were carried out using a light-weight dummy. The car ran over the dummy in several experiments at different velocities below 50 km/h. In all cases, the active hood had to be activated at a preprogrammed time before the collision took place, which ranges between 200 and 350 ms. The experiments were

TABLE I
TIME-TO-COLLISION ESTIMATION IN THREE DIFFERENT EXPERIMENTS

Impact velocity	Timer activation	Pre-programmed time	Ground truth
25 km/h	373 ms	350 ms	350 ms
34 km/h	308 ms	250 ms	235 ms
40 km/h	290 ms	250 ms	200 ms

recorded using a high-speed camera providing 1000 frames/s. Thus, the ground truth was obtained from recorded videos given that each frame corresponds to 1 ms. With the use of special software applications, digital videos recorded with high-speed cameras can be played frame by frame to precisely determine the exact time between the activation of the active hood and the instant the car-to-dummy collision occurred. This measurement can be done with an accuracy of 1 ms.

Normally, the specifications for hood activation are in the range of 200–350 ms. In previous versions of this paper, the authors achieved an accuracy of 120 ms in the time-to-collision estimation. This figure depends on the vehicle velocity and the number of correlated points for candidate selection. Let us take into account the fact that the exact desired activation time can occur while an image is being processed. The average algorithm execution time is 50 ms. During that time, the system remains blind. Thus, by issuing the activation signal at the end of frame processing, the resulting time-to-collision resolution is very low (50 ms in average) and depends on the processing time of the vision-based algorithm. To avoid this problem, a timer is programmed two or three frames before the activation time is reached. After that, the process is resumed. The timer interrupts the main process when the programmed time expires. At that moment, the vehicle hood is activated, and the vision process is overrun. After applying the pitch-correction method described in this paper, the accuracy in the issuing of the activation signal has been increased to some 50 ms (worst case). Table I shows the results for three collision-mitigation experiments at different velocities ranging between 20 and 50 km/h. The first column represents the vehicle velocity just before the collision. The second column shows the time between the activation of the timer and the real collision. The third column is the desired (preprogrammed) activation time, and the fourth column represents the ground truth provided by the high-speed camera recording the experiment.

V. CONCLUSION AND FUTURE WORK

In this paper, we have presented two pitch-compensation methods based on the YOZ projection map and the so-called virtual disparity image, respectively. Both of them operate over a nondense representation of the environment. The second method, which was based on the virtual disparity image, outperforms the first method; more realistic and steadier results are achieved, less computational time is needed, and a lower amount of road points is required for robust pitch estimation. Two main advantages are achieved by means of pitch compensation. First, the accuracy of the time-to-collision estimation in car-to-pedestrian accidents is increased. Second, the separation between road points and object points is improved, which results in lower false-positive and false-negative detection rates. The stereovision system has been mounted on two different prototype cars and tested on private circuits dealing with collision avoidance applications, by means of deceleration strategy, and collision-mitigation applications, where the time to collision is measured so that the vehicle hood can be activated in different ranges between 200 and 350 ms.

Nonetheless, further improvements are still necessary, particularly in collision-mitigation applications based on active hood, where activation of the vehicle hood must absolutely be precise to maximize the mitigation effect. This can be achieved by reducing the

computation time required to execute the algorithm. Some parts of the code can be implemented on a field-programmable gate array (FPGA)-based platform (i.e., FPGA-PCI cards [21]) programmed to process the correlation of different image regions at the same time and deliver the results to the microprocessor. This will significantly reduce the computational time and will allow for parallelization between the FPGAs and the microprocessor, which will be free to process other tasks during the FPGA's processing time.

REFERENCES

[1] M. M. Meinecke and M. A. Obojski, "Potentials and limitations of pre-crash systems for pedestrian protection," in *Proc. 2nd Int. Workshop Intell. Transp. Syst.*, 2005.

[2] M. M. Meinecke, M. A. Obojski, D. Gavrila, E. Marc, R. Morris, M. Tons, and L. Letellier, "Strategies in terms of vulnerable road user protection," *EU Project SAVE-U—Deliverable D6*, 2003.

[3] "U.T.D.W.F. for harmonization of vehicle regulations," *Pedestrian Safety Global Technical Regulation Preamble*. [Online]. Available: http://www.unece.org

[4] S. VDO, *Basic Technologies for Driver Assistance*. [Online]. Available: http://www.siemensvdo.com

[5] Siemens, *Recent Innovations. Pedestrian Protection*. [Online]. Available: http://www.autoliv.com

[6] K. C. Fuerstenberg, "Pedestrian protection using laserscanners," in *Proc. IEEE Intell. Transp. Syst.*, 2005, pp. 437–442.

[7] I. Parra, D. Fernández, M. A. Sotelo, L. M. Bergasa, P. Revenga, J. Nuevo, M. Ocaña, and M. A. García-Garrido, "Combination of feature extraction methods for SVM pedestrian detection," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 292–307, Jun. 2007.

[8] A. Shashua, Y. Gdalyahu, and G. Hayun, "Pedestrian detection for driving assistance systems: Single-frame classification and system level performance," in *Proc. IEEE Intell. Vehicle Symp.*, 2004, pp. 1–6.

[9] A. Broggi, M. Bertozzi, A. Fascioli, and M. Sechi, "Shape-based pedestrian detection," in *Proc. IEEE Intell. Vehicle Symp.*, 2000, pp. 328–333.

[10] D. M. Gavrila and S. Munder, "Multi-cue pedestrian detection and tracking from a moving vehicle," *Int. J. Comput. Vis.*, vol. 73, no. 1, pp. 41–59, Jun. 2007.

[11] G. Grubb, A. Zelinsky, L. Nilsson, and M. Rilbe, "3d vision sensing for improved pedestrian safety," in *Proc. IEEE Intell. Veh. Symp.*, 2004, pp. 19–24.

[12] F. Xu, X. Liu, and K. Fujimura, "Pedestrian detection and tracking with night vision," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 1, pp. 63–71, Mar. 2005.

[13] B. Fardi, U. Schuenert, and G. Wanielik, "Shape and motion-based pedestrian detection in infrared images," in *Proc. IEEE Intell. Vehicle Symp.*, 2005, pp. 18–23.

[14] M. Bertozzi, A. Broggi, A. Lasagni, and M. D. Rose, "Infrared stereo vision-based pedestrian detection," in *Proc. IEEE Intell. Vehicle Symp.*, 2005, pp. 24–29.

[15] D. Fernández, I. Parra, M. A. Sotelo, P. Revenga, S. Álvarez, and M. Gavilán, "3d candidate selection method for pedestrian detection on non-planar roads," in *Proc. IEEE Intell. Vehicle Symp.*, 2007, pp. 1162–1167.

[16] R. Labayrade, D. Aubert, and J. P. Tarel, "Real time obstacle detection in stereovision on non flat road geometry through v-disparity representation," in *Proc. IEEE Intell. Vehicle Symp.*, 2002, pp. 646–651.

[17] W. van der Mark and D. M. Gavrila, "Real-time dense stereo for intelligent vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 1, pp. 38–50, Mar. 2006.

[18] A. Broggi, A. Fascioli, I. Fedriga, A. Tibaldi, and M. D. Rose, "Stereo-based preprocessing for human shape localization in unstructured environments," in *Proc. IEEE Intell. Vehicle Symp.*, 2003, pp. 410–415.

[19] S. Nedevschi, R. Danescu, D. Frentiu, T. Marita, F. Oniga, P. Ciprian, R. Schmidt, and T. Graf, "High accuracy stereovision approach for obstacle detection on non-planar roads," in *Proc. IEEE INES*, 2004, pp. 211–216.

[20] N. Suganuma and N. Fujiwara, "An obstacle extraction method using virtual disparity image," in *Proc. IEEE Intell. Vehicle Symp.*, 2007, pp. 456–461.

[21] K. Electronic Development Products, *Dragon FPGA-PCI Board*. [Online]. Available: http://www.knjn.com/eu/