

# Vision-based Pedestrian Detection – Reliable Pedestrian Candidate Detection by Combining IPM and a 1D Profile

Guanglin Ma, Su-Birm Park, Stefan Müller-Schneiders, Alexander Ioffe and Anton Kummert

**Abstract**—This article presents the improvement of an Inverse Perspective Mapping (IPM) based obstacle detection algorithm and its utilization in a pedestrian candidate detection module of our vision based pedestrian detection system. A vertical 1D profile of the IPM detection on the region of interest termed Pedestrian Detection Strip (PDS) is created first and the candidates are chosen by applying a threshold on the profile. The usage of the vertical profile increases the robustness of the detection on low contrast images as well as distant pedestrians significantly. A low level pedestrian oriented segmentation and fast symmetry search on the leg region of pedestrians is also presented.

## I. INTRODUCTION

The vision based pedestrian detection on a moving vehicle is one of the most challenging but also interesting and active research topics in the field of active safety. The challenges of the pedestrian detection on a moving vehicle are: movement of the sensor, unpredictable road situations and background scenarios, variations in the size, appearance and pose of the pedestrians.

The vision sensors can be classified into visible light vision sensors and infra-red (far and near infra-red) sensors. Most vision based pedestrian detection activities have been made by using the visible light vision sensors in the last years. Recently more and more approaches with infra-red sensors [1]–[3] have been made with the decreasing cost of the infra-red technology. In [1], an attempt of combining a visible light vision system and an infra-red system (two infrared and daylight stereo systems) is made for pedestrian detection.

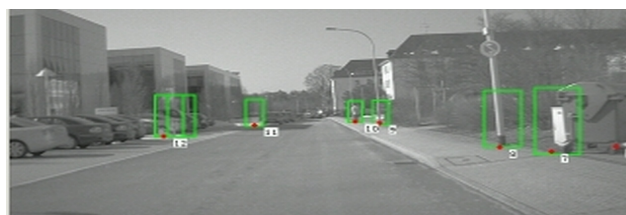
According to the number of sensors in use, there exist two big sets of vision systems. They are the monocular vision systems and the stereo vision systems. Many approaches on pedestrian detection have been presented with the stereo vision systems because of the capability of the easy acquisition of the depth information. In contrast few attempts have been made with monocular systems [4]–[6]. Although the stereo vision system is widely used in the research of pedestrian detection algorithms, however, the single camera approach is preferred by car manufacturers because of its lower cost.

IPM is widely used for lane detection because of its capability to remove the inherent perspective effect from the acquired images [7]–[10]. Few attempts have been made for obstacle detection [8], [11], [12] especially pedestrian detection. The algorithm described here is the utilization of the IPM based obstacle detection and its combination with a 1D profile for a fast and reliable pedestrian candidate detection.

G. Ma and A. Kummert are with the Faculty of Electrical Information and Media Engineering, University of Wuppertal, D-42119 Wuppertal, Germany (e-mail: guanglin.ma, kummert@uni-wuppertal.de).

S.B. Park and S. Müller-Schneiders are with Advanced Engineering, Delphi Delco Electronics Eurpoe, D-42119 Wuppertal, Germany (e-mail: su.birm.park, stefan.mueller-schneiders@delphi.com).

A. Ioffe is with Body & Security Electronics, Delphi Deutschland GmbH, D-51674 Wiehl, Germany (e-mail: alexander.ioffe@delphi.com).



(a) Candidate detection



(b) Candidate classification

Fig. 1: Pedestrian detection in two steps

The remainder of this paper is structured as follows: The next section describes the global structure of our vision based pedestrian detection system and the role of the pedestrian candidate detection and segmentation in the global project framework. Section III presents the improvement of the IPM algorithm and its combination with a 1D profile for the sake of robust detection and computational efficiency. A low level pedestrian segmentation algorithm is described in section IV. The experimental results are described in section V. Finally section VI ends the paper with a brief summarization of the proposed approaches.

## II. SYSTEM STRUCTURE

The detection of pedestrians in our system is done in two sub steps as illustrated in Fig. 1. First, the detection and segmentation module detects potential pedestrian candidates and generates a  $0.9\text{meter} \times 1.8\text{meter}$  bounding box around each candidate (Fig. 1(a)). The feature extraction module [13] extracts several features from each bounding box region. By using the extracted features, the classification module [13] makes its decision on each pedestrian candidate whether it shows a pedestrian or not (Fig. 1(b)). The output of the segmentation module and the classification module is handled by a low-level and a high-level tracking module, respectively. Here, a fast and robust detection of candidates plays a crucial role in a pedestrian detection system to achieve a good overall performance.

## III. OBSTACLE DETECTION BY COMBINING MODIFIED IPM DETECTION AND A 1D PROFILE

### A. IPM based obstacle detection

Fig. 2 illustrates the geometric properties of the camera image plane along with the corresponding transformed plane on the ground. In this paper the 3D world coordinates are



areas. The re-sampling process is a time consuming task, moreover, because of the scaling effect, there is no direct visual impression on the detection result. The detection result must be remapped to the image plane again in order to make a clear connection between the acquired image and the detection result.

In our algorithm, we omitted the re-sampling process on the ground plane and simply transformed the coordinate of each pixel in the image plane of time  $t_1$  to the image plane of the previous time instant  $t_0$ . Afterwards, a new image is built by interpolating pixels of image at time  $t_0$  on the transformed coordinates. The newly built image is the recovering of the image at time  $t_0$  to time  $t_1$  based on the assumption that the image acquired at time  $t_0$  is completely from a flat ground plane (No object above the ground plane). By subtracting the recovered image from the original image at time  $t_1$ , objects above the ground plane are detected. The coordinate transformation of a point at time  $t_1$  to time  $t_0$  can be described in the equation below:

$$P' = H^{-1} \cdot M \cdot H \cdot P \quad (3)$$

Where  $P$  and  $P'$  are the points at time  $t_1$  and time  $t_0$ , respectively, and  $M$  is the matrix for linear transformation of a point on the ground plane from time  $t_1$  to time  $t_0$  with the ego motion of the host vehicle.  $H$  and  $H^{-1}$  is the projection and back projection matrix respectively for a point between the ground plane and the image plane. The detailed mathematical equations are described in [14].

The advantage of the described method over conventional IPM detection is that:

- The sampling process on the ground plane is omitted and consequently the computational cost is reduced.
- Sub-sampling and over-sampling artifacts are avoided.
- Detection is directly done on the image plane instead of the ground plane and it gives intuitive visual impression on the detection result and it is also easier to further process the detection result.

#### D. 1D Vertical profile from the detection result

Often, an object consists of a homogeneous inner region and edges around it. With the IPM we usually get the detection on the edge region where big contrast in pixel intensity occurs. There are few detection on the homogeneous inner region. For the homogeneous inner region, there are the same disparities of pixel positions as the edge pixels. But because the pixel intensity is distributed homogeneously, there are low contrasts of the pixel intensities between two corresponding pixels and consequently there are few detections in that region. Despite of such an effect, in general pedestrians are detected reliably with the conventional IPM based detection algorithm.

In addition, there are also detections caused by the sensor noise which is distributed irregularly on the whole image plane. For obstacles from low contrast appearance (low contrast or lack of texture) or obstacles which have similar gray densities as the background, a reliable detection result is hard to achieve since it is difficult to figure out whether the detection is from real obstacles or just caused by noise. Another weak point of the IPM based detection is the detection on a stationary scene. For the camera with a very small motion, the detection result is also too weak

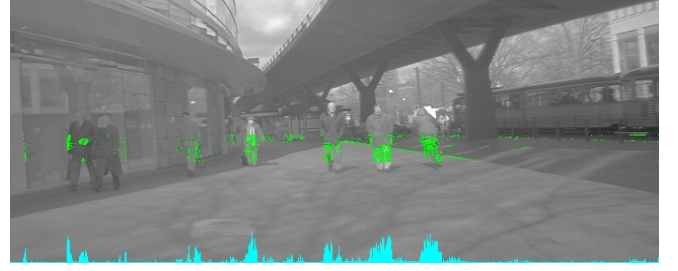


Fig. 4: 1D vertical profile of detection. Green blobs: IPM based detection results. Cyan graph: 1D profile of the IPM detection results.

to distinguish the noise on the detection. In order to get more reliable detection by overcoming the above mentioned weak points, a vertical profile of the binary detection result is generated.

For obstacles from a low contrast images or a very small ego motion, there are much smaller amount of detected blobs compare with normal situations. The detected blobs are distributed irregularly along the edge and there is no clear cue to distinguish the correct detections from both of the irregularly distributed sensor noise and the detections from small patterned areas like bushes etc. A 1D vertical profile of the binary detection result is created in order to distinguish a desired detection of a pedestrian with other unwanted detections (noise, detection on bushes etc). The idea is based on the fact that pedestrians have highly vertically oriented edges compare with their background. By integrating the detection result vertically, the detection on the vertically oriented objects is amplified. The 1D profile  $p$  is created with the following equation:

$$p_u = \sum_v I_{uv} \quad \forall u \quad (4)$$

Here  $u$  is the horizontal coordinate of the PDS and  $v$  is the vertical coordinate of the PDS. Since the image is binarized, the pixel intensity  $I$  either has a value 0 or 255. Fig. 4 illustrates the binary detection result (green blobs) and the corresponding vertical profile (cyan graph).

#### E. Stabilization over ego motion compensation

It has been previously mentioned that our detection algorithm relies on the assumption of an ideal ground plane (i.e. the ground plane is even and horizontal). In case this assumption is not a good model for the actual road (rolling, pitching of the camera, unevenness of the ground plane etc), the detection result will contain unwanted artifacts. In order to compensate for the detection artifacts caused by the pitching of the vehicle, a novel fast 2D image stabilization approach is used. The idea is to combine the stabilization algorithm which is presented in [15] and the IPM algorithm. By doing this, first the known vehicle motion is compensated and the image stabilization process is only performed to compensate the unknown motion of the camera.

### IV. LOW LEVEL PEDESTRIAN CANDIDATE SEGMENTATION

#### A. Foot point search and bounding box generation

The detection result which we get from the PDS is only the horizontal location of the detected objects on a 2D image plane. There is no information on the depth or the size of the detected object available. A segmentation algorithm is



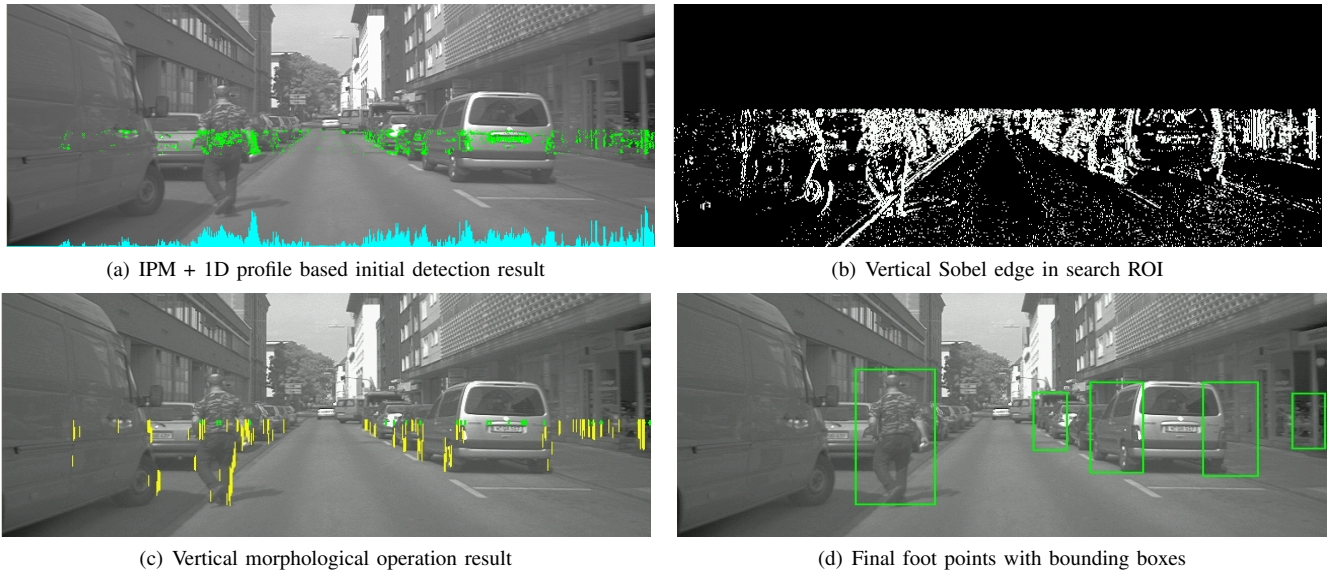


Fig. 5: Illustration of the foot point search.

mandatory to estimate the depth and the size information. In our system, the depth of the detected object is estimated by looking for the foot point of the detected object. First, the foot point of each detected object is searched on the image plane. Next, the foot point found is transformed to the ground plane under the assumption that the ground is flat. Attached to the transformed foot point on the ground plane a 1.8 meter high and 0.9 meter wide bounding box is defined. The defined bounding box is back transformed to the image plane again. This way the final ROI for the pedestrian is defined on the image plane. Although the size of the detected object is fixed to a 1.8 meter high and 0.9 meter wide bounding box which is not true in reality, but the size of the bounding box is defined to cover the most normal adult pedestrian inside of it and it is enough for our classification module to classify the pedestrians. More precise size of the detected pedestrian can be defined through the classification search process. A head point search algorithm is under development in order to acquire more accurate size information of the detected pedestrians.

It is obvious that there are two critical cues for accurate definition of the final bounding box: 1) the accuracy of the foot point search and 2) the pitching of the host vehicle. The pitching problem is compensated roughly by the fast stabilization algorithm described in section III-E. Although the foot point search with region growing described in [14] delivers good results in general, there are two challenging problems which need to be solved: 1) Near distance pedestrians: the closer the pedestrian to the camera, the more details of the pedestrian are captured in the image and the region growing algorithm is failed with too much detailed information. 2) Pedestrian stands on a patterned ground plane like shadow or road boundary (Fig. 5). In this case the region growing continues to go down along the edge of the shadow or the road boundary. Here again, we utilize the highly vertically oriented cue of the pedestrian for a reliable foot point search. A newly developed foot point search is working in the following way as illustrated in Fig. 5.

- 1) Vertical edges are detected on the ROI of the foot point search using Sobel edge detector and the edges are

binarized ( Fig. 5(b)) .

- 2) A vertical morphological operation is performed to eliminate the small vertical edges which are mainly caused by the non-pedestrian objects (Fig. 5(c)).
- 3) For each initial detection (Fig. 5(a)), a ROI of foot point search is defined. The lowest ending point of the morphological results (Fig. 5(c)) inside of the search ROI is defined as the foot point of the detection (red points in Fig. 5(d)). Finally the corresponding bounding boxes are drawn (Fig. 5(d)).

#### B. Symmetry search & bounding box merging

Because the final bounding boxes are generated from the foot points which are searched from the vertical edge of a pedestrian, the foot point search either ends up on the outer edge of the left leg or the outer edge of the right leg. They are not well aligned to the center of the pedestrian (especially for the walking pedestrians). A symmetry search is performed on the binary edge image in order to align the bounding box to the pedestrian.

Instead of performing the symmetry search on the whole pedestrian area, we narrow down the scope of the symmetry search regions to the lower leg region. This is based on the same observation as described in section III-B that normally the background of the lower leg region is a homogeneous ground plane. In contrast, the background of the upper body region can be very complicated and thus can affect the symmetry search a lot. Fig. 6 illustrates the idea. In the figure, the green boxes are our ROIs for symmetry search and the red box is the conventional symmetry search box which contains complicate background inside of it.

Furthermore, the distribution and the orientation of the symmetry density [13] is calculated out in order to filter out both of the false detections on the non-pedestrian area as well as the wrong symmetry search results which are caused by groups of pedestrians walking in parallel (Fig. 7(a)).

An example of our symmetry search result is shown in Fig. 7. It can be seen from the result that both the distant pedestrians walking parallel to the road and the crossing pedestrians in the near field are cropped correctly after

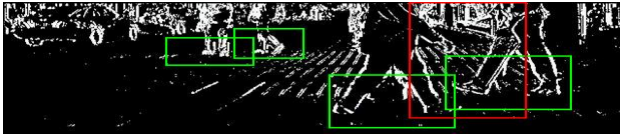


Fig. 6: Symmetry search on leg

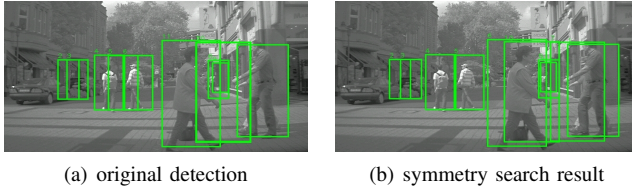


Fig. 7: Symmetry search result

symmetry search. Finally, the overlapped bounding boxes as in Fig. 7(b) are merged into one based on the distance and ratio of the overlapped area.

## V. EXPERIMENTAL RESULTS

Fig. 8 presents some of the detection results of the described algorithm with the comparison of the conventional IPM based detection algorithm. The left column of the figure shows the detected blobs (green blobs) with conventional IPM based algorithm and the 1D profile (cyan graph) of them. The center column shows the final pedestrian candidate detection result with the conventional IPM algorithm. The right column shows the final pedestrian candidate detection result from the presented algorithm. The test sequences were taken from low contrast situations like rainy, foggy or cloudy weather. From the presented examples it can be noticed that the system is able to detect pedestrians reliably even for the situations where the conventional IPM based detection is not capable like distant pedestrians, pedestrians in low contrast image frames or with very small ego motion of the host vehicle.

Fig. 8(a) shows a very cloudy weather condition. The image is very dark and consequently the contrast of the image is also very low. The weakly detected blobs (green blobs on left image) of the IPM detection are eliminated by the morphological operations and the conventional IPM based detection did not detect any object (center image). By creating a vertical 1D profile on the IPM detection, both pedestrians walking along the street and a woman loading goods to the car from a shopping trolley are detected correctly (right image). Two non pedestrians (parking cars one on the left hand side and another one on the right hand side of the driving way) are also detected. Distant pedestrians in Fig. 8(b) are also correctly detected by the presented algorithm.

An example of foggy weather conditions is presented in Fig. 8(c). A group of pedestrians are standing on the left hand side of the street, but because of the low contrast caused by the fog, they are not detected in the center image. After improving the detection by the 1D profile, both the group of pedestrians and a distant pedestrian at the very end of the street are detected. Some non-pedestrians along the roadside are also detected here.

The detection from the stopped car in Fig. 8(d) is mainly due to the ego-motion of the pedestrians and the yaw rate sensor noise (deviation of  $\pm 0.35$  degree in yaw rate).

Detection under heavy rain is shown in the Fig. 8(e). A pedestrian carrying an umbrella is crossing the street during heavy rain. Here again, the pedestrian is detected correctly after applying the 1D profile. A pole on the left hand side of the street and most cars on the right hand side of the street are also detected.

Processing on the images captured by a single monochrome VGA ( $640 \times 480$ ) camera mounted on the windshield of the host vehicle, the described algorithm is capable of processing more than 60 frames per second running on a 3GHz PC and can thus be considered as an extremely fast approach. Note that the processing speed is for the candidate detection algorithm only, i.e. the classification and tracking modules have been disabled during the measurement.

## VI. CONCLUSIONS

We have presented an approach for pedestrian candidate detection by utilizing the inverse perspective mapping and its combination with a 1D profile using a single monochrome camera. One of the common issues on camera detection is robustness against the environment. The conventional IPM based detection also has the difficulty on the bad environment like foggy, cloudy or rainy weather. Our method makes use of a combination of the conventional IPM detection with its vertical 1D profile to improve the detection on the vertical direction and to overcome its limitation on low contrast environment.

Another critical point of camera detection on a moving car is the efficiency on the computation load to make the system run online. Our novel pedestrian detection idea termed PDS not only improves the calculation time by a factor of six compared to the previous attempts but also sets the search region to a relatively simple background zone to make the detection more robust.

Finally, a vertical oriented pedestrian segmentation algorithm which is mainly composed of a foot point search algorithm and a fast symmetry search on the leg region is presented. Experimental results show that the combination of the IPM with a 1D profile outperforms the conventional IPM based detection. Most “false positives” occur on the cars, trees and poles along the roadside and can be easily eliminated through the subsequent classification process [13].

## REFERENCES

- [1] M. Bertozzi, A. Broggi, M. Felisa, G. Vezzoni and M. Del Rose, “Low-level Pedestrian Detection by means of Visible and Far Infrared Tetra-vision”, *IEEE Intelligent Vehicles Symposium 2006*, Tokyo, Japan, June 2006.
- [2] M. Yasuno, S. Ryosuke, N. Yasuda, and M. Aoki, “Pedestrian detection and tracking in far infrared images”, *Proceedings of IEEE Intelligent Transportation Systems 2005*, September 2005.
- [3] F. Xu, X. Liu and K. Fujimura, “Pedestrian detection and tracking with night vision”, *IEEE Transactions on Intelligent Transportation Systems*, March 2005.
- [4] D. M. Gavrila and V. Philomin, “Real-time Object Detection for Smart Vehicles”, *Proceedings of IEEE International Conference on Computer Vision*, pp. 87-93, Kerkyra, Greece, 1999.
- [5] A. Shashua, Y. Gdalyahu, and G. Hayun, “Pedestrian Detection for Driving Assistance Systems: Single-frame Classification and System level Performance.”, *Proceedings of IEEE Intelligent Vehicles Symposium 2004*, Parma, Italy, June 2004.
- [6] K. Wojciechowski, B. Smolka, H. Palus, R.S. Kozera, W. Skarbek and L. Noakes, “Pedestrian Detection using Derived Third-order Symmetry of Legs”, *Proceedings of International Conference on Computer Vision and Graphics (ICCVG) 2004*, Warsaw, Poland, September 2004.



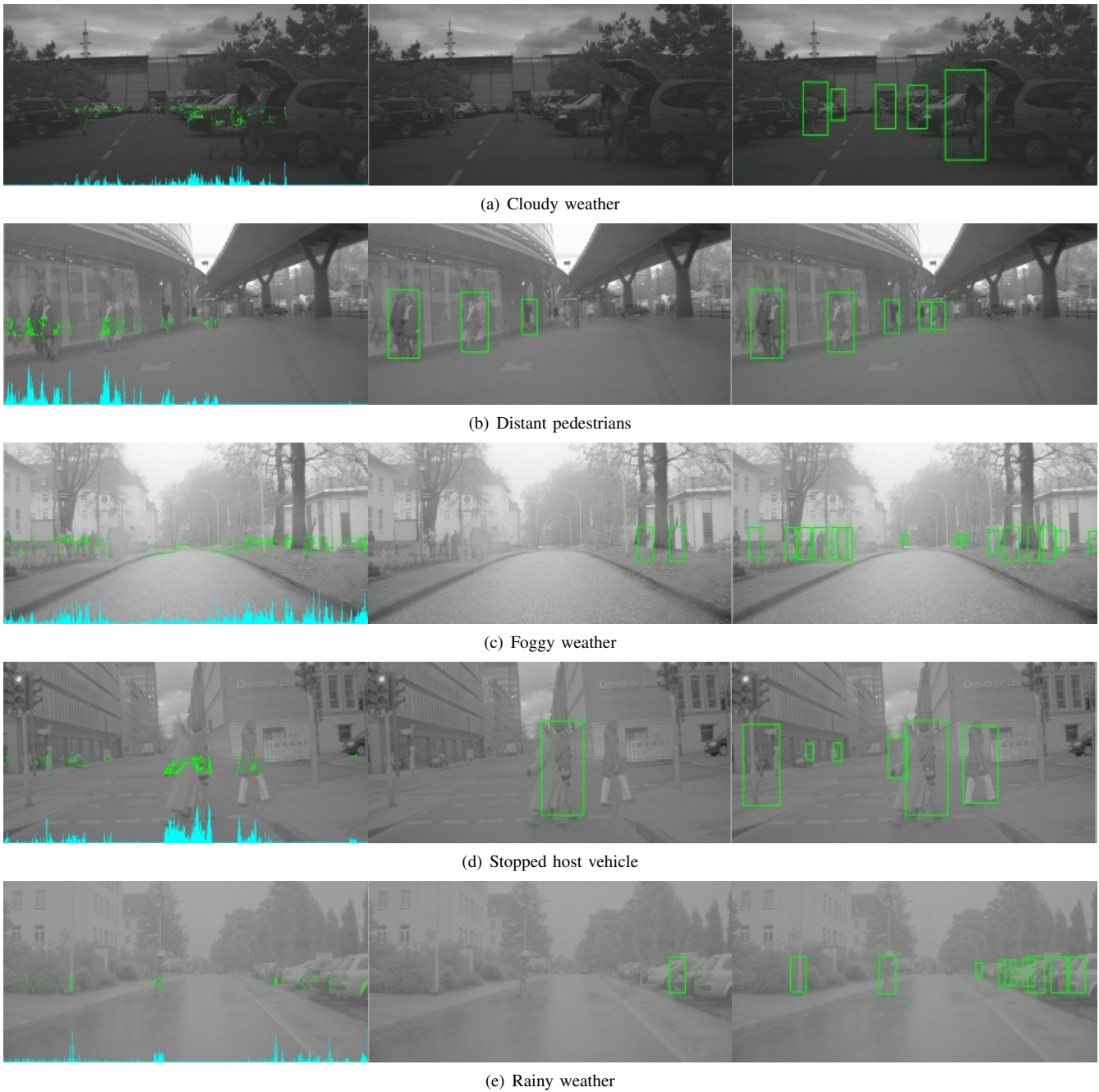


Fig. 8: Experimental results achieved with this system. Left column shows the detected blobs (green blobs) with conventional IPM based algorithm and the 1D profile (cyan graph) of them. Center column shows the final pedestrian candidate detection result with the conventional IPM algorithm. The right column shows the final pedestrian candidate detection result from IPM + 1D profile.

- [7] M. Bertozzi, A. Broggi, and A. Fascioli, "An extension to the inverse Perspective Mapping to handle not-flat roads", *IEEE International Conference on intelligent Vehicles*, 1998.
- [8] G.Y. Jiang, T.Y. Choi, S.K. Hong, J.W. Bae and B.S. Song, "Lane and Obstacle Detection Based on Fast Inverse Perspective Mapping Algorithm", *IEEE International Conference on Systems, Man, and Cybernetics*, October 2000.
- [9] P. Cerri and P. Grisleri, "Free Space Detection on Highways using Time Correlation between Stabilized Sub-pixel precision IPM Images", *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, Barcelona, Spain, April 2005.
- [10] A.M. Muad, A. Hussain, S.A. Samad, M.M. Mustaffa and B.Y. Majlis, "Implementation of Inverse Perspective Mapping Algorithm for the Development of an Automatic Lane Tracking System", *IEEE TENCON 2004*, Thailand, November 2004.
- [11] H.A. Mallot, H.H. Buelthoff, J.J. Little, and S. Bohrer, "Inverse perspective mapping simplifies optical flow computation and obstacle detection", *Biomedical and Life Sciences, Computer Science and Engineering*, Volume 64, Number 3 / January, 1991
- [12] H. Niemann, W. Kasprzka and P. Weierich, "Integrated motion and geometry based obstacle detection in image sequence of traffic scenes", *Proceedings of SPIE* Vol. 2736, p. 228-239, Enhanced and Synthetic Vision, 1996.
- [13] S. Schauland, A. Kummert, S.B. Park, U. Iurgel and Y. Zhang, "Vision-based Pedestrian Detection : Improvement and Verification of Feature Extraction Methods and SVM-based Classification", *Proceedings of the IEEE ITSC 2006*, Toronto, Canada, September 2006.
- [14] R. M. Haralick, "Monocular Vision Using Inverse Perspective Projection Geometry: Analytic Relations", *IEEE Intern. conference on Computer Vision and Pattern Recognition*, pp. 370-378, 1989, San Diego, USA.
- [15] Yeni. A., Ertürk. S., "Fast Digital Image Stabilization using One Bit Transform Based Sub-Image Motion Estimation", *IEEE Transactions on Consumer Electronics*, 51, (3), 917-921, August 2005.