
This item was submitted to [Loughborough's Research Repository](#) by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

Vehicle make and model recognition in CCTV footage

PLEASE CITE THE PUBLISHED VERSION

<http://dx.doi.org/10.1109/ICDSP.2013.6622720>

PUBLISHER

© IEEE

VERSION

AM (Accepted Manuscript)

PUBLISHER STATEMENT

This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Saravi, Sara, and Eran A. Edirisinghe. 2019. "Vehicle Make and Model Recognition in CCTV Footage". figshare. <https://hdl.handle.net/2134/24425>.

Vehicle Make and Model Recognition in CCTV Footage

Sara Saravi

Department of Computer Science
Loughborough University
Loughborough, United Kingdom
s.saravi@lboro.ac.uk

Eran A. Edirisinghe

Department of Computer Science
Loughborough University
Loughborough, United Kingdom
e.a.edirisinghe@lboro.ac.uk

Abstract— This paper presents a novel approach to Vehicle Make & Model Recognition in CCTV video footage. CPD is used to effectively remove skew of vehicles detected as CCTV cameras are not specifically configured for the VMMR task and may capture vehicles at different approaching angles. Also a novel ROI segmentation is proposed. A LESH feature based approach is used for vehicle make and model recognition with the novelty that temporal processing is used to improve reliability. A number of further algorithms are used to maximize the reliability of the final outcome. Experimental results are provided to prove that the proposed system demonstrates an accuracy in excess of 95% when tested in real CCTV footage with no prior camera calibration.

Keywords—VMMR; CCTV Footage; Make and Model Recognition; CPD; LESH; Local Energy Shape Histogram; Coherent Point Drift

I. INTRODUCTION

Several vehicle monitoring and security systems are based on the automated number plate recognition (ANPR). One of the available ways of circumventing monitoring and security systems that are based on ANPR is number plate forgery. ANPR systems are not sufficient alone to ensure proper security. One way to solve this problem would be to augment existing security systems that use ANPR with a vehicle make and model recognition (VMMR) system. The VMMR system could also provide improved surveillance and tracking in the event of crime. With a VMMR system, the amount of footage to be watched could be reduced to the few places where a particular make and model has been discovered thereby speeding up the capture of the criminal. The classification of vehicles has been a focus of interest in the past few years. Many techniques have been proposed in this area. [1] proposed the use of two dimensional linear discriminant analysis called 2D-LDA on image pixels to generate their feature vector. They extracted a static region of interest (ROI) from the vehicle's front view using the location of the license plate as reference. The make and model of the training image with the minimum distance to the test image, is assigned as the make and model of the test image. [2] presented a comparative study of the performance of Fourier, Wavelet, and Curvelet transform features for VMMR. The best recognition rates were achieved using Curvelet transform coefficients as features. [3] used Curvelet transform features with a Support Vector Machine

(SVM) in another proposed solution. [4] proposed Contourlet transform features in conjunction with SVM classifier. [5] proposed an improvement to the technique proposed by [4]. They proposed the use of a localized Contourlet feature extraction technique as opposed to the use of standard deviations of the Contourlet coefficients. They also used a SVM for classification. [6] proposed a new technique which detects salient regions called Patches. A local description for each patch was extracted using Local Energy Shape Histogram. [7] proposed a novel hierarchical technique by separating the make recognition and model recognition modules with the help of Probabilistic Neural Network. In much of the research conducted in VMMR, high quality static images or videos are used and the vehicles to be recognized are manually cropped out, tested, and results from this cropped datasets are provided. The goal of this work is to develop vehicle make and model system that works with low quality CCTV footage. The system should accept video data, extract the vehicles from the video, and return the make and model of the extracted vehicles. The performance of this complete system will be evaluated.

II. PROPOSED SYSTEM

The proposed VMMR system is designed to be used with video and not still images as in much of the previous research; hence, it is designed to take advantage of the temporal resolution of videos. This system will take the feature-based approach to solving the VMMR problem. From the literature, feature-based approaches have proven to be very robust having given a minimum classification accuracy of 59%, therefore a 59% classification rate will be considered as current system's success threshold; any accuracy rate above this minimum would imply proposed system can match some proposed systems. The proposed VMMR system can be divided into 5 stages. The stages are illustrated in Fig. 1.

A. Automatic Licence Number Plate (ALNP) Detection

This stage of the VMMR system is essential because it is used for the selection of the region of interest. The Segmentation of an appropriate region of interest for VMMR is a non-trivial issue due to the high variability in the design of vehicles. A solution to this segmentation problem is to select a region around a structure of reference that is common for all vehicles. The number plate is typically used as a structure of

reference; It is an excellent choice as every road worthy vehicle should have a number plate, and they are mostly rectangular in shape and located at the same position on vehicles (at the centre, and near bottom) with a few exceptions.

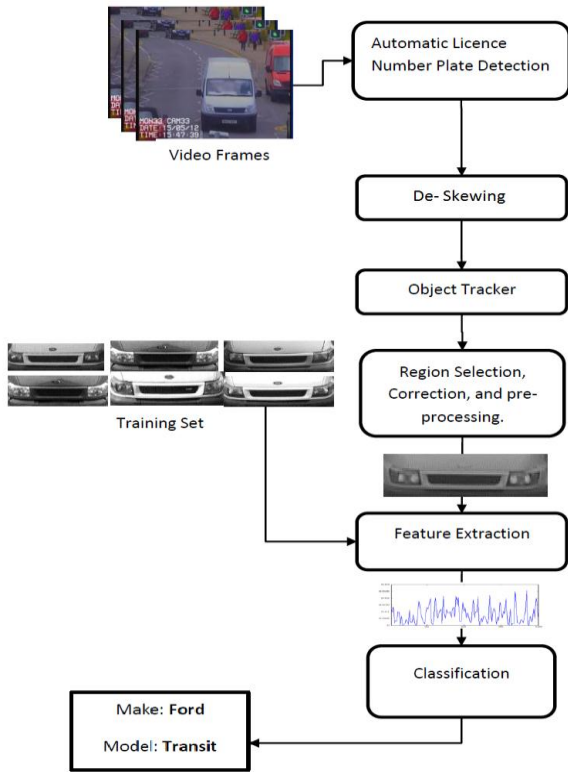


Figure 1. The Structure of the proposed VMMR System.

In the Automatic Licence Number Plate (ALNP) detection stage, the image of the vehicle is processed, the frame is scanned for the shapes and margins of connected components which are then processed based on template matching to eliminate the false ROI. If the size of region is smaller or larger than a specific threshold, it will be classed as false region and is discarded from processing. Also if the aspect ratio of any region is less than or greater than a certain threshold, then that region is also discarded for further processing. The remaining candidate regions are then checked based on their texture similarity to license plate-like areas. Then locations and the measurements of the four corners of the detected number plates are returned [8]. Each detected number plate is considered as the location of a vehicle which must be recognized. The disadvantage of using the number plate as a reference structure is that when the number plate is occluded the vehicle will not be detected. In the above number plate detection stage, once the algorithm locates a rectangular region, if it appears either skewed or rotated on the 2D image plane the rectangle obtained would be the one that contains the four edges of the rectangle. The lines connecting the two upper or two bottom corners of the number plate may not be horizontal, and the lines joining the two left-side and two right-side points may not be vertical (Fig. 2(b)). Under this condition if the remaining stages of the vehicle segmentation are followed the segmented area will not

symmetrically include the vehicle. This may lead to significant errors in VMMR rates. Therefore a de-skewing stage is introduced as detailed below.

B. De-Skewing

De-skewing is a very crucial part of the MMR system. Due to the road gradient and position of CCTV cameras, cars coming via a curve road (bend) or lopsided road will appear at a slight angle (skewed) which will result in a wrong detection: as the front view of the car will not appear straight or symmetric. Hence, correcting the angle of a car at an early stage of the VMMR system will benefit the car detection and recognitions. The De-Skewing algorithm which is proposed to tackle this problem is explained below.

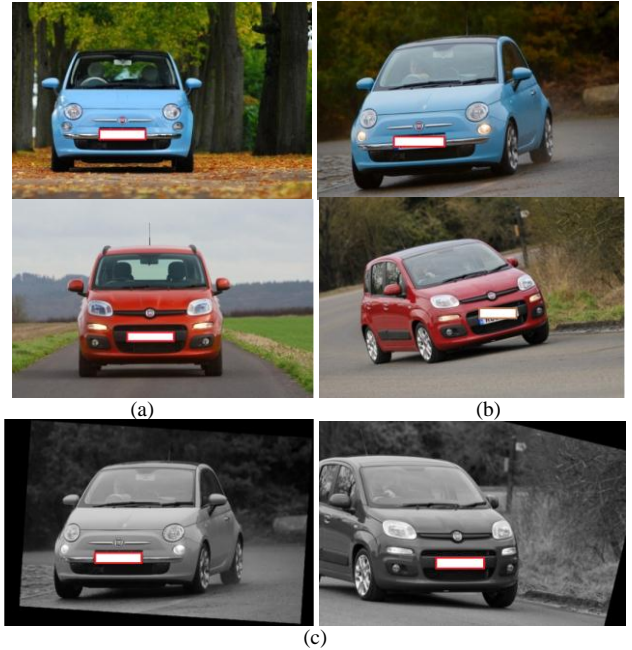


Figure 2. (a) Frontal view of template car, (b) Car approaching at an angle, (c) After registration

Coherent Point Drift (CPD) [9], is a probabilistic technique for non-rigid registration of point sets. The CPD technique concurrently finds the non-rigid transformation and the correspondence amongst two point sets using motion coherence, regardless of the transformation model. This technique can estimate complex non-linear non-rigid transformations and is strong in the presence of outliers and misplaced points. Non-rigid registration assumes that the underlying transformation - required to align point sets - is complex, locally non-linear and the insight is that the points close to one another tend to move coherently. A template of a car approaching the camera at a right angle to the camera plane is initially selected. (Fig 2(a)). Then the width and height of the detected number plates are returned from ALNP stage. Using the position of four corners of the number plate in video frame obtained from the ALNP detection stage (Fig. 2 (b)), and template point positions, the two images are subsequently registered using the CPD algorithm. As the result of registration, all points in video frame will be rotated to resolve angular mismatch, and the car which appeared to be

skewed in its original stage will be de-skewed and be ready for processing by the subsequent stages as described in the following sections.

C. Object Tracker

Object Tracking is the process of following an object through a sequence of images [10]. This stage has not been included in previous VMMR research; but it is a vital stage in the proposed VMMR system for the following reasons: With the object tracker, the system is able to maintain some information on each vehicle across multiple frames and not process each of the detections from every frame as a new vehicle. In the event that the number plate detection fails for a frame, it is possible to use previous predictions to predict the likely location of the number plate allowing continuity of detection. The object tracker could be used to reduce the number of false detections, if a certain number of consecutive detections are used to validate the object before it is tracked. The inputs to the object tracker are the measurements are received from the De-skewing section which are four corners of the corrected licence number plate. Using these measurements the centre location of the number plates are calculated and tracked. The object tracker is composed of two sub stages of Data Association and Kalman Filtering.

D. Region of Interest Selection, Correction, and Pre-Processing

The Region of Interest could be any part of a vehicle that contains sufficient unique details (or features) to distinguish that vehicle's make and model from that of others.

1) Region of Interest Selection

A region of interest (ROI) is selected from the frontal view of the vehicle using the number plate as a reference point. The ROI is defined as $Width = 2.6 \times W$, $Height = 0.32 \times Width$, where W is the width of the detected number plate and $Width$ and $Height$ are the width and height of the Region of Interest respectively (Fig. 3). The ROI is selected from the area above the number plate: nothing is selected from the area below the number plate. The Fig. 3 shows the ROI measurement. The ROI measurements are designed to encapsulate the front grill, manufacturer's logo, and headlamps. It is important to note that the assumption used for the ROI does not hold for all vehicles; if the number plate of the vehicle is not positioned at the center of the vehicle the assumption fails - some vehicles have their number plate closer to the bottom of the vehicle. In this scenario, the ROI selection will not completely encapsulate the headlamp and manufacturer's logo (Peugeot 207). But the detailed grill proved sufficient for recognition. When working with vehicle categories that are generally large (e.g. trucks) or small, the ROI will either fail to properly select the whole vehicle or select too large a region (Fig. 4). In this case, an attempt is made to adjust the ROI using a region correction algorithm. It is also essential that the vehicle has not more than 2 degrees variation from the horizontal plane, otherwise a poor region will be selected which could cause the known vehicle to be unrecognized. Considering the fact that the de-skewing algorithm described in section II.B has been already applied this assumption can be considered to be valid in practice.



Figure 3. The region of interest measurements



Figure 4. Scenarios in which the region selection might fail.

In this stage, the assumption is made that the vehicles are moving. The region correction algorithm is explained as follows. First using the current frame and the previous frame(s), Motion segmentation is used to find moving object(s). Then by using the measurements and location of the region selected in the ROI, if the segmented object obtained from step 1 is wider or smaller than the region selected, the difference (error) is computed. This difference is used to adjust (increase or decrease) the width of the region selected and a new height is computed as it was mentioned earlier. Motion segmentation fails when the object is stationary. This will also cause the region correction algorithm to reduce the region excessively. Finally as a solution to this a threshold is defined; if the width is reduced below this threshold, it is assumed the motion segmentation has failed and the original region of interest is used.

2) Pre-Processing

After the region correction, the Region of Interest is cropped out automatically, the cropped out image is resized to 64×152 pixels (i.e. normalised), and the same pre-processing is manually performed on all images in the training set which is used in the classification stage. This cropped image is converted to a grayscale image and is passed onto the feature extraction module.

E. Feature Extraction

Feature Extraction is a very important stage in any VMMR system. Features need to be distinctive enough to distinguish between a large number of vehicle makes and models; and at the same time, they also need to be invariant to changes in scale, colour, and illumination. These features should be able to model the shape of the local components (grill, headlamps) in the region of interest. The proposed system makes use of the Local Energy Shape Histogram (LESH) first introduced by [11], which performed satisfactorily when compared with other feature descriptors. It postulated that features are perceived at points in an image where the local frequency components have maximum phase congruency. The Local Energy indicates the

corners, contours, or edges of underlying shape in an image. The LESH feature vector is extracted from the region of interest, and normalized between [0, 1].

F. Classification

In this stage, features extracted from the region of interest are passed to a classifier which would determine the make and model. The support vector machine is a classifier which has proven very reliable in previous VMMR research; therefore it has been chosen for this system. The proposed system uses multiple frames from the video in the classification stage. When a vehicle is detected, features are extracted from them and classified; the resulting label is stored. This is repeated on a pre-specified number of frames for the same object; after which all the predicted labels for that object are used in a Winner-Takes-All voting and the vehicle is assigned the label with the majority of the votes. A random label is selected in the event of any ties.

1) Training the SVM

When training a SVM, the first step is to pre-process and crop the training images. Then LESH features are extracted from every image in the dataset and scaled appropriately. The next step is the selection of a kernel function for the SVM. The RBF kernel function was selected as the kernel function in the final implementation. The next stage is parameter selection. The different kernel functions have their different parameters, and it is important that the right parameters are selected to obtain the best possible accuracy. The parameters γ and C for the RBF kernel are selected using k-fold cross validation. The SVM is trained using the selected parameters and a data structure is generated which contains the parameters, support vectors, and other important data. The One versus All SVM is used in this implementation to provide best performance. A SVM will assign an input feature to its closest matching class but in a real world VMMR system this is undesirable because an unknown vehicle could be classified as any vehicle in the dataset. As a solution to this each class from the training set is assigned a threshold; when an unknown feature is classified, and the decision values returned from the SVM are below the threshold for that class, the feature is re-classified as an unknown vehicle. The decision thresholds are selected empirically.

III. EXPERIMENTAL RESULTS AND ANALYSIS

A number of experiments were performed to test and evaluate the performance of the implemented VMMR System. The results from each of these experiments are presented and discussed in this section.

A. Vehicle Database

The vehicle database used for this work consists of a training dataset of static images and the testing dataset of videos. These are discussed briefly below.

1) Training Dataset

The training dataset used comprised of frontal views of 22 different vehicle makes, with 7-10 images per make and model, making a total of 196 images (Fig. 5). Some of the images are taken in bright and clear conditions with about 5 degrees

variation from the horizontal plane and some are highly De-skewed which were tested with proposed De-skewing algorithm. The database contains some difficult types like the Renault Clio campus sport and Renault Megane which both have similar grill and headlights. The database also contains a few trucks and a jeep e.g. BMW X3, Ford Transit etc.

2) Testing Dataset

The Testing dataset includes 3 CCTV footage. Together these videos contain 24 vehicles that have the same make and model as the ones in the training dataset and also some unknown vehicles. The resolution is 720 x 576 pixels, frame rate is 25fps, duration of each video is between 1-2 minutes, and all the videos are stored in AVI file format (Fig. 6).

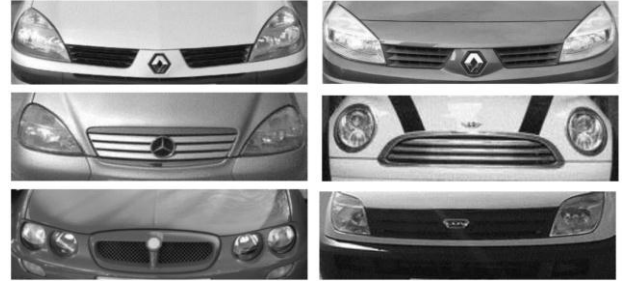


Figure 5. Samples of the images from the Training dataset showing 6 different vehicle make and models.

The testing dataset contains some of the challenges that a real world VMMR System might encounter. For example vehicles with their headlamps switched on, a vehicle which had stickers on the bonnet, and vehicle body or headlamps reflecting the sunlight.

3) Experiments

A number of experiments have been carried out to evaluate the proposed system; the results of these experiments are provided in the section that follows. All experiments are conducted using the testing dataset described in section III.A.2. To reduce the computational cost of running the system, a search area is defined; vehicles will be recognized only within that search area. It is important to note the following about selection of a search area:

- Results could differ depending on the location of the search area selected.
- The further the search area is from the camera, the poorer the recognition rate; this is because at some point the scale of the vehicles will be too small.

With this in mind, each experiment is performed using a fixed search area. The measurement of this search area is specified as the default search area in the implemented system, this search area will be used if no search area is manually specified, and this is done so each experiment can be replicated. The location of the search area is defined using the x and y coordinates of the top left edge of the search area and the x and y coordinates of the bottom right edge of the search area. The number of frames used for the classification is set at five for all experiments except otherwise stated. The accuracy

of each experiment is calculated as the ratio of the number of correct recognition to the total number of vehicles in the test dataset.



Figure 6. Samples of the video frames from the testing dataset showing 4 different vehicle make and models.

4) Experimental Results

The performance of the proposed system is presented in this section. The proposed system achieved a classification accuracy of 95.83%; and correct recognition on all classes except the Mercedes A140 (in the video the Mercedes A140 stops as soon as it enters the search area). The region correction fails since the vehicle is stationary and it is classified as a Vauxhall Meriva Breeze. Although the experiments are performed on a relative small number of cars, it is important to note that most VMMR experiments are performed using high resolution images but these experiments are performed using low resolution CCTV video and still acquired high recognition rates.

5) Alternate Features and Classifiers

In these experiments, the performance of 3 features vectors is compared against the proposed feature vector, and 4 classifiers are also tested. These experiments are done to check if a better performing feature extraction method or classifier could be selected. The features compared in these experiments are: 1) Sobel Response 2) Histogram of gradients 3) LESH-128 (4x4 partitions) 4) LESH-512 (8x8 partitions). The classifiers used in these experiment are: 1) KNN (K=1) 2) KNN (K=3) 3) SVM (One versus One) 4) SVM (One versus All). The classification rates for each feature and classifier combination is presented in the Table I. The original LESH (LESH-128) uses 4x4 partitions generating a 128-dimensional vector, and the modified LESH (LESH-512) is generated using 8x8 partitions generating a 512-dimensional vector. Histogram of gradients (HOG) [12] is a feature descriptor originally designed for pedestrian detection, but since development has been used in other object recognition tasks. It is invariant to both geometric transformation and changes in illumination. The HOG features are generated over 64 rectangular cells and using 9 bins thereby making a 576-dimensiional feature vector. The Sobel response was extracted by convolving the region of interest with both the horizontal and vertical Sobel convolution kernels, and then calculating the magnitude. The magnitude at each pixel in the ROI was used as the feature vector. The worst

performing feature vector was the Sobel response; it returned the lowest classification rate in comparison with the other features irrespective of the classifier used. This was expected as gradient based edge detectors are typically sensitive to noise. Classification rate of different features

TABLE I. CLASSIFICATION RATE OF DIFFERENT FEATURES USING MULTIPLE CLASSIFIERS.

	LESH-128	LESH-512	HOG	Sobel Response
<i>KNN (K=1)</i>	70.83	70.83	62.50	25.00
<i>KNN (K=3)</i>	70.83	70.83	62.50	33.33
<i>One Versus One SVM</i>	70.83	91.67	58.33	54.17
<i>One Versus All SVM</i>	75.00	95.83	66.67	62.50

The best classification accuracy is obtained by using the proposed feature vector (LESH-512) with the proposed classifier (One versus All SVM). This is to be expected because the LESH uses the local energy model which is able to extract edges, contours, and corners at the same time thus making it a very robust feature and using 8x8 partitions makes LESH-512 more discriminating. The One versus One SVM perform closely to the One versus All SVM and could outperform it if the number of classes is increased significantly. It is important to note that the results of these experiments could vary considerable using another dataset. The tested features like HOG have a number of parameters that could also influence the results. A few parameters were tested and the best results are provided.

6) Number of Frames for Classification

The experiments are aimed at finding the optimal number of frames for the classification process. It should be mentioned that using smaller or larger number of frames could lead to misclassifications. It is also possible that the vehicle could leave the search area without getting classified if a very large number is specified.

For the testing dataset used, 1, 5, or 7 Frames gave the best accuracy. At 13 Frames, some vehicles began to exit the search area unclassified causing the rate of recognition to decline. For a real world system, the use of a single frame for classification might be less accurate than using 5 or 7 frames but using the available dataset it works well (Fig. 7). It should be noted that the number of frames that is used to make the best judgment in VMMR will depend on many practical factors affecting the video capture setup, quality of footage, the path taken by the vehicle, size of vehicle, location of analysis window etc. However the above experiments reveal that VMMR in video, attempted for the first time, within the research context of this paper, provides a number of additional advantages as compared to processing single frames to make judgment.

7) Effect of Region Correction

The region correction algorithm implemented used approximate median background subtraction, which is a complex algorithm in terms of the memory expended. In these experiments, the region correction module is deactivated and the system is tested. The significance of these experiments is to show experimentally the importance of the region correction

stage. The region correction was deactivated from the settings dialog and the testing dataset was run through the system. The system achieved a 66.67% without the region correction having 8 vehicles wrongly recognized amongst these were the BMW X3, LDV Maxus, Ford Transit, Vauxhall Astra, Renault Megane, Ford Focus (2002), Mercedes Sprinter, and Mercedes A140. It is expected that all trucks and jeeps in the testing dataset are misclassified because a smaller region of interest was selected but the Volkswagen Tr-porter is still classified correctly. The ALNP detection implementation sometimes fails to return the correct width of the number plate which is essential for the region selection; this leads to a badly selected region and eventually leads to the smaller vehicles being misclassified. These experiments proves that the region correction is significant not just for larger vehicles but for some of the smaller vehicles also.

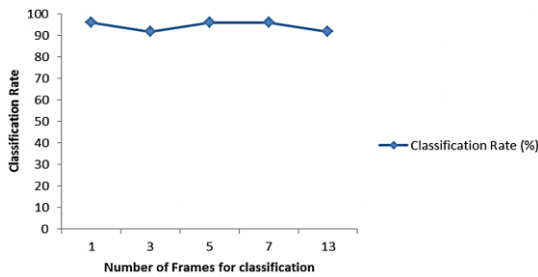


Figure 7. Classification results from the proposed system using varied number of frame(s) for classification.

8) Real Time Performance

In these experiments, performance of the proposed system in real time is tested. This is essential for a real world system. Although, the system is implemented in MATLAB which is typically slow, it should still use a relatively small amount of time per frame. This would prove that when the system is re-written in a language that can run quicker, it could achieve real-time speed. The difference between the start and end of processing gives the time it took to process that frame.

The proposed system used an average of 1.011 seconds per frame when a vehicle was in the search area. This is far from the desired 25 frames per second required for real-time performance. It is important to note that the average time gained from these experiments would vary depending on the CPU speed and memory of the computer used to run the experiments. It is further worth mentioning that most CCTV cameras capture at 5 fps and processing a frame within 0.2 S is thus the ultimate goal towards achieving real time performance.

IV. CONCLUSION

Vehicle Make and Model Recognition is a developing area of research with multiple applications. It is mainly investigated as a potential upgrade to automatic number plate recognition systems used in vehicle monitoring and security systems, due to number plate cloning becoming more common. It could also be applied in vehicle marketing research and in more robust vehicle tracking.

In this paper, a complete VMMR System has been presented. The proposed VMMR system accepts a video and returns the makes and model of the vehicles detected in that video. The VMMR system began by automatically detecting a number plate from a video frame, the location of the number plate was used subsequently to select a region of interest. It was important to treat the same vehicle across multiple frames as a single vehicle; hence an object tracker was used to track the detected number plate across multiple frames. A static region of interest was selected above the number plate and motion segmentation was used to adjust this static region if necessary. LESH Features were extracted from the region of interest and a SVM was used to classify the feature, assigning it a label. The final label was decided using multiple frames from one vehicle object. A number of experiments were carried out on the proposed system using a test set that comprised of low quality CCTV videos. The proposed system achieved a best classification rate of 95.83%. Although the proposed system shows promising performance much can be done to improve the system.

REFERENCES

- [1] Zafar, I., E. A. Edirisinghe, S. Acar, and H. E. Bez. (2007). Two Dimensional Statistical Linear Discriminant Analysis for Real-Time Robust Vehicle Type Recognition. *In Proceedings SPIE, Electronic Imaging*. 6496(649602).
- [2] Kazemi, F. M., S. Samadi, H.R. Poorreza, and M. Akbarzadeh-t. 2007a. Vehicle Recognition Based on Fourier, Wavelet and Curvelet Transforms - a Comparative Study. *In: Information Technology, 2007. ITNG '07. Fourth International Conference on*. Las Vegas, 939-940.
- [3] Kazemi, F. M., S. Samadi, H. R. Poorreza, and M. Akbarzadeh-t. 2007b. Vehicle Recognition Using Curvelet Transform and SVM. *In: Information Technology, 2007. ITNG '07. Fourth International Conference on*. Las Vegas, 516-521.
- [4] Rahati, S., R. Moravejian, E. Mohamad, and F. Mohamad. (2008). Vehicle Recognition Using Contourlet Transform and SVM. *In: Information Technology: New Generations, 2008. ITNG 2008. Fifth International Conference on*. Las Vegas, 894-898.
- [5] Zafar, I., E. A. Edirisinghe, and S. Acar. (2009). Localized contourlet features in vehicle make and model recognition. *In Proceedings SPIE, Electronic Imaging*. 7251(725105).
- [6] Sarfraz, M. S. and M. H. Khan. (2011). A Probabilistic Framework for Patch based Vehicle Type Recognition. *In: Leonid MESTETSKIY and José BRAZ, (eds). VISAPP, SciTePress*, 358-363.
- [7] Pearce, G. and N. Pears. (2011). Automatic make and model recognition from frontal images of cars. *In: Advanced Video and Signal-Based Surveillance (AVSS), 2011 8th IEEE International Conference on*. Klagenfurt, 373 - 378.
- [8] M. S. Sarfraz, et al., Real-time automatic license plate recognition for CCTV forensic applications, *Journal of Real-Time Image Processing*. DOI 10.1007/s11554-011-0232-7
- [9] M. S. Sarfraz, et al., Real-time automatic license plate recognition for CCTV forensic applications, *Journal of Real-Time Image Processing*. DOI 10.1007/s11554-011-0232-7
- [10] Yilmaz, A., O. Javed, and M. Shah. (2006). Object tracking: A survey. *ACM Computing Surveys*. 38(4).
- [11] Sarfraz, M. S. and O. Hellwich. (2008). An Efficient Front-end Facial Pose Estimation System for Face Recognition. *In International Journal of Pattern Recognition and Image Analysis, distributed by Springer*. 18(3), 434-441.
- [12] Dalal, N. and B. Triggs. (2005). Histograms of Oriented Gradients for Human Detection. *In: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*., 886-893.