

Vehicle Type Classification with Geometric and Appearance Attributes

Ghada S. Moussa

Abstract—With the increase in population along with economic prosperity, an enormous increase in the number and types of vehicles on the roads occurred. This fact brings a growing need for efficiently yet effectively classifying vehicles into their corresponding categories, which play a crucial role in many areas of infrastructure planning and traffic management.

This paper presents two vehicle-type classification approaches; 1) geometric-based and 2) appearance-based. The two classification approaches are used for two tasks: multi-class and intra-class vehicle classifications. For the evaluation purpose of the proposed classification approaches' performance and the identification of the most effective yet efficient one, 10-fold cross-validation technique is used with a large dataset. The proposed approaches are distinguishable from previous research on vehicle classification in which: i) they consider both geometric and appearance attributes of vehicles, and ii) they perform remarkably well in both multi-class and intra-class vehicle classification. Experimental results exhibit promising potentials implementations of the proposed vehicle classification approaches into real-world applications.

Keywords—Appearance attributes, Geometric attributes, Support vector machine, Vehicle classification.

I. INTRODUCTION

VEHICLE classification (VC) has emerged as a major component of many contemporary traffic management and operational systems. VC has a large number of real world applications in numerous fields such as traffic surveillance, toll plaza, traffic congestion avoidance, and accident prevention, etc. Vision-based VC systems have quickly gained recognition over electronic sensors such as inductive loops due to their cost-effectiveness, portability, ease of maintenance and visualization capabilities [1].

In recent years, research in the area of vehicle detection and classification via still images has attained an enormous amount of attention [2], [3]. However, this is a challenging task due to the ever increasing number of vehicle models and sizes (even within a single category), they are generally textureless. In addition, occlusion, shadow and illumination make the classification task even more challenging [3]. The requirement to distinguish between intra classes (sub classes) such as pick up, sport utility, and van even complicates the task, though it is more common for road users to use such classes in both urban and rural areas. Therefore, constructing a robust vehicle classification system to cope with those issues is desired in real-world applications.

G.S. Moussa, Ph.D., is with Civil Engineering Department, Assiut University, Assiut 71515, Egypt, (phone: 111-760-5780; fax: 88-233-2553; e-mail: ghada.moussa@gmail.com).

In this study, two vehicle classification approaches are presented, using the Support Vector Machine (SVM) algorithm: 1) geometric-based approach and 2) appearance-based approach. The two classification approaches are used for two tasks: multi-class and intra-class vehicle classifications. In the multi-class vehicle classification task, vehicles are categorized into three main classes: small, medium, and large size vehicles. Whereas in the intra-class vehicle classification task, the medium size vehicles are categorized into intra classes (sub classes): pickup (PU), sport utility vehicle (SUV), and van (VAN).

The novel contribution of this work is that by adopting SVM algorithm, we aim to discover the potential of using geometric and appearance attributes to achieve high levels of performance for multi-class and intra-class classifications.

The rest of the paper is organized as follows. The next section reviews previous research on vehicle classification. Section III outlines the proposed work methodology. The two proposed classification approaches are presented in Sections IV and V respectively. Experimental results and data analysis are presented in Section VI. Finally, conclusions are summarized in Section VII.

II. RELATED WORK

A. Review Stage

Numerous approaches have been developed for vehicles' classification. Each approach follows different procedures and uses different attributes for vehicles detection and classification. Generally, the classification accuracy of detected vehicles depends crucially on the combination of the extracted features of vehicles and the type of a classifier used for the classification [4]. Most machine-learning classification methods in the literature can be categorized into two main approaches based on the extracted features: geometry-based and appearance-based classification [5].

In geometry-based approaches, geometric measurements such as width, length, height, area... etc. are used as features for the classification. Therefore, those methods are usually tailored for specific object classes through user-determined choices of measurements. Avery et al. presented a length-based vehicle classification algorithm [6]. They used streams of images captured from un-calibrated video camera to compare different vehicles' lengths for estimating truck volumes while eliminating the needs of different complex calibration systems. They reported 92% accuracy for truck classification under certain conditions. Zhang et al. developed a length-based vehicle detection and classification system for

the collection of truck related data [7]. They reported an accuracy of 97% for truck classification. Their system faced some problems with the longitudinal vehicle occlusions, camera movements and head-light reflection. Moussa and Hussain developed a laser intensity vehicle classification system based upon images obtained from range sensors for vehicle classification [8]. Their system extracts features of laser intensity images and recalls its trained random neural network for classification of vehicles. Vehicles are automatically classified into five categories. They reported 94% classification accuracy.

In most appearance-based methods, vehicle images are represented as vectors in some high-dimensional space. Ma and Grimson used edge-based features and modified scale invariant feature transform (SIFT) descriptors for vehicle classification [9]. They reported classification rates of 98% and 96% for car vs. minivan and for car vs. taxi respectively. Zhang et al. presented an appearance learning-based method to distinguish between moving objects such as cars, vans, trucks, people and bikes using multi-block local binary patterns [10]. Morris and Trivedi used blob features followed by linear discriminate analysis (LDA), fuzzy C-means clustering and a weighted k-Nearest Neighbor (wkNN) classifiers to classify vehicles in eight classes [11]. They reported a classification accuracy of up to 94% with a rejection of low confidence objects. Recently, Moussa developed a multi-type vehicle classification system based on the bag-of-words paradigm

using SIFT features and four well-known classifiers; Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), k-Nearest Neighbor (KNN), and Decision Tree to classify vehicles into four different categories [12]. The reported classification accuracies of the system using LDA, SVM, KNN, and decision tree are 90.6%, 95.7%, 82.9%, and 76% respectively. In which, SVM algorithm outperformed other classification algorithms in terms of both accuracy and robustness alongside a considerable reduction in execution time.

III. METHODOLOGY OVERVIEW

In this work, two vehicle classification approaches are developed (geometric-based and appearance-based) to carry out two classification tasks: multi-class and intra-class vehicle classifications. In the multi-class vehicle classification task, vehicles are categorized into three main classes: small, medium, and large size vehicles. Whereas in the intra-class vehicle classification task, the medium size vehicles are categorized into intra classes (sub classes): pickup (PU), sport utility vehicle (SUV), and van (VAN). The framework that outlines the presented work is depicted in Fig. 1.

The two developed classification approaches (geometric-based and appearance-based) are discussed in details in Sections IV and V.

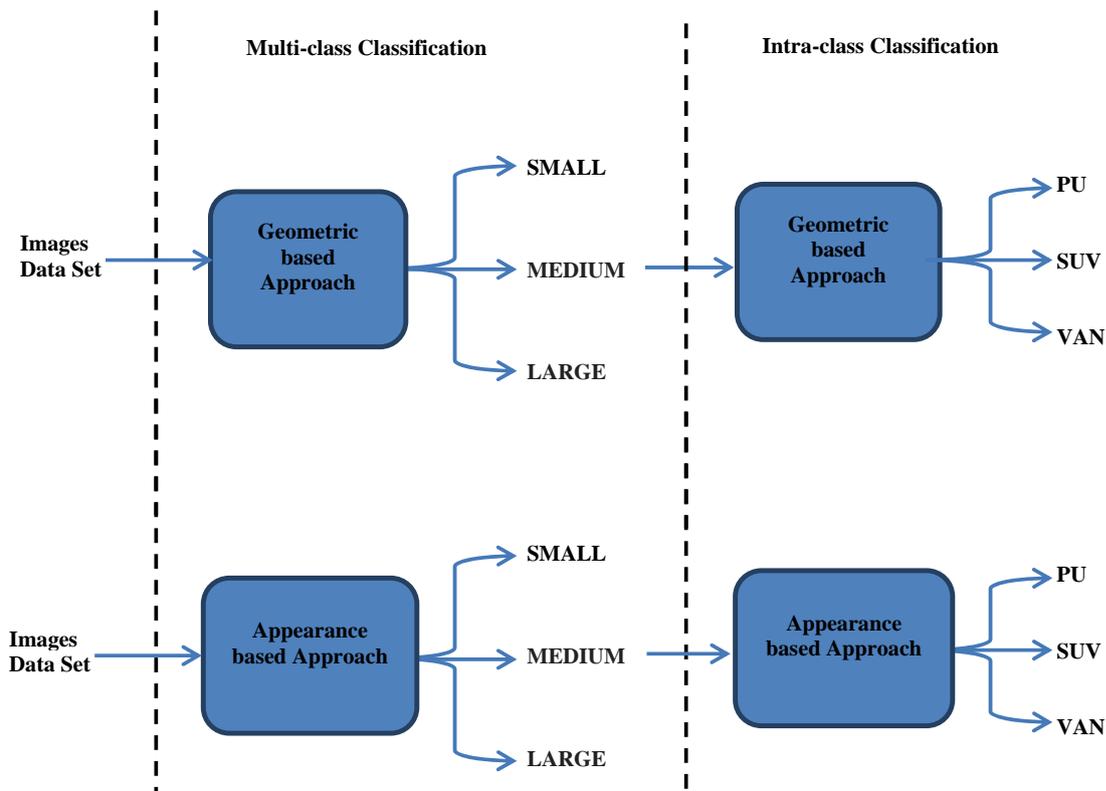


Fig. 1 Framework of presented work

IV. GEOMETRIC-BASED VEHICLE CLASSIFICATION APPROACH

The geometric-based vehicle classification approach is accomplished via two main steps: feature extraction and classification as shown in Fig. 2.

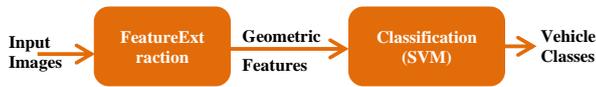


Fig. 2 The block diagram of the proposed geometric-based classification approach

A. Feature Extraction

The feature extraction expresses the visual content of the images, which should ideally quantify certain significant characteristics of vehicles in the images. Simple dimensional measurement-based features to describe the vehicle are extracted from the images. The extracted features are: width, length and height. The reasons for selecting such features are their low computational cost and storage requirements. For more details about extracting these features, see [8].

B. Classification

The Support Vector Machine (SVM) algorithm is used for vehicle classification. The extracted geometric features are fed directly into the SVM to classify vehicles to their corresponding classes. SVM is a supervisory classifier originated from the statistical learning theory and attempts to identify a set of support vectors. SVM separate other learning systems in its decision surface which is an optimal hyperplane in a higher dimensional feature space. This hyperplane minimizes the risk of misclassification [3] and [13]. The SVM algorithm identifies a set of support vectors and the final decision of the SVM is based on a “max. voting”, in which the category that corresponds to the highest confidence is the winner.

V. APPEARANCE-BASED CLASSIFICATION APPROACH

The appearance-based classification approach consists of three main steps. First, the extraction of local appearance features followed by using the well-known bag-of-words paradigm (BoW), then the classification using Support Vector Machine (SVM). The block diagram of the proposed appearance-based classification approach is depicted in Fig. 3.



Fig. 3 The block diagram of the proposed appearance-based classification approach

The three steps are explained as follows;

A. Feature Extraction

In this study, the two main types of appearance features are extracted from input images: Scale Invariant Feature Transform (SIFT) and Speeded-Up Robust Feature (SURF).

Scale Invariant Feature Transform (SIFT), introduced by David Lowe in 1999, is a widely used feature descriptor algorithm [14]. Most state-of-the-art object recognition systems use SIFT to represent images and it has been proven to be the most powerful and successful among local image descriptors [1], [13]. Using SIFT, each input image is represented by a set of relatively invariant local features in which, feature points in the image are detected using Harris-Laplace salient point detector. Then, descriptors for each feature point in the image are computed. Each feature point in the image is represented by a 128-dimensional orientation vector.

The second appearance feature used in this study is Speeded-Up Robust Feature (SURF) which introduced by Bay et al. in 2006 [15]. SURF is also a well-known descriptor and is similar, in some concepts, to SIFT. In which, both focus on the spatial distribution of gradient information. The SURF descriptors are constructed by extracting square grid around the interest points, then dividing each grid cell into sub-grids.

The resulting descriptor vector for each feature point in the image is a 64-dimensional orientation vector.

B. Bag-of-Words Modeling

After feature extraction, each input image is represented by a set of feature descriptors (SIFT or SURF). The dimensions of these descriptors are very high (more than thousands). Therefore, it is unreasonable to feed them directly to the classifier, but they have to go through a feature representation. In which, an efficient modeling method is used to transform these high dimensional descriptors to more compact and informative representations. For that reason, the bag-of-words (BoW) paradigm is adapted here. The BoW paradigm was pioneered by Csurka et al. and has received much attention in object recognition [16]-[18]. The main idea of the BoW paradigm is to treat image features as words. In which, after feature extraction, each input image is represented by a set of feature descriptors (SIFT or SURF). A visual vocabulary dictionary is constructed by applying a k-means clustering algorithm on the training data images, and each cluster center is considered as a “visual word” in the visual vocabulary dictionary. Then, all feature descriptors extracted from an image are quantized to their closest “visual word”. For each image, the number of feature descriptors assigned to each “visual word” is then accounted into a histogram. Therefore, each image is represented as a histogram of frequencies of

visual words that are in the image. For more details about using BoW in appearance-based classification approach, see [12].

C. Classification

As in the geometric-based approach, the Support Vector Machine (SVM) algorithm is applied for vehicle classification.

VI. RESULTS AND DISCUSSIONS

This section aims to assess the performance of our proposed vehicle classification approaches (geometric-based and appearance-based) and identify the effective yet efficient approach for the two classification tasks: multi-class and intra-class vehicle classifications.

To evaluate the performance of the proposed classification approaches, we applied the 10-fold cross-validation technique. The cross-validation technique is a strategy to compare the performances of different predictive modeling procedures [19]. A large data set of vehicles' images with their ground truth data is used for the evaluation process; see Table I. Using 10-fold cross-validation technique, the dataset is randomly split into 10 equal size subsamples. From the 10 subsamples, a single subsample is reserved as the validation data for testing the model, and the remaining 9 subsamples are used as training data. The process is repeated 10 times, each of the 10 subsamples used exactly once as the validation data, then the results are averaged. The 10-fold cross-validation method outperforms the traditional repeated random sub-sampling in that all observations are used for both training and validation,

and each observation is used for validation exactly once. The variable that used for the validation is the True Positive Rate (TPR) that can be defined as:

$$TPR = \frac{\text{number of true positive}}{\text{total number of samples}}$$

TABLE I
DATASET OF VEHICLES' IMAGES USED IN THE EXPERIMENTS

Category	Small	Medium			Large
		PU	SUV	VAN	
Subtotal	1980	1250	280	620	800
Total	4930				

The proposed vehicle classification approaches have been implemented with a Matlab software package to facilitate obtaining the results. Matlab (matrix laboratory) is a high-level programming language and interactive environment for numerical computation, visualization, and programming, developed by MathWorks. Using Matlab enables exploring multiple approaches while reaching a faster solution than traditional programming languages (such as C/C++ or Java). Matlabtoolbox includes many algorithms for feature extraction, object detection, object tracking, video processing, and video analysis and more [20].

Those recognition rates (TPRs) for multi-class and intra-class classification tasks are listed in Tables II and III respectively. The highlighted cells represent the highest recognition rates for different classes.

TABLE II
THE TPR FOR MULTI-CLASS CLASSIFICATION USING GEOMETRIC-BASED AND APPEARANCE-BASED APPROACHES

	Geometric-based approach							Appearance-based approach		
	Width	Length	Height	Width & Length	Width & Height	Length & Height	Width & Length & Height	SIFT	SURF	SIFT & SURF
Small	0.780	0.975	0.846	0.970	0.995	0.990	0.992	0.955	0.614	0.947
Medium	0.284	0.198	0.919	0.244	0.930	0.714	0.879	0.963	0.826	0.967
Large	0.969	0.838	0.969	0.875	0.975	0.956	0.956	0.956	0.613	0.950
Total	0.594	0.614	0.898	0.638	0.963	0.864	0.937	0.958	0.706	0.956
Average*	0.678	0.670	0.911	0.696	0.967	0.887	0.943	0.958	0.684	0.955

*un-weighted average

TABLE III
THE TPR FOR MULTI-CLASS CLASSIFICATION USING GEOMETRIC-BASED AND APPEARANCE-BASED APPROACHES

	Geometric-based approach							Appearance-based approach		
	Width	Length	Height	Width & Length	Width & Height	Length & Height	Width & Length & Height	SIFT	SURF	SIFT & SURF
PU	0.534	0.815	0.896	0.490	0.434	0.940	0.816	0.968	0.852	0.960
SUV	0.540	0.507	0.179	0.684	0.738	0.196	0.570	0.786	0.196	0.696
VAN	0.148	0.336	0.452	0.157	0.148	0.403	0.187	0.935	0.718	0.976
Total*	0.423	0.640	0.674	0.446	0.422	0.688	0.632	0.935	0.728	0.930
Average*	0.407	0.553	0.509	0.443	0.440	0.513	0.524	0.896	0.589	0.877

*un-weighted average

From Table II, "Width & Height geometric attribute" and "SIFT appearance attribute" have the highest TPRs for multi-class classification, with an un-weighted average TPR of 0.967 and 0.958 respectively. The TPRs for (Small, Medium, Large) classes using "Width & Height geometric attribute" and "SIFT appearance attribute" are (0.995, 0.930, 0.975) and

(0.955, 0.963, 0.956) respectively which implies that both "Width & Height geometric attribute" and "SIFT appearance attribute" can be practically applied to multi-class classification systems, with high recognition accuracies. From Table III, "SIFT appearance attribute" has the highest TPRs almost for all classes in the intra-class classification, giving an

un-weighted average TPR of 0.896. This means that “SIFT appearance attribute” can extract richer information from vehicle images and obtain more discriminative recognition rate than what geometric attributes do. The resulting TPR of using “SIFT appearance attribute” for (PU, SUV, VAN) classes are (0.968, 0.786, 0.935) respectively.

Moreover, the boxplots in Figs. 4 and 5 show the stability and accuracy of TPR for 10 runs for multi-class and intra-class classifications respectively, for the best performance of geometric-based and appearance-based approaches. In Fig. 4, the mean and standard deviation of TPR for multi-class classification using geometric-based and appearance-based approaches are (0.967 ± 0.006) and (0.958 ± 0.007) respectively. While for intra-class classification, the mean and standard deviation of TPR using geometric-based and appearance-based approaches are (0.553 ± 0.022) and (0.896 ± 0.016) respectively, as shown in Fig. 5.

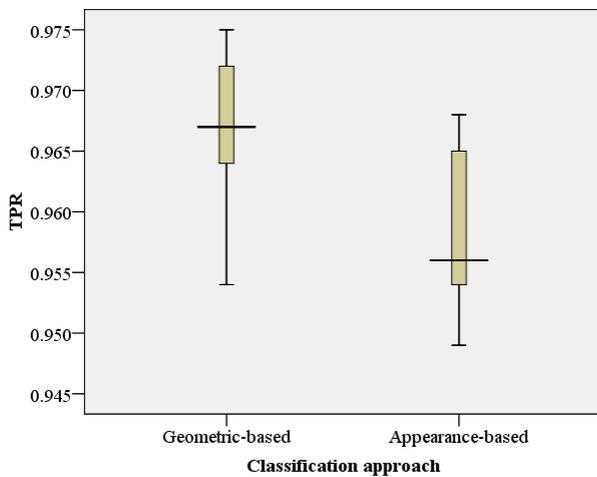


Fig. 4 Boxplot of TPR for multi-class classification using the best performance of geometric-based and appearance-based approaches

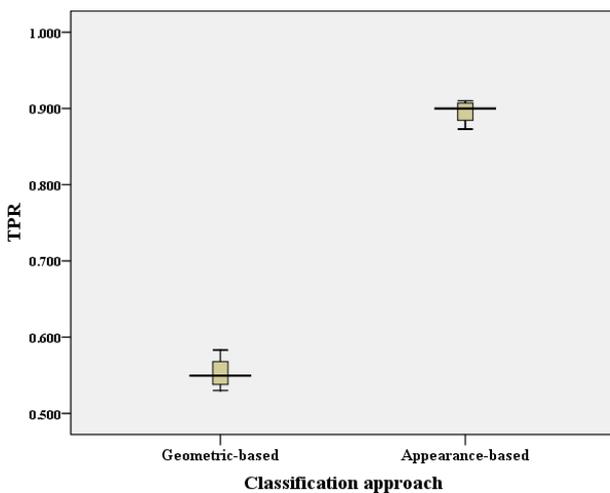


Fig. 5 Boxplot of TPR for intra-class classification using the best performance of geometric-based and appearance-based approaches

The experimental results demonstrate that for multi-class classification both geometric-based approach and appearance based approach bring almost similar recognition rates. However, the appearance-based approach is the best choice for intra-class classification. This means that, for intra-class classification, appearance attributes can extract richer information from vehicle images and obtain more discriminative recognition rate than geometric attributes do. This might be due to the strong similarity in geometry between these classes (PU, SUV, and VAN).

To measure the similarity between classes in the classification process, the confusion matrix is used. In the confusion matrix, the diagonal entries represent correct classifications, whereas the off-diagonal entries represent incorrect ones. The associated confusion matrixes for multi-class (using geometric-based approach) and intra-class classifications (using appearance-based approach) are given in Tables IV and V respectively.

TABLE IV
CONFUSION MATRIX FOR MULTI-CLASS CLASSIFICATION USING GEOMETRIC-BASED APPROACH

	Small	Medium	Large
Small	197	1	0
Medium	14	200	1
Large	0	2	78

TABLE V
CONFUSION MATRIX FOR INTRA-CLASS CLASSIFICATION USING APPEARANCE-BASED APPROACH

	PU	SUV	VAN
PU	121	2	2
SUV	4	22	2
VAN	2	2	58

VII. CONCLUSIONS

In this work, two classification approaches are presented: 1) geometric-based approach and 2) appearance-based approach. The two classification approaches are used for two tasks: multi-class and intra-class vehicle classifications. In the multi-class vehicle classification task, vehicles are categorized into three main classes: small, medium, and large size vehicles. Whereas in the intra-class vehicle classification task, the medium size vehicles are categorized into intra classes (sub classes): pickup (PU), sport utility vehicle (SUV), and van (VAN). 10-fold cross-validation technique is used to evaluate the performance of our proposed classification approaches for the two tasks (multi-class and intra-class), using a large dataset.

The proposed geometric-based approach is simpler than the appearance-based one. The simplicity comes from two issues: Firstly, the extracted features are simple dimensional measurements with lower computational cost and storage requirements than what the appearance-based features require. Secondly, the extracted features (simple dimensional measurements) are fed directly to the classifier without a need to feature representations yet the appearance-based features need such representations.

Several remarks can be made from the experimental results. First, “Width & Height geometric attribute” and “SIFT appearance attribute” bring almost similar recognition rates for multi-class classification which implies that both geometric-based and appearance-based approaches can be practically used for multi-class classification, with high recognition accuracies. Second, “SIFT appearance attribute” is the best choice for intra-class classification. This means that, for intra-class classification, appearance-based approach can extract richer information from vehicle images and obtain more discriminative recognition rate than what geometric-based approach does. This might be due to the strong similarity in geometry between these classes (PU, SUV, and VAN).

Finally, the presented work distinguishes itself from other classification methods in which: i) both geometric-based and appearance-based attributes are considered, and ii) the proposed classification approaches perform remarkably well in both multi-class and intra-class vehicle classification.

ACKNOWLEDGMENT

The author would like to express her gratitude to Dr. Khaled Hussain, Chairman of Multimedia Department, Faculty of Computers and Information, Assiut University. His valuable advice and recommendations contributed a lot to this research work.

REFERENCES

- [1] Kafai, M. and Bhanu, B., Dynamic Bayesian Networks for Vehicle Classification in Video, *IEEE Transactions On Industrial Informatics*, vol. 8, no. 1, February 2012.
- [2] W.H. Lin, J. Dahlgren, and H. Huo, “An Enhancement to Speed Estimation Using Single Loop Detectors,” *Proceeding of Intelligent Transportation Systems*, 2003, pp. 417–422.
- [3] H.J. Cho and M.T., “A Support Vector Machine Approach To CMOS-Based Radar Signal Processing For Vehicle Classification And Speed Estimation,” *Mathematical and Computer Modelling*, Volume 58, Issues 1–2, July 2013, Pages 438–448.
- [4] E. Dallalzadeh, D.S. Guru, S. Manjunath and M. G.Suraj, “Classification of Moving Vehicles in Traffic Videos, *Advances in Computer Science and Information Technology*,” *Computer Science and Engineering, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering* Volume 85, 2012, pp 211-221
- [5] Z. Chen and T. Ellis, “Multi-shape Descriptor Vehicle Classification for Urban Traffic,” *International Conference on Digital Image Computing: Techniques and Applications*, pp. 456–461, Dec. 2011.
- [6] R.P. Avery, Y. Wang, G.S. Rutherford, “Length-Based Vehicle Classification Using Images from Uncalibrated Video Cameras,” in: *Proceedings of the 7th International IEEE Conference on Intelligent Transportation System*, pp.737-742, 2004.
- [7] G. Zhang, R.P. Avery, Y. Wang, “A Video-Based Vehicle Detection and Classification System for Real-Time Traffic Data Collection Using Uncalibrated Video Cameras,” *Transportation Research Record: Journal of the Transportation Research Board*, 1993: 138-147, 2007.
- [8] G. Moussa and K. Hussain, “Laser Intensity Automatic Vehicle Classification System,” *North American Travel Monitoring Exposition and Conference (NATMEC)*, Washington, DC, USA, August 6-8, 2008.
- [9] X. Ma, W. Eric, and L. Grimson, “Edge-Based Rich Representation for Vehicle Classification”, *Proc. Int. Conf. Computer Vision*, vol. 2, pp. 1185- 1192, 2005.
- [10] L. Zhang, S.Z. Li, X. Yuan, S. Xiang, “Real-Time Object Classification in Video Surveillance Based On Appearance Learning,” in: *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 1-8, 2007.
- [11] B. Morris, M.M Trivedi, “Learning, Modelling and Classification of Vehicle Track Patterns from Live Video”, *IEEE Trans on Intelligent Transport Systems*, vol.9, no.3, pp. 425-437, 2008.
- [12] G. Moussa, “A Reliable Multi-Type Vehicle Classification System,” submitted for publication.
- [13] Buch, N., Velastin, S. A. and Orwell J. 2011, A Review of Computer Vision Techniques for the Analysis of Urban Traffic, *IEEE Transactions On Intelligent Transportation Systems*, Vol. 12, No. 3, September 2011.
- [14] Lowe, D.G. “Distinctive Image Features from Scale-Invariant Keypoints”, *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110-110., 2004.
- [15] H. Bay, T. Tuytelaars, and L. Van Gool. SURF: Speeded up Robust Features. In *ECCV*, 2006.
- [16] Csurka, G. Dance, C. R. Fan, L. Willamowski, J. and Bray C., “Visual Categorization with Bags of Keypoints,” in *Proc. European Conference on Computer Vision (ECCV)*, May 2004
- [17] Sivic, J. Russell, B. C Efron, A. A. Zisserman, A. and Freeman, W. T. “Discovering Objects and Their Location in Images,” in *Proc. IEEE Int. Conf. on Computer Vision (ICCV)*, vol. 1, October 2005, pp. 370–377.
- [18] Creusen I.M., Wijnhoven R.G.J., and With P.H.N. de (2009). Applying Feature Selection Techniques for Visual Dictionary Creation in Object Classification. *Proceedings of the 2009 International Conference on Image Processing, Computer Vision and Pattern Recognition*, 13-16 July 2009, Las Vegas, Nevada, (2, pp. 722-727).
- [19] G. J. McLachlan, K-A Do, C. Ambrose, *Analyzing Microarray Gene Expression Data*. Wiley, 2004.
- [20] Matlab 2013, *Computer Vision System Toolbox User’s Guide*, The Math Works, Inc. http://www.mathworks.com/help/pdf_doc/vision/vision Ug.pdf, last access January 2014.