SVM-BASED OBSTACLE CLASSIFICATION IN VISIBLE AND INFRARED IMAGES

Anca Apatean[1,2], Alexandrina Rogozan[2] and Abdelaziz Bensrhair[2]

[1]Technical University of Cluj-Napoca, Communications Department, 26-28 Baritiu st., RO-400027, Cluj-Napoca, Romania
[2]National Institute of Applied Sciences (INSA de Rouen), LITIS Laboratory, 12 Av. de l'Université, BP 76801, Rouen, France
phone: +40 264 401 568, fax: +40 264 591 340,
email: anca.apatean@com.utcluj.ro, arogozan@insa-rouen.fr, abdelaziz.bensrhair@insa-rouen.fr

ABSTRACT

This paper describes a Support Vector Machine (SVM)-based obstacle recognition system that can recognize both vehicles and pedestrians using bimodal vision. Different techniques were investigated in order to recognize the detected obstacles by the extraction of a compact and pertinent numeric signature from visible and infrared spectrum. A bi-objective optimization (using error classification rate and classification time) is employed to assure the SVM-parameters selection. A comparative study of individual visual obstacle recognizers versus fusion-based (at the feature, kernel and matching-score level) systems is performed. An important advantage of the fusion-based systems is their possibility to adapt to the environmental conditions due to a weighting parameter which establishes the importance of each sensor in a specific situation.

1. INTRODUCTION

Combining information from different sources contributes to forming a more complete image of an object to be detected or recognized in a road scene. Our purpose is to develop methods for an obstacle recognition system which combines visible (VIS) and infrared (IR) information in order to improve the road security. A stereo vision system has been developed in our laboratory [1] for the obstacle detection task and our efforts[1] aim to continue this work. Our main purpose is to develop approaches to reduce the number of false alarms and to recognize the detected obstacles (like pedestrians, cars) by the extraction of a compact and pertinent numeric signature, followed by a Support Vector Machine (SVM) classification.

Almost all obstacles categorization systems developed until now use an object detection step followed by an object recognition or a hypothesis verification step. During the detection step a rectangular region of interest called bounding box (BB) is found and it is associated with a potential obstacle; then, the recognition or verification process follows, where the false alarms are removed and the type of the object is determined.

Many machine learning algorithms have been tested during the last years, in order to solve the obstacle categorization problems. Pedestrians are detected by multiple methods, using for example: shifted windows of various sizes over the image [2,3], symmetries and BB generation [4], an N-Cut segmentation [5], a shape finding method based on a hierarchy of templates [6]. Features on which the obstacle classification module is usually based, are: Haar wavelets [2,7], Sobel edge features [3], set of simple appearance filters (similar to the Haar wavelets) [8], texture features obtained by high-pass filtering [9], Gabor features [7], vertical histograms and aspect ratio [4], graphs which model the pedestrian shape [5], global features derived from a PCA analysis [6,7]. The obstacle template is learned and then classified with an SVM [2,5,7,9], or even with two SVMs (for different obstacle poses) [3], an Adaboost cascade classifier [8], a neural network [6,7,9] or a polynomial classifier [6].

Considering the price and the lack of interference problems, we chose two complementary vision sensors because the system must work well even under difficult conditions, like poor illumination or bad-weather situations (dark, rain, fog).

For an obstacle categorization system, the use of a single sensor cannot provide complex information about the environment in any weather conditions and any illumination situation. Different sensor inputs, data or even algorithms could be combined together by the fusion process in order to provide complementary information and to increase the system's performances. We want to compare various solutions prior to implementing our final system because this would help in choosing the best solution for a given scenario. For example, we have to decide whether first to fuse data and then to detect/recognize obstacles in the fused data (*low-level fusion*) or first to detect/recognize obstacles in each image separately and then to fuse the decisions (*high level fusion*). In this paper kernel-based and matching-score fusion techniques are compared for a pedestrian-vehicle SVM-based classification problem. In order to ensure the adaptation of the system to the environmental conditions, kernels or matching-scores could be weighted (with a sensor weighting coefficient) according to the importance of the sensor in a specific environmental situation.

Different global texture features have been extracted from the visible and infrared images in order to compute a fused feature vector encoding both types of data. Then, the feature selection process follows where just the most relevant features are retained for time reduction reasons.

[1] Part of this work has been supported by the Romanian Ministry of Education and Research under Grant "Researches in Processing and Fusioning Data Provided by Multiple Optical Sensors. Applications for Objects Detection and Classification", CNCSIS TD 521/2007-2008.

2. FEATURES SELECTION FROM VISIBLE AND INFRARED IMAGES

Our first goal is to assure a real time processing for the proposed system. Because of the great number of possible appearances and shapes of obstacles on the road scene, a robust classification system should have the ability to learn multiple instances of the same object. This is generally achieved by learning the classifier with a great number of items per each class. But this growth of the image-database dimension will have an unfortunate contribution to the increasing processing time. Thus, we have to ensure a processing time as short as possible, while still maintaining a high accuracy rate. The reduction of the computational time corresponding to the feature extraction step and to the classification mechanism is quite mandatory. By using a feature selection process we intend to investigate the attribute's pertinence, to evaluate different combinations of such features and finally to select the best suited ones (from the recognition rate and classification time point of view). The founding of one convenient set of features is of the same importance as the founding of the most performing classifier.

If the reduction of the feature vector is achieved with slight decrease of the accuracy rate (compared to the initial feature vector), we can consider the system as quite robust. Since feature extraction is desired to be fast, the performances of the entire system depends heavily on the choosen features. Therefore, to obtain a fast recognition system, the most representative features for each modality VIS and IR should be retained. In order to encode both visible and infrared information in one single feature vector, we computed a fused feature vector on which the feature selection mechanism has been applied. Futher on, we investigated different types of features in order to find the best feature combination for the fusion process.

We computed features like width and height of the BB enclosing the object, followed by 169 features: 64 Haar and 32 Gabor wavelets, 7 statistical moments (mean, median, mode, variance, standard deviation, skewness and kurtosis), 8 DCT (discrete cosine transform), 16 cooc (obtained from the gray level co-occurrence matrix GLCM), 14 rle (run-length encoding) and 28 laws (a set of convolution masks) features. For more detailed explanations about how the features were extracted and how the feature vectors were obtained, please consult [10] and [11]. From the VIS and IR images, we retained 2 individual feature vectors: VIS_{171} and IR_{171}. In order to compute a fused feature vector (so a feature fusion case could be also considered), we fused 171 VIS and 169 IR features. In this case of *feature-fusion*, the feature sets extracted from the VIS and IR images are fused in order to create a new feature set which will represents the object:

$$(x_1,..x_k,x_{k+1},...,x_n)_{Fused} = (x_1,...,x_k)_{VIS},(x_{k+1},...,x_n)_{IR} \quad (1)$$

where n=340 and k=171. The obtained fused feature vector will be referenced in the paper as $VisIr_{340}$.

Most methods for feature selection involve searching the space of attributes in order to find a subset of relevant features that is most likely to predict the correct class. We used a Feature Selection Subset Evaluation (CfsSubsetEval [12]) method which is based on correlation and is combined with the best first search method. We applied the CfsSubsetEval method on the fused feature vector $VisIr_{340}$ and the evaluation takes into consideration the entire learning set. Thus, 42 features were selected and were grouped in 3 feature vectors: $VisIr_{42}$, VIS_{17} and IR_{25}. The feature vector $VisIr_{42}$ was used in feature-fusion and kernel-fusion cases, with respect to equation (1), with k=17 and n=42. The other two corresponding feature vectors were used either for a monomodal classification (with no fusion scheme), or for obtaining the VIS and IR scores in the matching-score fusion case. We used Weka (a very powerful tool including a collection of machine learning algorithms for data mining tasks) in order to evaluate the feature selection task.

3. CLASSIFICATION BASED ON SUPPORT VECTOR MACHINES

We used an SVM classifier in order to develop different obstacle classification schemes, considering both visible and infrared information. The high performance and robustness of the system will be assured by the fusion of these 2 types of information, weighted in such a manner to allow the adaptation of the system to the environmental conditions. Therefore, our main purpose is to obtain a robust model which has to incorporate the information related to the environmental context. One first possibility is the one presented in [11] where the single kernel was replaced by a multiple kernel for the SVM classifier. Because our multiple kernel is defined as a combination (a linear one) of single kernels, we can consider this first possibility as a kernel-fusion scheme. The second fusion system considered in this paper take into account the scores obtained from individual classifiers on each modality.

A comparative study of individual visual and infrared obstacle recognizers versus fusion-based systems is performed and the obtained results are presented in section 4.

3.1 Support Vector Machines

SVMs classifiers are based on kernel functions, which define similarities between pairs of data: $K(x,z) = \langle \phi(x), \phi(z) \rangle$, $\forall x, z \in R^n$. Let the instance-label pairs (x_i, y_i), $i = 1,...,m$ be the m training data, where $x_i \in R^n$ represents the input vector and $y_i \in \{1,-1\}$ the output label associated to the corresponding item x_i. To obtain an SVM classifier with kernels, one has to solve the following optimization problem: $\max_{a \in R^m} \sum_{i=1}^{m} a_i - \frac{1}{2}\sum_{i,j=1}^{m} a_i a_j y_i y_j K(x_i, x_j)$, subject to $\sum_{i=1}^{m} a_i y_i = 0$, $0 \le a_i \le C$, $\forall i \in 1,2,...,m$, with $K(x_i, x_j)$ the kernel function and $a_i \ge 0$ the Lagrange coefficients [13]. The coefficient C>0 is the penalty parameter that controls the trade off between maximizing the margin hyperplanes (which separates the classes) and classifying without errors. The optimal separating hyper-plane is used to classify the un-labelled input data x_k using the following decision function:

$$y_k = sign\left(\sum_{x_i \in S} a_i y_i K(x_i, x_k) + b\right) \qquad (2)$$

where S is the set of support vector items x_i and the offset value b is calculated based on vector a and the training set. For the SVM classifier, we considered two types of kernels: polynomial (POL) and Radial Basis Function (RBF), the most used in the literature and those having only one parameter. The POL kernel has the degree d (with $d=1$ for the linear kernel) and the RBF kernel has the bandwidth γ. Different parameters of these kernels have been tested.

We compared the system performances in the case of using no fusion scheme (using the feature vectors corresponding to one modality VIS_{17} or IR_{25}), with the feature-fusion case (using $VisIr_{42}$), with the case corresponding to the kernel-fusion and with the case of matching-score fusion (where the kernels and the scores are evaluated using the feature vectors VIS_{17} and IR_{25}).

3.2 Kernel based fusion

The kernel methods represent data by means of a kernel function, which defines similarities between pairs of items. Our goal in the kernel-fusion case is to find a kernel that best represents all of the information available for the two types of images. Generally, classical kernel-based classifiers use only a single kernel (SK), while the applications from the real world need a combination of kernels in order to perform a better adaptation to the heterogeneous and multi-sensorial data. A common approach is to consider that the kernel function $K(x_i, x_j)$ is a linear combination of the basic kernels [11]:

$$MK(x_i, x_j) = \alpha \cdot SK_{VIS}(x_i^{1,k}, x_j^{1,k}) + (1-\alpha) \cdot SK_{IR}(x_i^{k+1,n}, x_j^{k+1,n}) \quad (3)$$

The obtained multiple kernel will be the sum of two independent kernels, each one corresponding to one modality VIS or IR and weighted with a value representing the importance of that domain in the context. For our problem, the kernel functions SK_{VIS} and SK_{IR} could be either RBF, or Polynomial, and could work with different hyperparameters. Generally, the POL kernel has a degree $d \in \{1,2,...,15\}$ (with $d=1$ for the linear kernel) and the RBF kernel has the bandwidth γ, of the form $\gamma = q \cdot 10^t$ with $q \in \{1,2,...,9\}$ and $t \in (-5,-1)$. The results found in literature indicate that these discrete spaces of parameters are the most suitable for an efficient classification. A proper choice of these parameters is crucial for SVM to achieve good classification performance. The values C and d or γ are called hyperparameters and they need to be determined by the user. They are usually chosen by optimising a validation measure (such as the k-fold cross validation error) on a grid of values (e.g. uniform grid in the (C, d) or (C, γ) space).

Because kernel-fusion approach (which is MK from [11]) uses a linear combination of simple kernels for the feature vectors VisIr, the MK has the following parameters: the kernel type (RBF or POL), the context adjustment value α, the penalty parameter C and the kernel parameters, according to each domain. Thus, our MK is entirely described by the parameters set: (kernel, α, p_{VIS}, p_{IR}, C), where p_{VIS} and p_{IR} are the parameters corresponding to the kernel from the corresponding domain.

3.3 Matching-scores based fusion

For a *matching-score fusion*, multiple classifiers output a set of matching scores which represent the probabilities that one object belongs to different possible classes, based on different modalities. The matching scores generated by the VIS and IR modalities for an object can be combined by the weighted parameter α in order to obtain a new match score which is then used to make the final decision. If in equation (2) we would consider the value of the respective sum instead of it's sign, we can obtain the scores of the classifier. Consider an input pattern X, which could be classified into one of M possible classes $\{y_1, y_2, \ldots, y_M\}$ based on the evidence provided by N different classifiers. The input pattern X could be considered to have multiple feature vectors (x_j, with j=1,…,N), with one feature vector for each classifier [14]. Given the feature vectors x_j, j=1,…,N, the input pattern X should be assigned to the class y_n if $\sum_{j=1}^{N} P(y_n \mid x_j) \geq \sum_{j=1}^{N} P(y_k \mid x_j)$, where k=1,…,M. In order to consider the sensor weighted parameter, we will assign X to the class y_n if $\sum_{j=1}^{N} \alpha_j \cdot P(y_n \mid x_j) \geq \sum_{j=1}^{N} \alpha_j \cdot P(y_k \mid x_j)$, where α_j is the weighted value associated to each classifier, with $\sum_{j=1}^{N} \alpha_j = 1$. In that way, for our bimodal problem with the classes pedestrian and vehicle, a new object X having the vectors x_{VIS} and x_{IR} will be assigned, for example, to the class pedestrian (P) if respects the relation:

$$\alpha P(P \mid x_{VIS}) + (1-\alpha) P(P \mid x_{IR}) \geq \alpha P(V \mid x_{VIS}) + (1-\alpha) P(V \mid x_{IR}) \quad (4)$$

Generally, a normalization step is necessary before the matching scores originating from different classifiers to be combined in the fusion process. In order to normalize the SVM scores, a min-max normalization method was used:

$$n = \frac{s - \min(S)}{\max(S) - \min(S)} \qquad (5)$$

where s represents the score from the set of all scores of that classifier (in the validation set) before normalization, and n is the corresponding normalized score. This method maps the raw scores to the [0,1] range. Each simple kernel is involved with a weight that represents its relative importance for classification. The kernel selection process and the optimization of the hyper-parameters are described in the following.

4. EXPERIMENTS AND RESULTS

The VIS - IR image database used for our experiments is provided by the Artificial Vision and Intelligent Systems Laboratory (VisLab) of Parma University. This image database has been obtained using an experimental vehicle equipped with two CCD and two far-infrared cameras. For more details about the aspects of the video acquisition module and the procedure used to calibrate the cameras, please refer to [15]. To the best of our knowledge, even many research groups used this database, they performed a pedestrian detection and a pedestrian/non-pedestrian classification. Our goal being to assign the correct label to a bounding box which may contain a possible road obstacle, we performed the discrimination of humans and vehicles. Examples of different VIS and IR objects (correlated each other) belonging to the class pedestrian or to the class vehicle are presented in Fig.1. We created a preliminary database containing 486 samples of road obstacles grouped in two sets: the learning set (80%, 389 objects) and the test set (the remaining 20%, 97 objects). On the learning set, 10-folds crossvalidation has been performed in order to optimize the parameters set and to choose the proper kernels. In the testing step, for the selected kernels from the crossvalidation step, we computed the accuracy rate in order to select the best suited kernel for our system.

In Weka there is a collection of machine learning algorithms, but there is no algorithm to treat the fusion problem. Thus, we implemented our fusion schemes, starting from a similar toolbox of classification developed in Matlab [16].

Having in mind that it is not known beforehand which parameters for the SVM kernels (C and γ or d) gives the best solution for one problem, there must be done a model selection (parameter search) that could identify good C, γ or d. The result of this parameter selection process is that the classifier will be able to predict accurately unknown data. For each RBF and POL kernels, a number of 220 combinations (C; γ or d) revealed in Table I were experimented in order to find proper kernels for the kernel and matching score fusion cases. The parameter α was tested with 11 values $\alpha \in \{0.0, 0.1, 02, ..., 0.9, 1.0\}$, obtaining 11 possibilities for the fused information in the *kernel fusion* and the *matching-score fusion* cases, as equations (3) respective (4) show. For the optimization process, different values among a discrete set (which covers the domains early mentioned) are used: the penalty parameter $C \in \{0.1, 1, 10, \quad 20, 30, \dots, 200\}$, $\gamma \in \{5 \cdot 10^{-5}, 10^{-4}, ..., 1\}$ and $d \in \{1, 2, ..., 10\}$.

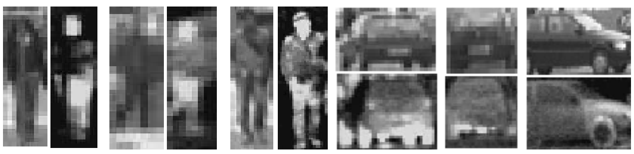

Figure 1. Examples of correlated images (visible and infrared) for pedestrians and cars

Table I. Different combinations of parameters providing 220 RBF and 220 POL kernels for SVM

Kernel		**The penalty parameter C**					
RBF γ	**POL** d	**0.1**	**1**	**10**	**20**	**…**	**200**
$5*10^{-5}$	**1**	K1	K2	K3	K4	…	K22
$1*10^{-4}$	**2**	K23					
$5*10^{-4}$	**3**	K45			…		
$1*10^{-3}$	**4**	K67					
…	**…**	…					
$1*10^{-0}$	**10**	K199	K200	K201	K202	…	K220

In the crossvalidation process, the experiment was conducted in order to select proper single kernels which will be used in the test step. This means we are looking for the SVM parameters sets for which *good performances* are obtained in the crossvalidation process. Table II contains the mean classification time and the error classification rate (which is calculated as 100 minus accuracy rate [%]) obtained in the 10-folds crossvalidation process. What we mean by the term *good performances*: for the SVM kernel function, different combinations of hyperparameters could be revealed; (1) Ones could provide good accuracy rates, (2) others could present the advantage of a small processing time and (3) others can provide also good accuracy rates and a small classification time. In the crossvalidation process, we are searching for this last type of kernels. The proposed method is as it follows: first, the computed accuracy rates are used to obtain the corresponding error classification rates (ERR); the second step is to perform the min-max normalization (equation 5) on both the error and the classification time. In that way, we obtained 2 values belonging to the [0,1] range, which were then multiplied. By this multiplication for each of the 220 RBF kernels and 220 POL kernels, we obtained 2 sets of values (one for RBF and one for POL) belonging to [0,1] which were then sorted from the smallest to the highest. From these 2 lists of values (where each value correspond to a kernel), we have two possibilities to choose the proper kernels: (1) to consider the first n kernels from the list (when n could be any integer value) or (2) to consider a threshold as a maximum value. Because there is for no use a kernel providing for example an accuracy rate of 70% but a classification time (after the min-max normalization) of 0.001, we performed the kernel selection algorithm just on the kernels having an accuracy

Table II. Mean classification time and error classification rate in the 10-folds crossvalidation step

Classifier: SVM / **Vector**	**The feature vectors**	**SK**			
		RBF		**POL**	
		ERR [%]	Time [msec]	ERR [%]	Time [msec]
Initial vector	VIS_{171}	2.9	0.123	2.9	0.411
	IR_{171}	2.9	0.096	2.9	0.318
	$VisIr_{340}$	2.9	0.195	2.9	0.468
Feature Selection	VIS_{17}	11.3	0.053	11.8	0.259
	IR_{25}	5.5	0.048	6.6	0.149
	$VisIr_{42}$	4.7	0.049	5.3	0.158

rate above 80% and a classification time (after the normalization) below 0.8. By considering the second approach, with a threshold value of 0.2, we obtained a number of 126 kernels for the RBF and 143 for the POL kernels. Next, these selected kernels were used in the test step.

For each RBF and POL kernel selected in the validation process, a number of 4 situations were obtained, from which 3 are based on fusion. One of them uses monomodal feature vectors (Vis and Ir), in which case no fusion scheme was used. Another possible situation is the feature-fusion case, where the fused feature vector was used. Other situations (each with 11 possible vectors for testing) are the kernel-fusion and the matching score fusion cases. The maximum accuracy rates obtained in all these cases are shown in Table III. In order to compare the results obtained from different classifiers, we added in Table III the accuracy rates provided by the k-nearest neighbour, with k=1 and k=3. We can remark that the accuracy is greater in the case of fused modalities than in the case of separate feature-vectors (where no fusion scheme was considered): the Vis feature vector gives the lowest accuracy rates, maybe also because it has the smallest number of features. All these fusion-based methods offer reasonable accuracy rates compared to the no-fusion cases. The main difference between the kernel-fusion approach and the matching-score one is that in the matching-scores fusion case two different values could be considered for the complexity parameter C, because the final decision of the fusion system is taken after the classifiers provide their individual solution. From Table III, we can remark that the best accuracy rate is obtained by the match-scores fusion (RBF-RBF fusion, γ_{VIS}=0.01, γ_{IR}=0.5, C_{VIS}=100, C_{IR}=50, $\alpha = 0.3$), and is higher even than the values obtained with the initial feature vectors (Table II, the ERR 2.9 is corresponding to an accuracy rate of 97.1).

5. CONCLUSIONS AND PERSPECTIVE WORK

The purpose of this paper is to investigate if different levels of VIS-IR fusion are efficient, especially if they are associated with the feature selection and the SVM-based classification. The experiments show that we can consider the fusion at different levels to provide better results for our problem than the monomodal systems (based on VIS_{17} and

Table III. Maximum Accuracy rates obtained with the test set

Classification method	Kernel type	Vis_{17}	Ir_{25}	$VisIr_{42}$
SK (no fusion)	RBF	89.7	94.9	X
	POL	87.6	95.9	X
Feature fusion	RBF	X	X	***96.9***
	POL	X	X	***96.9***
Kernel fusion	RBF+RBF	X	X	***96.9***
	POL+POL	X	X	94.9
	RBF+POL	X	X	***96.9***
	POL+RBF	X	X	95.9
Match-scores fusion	RBF+RBF	***97.4***		X
	POL+POL	95.9		X
	RBF+POL	96.9		X
	POL+RBF	96.2		X
K-NN with k=1	-	93.8	94.6	94.6
K-NN with k=3	-	94.0	96.7	96.9

IR_{25}). Additionally, these monomodal systems do not allow the system adaptation to the VIS-IR context, like a fusion scheme does through the weighted parameter α. Different types of kernels with different values for the hyperparameters could be tested in order to enlarge the searching space and to find the best global solution for our problem.

As further improvements, we intend to integrate these fusion schemes in an entire obstacle-detection and classification system.

The purpose of the fusion system is to utilize the information from both sensors in order to classify the detected obstacle. The weighted value α allows the system adjustment to the VIS or IR domain according to the context. The VIS-IR weighted parameter α should be adapted to the system based on the illumination and weather conditions. By now, we do not have knowledge about any image database containing different illumination or weather situations, but we intend to develop approaches to simulate these difficult conditions. When such an image database will be available, the α parameter would characterize a specific situation and it will be determined in the validation step. In that way, multiple classifiers will be available for different situations. In the test step, the sensor weighted parameter α will be determined by some statistical measurements.

REFERENCES

[1] A. Bensrhair et al.,"Stereo visionbased feature extraction for vehicle detection", *IEEE Intelligent Vehicles Symp.*, vol.2, pp 465-470, 2002.

[2] C. Papageorgiou et al.,"A trainable pedestrian detection system", *IEEE Int. Conference on Intelligent Vehicles*, 241-246, Germany, 1998.

[3] G. Grubb et al., "3D Vision Sensing for Improved Pedestrian Safety", *IEEE Intelligent Vehicle Symposium*, pp. 19-24, Parma, Italy, 2004.

[4] A. Broggi et al.,"Shape-based Pedestrian Detection", October 17, 2000.

[5] F. Suard et al., "Pedestrian Detection using Stereo-vision and Graph Kernels", *IEEE Intelligent Vehicles Symp.*, pp. 267-272, June 2005.

[6] D. M. Gavrila et al., "A multi-sensor approach for the protection of vulnerable traffic participants - the PROTECTOR project", *IEEE Instrumentation and Measurement Technology Conference*, vol. 3, pp. 2044-2048, 2001.

[7] Z. Sun et al.,"Monocular precrash vehicle detection: features and classifiers", *IEEE Transactions on Image Processing*, vol. 15, no.7, pp 2019-2034, July 2006.

[8] P. Viola et al.,"Detecting pedestrians using patterns of motion and appearance", *Int.Conf. Computer Vision*, pp.734-741, France, 2003.

[9] L. Zhao et al.,"Stereo and Neural Networkbased Pedestrian Detection", *IEEE Transactions on Intelligent Transportation Systems*, 1(3), 2000.

[10] A. Apatean, A. Rogozan, A. Bensrhair, "Objects Recognition in Visible and Infrared Images from the Road Scene", IEEE *International Conference on Automation, Quality and Testing, Robotics*, AQTR

[11] A. Apatean, A. Rogozan, A. Bensrhair, "Obstacle Recognition using Multiple Kernel in Visible and Infrared Images", *IEEE Intelligent Vehicles Symposium*, pp.370-375, Xi'an, China, June 02-05, 2009.

[12] I.H. Witten, E. Frank, Data Mining – Practical Machine Learning Tools and Techniques, second edition, Elsevier: San Francisco.

[13] O. Chapelle et al., "Choosing multiple parameters for support vector machines", *Machine Learning,* 46(1/3), 131-159, 2002.

[14] A. Ross, K. Nandakumar, A.K. Jain, *Handbook of Multibiometrics*, Springer, Verlag, 2006.

[15] Hammoud, R.I., (Ed.), *Augmented Vision Perception in Infrared Algorithms and Applied Systems*, Springer 2009, ISBN: 978-1-84800-276-0, pp. 371-401

[16] http://upload.pobrun.com/40/uploads/files/simplesvm/documentationSimpleSVM.pdf