

# Bounding Box Accuracy in Pedestrian Detection for Intelligent Transportation Systems

David Fernández, Ignacio Parra, Miguel Ángel Sotelo, Pedro A. Revenga

Department of Electronics  
University of Alcalá  
Alcalá de Henares, Madrid, Spain  
llorca,parra,sotelo,revenga@depeca.uah.es  
<http://www.depeca.uah.es>

*Abstract*— This paper describes a stereo-vision-based pedestrian detection system for Intelligent Transportation Systems. The basic components of pedestrians are first located in the image and then combined with a SVM-based classifier. Generic obstacles are located using a subtractive clustering attention mechanism based on stereo vision. A by-components learning approach is proposed and different feature extraction methods are tested in order to better deal with pedestrian variability and justify what features are better to be learnt for pedestrian detection. Candidate selection mechanisms usually yield pedestrians with inaccurate bounding boxes. Then a decrease in detection rate takes place if the SVM classifier is trained only with well-fitted pedestrians. Using several off-line databases containing thousands of pedestrians samples the effect of bounding box accuracy is studied. A multi-candidate generation mechanism is also developed in order to enhance the single frame performance, decreasing the number of false positives due to inaccurate bounding boxes.

## I. INTRODUCTION

The most successful human detection systems from a moving vehicle are being accomplished through computer vision as main sensor. Using the same sensor humans use for driving is not a triviality. It provides the main clues for pedestrian detection although other sensors, such as laser-scanners, have also been tested [1]. So far, several approaches have been proposed, but only a few have achieved the challenge of a “global design”: candidate selection, classification and tracking. These ones are the main pieces of the vision-based pedestrian detection systems installed on-board of fast-moving vehicles. This study deal with the link between the candidate selection and classification steps, as a crucial factor for the optimal performance, which has not been studied in the literature.

Various approaches have been proposed as specific classification studies. For example in a recent work [3] a global representation based on gradient orientation histograms is proposed. They use the original MIT pedestrian database [4] adding another 1800 pedestrian images. As in [4], whose monocular candidate selection mechanism releases almost 16,000 windows per frame, their results are only valid in case there were a candidate selection mechanism that generates exactly the same kind of data they are using in their experiments, without the “brute force” procedure, which makes computational cost unacceptable. Otherwise a different behavior would be expected.

The candidate selection can be implemented by performing an object segmentation either in the 3D scene (stereo) or in the 2D image plane (monocular). Only a few authors succeed in dealing with the problem of monocular pedestrian detection, to some extent. In [6] a monocular attention mechanism generates up to 75 windows per frame which are fed to the classifier as potential pedestrians. Their training dataset contains a large amount of samples, whose negatives were generated automatically using their attention mechanism. It means, as they say, that the negatives are not random image fragments. We want to stress that not only the negatives samples should be generated with the attention mechanism, but also the positive samples in order to enhance the single frame performance.

A monocular pedestrian detection system, using vertical symmetries, is proposed in [7]. After a refinement process based on maximizing the symmetry and density of vertical edges, along with a stereo refinement, they fit the bounding boxes by releasing candidates which might be pedestrians [8],[9]. Then a bounding boxes filtering, called classification by other researchers, based on a number of filters, eliminates false positives. In a later work [10] the stereo vision is used for both the bounding box validation and the pedestrian position estimation.

Among the frameworks that use stereo vision for candidate selection we emphasize the next three. In [11] a stereo vision system to generate 3D representation of the scene with disparity maps, is propounded. The candidates are classified as pedestrian or non-pedestrian using a trained neural network. Since this is the first stereo approach in the literature, the segmentation algorithms are very basic. In [12] an obstacle detection procedure is done by using a multiplexed depth map, and selecting regions of interest whose number of depth features exceeds a percentage of the window area. Then they extract edge images and match them to a set of learned examples using chamfer distance [13]. In order to extract information from 3D scene in [14] a segmentation based on v-disparity maps is performed. The information for performing generic obstacles detection is defined with vertical lines. This implies managing very little information to detect obstacles, which may work well for big objects detection, such as vehicles [15], but might not be enough for small, thin objects detection, such as

pedestrian, especially in city traffic due to the heavy disparity clutter.

Infrared cameras have also been tested with both monocular [16][17] and stereo [18][19] vision. These cameras provide better visibility at night but they also provide images with lower resolution, where the appearance of pedestrians is not clear compared to that of day-time images. In addition the use of infrared cameras is quite an expensive option that makes mass production an intractable problem nowadays.

In this work a stereo vision based method for pedestrians detection in the framework of ITS is described. We aim at optimal single frame classification by optimizing the performance between both the candidate selection and the candidate classification methods. Candidate selection mechanisms tend to produce pedestrian candidates that are not exactly similar to the pedestrian examples which were used for training, in the sense that on-line candidates extracted by the attention mechanism might contain some parts of the ground or, on the contrary, might cut pedestrians feet, arms or heads. This results in significant differences between candidates and examples. As a consequence of that, a decrease in detection rate takes place. No authors have noticed and examined this detail.

Two alternatives are proposed in this work. On the one hand the classifier is trained taking into account the possibility of badly bounded pedestrians. This would imply having a candidate selection approach with an effective candidate depth measure, which can be accomplished only by using stereo vision since monocular systems need to apply scale constraints on the depth computation and flat terrain assumptions. On the other hand a multiple candidate selection strategy could be implemented by generating several bounding boxes around each selected candidate in an attempt to compensate for the effect of the candidate bounding box accuracy in the recognition step. For this alternative a low amount of candidates per frame is advisable in order to keep a reasonable candidates per frame ratio.

The rest of the paper is organised as follows: section II provides the system description including both the candidate selection and candidate validation methods. The implementation and comparative results achieved up to date are presented and discussed in section III. Finally, section IV summarizes the conclusions and future work.

## II. SYSTEM DESCRIPTION

This section includes the description of the candidate selection mechanism, the pedestrian detection using SVM and the multi-candidate generation strategy for single frame classification improving.

### A. Candidate Selection

Generic obstacles are detected by using a stereo approach. Firstly the intensities of the left and right images are normalized and the radial and tangential distortions are compensated for. Then, relevant points in the left image are extracted using a Canny algorithm with adaptive thresholds. A non-dense 3D map is then created after solv-

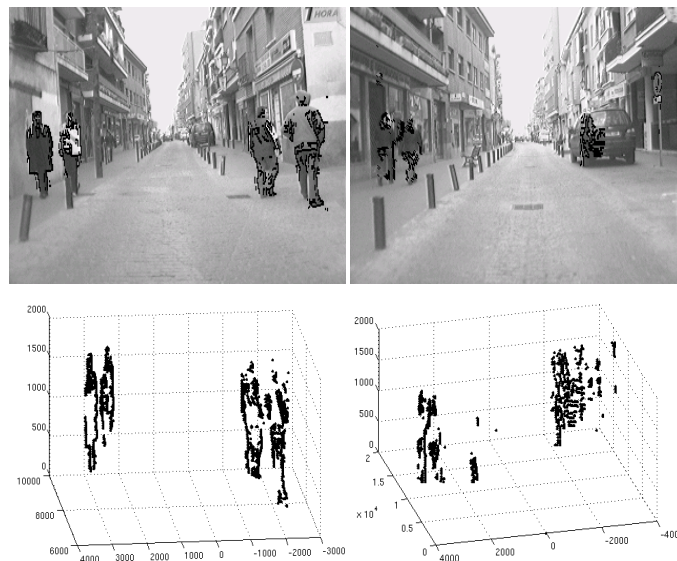


Fig. 1. Upper row: original images, Lower row: 3D maps

ing the correspondence problem. The matching computational cost is further reduced in two ways. First, the matching searching area is greatly decreased by using the epipolar geometry. Second, pixels in the right image are considered for matching only if they are also relevant points. Otherwise, they are discarded and correlations are not computed for that pixel. Among the wide spectrum of matching techniques that can be used to solve the correspondence problem we implemented the *Zero Mean Normalized Cross Correlation* [20]. A post-processing is applied in order to add robustness and reduce noise. The right image correlated points are also correlated over the left image. If the new left matched points are not the same than the original ones these correspondences are considered as noise. In case different left image points would be correlated over the same right image point two strategies could be taken: maximum correlation criterion or minimum disparity criterion. The second one is used so as the noise due to structured backgrounds, which usually produces close 3D points, is avoided. In spite of previous statements the correspondences always yield some outliers, which will be rejected according to a filtering criteria. Firstly, road lane markings are detected [21]. The area contained by the limits of the lanes determines the zone of the real 3D scene where pedestrians are searched for. In case that no lane markings are detected, a basic rectangular ROI is used instead, covering the front part ahead of the ego-vehicle. Secondly, road surface points as well as high points (points with a Y coordinate above 2 m) are removed. Finally, the XZ map (bird's eye view of the 3D scene) is filtered following a neighbourhood criterion. As depicted in figure 1, the appearance of general obstacles in 3D space is represented by an uniformly non-dense distributed set of points.

Data clustering techniques are related to the partitioning of a data set into several groups in such way that the similarity within a group is larger than among groups. In this application the number of clusters is unknown, since

no a priori estimate of the number of general obstacles in scene can be reasonably made. The effects of outliers have to be reduced or completely removed and it is necessary to define specific space characteristics in order to group different pedestrians in the scene. For these reasons, we propose a *Subtractive Clustering* method [22], which is based on a density measure of data points. The idea is to find high density regions in 3D space. The point with highest number of neighbours is selected as the cluster centre. The data points within a prespecified neighbourhood radius are then removed (subtracted) and then the algorithm looks for a new point with the highest number of neighbours. In our approach this algorithm is executed using a 3-dimensional neighbourhood radius  $r_a = (r_{ax}, r_{ay}, r_{az})$ , so that only volumes with a shape similar to a pedestrian will be selected. After applying subtractive clustering to a set of input data, each cluster finally represents a candidate.

### B. Pedestrian Detection Using SVM

The use of Support Vector Machine (SVM) in the field of pedestrian recognition has become a common approach for many researches in the last years since it was first proposed in [5] and [4]. It is well known that SVM provides a method to calculate the hyperplane that optimally separates two high-dimensional classes of objects. Many authors have proved that the by-components approach yields better results than the holistic one [4],[6]. The by components approach suggests the division of the candidate body into several parts, with the aim of breaking the class variability. Thus, each pedestrian body part is independently learnt by a specialized classifier in a first learning stage. The body local parts are then integrated by a second classifier in a second learning stage. By using independent classifiers for each body part the learning process is simplified, due to the fact that the appearance of pedestrians in the scene presents a high intraclass variability. After extensive trials we propose a total of 6 different sub-regions for each candidate region of interest, which has been re-scaled to a size of  $24 \times 72$ . This solution constitutes a trade-off between exhaustive sub-region decomposition and the holistic approach. The optimal location of the six sub-regions, empirically achieved after hundred of trials, has been chosen in an attempt to detect coherent pedestrian features, as depicted in figure 2. A single classifier trained with samples extracted from daytime images is used.

Recognition performance depends crucially on the features that are used to represent pedestrians. It is not obvious to know what features are better to be learnt for pedestrian detection and only a few authors perform a comparative study [3],[25]. In addition, using a by components approach a new problem appears: what features are better for each component?. In a first intuitive approach, some features seem to be more suitable than others for representing certain parts of human body. Thus, limbs are long elements that tend to produce straight lines in the image, while the torso and head are completely different parts, not so easy to recognize. For example edge clues have been used in [13], [14], [11] and [7]. Based on previous works, 6 different fea-

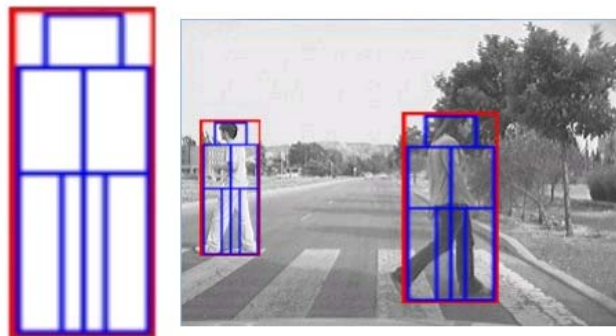


Fig. 2. Decomposition of a candidate ROI into 6 sub-regions.

ture extraction methods are suggested: *Canny image*, *Haar Wavelets*, *Cooccurrence matrices*[5], *Histogram of Oriented Gradients (HON)*[3], *Texture Unit Number (NTU)*[24] and *Gradient Magnitude and Orientation*. For better invariance to illumination and shadowing, all feature vectors are normalized.

### C. Multi-Candidate Generation

As we saw above, the candidate selection method yields generic obstacles with a 3D shape similar to pedestrians. The 2D candidates are selected by projecting the 3D points over the left image and computing the box that bounds these points. Two bounding box boundaries are defined, one for maximum width and height, and one for minimum ones, taking into account people taller than 2 m and kids with a height lower than 1 m. The 3D candidate position is always known thanks to the stereo candidate selection approach (subtractive clustering) whose outputs are the 3D cluster center coordinates, but the 2D bounding box could not be perfectly fixed due to several effects: body parts which are partially occluded or camouflaged with the background, 3D close objects which have been subtracted along with a pedestrian (for example, pedestrians beside traffic signals, trees, cars, etc.), low contrast pedestrians defined with a few number of 3D points, etc. These badly bounded pedestrians will be classified as non pedestrians if the positives samples used to train the classifier are well-fitted. Figure 3 depicts some examples of bad-fitted candidates. Note that this problem also appears with 2D candidate selection mechanisms [6] with the main drawback of losing the actual pedestrian depth.

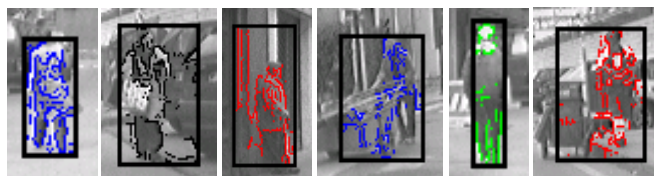


Fig. 3. Some bad-fitted candidates.

Two strategies are proposed to solve the “badly bound-

ing effect” and both of them need to know the 3D pedestrian depth without monocular lacks. The first one consists in training the classifier with also bad-fitted pedestrians in an attempt of absorbing the extra information due to larger bounding boxes and the loss of information due to smaller ones. In other words, training the classifier with the positive samples yielded by the candidate selection method. For that purpose, it is necessary to execute the candidate selection process with an off-line validation to distinguish pedestrians from non-pedestrians. In [6] and [4] the same procedure is only applied for non-pedestrian samples.

The other strategy consist in, using a well-fitted trained classifier, performing a multi-candidate generation for every extracted candidate, trying to hit the target and add redundancy. Three window sizes are defined: the one generated by the candidate selection method, a 20% oversized window and a  $-20\%$  downsized one. These three windows are moved 5 pixels for each direction: top, down, left and right. Thus a total of 15 candidates are generated as we can see in Figure 4.

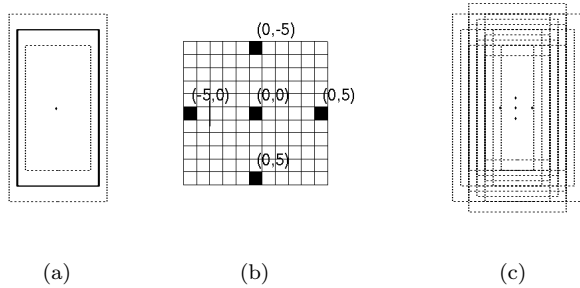


Fig. 4. Multi-candidate generation approach: (a) Oversized and downsized windows, (b) spatial centers for each window, (c) 15 candidates generated

The multi-candidate strategy yields a pedestrian when more than 5 candidates have been classified as pedestrians. This number has been defined after comprehensive experiments. On average the candidate selection mechanism generates 6 windows per frame, which yields a total of 90 candidates per frame, after the multi-candidate process. In case the attention mechanism generated more generic candidates per frame this approach would become impractical.

### III. EXPERIMENTAL RESULTS

The system was implemented on a Pentium IV at 2.4 Ghz, using  $320 \times 240$  pixel images. The complete algorithm runs at an average rate of 15 frames/s, depending on the number of pedestrians being tracked and their positions. The average rate has a strong dependency on the number of pixels being matched because of the correlation computational cost. The candidate selection system has proved to be robust in various illumination conditions, different scenes and distances up to 25 m. Once candidates are validated by the SVM classifier, a Kalman filter tracks the position of each pedestrian in the 3D scene. The probability of a candidate region being classified as pedestrian

is modelled as a Bayesian random variable, whose value is recomputed at each frame as a function of the output provided by the single-frame classifier. Two studies have been carried out following the guidelines described in II-C with the aim of studying and decreasing the influence of inaccurate candidates selection.

#### A. Off-line study

With the purpose of showing the influence of the bounding box accuracy in the global performance, we devised an experiment in which a SVM classifier was off-line trained using a training set of 3000 well fitted or tightly bounded candidates (i.e., the bounding box of candidates fits the real position of the corresponding pedestrians in the image), while a different SVM classifier was off-line trained using a training set containing 2000 badly bounded candidates. Next, the system is evaluated using a test set containing 1000 off-line badly bounded candidates. The results of the experiment are illustrated in Figure 5. Figure 5(a) depicts the performance obtained after testing a set of badly bounded samples using a classifier trained on badly bounded samples. Figure 5(b) shows the performance obtained after testing a set of badly bounded samples using a classifier trained on well fitted ones.

Practical results show that the performance of every feature extraction method has a remarkable decrease when no badly bounded samples are used for the training process. In Figure 5(b) all methods exhibit much worse figures as long as none of the proposed extractors succeeds in providing a detection rate (DR) above 83% (for the case of HON, which is the best performing one) at a false positive rate (FPR) of 5%. Therefore it is clear that every feature extraction method reduce their performance when the selected candidates are not exactly bounded as the candidates used in the training step, which demonstrates that not only the negatives samples should be generated with the attention mechanism, but also the positive samples in order to enhance the single frame performance.

#### B. On-line study

A well fitted database is used to train a SVM model which is evaluated in a set of sequences recorded in real traffic conditions. Two strategies are compared. In the first one, which we call “single method”, candidates obtained from the candidate selection mechanism are directly classified using the best feature extraction method (HON). The second way, which we call “multi-candidate method” (MC) consists in applying the multi-candidate generation for every candidate yielded by the candidate selection method (which results 15 candidates per every one “subtracted”), using also HON as feature extraction method. The results obtained in the experiment are listed in Table I. For each row in the table the following information is provided: type of method, number of detected pedestrians (only pedestrians below 25 m are considered), number of missed pedestrians, and number of false alarms issued by the system.

The analysis of the results reveals that detection rate is improved using multi-candidate generation for the sin-

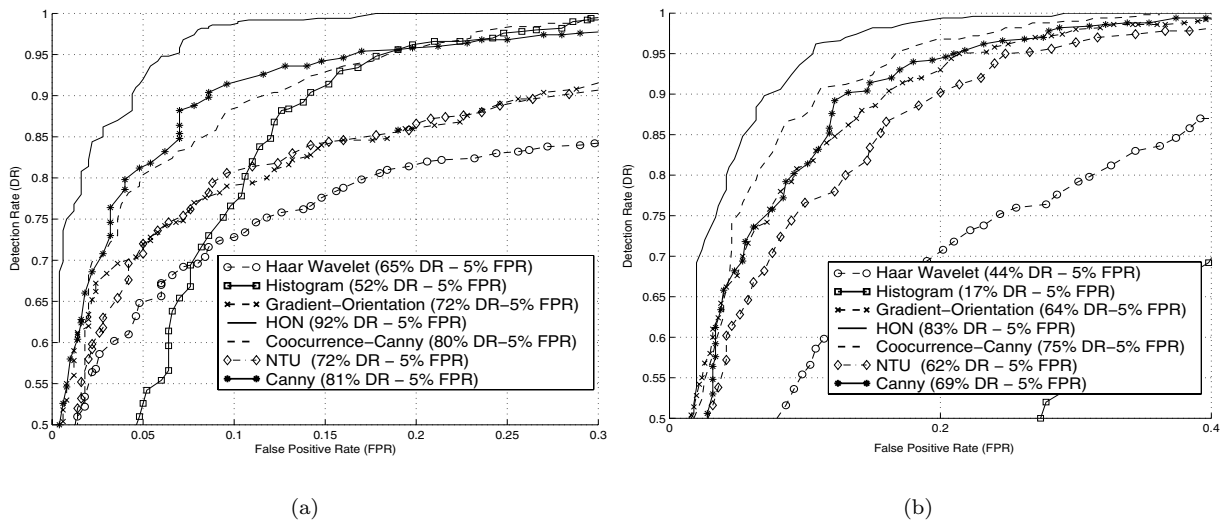


Fig. 5. Off-line Receiver Operating Characteristic (ROC) for bounding box accuracy: (a) Classification of badly bounded samples using training set containing badly bounded samples, (b) Classification of badly bounded samples using training set containing only well-fitted samples

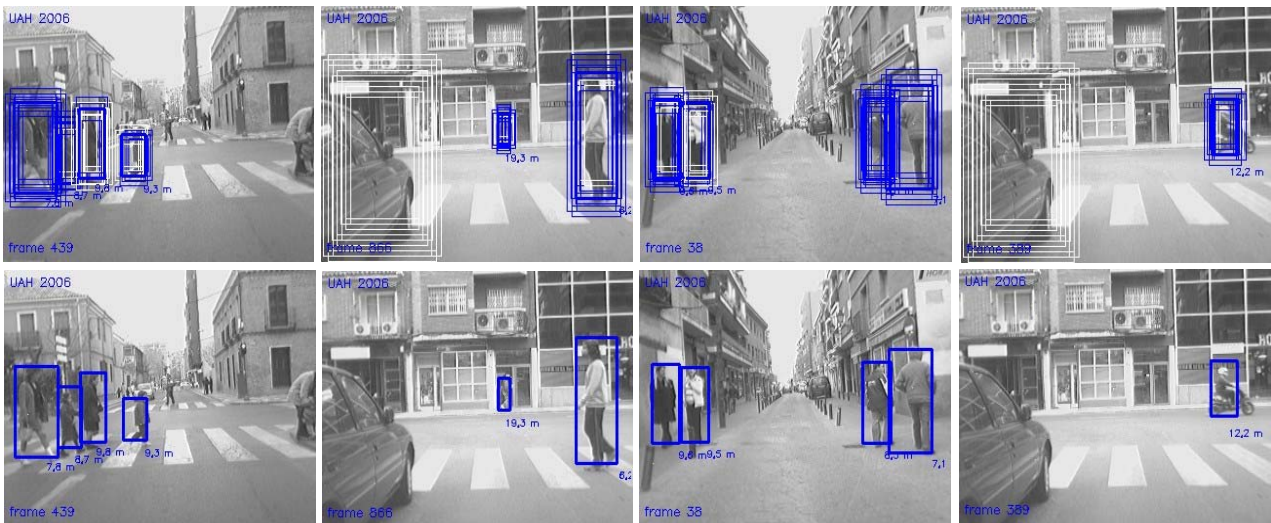


Fig. 6. Upper row: multi-candidate generation. Lower row: results after classifying the 15 candidates

Method	Detected	Missed	False alarms
Single	138	10	9
Multicandidate	143	5	8

TABLE I

GLOBAL PERFORMANCE EVALUATED IN A SET OF SEQUENCES WITH A TOTAL DURATION OF 20 MINUTES

gle frame classification while the false positive rate remains practically unchanged. The single and MC methods exhibit a ratio of 9 and 8 false alarms, respectively, in 20 minutes of operation. This yields ratios of 27 and 24 false alarms per hour, respectively. The MC method has not an appreciable improvement in the FPR because false alarms are usually

due to candidates with features similar to pedestrians, like motorbikes, fences, windows reflections, etc. These candidates are not better classified by using multi-candidate generation since at least 5 out of the 15 candidates continue being classified as pedestrians.

The DR is 93.24% using the single method, and 96.62% using the multi-candidate one. Five non-detected pedestrians by the single method have been detected using the multi-candidate approach. Several cases are corrected by this approach. For example, kids, which are usually selected as candidates with very few points, are better detected by using MC method. When several people are together in the same area the candidate selection method usually yields bounding boxes which fall between two people due to the 3D approach. Thanks to the MC method these pedestrians are well classified. Let us clarify the fact

that the other missed pedestrians were partially occluded or completely out of the vehicle path. There is other important effect due to the use of MC classification: pedestrians are classified at larger distances, that is, before in time, with the improvement of increasing the time used in anticipating actions. Figure 6 depicts typical images from our test sequences. The number below each bounding box represent range. Right image shows a motorcyclist which is detected as a pedestrian (false positive). In the left image two kids are detected with a correct range.

#### IV. CONCLUSION

In this work, we have carried out a global vision-based pedestrian detection system for ITS. Candidates are first selected by an attention mechanism based on stereo vision and subtractive clustering. The stereo approach allows us to know where the pedestrians are without scale constraints and to reduce the average number of false candidates detected per frame. This also helps to deal with the problem of bad fitted pedestrians which could be classified as non-pedestrians, since the candidates range is known and bounding boxes can be translated, oversized and downsized in order to improve the single frame classification, which is not possible using a single camera. In order to reduce the variability of pedestrians, the learning process has been simplified by decomposing selected candidates into 6 local subregions that are easily learned by individual SVM classifiers. Several features extraction methods are tested because it is not obvious what features are better for representing pedestrians. We have carried out a comparative study and selected the extractor with best performance (HON). Single frame detection rate is improved by using multicandidate generation strategy which absorbs the false negatives that have been bad classified by the single strategy. In order to improve also the false positive rate other methods have to be proved like, for example, multiclassification with SVM with different models for motorbikes, cars, etc. Although experimental results show that progress is being done in the right direction, further improvement needs to be made before deploying a really robust vision-based pedestrian detection system for assisted driving in real traffic conditions. For this purpose, motion-based and position dependant features will be incorporated, aiming at enhancing the shape-based pedestrian detection algorithm developed in this work. In order to reduce the correlation computational cost, SIMD code optimization should be carried out.

#### Acknowledgments

This work has been supported by grants DPI2002-04064-C05-04 and DPI2005-07980-C03-02 from the Spanish Ministry of Education and Science and FOM2002-002 from the Spanish Ministry of Public Works.

#### REFERENCES

[1] Fuerstenberg, K.C., Dietmayer, K.J., Willhoeft, V., "Pedestrian recognition in urban traffic using a vehicle based multilayer laser-scanner." *In Proc. IEEE Intelligent Vehicles Symposium*, Versailles, France, June 2002 (2002).

[2] Viola, P., Jones, M. and Snow, D. "Detecting Pedestrians using Patterns of Motion Appearance" *IEEE International Conference on Computer Vision* pp. 734-741, 2003.

[3] Dalal, N. and Triggs, B., "Histograms of Oriented Gradients for Human Detection" *In. Proc. IEEE Computer Vision and Pattern Recognition* Vol. 1, pp. 886-893, 2005.

[4] Mohan, A., Papageorgiou, C. and Poggio, T., "Example-based object detection in images by components" *IEEE Transactions on Pattern Analysis and Machine Intelligence* pp. 349-361, 2001.

[5] Oren, M., Papageorgiou, C., Sinha, P., Osuna, E. and Poggio, T., "Pedestrian Detection Using Wavelet Templates" *In. Proc. IEEE Computer Vision and Pattern Recognition* Puerto Rico, June 1997

[6] Sashua, A., Gdalyahu, Y. and Hayun, G., "Pedestrian detection for driving assistance systems: single frame classification and system level performance" *In Proc. IEEE Intelligent Vehicles Symposium* pp. 1-6, Parma, Italy, 2004.

[7] Broggi, A., Bertozzi, M., Fascioli, A. and Sechi, M., "Shape-based Pedestrian Detection" *In Proc. IEEE Intelligent Vehicles Symposium* Dearbon (MI), USA, October, 2000.

[8] Bertozzi, M., Broggi, A., Fascioli, A. and Lombardi, P., "Vision-based Pedestrian Detection: will Ants Help?" *in Proc. IEEE Intelligent Vehicles Symp*, vol. 1, June 2002, pp. 1-7.

[9] Bertozzi, M., Broggi, A., Chapuis, R., Chausse, F., Fascioli, A. and Tibaldi, A., "Shape-based pedestrian detection and localization" *In Proc. IEEE Intelligent Transportation Systems*, vol.1, pp. 328-333, Oct. 2003.

[10] Lefe, D., Mousset, S., Bensrhair, A. and Bertozzi, M., "Cooperation of Passive Vision Systems in Detection and Tracking of Pedestrians" *In Proc. IEEE Intelligent Vehicles Symposium* Parma, Italy, 2004.

[11] Zhao, L. and Thorpe, C., "Stereo and neural network-based pedestrian detection" *In Proc. IEEE Intelligent Transport Systems* vol 1. No 3, September 2000.

[12] Gavrila, D.M., Giebel, J. and Munder, S., "Vision-Based Pedestrian Detection: The PROTECTOR System" *In Proc. IEEE Intelligent Vehicles Symposium* Parma, Italy, 2004.

[13] Gavrila, D.M. and Philomin, V. "Real-time object detection for smart vehicles" *In Proc. Seventh IEEE International Conference Computer Vision* pp. 87-93 vol.1, Greece, 1999.

[14] Grubb, G., Zelinsky, A., Nilsson, L. and Rilbe, M., "3D Vision Sensing for Improved Pedestrian Safety" *In Proc. IEEE Intelligent Vehicles Symposium* Parma, Italy, 2004.

[15] Labayrade, R., Aubert, D. and Tarel, J.P., "Real Time Obstacle Detection in Stereovision on Non Flat Road Geometry Through V-disparity Representation" *In Proc. IEEE Intelligent Vehicle Symposium* Versailles, France, June 2002.

[16] Bertozzi, M., Broggi, A., Grisleri P., Graf, T. and Meinecke, M., "Pedestrian detection in infrared images" *In Proc. IEEE Intelligent Vehicle Symposium* pp.662 667, Columbus, 2003,

[17] Xu, F., Liu, X. and Fujimura, K., "Pedestrian detection and tracking with night vision" *In IEEE Transactions on Intelligent Transportation Systems* vol 6, No. 1, March 2005.

[18] Liu, X. and Fujimura, K., "Pedestrian detection using stereo night vision" *In IEEE Transactions on Vehicular Technology* vol 53, No. 6, November 2004.

[19] Bertozzi, M., Broggi, A., Lasagni A., and Del Rose, M., "Infrared Stereo Vision-based Pedestrian Detection" *In Proc. IEEE Intelligent Vehicle Symposium* Nevada, USA, June 2005.

[20] Boufama, B., "Reconstruction tridimensionnelle en vision par ordinateur: Cas des cameras non etalonnees" *in PhD thesis*. INP de Grenoble, France 1994.

[21] C. Hernández, "Sistema de asistencia a la conduccion de vehiculos de carretera mediante la deteccion y aviso de la salida de carril" *in master Thesis*. University of Alcal, 2005

[22] Chiu, S., "Fuzzy model identification based on cluster estimation" *in J. of Intelligent and Fuzzy Systems* Vol.2, No 3, pp. 267-278, 1994.

[23] Haralick, R. M., "Statistical and structural approaches to texture" *in Proceedings of the IEEE* Vol. 67, No 5, May, 1979.

[24] He, D and Wang, Li. "Texture Unit, Texture Spectrum, and Texture Analysis" *in IEEE Transactions on Geosciences and Remote Sensing* Vol. 28, No 4, pp. 509-512, 1990.

[25] Parra, I., Fernández, D., Sotelo, M. A., Revenga, P., Bergasa, L. M., Ocaña, M., Nuevo, J., and Flores R. "Stereo Vision-based Pedestrian Recognition for ITS Applications" *WSEAS Transactions on Information Science and Applications* Vol. 3, pp. 554-561, 2006.