A Pedestrian Detector Using Histograms of Oriented Gradients and a Support Vector Machine Classifier

M. Bertozzi, A. Broggi, M. Del Rose, M. Felisa, A. Rakotomamonjy and F. Suard

Abstract— This paper details filtering subsystem for a tetra-vision based pedestrian detection system. The complete system is based on the use of both visible and far infrared cameras; in an initial phase it produces a list of areas of attention in the images which can contain pedestrians. This list is furtherly refined using symmetry-based assumptions. Then, this results is fed to a number of independent validators that evaluate the presence of human shapes inside the areas of attention.

Histogram of oriented gradients and Support Vector Machines are used as a filter and demonstrated to be able to successfully classify up to 91% of pedestrians in the areas of attention.

I. Introduction

A reliable pedestrians detection system is interesting for a number of applications, such as video surveillance or automatic driver-assistance systems for vehicles; moreover, the heavy investments made by almost all car manufacturers on this kind of research proves that particular attention is now focused on improving road safety, especially for reducing the high number of pedestrians being injured every year.

Also military organizations, like the U.S. Army, are actively involved on this research topic for robotic applications like obstacle detection, automatic path following, or anti-tamper surveillance [13], [17]. For military applications, the use of vision as a primary sensor for the detection of human shapes is widely exploited, since cameras do not emit signals that can be detected by third parties.

Unfortunately, vision-based pedestrian detection, especially when performed from a moving vehicle, is a difficult task. Pedestrians are *deformable* obstacles and can therefore assume several different poses and wear different clothes with different colors changing their appearance. Moreover, the automotive environment is often barely unstructured, incredibly variable, and apparently moving, due to the fact that the camera itself is in motion; therefore, really few assumptions can be made on the scene.

This paper describes a tetra-vision based pedestrian detection system to be installed on an autonomous military vehicle.

This system is able to detect all obstacles appearing in the scene and is based on the simultaneous use of two stereo camera systems: two far infrared cameras and two daylight

cameras [6]. The first stages of such system provide a reliable detection of image areas that potentially contain pedestrians. Unfortunately, some of the detected areas do not contain pedestrian or contain more than one pedestrian; therefore a refinement and a preliminary filtering of these areas is mandatory. This refinement and filtering stage is based on the assumption that a pedestrian features a high symmetry and edge density content. The process is only performed on far-infrared images in order to reduce the impact of the background on the whole process. Following filters are used to furtherly analyze the list of refined areas of attention in a multi-vote fashion. A final validation stage inputs all the votes coming from the different filters and evaluates how much the obstacle is likely to be a pedestrian.

This paper details the subsystem developed for extracting the object shape, in order to provide a vote based on it. However, the evaluation of the object shape is currently under development, therefore only results on shape extraction are given.

When a shape is extracted, a function has been implemented to detect the presence of a pedestrian. To analyze the shape, it extracts the corresponding image and describe it thanks to histograms of oriented gradient. This method has been introduced by [11] and aims at splitting the image into small regions and computing histograms of oriented gradients for each region and merging all histograms to describe the image. The idea consists in describing the shape of the object thanks to the information of gradient orientation. This characterization is finally used with a linear Support Vector Machine classifier [32], which is the state of the art within the classification methods and has proved to be very efficient in such case of high dimensional data.

This paper is organized as follows: section II describes related work in pedestrian detection systems based on artificial vision. The pedestrian detection system is discussed in III and the submodule for the analysis of the areas of attention is detailed in section IV. Results are discussed in paragraph V and, finally, section VI ends the paper with some final remarks.

II. Related work

For the U.S. Army the use of vision as a primary sensor for the detection of human shapes is a natural choice since cameras are non-evasive sensors and therefore do not emit signals.

Vision-based systems for pedestrian detection have been developed exploiting different approaches; like the use of monocular [19], [25] or stereo [26], [29] vision. Daylight
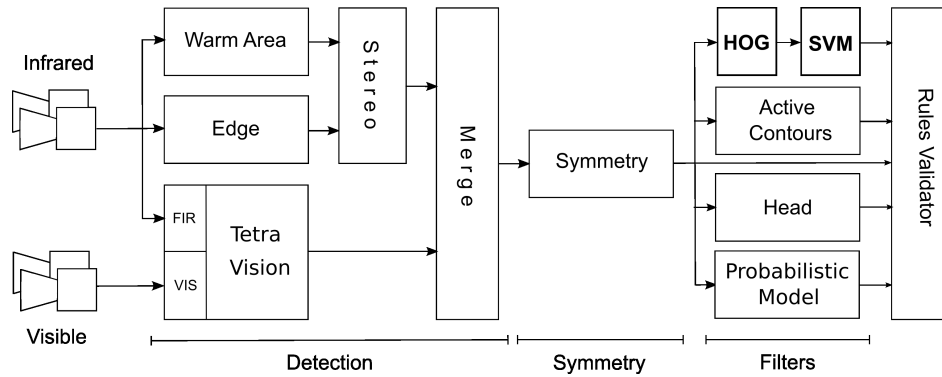
Fig. 1. Overall algorithm flow. This paper describes the Histogram of Oriented Gradients (HOG) and Support Vector Machines (SVM) stages.

cameras are widely used for pedestrian detection, but, in the last years, thanks to the decreasing costs and to the improved technology, also infrared devices have been employed for vision based pedestrian detection [2], [12].

Many systems based on the use of a stationary camera employ simple segmentation techniques to obtain foreground region; but this approach fails when the pedestrians must be detected from moving platforms. Most of the current approaches for pedestrian detection using moving cameras treats the problem as a recognition task: a foreground detection is followed by a recognition step to verify the presence of a pedestrian. Some systems use motion detection [10] or stereo analysis [27], [29] as a means of segmentation.

Other systems, substitute the segmentation step with a focus-of-attention approach, where salient regions in feature maps are considered as candidates for pedestrians. recognition approach is used to validate the presence of pedestrians. In the GOLD system [6] a DSI-based approach is used to detect potential pedestrians and the recognition is obtained by a number of independent validators. A similar approach is used in [15] where region of interests are preliminary selected using a stereo vision approach and following shape-based and texture-based stages are used for the validation. In [9] the local image entropy directs the focus-of-attention and is followed by a model-matching module.

The most critical problem is to quickly and reliably determine whether a candidate region contains a pedestrian. The computer vision approaches dealing with the detection and recognition of pedestrians can be divided into three categories: model-based methods that use models set of the human body, holistic methods which model the appearance or shape of the entire body, and motion-based recognition methods which use a model of the movements of the whole body or individual parts of it. For examples, [9], [10] use the periodicity of human gait or gait patterns for pedestrian detection. The model-based approaches seems to be more reliable than shape based ones, but they require temporal information and are unable to correctly classify pedestrians that are still or have an unusual gait pattern.

Conversely, shape-based approaches exploit the pedestrians' appearance, so can detect both moving and stationary people [19], [23]. In model-based approach, the challenge is

to model the several variations in the shapes, pose, size and appearance of humans and their backgrounds. Basic shape analysis methods consist in matching a template onto candidate foreground regions. In [14] a tree-based hierarchy of human silhouettes is constructed and the matching follows a coarse-to-fine approach. In [22], [28] probabilistic templates are used to take in account the possible variations in human shape.

The method proposed in this paper falls in the category of holistic approaches. The advantage of this choice is its general applicability; it learns the shape model from examples, and hence is not specific to humans and can be readily applied to the recognition of other classes of objects. Learning based classifiers, such as Neural Network [21], [30], [31] and Support Vector Machine (SVM) [1], [7], [20], [32]–[34], have mostly been used for the validation. As an example, [18] uses a neural net to discriminate pedestrians based on the disparity map obtained from a stereo pair. In [16] a vertical and horizontal 3×3 Sobel edge detector is used to extract pedestrian shape features and to train two SVM-based classifiers for detecting pedestrians in front/rear pose respectively. In [34] SVM-based classifier is trained using grayscale features in a infrared image.

III. SYSTEM DESCRIPTION

The active contours-based algorithm described in this work has been developed as a part of a tetra-vision based pedestrian system [3], [6]. The whole architecture is based on the simultaneous use of two far infrared and two daylight cameras. Thanks to this approach, the system is able to detect obstacles and pedestrians when the use of infrared devices is

Fig. 2. Examples of typical scenarios in FIR and visible images.

more appropriate (night, low illumination conditions...) or, conversely, in the case visible cameras are more suitable for the detection (hot, sunny environments...).

In fact, FIR images convey a type of information that is very different from those in the visible spectrum. In the infrared domain the image of an object depends on the amount of heat it emits, namely, it is generally related to its temperature (see fig. 2). Conversely, in the visible domain, objects appearance depends on how the surface of the object reflects the incident light as well as on the illumination conditions.

Since humans usually emit more heat than other objects like trees, background, or road artifacts, the thermal shape can be often successfully exploited for pedestrian detection. In such cases, pedestrians are in fact brighter than the background. Unfortunately, other road participants or artifacts emit heat as well (cars, heated buildings...). Moreover, infrared images are blurred, have a poor resolution and the contrast is low compared with rich and colorful visible images.

Consequently, both visible and far infrared images are used for reducing the search space.

Figure 1 depicts the overall algorithm flow for the complete pedestrian system. Different approaches have been developed for the initial detection in the two image domains: warm areas detection, vertical edges detection, and an approach based on the simultaneous computation of disparity space images in the two domains [3], [6].

These first stages of detection output a list of areas of attention in which pedestrians can be potentially detected. Each area of attention is labelled using a bounding box. A symmetry-based approach is furtherly used to refine this rough result in order to resize bounding boxes or to separate bounding boxes that can contain more pedestrians [5].

These two steps in the processing, barely take in account specific features of pedestrians; in fact, only symmetrical and size considerations are used to compute the list of bounding boxes. Therefore, independent validators are used to evaluate the presence of human shapes inside the bounding boxes. These stages exploit specific pedestrian characteristics to discard false positives from the list of bounding boxes.

The algorithm described in this paper is used as one of the independent validators as shown in figure 1. Currently, a final rules validator is under development, in order to take into account the behavior and characteristics of each independent validator.

IV. BOUNDING BOXES ANALYSIS

In this part, we will describe in details the method to classify the content of bounding boxes. As we said in previous section, the result of stereovision process is a list of bounding boxes. Each bounding box produces one image with an object.

First we will describe the image with a descriptor which extract the discriminant information contained in this image.

Then we analyze the descriptor with a classifier, in our case the Support Vector Machines.

A. HoG descriptor

In 2005, Shashua et al. [24] had already introduced a descriptor using the information of gradient orientation. But he proposed to computed local histograms within small regions corresponding to the human morphology. Dalal et al. [11] has extended the use of histograms but with a dense approach.

The computation of a descriptor is done according the following steps:

1) compute horizontal G_H and vertical G_V gradient of image by filtering image with $[-1\ 0\ 1]$
2) compute both norm and orientation of the gradient:
   - $N_G(x,y) = \sqrt{G_H(x,y)^2 + G_V(x,y)^2}$ (figure 3-b)
   - $O_G(x,y) = atan\left(\frac{G_H(x,y)}{G_V(x,y)}\right)$ (figure 3-c)
3) split image into cells (figure 3-d),
4) compute one histogram for each cell (figure 3-e),
5) normalize all histograms within a block of cell.

The last step is a specificity of this descriptor, since the normalization can reduce the illumination variability. The final descriptor is obtained by grouping all normalized histograms into a single vector.

B. SVM classifier

The recognition system is based on a supervised learning technique. Hence, we have used a set of training image examples with and without pedestrians, and described by their HoG, to learn a decision function. In our case, we have used a Support Vector Machines classifier.

The Support Vector Machines classifier is a binary classifier algorithm that looks for an optimal hyperplane as a decision function in a high-dimensional space [4], [8], [32]. Thus, consider one has a training data set $\{\mathbf{x}_k, y_k\} \in \mathcal{X} \times \{-1, 1\}$ where $\mathbf{x}_k$ are the training examples HOG feature vector and y_k the class label. At first, the method consists in mapping $\mathbf{x}_k$ in a high dimensional space owing to a function Φ. Then, it looks for a decision function of the form: $f(\mathbf{x}) = \mathbf{w} \cdot \Phi(\mathbf{x}) + b$ and $f(\mathbf{x})$ is optimal in the sense that it maximizes the distance between the nearest point $\Phi(\mathbf{x}_i)$ and the hyperplane. The class label of $\mathbf{x}$ is then obtained by considering the sign of $f(\mathbf{x})$. This optimization problem can be turned, in the case of L_1 soft-margin SVM classifier (misclassified examples are linearly penalized), in this following way:

$$\min_{\mathbf{W},\xi} \frac{1}{2}\|\mathbf{w}\|^2 + C\sum_{k=1}^{m} \xi_k \qquad (1)$$

under the constraint $\forall k, \quad y_k f(\mathbf{x}_k) \geq 1 - \xi_k$. The solution of this problem is obtained using the Lagrangian theory and it is possible to show that the vector $\mathbf{w}$ is of the form:

$$\mathbf{w} = \sum_{k=1}^{m} \alpha_k^* y_k \Phi(\mathbf{x}_k) \qquad (2)$$
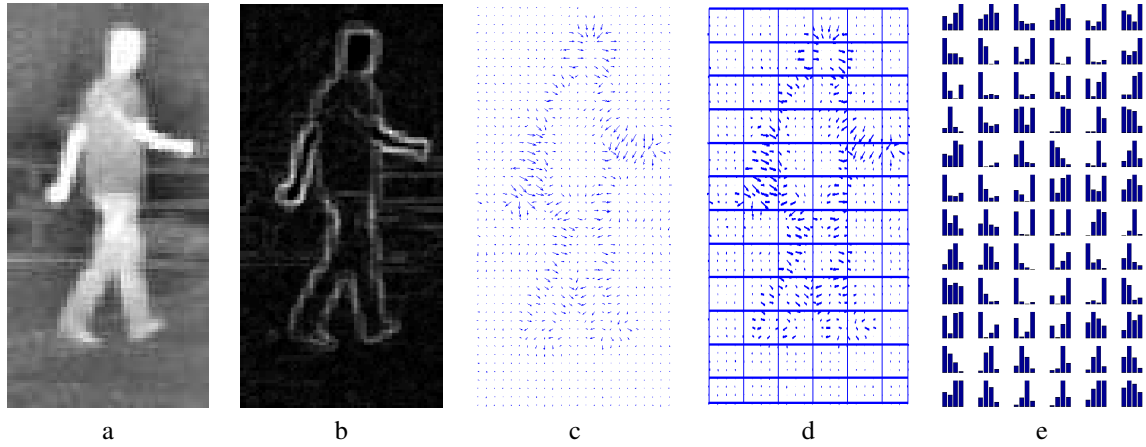
Fig. 3. Image characterization using HoG: original image (a), gradient norm (b), gradient orientation (c), cell splitting (d) and histogram computation (e).

where α_i^* is the solution of the following quadratic optimization problem:

$$\max_{\alpha} W(\alpha) = \sum_{k=1}^{m} \alpha_k - \frac{1}{2} \sum_{k,\ell}^{m} \alpha_k \alpha_\ell y_k y_\ell K(\mathbf{x}_k, \mathbf{x}_\ell) \quad (3)$$

subject to $\sum_{k=1}^{m} y_k \alpha_k = 0$ and $\forall k, 0 \leq \alpha_k \leq C$, where $K(\mathbf{x}_k, \mathbf{x}_\ell) = \langle \Phi(\mathbf{x}_k), \Phi(\mathbf{x}_\ell) \rangle$. According to equation (2) and (3), the solution of the SVM problem depends only on the Gram matrix K.

V. RESULTS

In this last part, we will present some results of our system. We evaluated the HoG method with 2 video sequences. The first: *Tetravision05*, was taken during day with good luminosity conditions. The second: *Tetranight01*, was taken during night. For each sequence, we used both a visible and infrared stereovision system.

Thanks to the tetravision system which was described previously, a list of bounding boxes can be extracted from all sequences. For each bounding box, we extracted the corresponding image and labeled it manually as pedestrian or non-pedestrian. An image is labeled as a pedestrian if it contains only one person, which is centered. The size of the pedestrian should also be the size of the bounding box. Figure 4 shows some examples of pedestrians and non-pedestrians. We can note that we can have the same pedestrians images, since we are dealing with video sequences, but under various conditions and postures.

	Tetravision05		Tetranight01	
	FIR	VIS	FIR	VIS
pedestrian	2255	1860	1678	1359
non-pedestrian	20246	20520	2933	3262

All images are then resized to the same size : 128×64 pixels. This operation is due to the fact that the descriptor works on images with the same size. This particular size can be explain that it is a multiple of 2, so that we can try different combinations of the descriptor, concerning the cells and the blocks.

As we can see, this descriptor involves different kind of parameters, like the number of cells, the number of blocks, the vote for histogram.

To evaluate the most efficient set of parameters, we have set up a complete test, where we tried different value for each parameter. This test has been realised with 4400 infrared images : 2200 pedestrians, and 2200 non-pedestrians. These images are obtained by selecting manually in original images different boxes contained a pedestrian or any kind of object.

We tested a large variety set of parameters :

- Size of cell : 4×4, 8×8 or 16×16 pixels,
- size of block : 1×1, 2×2 or 4×4 cells,
- overlap of block : 1, 2,
- number of bins for histogram : 4, 8 or 16,
- vote method for histogram : weigthed with gradient magnitude or no,
- normalization factor for block : no, L1 or L2,

and we retained the optimal set of parameters :

- size of block=2,
- number of bins = 8,
- size of cell = 8,
- overlap of blocks = 1,
- adding values in histogram = normalized,
- normalization factor for block = L2.

Concerning the parameters of the SVM classifier, we use a linear kernel, so that we had no parameter to deal with. The only parameter is the weight for misclassified point C, which has been fixed at 1000, after having compared various values.

We evaluated independently each category: tetravision visible, tetravision infrared, tetranight visible and tetranight
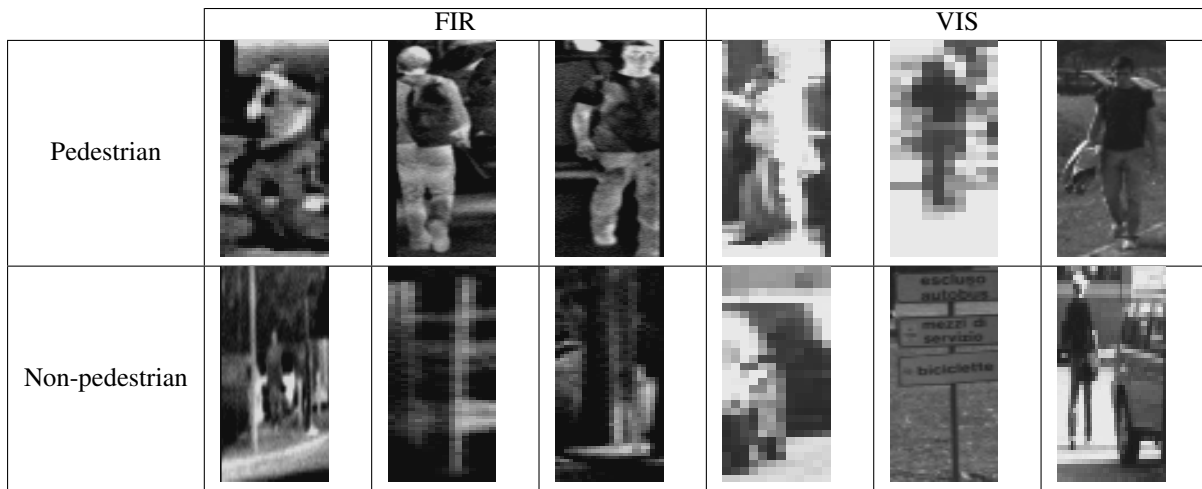
	FIR	VIS
Pedestrian		
Non-pedestrian		

Fig. 4.   Examples of pedestrian and non pedestrian images extracted for the tetravision sequence.

AUC		10	50	100	500
Tetranight	FIR	0.9554	0.9602	0.9662	0.9704
	VIS	0.9364	0.9447	0.9523	0.9550
Tetravision	FIR	0.7304	0.8374	0.8622	0.8935
	VIS	0.7416	0.8460	0.8618	0.8977

TABLE I

VALUE OF AREA UNDER CURVE FOR EACH SEQUENCE, WHEN THE SIZE

OF THE LEARNING SET VARIES.

Recognition rate		10	50	100	500
Tetranight	FIR	0.8727	0.8993	0.9060	0.9144
	VIS	0.8546	0.8832	0.8893	0.8999
Tetravision	FIR	0.6583	0.7555	0.7784	0.8121
	VIS	0.6702	0.7603	0.7822	0.8165

TABLE II

RECOGNITION RATE FOR EACH SEQUENCE.

infrared. We extracted randomly a subset of images to build a learning set. We learn the linear Support Vector Machine classifier on this subset and test on a random subset of 500 images of pedestrian and 500 non-pedestrian. It should be noticed that images of the test dataset are not images used during the learning step.

To evaluate the performance of our system, we compute the rate of true positives against the rate of false positives, and we compute the area under the curve (AUC) which is obtained. The table I shows our results. We also determined the good recognition rate on the table II. We also evaluate the performance when the size of the learning set varies.

To improve the reliability of these results, each test was iterated 10 times, and we renewed randomly the learning set and test set for each iteration. The given results are the average of all tests.

We can note that we obtained very good results, particularly during the night (Tetranight) with up to 91% of good recognition rate for the infrared night sequence. We can

explain this by the fact that with the infrared images the pedestrian is warmer than the background. Concerning the visible images, all pedestrian detected are pedestrian located in front of the car, so images are well defined. Moreover, due to the headlight, the pedestrian could be easy detached from the background.

During the day, shapes are less contrasted, so that the characterization is not facilitated. This fact could explain the lower performance achieved during day.

If we consider the generalization capacity of this method, the performance is quite optimal for a learning set containing only 50 pedestrian and 50 non-pedestrian. When the the learning set becomes larger, the system performs better.

VI. CONCLUSION

This paper presented a subsystem for the recognition stage of pedestrian detection system. This module is based on use of histogram of oriented gradients combined with Support Vector Machines classifier and works on both infrared and daylight images.

Results demonstrated that this approach is promising; in fact, the system is able to correctly evaluate up to 91% of pedestrians in the nocturnal sequence and up to 80% in daylight scenes.

Different ideas are currently under evaluation for improving the result. Particularly, different kind of representation to characterize the image by means of color or texture can be combined. Also the comparison of the descriptors can be improved using various kernel formulations, more adapted to the oriented gradient descriptor. Finally, an optimization of the computation time of the recognition step in order to assemble and test under real conditions our system is currently under development.

Moreover, this is not to be considered as the final result, since other filters are used as well to independently evaluate each area of attention. The results of all filters is fed to a final neural network-based decision step.

REFERENCES

[1] L. Andreone, F. Bellotti, A. D. Gloria, and R. Laulette. Svm-based pedestrian recognition on near-infrared images. In *Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis*, 2005.

[2] M. Bertozzi, A. Broggi, M. Del Rose, and A. Lasagni. Infrared Stereo Vision-based Human Shape Detection. In *Procs. IEEE Intelligent Vehicles Symposium 2005*, pages 23–28, Las Vegas, USA, June 2005. doi:10.1109/IVS.2005.1505072.

[3] M. Bertozzi, A. Broggi, M. Felisa, G. Vezzoni, and M. Del Rose. Low-level Pedestrian Detection by means of Visible and Far Infra-red Tetra-vision. In *Procs. IEEE Intelligent Vehicles Symposium 2006*, pages 231–236, Tokyo, Japan, June 2006. doi:10.1109/IVS.2006.1689633.

[4] B. E. Boser, I. Guyon, and V. Vapnik. A training algorithm for optimal margin classifiers. In *Computational Learing Theory*, pages 144–152, 1992.

[5] A. Broggi, M. Bertozzi, M. Del Rose, and M. Felisa. A Symmetry-based Validator and Refinement System for Pedestrian Detection in Far Infrared Images. In *Procs. IEEE Intl. Conf. on Intelligent Transportation Systems 2007*, Seattle, USA, Sept. 2007.

[6] A. Broggi, M. Bertozzi, M. Felisa, P. Grisleri, S. Ghidoni, G. Vezzoni, C. Hilario Gómez, and M. Del Rose. Pedestrian Detection by means of Far-infrared Stereo Vision. *Computer Vision and Image Understanding*, 106(2):194–204, 2007, doi:10.1016/j.cviu.2006.07.016.

[7] S. Cheng and M. M. Trivedi. Turn-intent analysis using body pose for intelligent driver assistance. In *IEEE Pervasive Computing*, pages 23–37, Oct. 2006.

[8] N. Cristianini and J. Shawe-Taylor. *Introduction to Support Vector Machines*. Cambridge Univeristy Press, 2000.

[9] C. Curio, J. Edelbrunner, T. Kalinke, C. Tzomakas, and W. von Seelen. Walking Pedestrian Recognition. *IEEE Trans. on Intelligent Transportation Systems*, 1(3):155–163, Sept. 2000.

[10] R. Cutler and L. S. Davis. Robust Real-time Periodic Motion Detection, Analysis and Applications. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(8):781–796, Aug. 2000.

[11] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In C. Schmid, S. Soatto, and C. Tomasi, editors, *International Conference on Computer Vision and Pattern Recognition*, volume 2, pages 886–893, June 2005.

[12] J. W. Davis and V. Sharma. Robust Background-Subtraction for Person Detection in Thermal Imagery. In *Procs. Intl. IEEE Wks. on Object Tracking and Classification Beyond the Visible Spectrum*, Washington D. C., USA, 2004.

[13] M. Del Rose and P. Frederick. Pedestrian Detection. In *Procs. Intelligent Vehicle Systems Symposium*, Traverse City, USA, 2005.

[14] D. M. Gavrila. Pedestrian Detection from a Moving Vehicle. In *Procs. of European Conference on Computer Vision*, volume 2, pages 37–49, June–July 2000.

[15] D. M. Gavrila and S. Munder. Multi-cue Pedestrian Detection and Tracking from a Moving Vehicle. *Image and Vision Computing Journal*, 2007. In press.

[16] G. Grubb, A. Zelinsky, L. Nilsson, and M. Rilbe. 3d vision sensing for improved pedestrian safety. In *IEEE Intelligent Vehicles Symposium*, 2004.

[17] R. Kania, M. Del Rose, and P. Frederick. Autonomous Robotic Following Using Vision Based Techniques. In *Procs. Ground Vehicle Survivability Symposium*, Monterey, USA, 2005.

[18] C. T. L. Zhm. Stereo- and neural network-based pedestrian detection. In *IEEE Intelligent Transportation Systems*, 1999.

[19] M. Mählisch, M. Oberländer, O. Löhlein, D. Gavrila, and W. Ritter. A multiple detector approach to low-resolution fir pedestrian recognition. In *Procs. IEEE Intelligent Vehicles Symposium 2005*, pages 23–28, Las Vegas, USA, June 2005.

[20] S. Munder and D. Gavrila. An experimental study on pedestrian classification. In *IEEE Trans. Pattern Anal. Machine Intell*, pages 1863–1868, Nov. 2006.

[21] H. Nanda, C. Benabdelkedar, and L. Davis. Modelling pedestrian shapes for outlier detection: a neural net based approach. In *IEEE Intelligent Vehicles Symposium*, pages 428 – 433, June 2003.

[22] H. Nanda and L. Davis. Probabilistic Template Based Pedestrian Detection in Infrared Videos. In *Procs. IEEE Intelligent Vehicles Symposium 2002*, Paris, France, June 2002.

[23] 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. In *Procs. IEEE Intl. Conf. on Intelligent TransportationSystems 2006*, pages 97–102, Toronto, Sept. 2006.

[24] A. Shashua, Y. Gdalyahu, and G. Hayon. Pedestrian detection for driving assistance systems: Single-frame classification and system level performance. In *Proceedings of IEEE Intelligent Vehicles Symposium*, 2004.

[25] A. Shashua, Y. Gdalyahu, and G. Hayun. Pedestrian Detection for Driving Assistance Systems: Single-frame Classification and System level Performance. In *Procs. IEEE Intelligent Vehicles Symposium 2004*, Parma, Italy, June 2004.

[26] H. Shimizu and T. Poggie. Direction Estimation of Pedestrian from Multiple Still Images. In *Procs. IEEE Intelligent Vehicles Symposium 2004*, Parma, Italy, June 2004.

[27] M. Sotelo, I. Parra, D. Fernandez, and E. Naranjo. Pedestrian Detection Using SVM and Multi-Feature Combination. In *Procs. IEEE Intl. Conf. on Intelligent Transportation Systems 2006*, pages 103–108, Toronto, Sept. 2006.

[28] C. Stauffer and W. E. L. Grimson. Similarity Templates for Detection and Recognition. In *Procs. IEEE Intl. Conf. on Computer Vision and Pattern Recognition*, volume 1, pages 221–228, 2001.

[29] F. Suard, V. Guigue, A. Rakotomamonjy, and A. Bensrhair. Pedestrian detection using stereo-vision and graph kernels. In *Intelligent Vehicles Symposium, Las Vegas, Nevada*, pages 267–272, June 2005.

[30] M. Szarvas, A. Yoshizawa, M. Yamamoto, and J. Ogata. Pedestrian detection with convolutional neural networks. In *IEEE Intelligent Vehicles Symposium*, 2005.

[31] K. Tabb, S. George, R. Adams, and N. Davey. Human shape recognition from snakes using neural networks. In *Computational Intelligence and Multimedia Applications*, pages 292 – 296, Sept. 2000.

[32] V. Vapnik. *The Nature of Statistical Learning Theory*. Springer, N.Y, 1995.

[33] Y. Wi, T. Yu, and G. Hua. A statistical field model for pedestrian detection. In *Computer Vision and Pattern Recognition*, 2005.

[34] F. Xu, X. Liu, and K. Fujimura. Pedestrian detection and tracking with night vision. In *IEEE Intelligent Transportation Systems*, pages 63–71, Sept. 2005.