Vision Based Road Traffic Density Estimation and Vehicle Classification for Stationary and Moving Traffic Scenes during Daytime

Daniel Arani Osuto[1*], Heywood Ouma Absaloms[2] and Edward Ng'ang'a Ndungu[1]

[1]*Department of Telecommunication and Information Engineering, Jomo Kenyatta University of Agriculture and Technology, P.O. BOX 62000-00200, NAIROBI, KENYA.*

[2]*Department of Electrical and Information Engineering, University of Nairobi, P.O. BOX 30197-00100, NAIROBI, KENYA.*

*Corresponding Author – Email: danielosuto@gmail.com

Abstract Automatic road traffic density estimation and vehicle classification are very important aspects of today's Intelligent Transportation Systems (ITSs). Traditionally loop sensors have been used for this purpose, but lately vision based systems have been preferred due to their advantages and the problems associated with loop sensors. Many vision based vehicle detection and classification algorithms for free flowing traffic have been proposed. These systems are largely dependent on either motion detection or more generally background modelling and subtraction. There is little reported of traffic scenes with very slowly moving or stationary vehicles for which motion detection based approaches are impractical. This paper presents a novel vision based road traffic density estimation and vehicle classification approach that is independent of motion detection and background modelling and subtraction. It combines selected image processing, computer vision and pattern recognition algorithms to obtain the traffic parameters. The approach is applied to both standstill or slow moving traffic, and free flowing traffic under different illumination conditions during the day. The approach does not require camera calibration, therefore, it can work with already installed video surveillance systems, making it economical and convenient. The algorithm is based on image segmentation using a Laplacian of Gaussian edge detector (LoG), morphological filtering of the edge map objects and classification into small, medium and large vehicles on the basis of size using a nearest centroid minimum distance classifier. The proposed approach can be used for both stationary and fast moving traffic in contrast to motion detection based approaches. The algorithm was implemented in MATLAB R2015a and average detection and classification accuracies of 96.0% and 89.4% respectively were achieved for fast moving traffic, while for slow moving traffic, 82.1% and 83.8% respectively were achieved

Keywords Laplacian of Gaussian edge detector, Road traffic density estimation, Stationary traffic, Vehicle classification.

1. Introduction

Intelligent Transportation Systems (ITSs) have become increasingly popular in the world today. Many important road traffic parameters such as traffic density and types of vehicles on the roads can be obtained automatically by these systems. One aspect of these systems that has attracted much attention among researchers in the last two decades is the use of surveillance videos to obtain the required road traffic parameters. Consequently, many approaches have been proposed [1]. The overwhelming majority of these approaches are dependent on motion detection or background modelling and subtraction to detect vehicles. This limits their application to free flowing traffic scenes or scenes with static backgrounds. In cases where the traditional static background

subtraction is used, changing scene conditions are not factored in [2], and segmentation results may thus not always be reliable. Dynamic background modelling is excellent in handling changing scene conditions to a large extent [2], [3]. Unfortunately, this method cannot be used for stationary traffic, a common problem in the developing world and the main focus of this research. In addition to this, many of the proposed approaches [1] neither consider the possibility of having both pedestrians and vehicles in the same traffic scene, nor performance in different illumination conditions. In this paper, the background modelling and subtraction methods are avoided. Instead, a combination of simple but robust image processing and computer vision algorithms are used to extract the vehicles. This enables the algorithm to effectively handle stationary traffic scenes.

The main novelty of the proposed approach is detection of vehicles using traffic surveillance videos in which the traffic is either slow-moving or at standstill thereby making meaningful background modelling extremely difficult at best or impossible.

2. Literature Review

Vision based vehicle detection and classification has long been explored. Although initial efforts were not so successful, lately much improvement has been achieved. Commercial software for this purpose exist but are dogged with problems such as inability to handle vehicle occlusions from the camera's view [4], [5], limited functionality in severe weather conditions [6], [7], shadows and night detection [8], [9]. Inter-system compatibility is one other drawback associated with today's video analytic algorithms thus severely limiting their deployment, as they do not generally work with already installed hardware unless the hardware is from the same vendor as the algorithms. Open platforms have been formulated [10], but so far remain at the specifications stage and are not yet standardized. Each vendor understands these specifications differently and as a result, the integration of their products remains at the very basic level.

Ambardekar, et al. [11] and, Sivaraman and Trivedi [1] give good general overviews of the state of the art in vehicle detection and classification algorithms. In general, some of the research work done in this area focus on given problems that have long been identified in earlier works while the majority use standard algorithms to develop different approaches for detecting

and classifying vehicles.

Other than the stated problem in the abstract, there are three main problems currently being addressed: occlusions, shadows and different weather conditions.

2.1 Occlusions

Vehicle occlusions have long been identified as a major bottleneck in vision based vehicle detection and classification systems. Even though many challenges still remain, good results for handling partial occlusions have been published [4], [5], and [12]. In [4], a probability-based background extraction and segmentation algorithm was used to detect partially occluded vehicles in a sequence of images by evaluating their convexity and analyzing the occlusion regions before classifying them on the basis of their normalized sizes. The approach showed good ability to handle partial occlusions and classify vehicles. Pang, et al. [5] resolved partial occlusions between two vehicles by estimating the background using a running-average method and then used a texture-based segmentation to obtain shape contours of vehicles that were used as the basis of detecting the occlusions. The approach only failed for very severe occlusion cases. Habibu Rabiu [12] handled occlusions in cluttered urban intersections. He combined background subtraction for detection, the Kalman filter for tracking and a Linear Discriminant Analysis (LDA) classifier for classification with a good degree of success.

2.2 Shadows

Shadows have also been a bottleneck for vision based vehicle detection and classification systems. Shadows cause three main problems for these systems. Firstly, for free flowing traffic, the shadows move with the vehicles and the algorithm can easily 'see' them as independent vehicles. Secondly, the shadows can 'join' adjacent vehicles to make bigger vehicles leading to erroneous counting and classification results; and thirdly, the shadows cause non-uniform illumination in the scene thus making segmentation difficult. Yu, et al. [8] proposed a vehicle tracking and classification system that takes into account size variations and shadows. Vehicle detection was achieved through background subtraction and the Kalman filter was used for tracking. The proposed shadow removal algorithm was based on the assumption that a vehicle can only be in one lane at any given time and that the distance between vehicles in adjacent lanes is uniform throughout such that vehicles in adjacent lanes can be separated optimally using a

straight line. This may not always hold in real world traffic scenes.

2.3 Different Weather Conditions

Severe weather conditions pose a great challenge to visual vehicle detection and classification systems. An excellent system in one condition may fail completely when subjected to different weather conditions. This calls for systems that are capable of adapting to different weather conditions. In [13], an adaptive video-based traffic management system for counting vehicles was developed. The system was able to adapt to changing weather and illumination conditions and partially addressed the problem of occlusions. Buch, et al. [14] performed per frame vehicle detection and classification using 3D models under three different weather conditions: sunny condition, overcast condition and overcast changing to sunny condition. Sunny condition was reported to have classification precision results of 100% followed by overcast condition at 95.6% and overcast changing to sunny condition at 81.2%. Mishra, et al. [15] used background subtraction and blob tracking to detect vehicles and a kernel Support Vector Machine (SVM) classifier to classify vehicles in heterogeneous traffic scenes. Different times of the day were considered and it was shown that the performance started to deteriorate after 4.00 p.m. due to excessive reflections from the road surface. Since the system relied on motion detection, it was noted not to work for 'stop-go traffic'. Finally, Vujovic, et al. [16], conducted a series of experiments under different weather conditions to assess the impact of such conditions on the quality of video surveillance systems. They established that weather conditions should be considered in the development of any video based traffic management system.

From this sample of published works, it is clear that there is no single approach that works well for varied weather scenarios.

2.4 General Approaches

The overwhelming majority of systems that have been proposed use standard image processing, computer vision and pattern recognition algorithms to handle general visual vehicle detection and classification problems. A vast majority of them are dependent on motion detection for their effective functionality, and cannot therefore be adopted for stationary traffic scenes. Kwigizile, et al. [17] used probabilistic neural networks to classify vehicles according to the F-scheme guidelines provided by the Federal Highway Administration (FHWA) in the USA. However, the proposed approach assumed that the probability density function of each class is known a priori. Besides, it is extremely difficult to extract the required data for the F-scheme classification from a natural traffic scene.

In [18], principal component analyses of front and back views of vehicles were used to classify the vehicles into either cars or trucks. Most importantly, no motion information is required to extract the vehicles, although it does not use a natural traffic scene. The proposed method, however, assumes that the backlights of all vehicles are at the same height in the computation of "Eigen-back views" and makes use of a static background in the extraction of vehicles. This assumption may not hold in a natural traffic scene.

Avery, et al. [9] used images from uncalibrated video cameras to count and classify trucks and heavy vehicles on the basis of length. The vehicles were extracted using a dynamic background subtraction method and then their lengths were extracted only when they reached a particular point in the scene while traveling in a straight line. In this way, reliable lengths for classification were obtained. The limitation of this approach is that it cannot work in heterogeneous traffic scenes where the vehicles are not moving in a straight line. Similarly, Pancharatnam and Sonnadara [19] used adaptive background subtraction to detect moving vehicles and tracked them using their bottom coordinates before counting and classifying them on the basis of size into large, medium and small classes. Although good results were reported, similar to [8], the system relies on the assumption that a vehicle will only occupy one lane at a time for its effective performance. It is shown that great errors occur when this is not true.

In [20], a video-based vehicle detection and classification system for real time traffic data collection using uncalibrated video cameras is proposed. The system eliminates the need for complicated camera calibrations. It strikes a good balance between algorithm complexity and effectiveness in real time applications. The paper also notes one other critical limitation of all background modelling and subtraction based algorithms (both static and dynamic) for foreground extraction: they do not account for transient lighting changes in the scene. This is confirmed by the results in [14]. Ince [21] used invariant moments and shadow aware foreground masks to count vehicles and classified them using a perspective projection of the scene geometry. The algorithm was tested on real world data and showed to be computationally efficient.

In this paper, the stated problem of stationary traffic scenes and the problem of shadows shall be addressed.

3. Overview of the Proposed System

A multi-stage vehicle extraction, counting and classification system for handling slow-moving or standstill traffic and free flowing traffic is proposed. It combines selected image processing and computer vision algorithms to obtain the traffic density and uses the nearest centroid minimum distance classifier to classify the vehicles into small, medium and large classes. Fig. 1 illustrates key stages of the algorithm.

First, the vehicles are extracted from the video frames and their negatives using the Laplacian of Gaussian edge detection method and mathematical morphology. The vehicles so obtained are counted and their number used to calculate the traffic density as the number of vehicles per unit area of the road at any given time. The dimensions of the vehicles are also extracted and fed into the classifier for classification. The key stages of the algorithm are explained next.

3.1 Negative transformation

For each grey frame extracted from the video, its negative is computed. This is to ensure that as much relevant edge detail as possible is extracted in the segmentation stage, thus minimizing spurious edge discontinuities. Fig. 2 shows an image from a typical traffic scene and its negative.

(a)   (b)

Fig. 2. (a) Original Frame (b) Negative Frame

3.2 ROI Mask Modeling

One of the extracted frames is used to model a Region of Interest (ROI) polygon. This polygon is ultimately used to mask the processed binary frames so as to limit the counting and classification of vehicles to those found only within the region of interest. The size of this polygon is chosen empirically such that the vehicle intra-class variations are minimized. Fig. 3 shows a traffic scene and the modelled ROI mask for the free flowing

traffic.

(a)   (b)

Fig. 3 (a) Frame (b) ROI Mask

3.3 Top-hat Transformation

The segmentation performance is improved by compensating for non-uniform illumination of the scene using the morphological top-hat transformation prior to the segmentation stage. This is computed as

$$g(x, y) = f(x, y) - (f(x, y) \circ b(x, y)) \qquad (1)$$

Where g(x, y) is the uniform background frame, f(x, y) is the input frame and $f(x, y) \circ b(x, y)$ is the morphological opening of $f(x, y)$ using a structuring element (SE), $b(x, y)$. The size of this structuring element is chosen such that it is larger than any object of interest in the scene so as to avoid deletion of any vehicle in the subtraction process. This transformation also helps to minimize the effects of shadows.

3.4 Image Smoothing and Blurring

The uniform background frame is smoothed using a median filter to remove random noise and then aggressively blurred using a Gaussian filter so as render 'noise' edges into the background and therefore reduce the chances of their detection. This also minimizes the effects of shadows in the traffic scene. Finally, the blurred frame is contrast enhanced linearly so as to emphasize the edges while preserving the mean intensity values of the frames using a contrast stretching algorithm prior to segmentation.
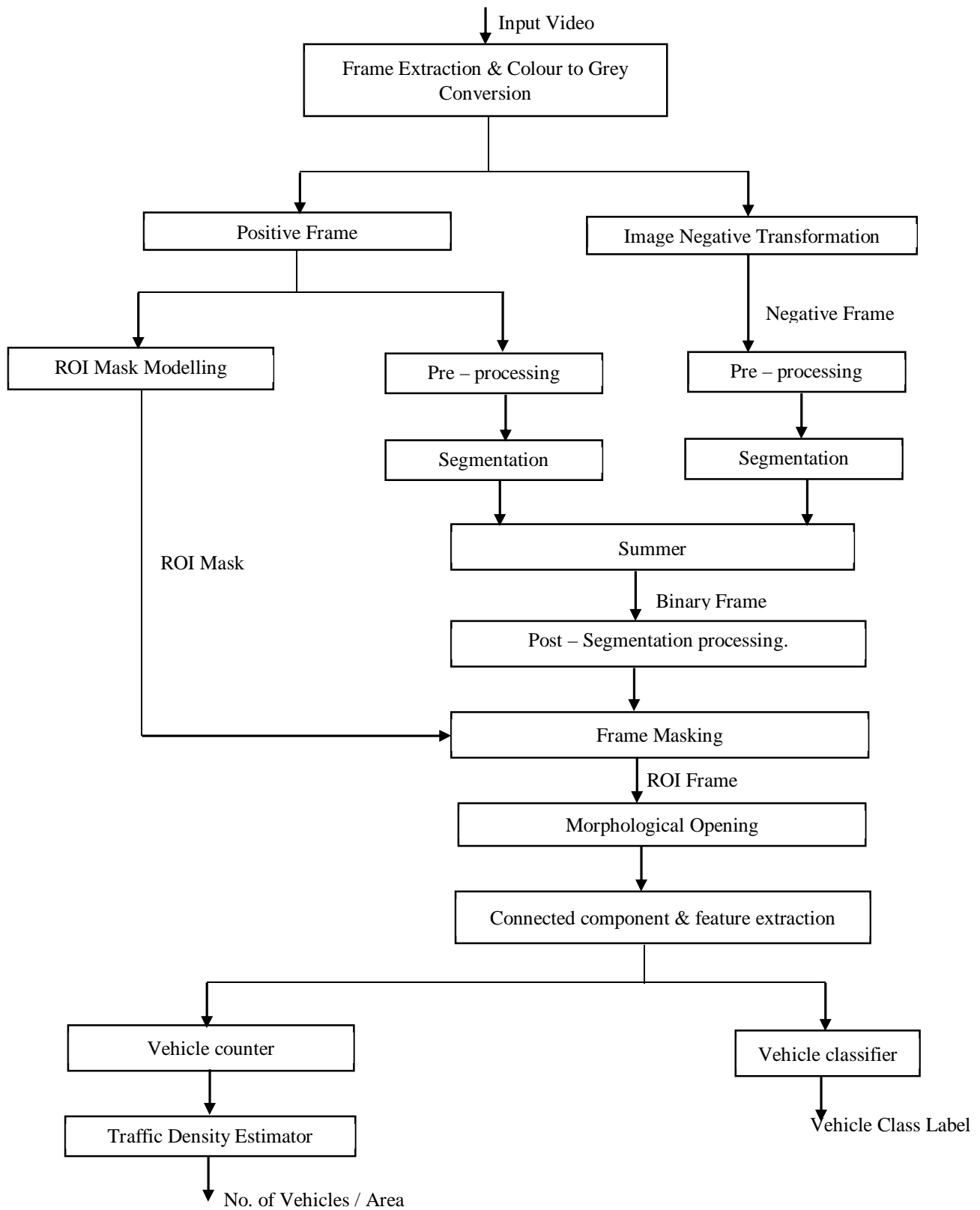
Input Video

Frame Extraction & Colour to Grey Conversion

Positive Frame

Image Negative Transformation

Negative Frame

ROI Mask Modelling

Pre – processing

Pre – processing

Segmentation

Segmentation

Summer

Binary Frame

Post – Segmentation processing.

ROI Mask

Frame Masking

ROI Frame

Morphological Opening

Connected component & feature extraction

Vehicle counter

Vehicle classifier

Traffic Density Estimator

Vehicle Class Label

No. of Vehicles / Area

Fig. 1. Summary of the Algorithm

3.5 Image Segmentation

To extract objects in both the frame and its negative, the Laplacian of a Gaussian (LoG) edge detection method is used due to its excellent edge detection properties and relative simplicity [22]. This preserves generality unlike the trial and error thresholds normally used in many of the reported approaches. The LoG of a two dimensional image is computed as

$$\nabla^2 G(x, y) = \left[\frac{x^2 + y^2 - 2\sigma^2}{\sigma^4} \right] e^{-\frac{x^2+y^2}{2\sigma^2}}$$

(2)

Where ∇^2 is a Laplacian operating on the Gaussian smoothed image, $G(x, y)$ and σ is the standard deviation of the image pixel intensities. Fig. 4 and Fig. 5 show video frames so segmented.

(a) Original frame

(b) Segmented frame

Fig.4. Free flowing traffic scene

This approach enables the system to exploit the fact that shadows are semi-transparent and therefore by appropriately enhancing and segmenting the frames, their effects can be greatly reduced. In this way, the complex and often ineffective shadow removal algorithms are avoided.

(a) Original frame

(b) Segmented frame

Fig. 5. Slow moving traffic scene

3.6 Summation

After segmentation, the two branches are added to eliminate double counts and to ensure that as many objects are detected as possible. This addition is possible since the frame and its negative are spatially registered and therefore the objects which occur simultaneously in both the frame and its negative reinforce each other. The output of the summer give the complete edge map, and therefore the binary image of the frame.

3.7 Post-processing and Feature Extraction

The obtained binary frame is then subjected to morphological filtering. First, the segmented binary frame is closed so as to eliminate any spurious disjoints between connected components. Then the holes in the connected components are filled to ensure that true sizes of objects are used in subsequent stages. Next, a skeletonizing algorithm is run once before pruning the image to get rid of spurious branches that result after segmentation. Then, the processed frame is masked using the modelled ROI mask so as to limit the counting and classification to the objects found in the region of interest only. In this way, objects that are not of interest such as roadside buildings and vehicles moving on lanes that are not of interest such as in Fig. 7(a) are effectively eliminated as shown in Fig. 7(b). Finally, the irrelevant small objects within the ROI such as pedestrians are

deleted using a morphological opening operation. The morphological opening is done after masking the frame so as to ensure that unwanted objects on the border of the region of interest are deleted as well.

The consequence of this processing is that the shapes of the vehicles are not preserved, and therefore, cannot be used for classification. Instead, areas of the bounding boxes of the resulting connected components are extracted and used as inputs of the nearest centroid minimum distance classifier which assigns the vehicles appropriate class labels.

3.8 Vehicle Counting and Traffic Density Estimation

The resulting connected components in the fully processed frame represent vehicles on the road at that time. These components are counted to give the total number of vehicles on the given section of the road at the given time. Fig. 6 shows the result of the count for the frame shown in Fig. 4, and Fig. 7 (b) shows the result of the count for the frame shown in Fig. 5. Figure 7 (a) shows the ROI polygon used for the slow moving traffic scene. With this value, the road traffic density can be calculated as

$$\text{Traffic Density} = \frac{\text{Number of vehicles}}{\text{Area of traffic scene}} \qquad (3)$$

From Fig. 6, it can be seen that the vehicles were well detected and that their shadows were rendered into the background.

Fig. 6. Post-processed frame

Fig. 7(a). ROI polygon

Fig. 7(b). Post-processed frame

It should be noted that, object 5 in Fig. 7(b) consists of two vehicles which are represented as one vehicle as can be seen in Fig. 7(a). This is due to the fact that the car is occluded from the camera's view by the larger vehicle in front of it. This scenario highlights the importance of proper installation of cameras meant for traffic management systems.

3.9 Vehicle Classification

A Euclidean distance based nearest centroid minimum distance classifier is used to classify the vehicles into three classes on the basis of their dimensions: small, medium and large vehicles.

For both convenience and practical reasons, five-fold cross-validation technique was used. Using this method, the dataset of the extracted vehicles is split randomly into five approximately equal subsets for cross-validation. Each subset contains all the three classes, but not necessarily in equal portions. At each of the five validation trials, one subset is used for testing while the other four are used for training. Classification accuracies

for the five trials are averaged to obtain the classification accuracy of the algorithm for a given dataset.

3.9.1 Algorithm Training

In order to use the nearest centroid minimum distance classifier, the feature vectors of the vehicles present in the four training subsets are averaged for each class at each trial. Therefore, in the training set, each class is represented by its mean vector.

3.9.2 Classification

To classify a given unlabeled vehicle, the Euclidean distance between its feature vector and each of the vectors representing the three classes is calculated. Then the vehicle is assigned to the class of the nearest centroid. This can be simplified by evaluating the decision functions of all the three classes for this classifier as [22]:

$$d_j(x) = x^T m_j - \frac{1}{2} m_j^T m_j \qquad j = 1, 2, 3. \qquad (4)$$

Where $d_j(x)$ is the decision function of class w_j, x is the unknown feature vector and m_j is the mean vector representing class w_j; x is assigned to class w_j if one of the three decision functions, $d_j(x)$ yields the largest numerical value.

4. Experimental Results

Video data from a road section was collected using a 5MP camera mounted above the road under which the subject vehicles passed. In order to assess the performance of the system under various illumination levels across the day, the data was collected at 0630hrs before the sun is up; 1230hrs when the sun is overhead and the shadows are negligible, and 1630hrs when both reflections from the road surface [15] and shadows are strongest. Data was also collected from a traffic scene that involved very slow moving traffic so as to assess the performance of the proposed system on such traffic scenes or on stationary ones. Each collection period lasted 20 minutes, resulting in at least 36000 frames each time. The camera was installed anew each time just before data collection due to 'external factors'.

At 0630 - 0650hrs, there were a total of 220 vehicles in the video data, 209 of which were correctly detected. This translates to 95% detection accuracy. In order to obtain as many vehicles as possible for classification, manual adjustments were done to the vehicle detection algorithm for frames whose vehicles were not correctly detected and as a result, 216 vehicles of the 220 were

extracted for classification, while the other 4 were over-segmented and were therefore not included in the classification. For this dataset, a classification accuracy of 81.7% was achieved. The same was done for the other two datasets from the other two time periods. The results are summarized in Table 1. The three datasets from the three time periods of the day were added to form an overall dataset and then 5-fold cross-validated as explained in section 3.9. Tables 2 – 6 show the confusion matrices for each of the 5 subsets of the overall dataset used as a testing set.

For the slow moving traffic scene, there were a total of 246 vehicles in the video data, 202 of which were correctly detected. This translates to 82.1% detection accuracy. After manual manipulations on the frames whose vehicles were not correctly detected as explained above, 224 vehicles were extracted for classification. On this dataset, a classification accuracy of 83.8% was achieved. These results are generally poorer than those for the free flowing traffic scene as shown in Table 1. The reason for this is that occlusions were more severe in the slow moving traffic scene than for the free flowing traffic scene. Consequently, at times, two or even more vehicles could be detected as one vehicle rather than as separate vehicles as shown in Fig. 7.

The relatively low camera position was the main cause of detection errors since it was difficult to 'see' the spaces between the vehicles on the same lane when the vehicles involved were very close together as seen in Fig. 7. It was also the main cause of misclassification. For small vehicles, for example, due to their size, their entire tops were visible while this was not true for the other classes where only the fronts and some part of the tops were visible; resulting in the usage of different dimensions for different vehicles in classification. This greatly reduced the ability of the classifier to distinguish between small and medium vehicles as is evident from the given confusion matrices. From the confusion matrices, it is of interest to note that for the overall dataset, only two large vehicles were misclassified as small vehicles as shown in Tables 3 and 4, and that there was no small vehicle which was misclassified as a large vehicle. This was due to the fact that the visible parts of large vehicles were generally larger than the visible parts of most small vehicles, making it easier for the classifier to distinguish between the two classes. Attempts to raise the camera position were not successful due to practical limitations.

Table 1: Summary of the results

Time of the day	Total no. of vehicles	Total no. of correct detections	Total no. of wrong detections	Total no. of vehicles used for classification	No. of small vehicles used for classification	No. of medium vehicles used for classification	No. of large vehicles used for classification	Detection accuracy	Classification accuracy
0630hrs	220	209	11	216	132	56	28	95.0%	81.7%
1230hrs	306	291	15	298	182	65	51	95.1%	88.0%
1630hrs	438	425	13	438	280	78	80	97.0%	93.8%
Overall	964	925	39	952	594	199	159	96.0%	89.4%
Slow moving traffic	246	202	44	224	114	73	37	82.1%	83.8%

Table 2. Confusion matrix for test subset 1

		Predicted class labels		
		Small vehicle	Medium vehicle	Large vehicle
Actual class labels	Small vehicle	117	1	0
	Medium vehicle	8	23	8
	Large vehicle	0	0	31
Classification accuracy			91.0%	

Table 4. Confusion matrix for test subset 3

		Predicted class labels		
		Small vehicle	Medium vehicle	Large vehicle
Actual class labels	Small vehicle	115	3	0
	Medium vehicle	9	23	8
	Large vehicle	1	0	31
Classification accuracy			89.0%	

Table 3. Confusion matrix for test subset 2

		Predicted class labels		
		Small vehicle	Medium vehicle	Large vehicle
Actual class labels	Small vehicle	114	4	0
	Medium vehicle	10	21	8
	Large vehicle	1	1	29
Classification accuracy			87.2%	

Table 5. Confusion matrix for test subset 4

		Predicted class labels		
		Small vehicle	Medium vehicle	Large vehicle
Actual class labels	Small vehicle	112	8	0
	Medium vehicle	10	25	5
	Large vehicle	0	0	32
Classification accuracy			88.0%	

Table 6. Confusion matrix for test subset 5

		Predicted class labels		
		Small vehicle	Medium vehicle	Large vehicle
Actual class labels	Small vehicle	117	4	0
	Medium vehicle	8	29	4
	Large vehicle	0	0	33
Classification accuracy		91.8%		

Overall classification accuracy = mean $(91.0\%, 87.2\%, 89.0\%, 88.0\%, 91.8\%) = 89.4\%$.

As regards to the classification accuracies related to the times of the day, it may seem that they improve as the day matures. This is not true. The reason behind this could be the number of training samples and the accuracy with which the data was collected. For the morning dataset, for example, the set was much smaller than any of the other two. Also, the fact that the camera was removed after each event and installed anew the next time data was to be collected meant that the regions of interest (ROI) were not exactly the same for the different time periods. This could also cause errors and was the main reason for coming up with the overall dataset so as to be able to get a reasonable average of the classification accuracy of the algorithm across the given time periods. It is therefore, clear that the variations are related directly to object extraction rather than the classification itself.

5. Conclusions

This paper attempted to solve the problem of vehicle detection, counting and classification in natural traffic scenes using video surveillance systems for both free flowing and slow moving or stationary traffic scenes. Stationary or slow moving traffic scenes have little reported about them and the majority of the proposed systems make use of motion detection based approaches and are therefore inappropriate for these scenes. This is despite the fact that slow moving or stationary traffic is

the main problem facing traffic management authorities in most towns around world.

The proposed algorithm detected vehicles with a good degree of accuracy under different illumination conditions during the day for both free flowing and stationary traffic scenes. The shadows were also well handled with a good degree of success as shown in Fig. 4 and Fig. 6. The vehicle detection algorithm used a novel approach in which the vehicles were simultaneously extracted from the traffic video data frames and their negatives using the Laplacian of a Gaussian edge detector. Edge linking was achieved through mathematical morphology and summation of the positive and negative edge maps. However, despite the success of the algorithm, it was noted that over-segmentation occurred for very large trucks: with cabins and their trailers being detected as separate vehicles. The algorithm also had problems with occluded vehicles. To minimize these problems and classification errors, it is suggested to raise the position of the camera to be high enough with respect to the ROI.

The results obtained by the proposed system are comparable to those published in the literature [8], [9], [11-15] and [17-19]; where average detection and classification accuracies of between 80% - 100% are reported. However, it should be noted that a one to one comparison between the performances of any two systems is only realistic if, among others, the same dataset was used to test them [23], [24], and [25].

6. Recommended Future Work

This work could be extended to incorporate the specifications of the cameras into the algorithm. For example, it was noted after trying different cameras that the camera resolution has an implicit relationship with the sizes of the required structuring elements and the appropriate size of the region of interest. This can be investigated more so as to come up with recommendations for real world applications. This need has also been identified in [21]. In addition to this, an approach for determining optimal sizes of structuring elements with respect to inter-vehicle distances need to be developed. Other weather conditions such as rainy condition could also be investigated and the algorithm developed in such a way that it is able to adapt to such conditions. Besides this, occlusions in traffic scenes which have remained a major problem for vision based traffic management systems for many years need to be solved.

References

[1] S. Sivaraman and M. M. Trivedi, "A Review of recent developments in vision-based vehicle detection," in 2013 IEEE Intelligent Vehicles Symposium (IV), Gold Coast, Australia, 2013.

[2] K. Kim, T. H. Chalidabhongse, D. Harwood and L. Davis, "Real-time foreground–background segmentation using codebook model," Elsevier, 2005.

[3] C. Stauffer and W. E. Grimson, "Adaptive background mixture models for real-time tracking," in IEEE International Conference on Computer Vision and Pattern Recognition, 1999.

[4] V. Heidari and M. R. Ahmadzadeh, "A method for vehicle classification and resolving vehicle occlusion in traffic images," Pattern Recognition and Image Analysis (PRIA), 2013 First Iranian Conference on, pp. 1-6, 6-8 March 2013.

[5] C. C. C. Pang, W. W. L. Lam and N. H. C. Yung, "A Novel method for resolving vehicle occlusion in a monocular traffic-image sequence," IEEE Transactions on Intelligent Transportation Systems, vol. 5, no. 3, pp. 129-141, 2004.

[6] A. Sharma, D. Bullock, S. Peeta and J. Krogmeier, "Detection of inclement weather conditions at a signalized intersection using a video image processing algorithm," The Civil Engineering at DigitalCommons@University of Nebraska - Lincoln, vol. 32, pp. 150-155, 24-27 September 2006.

[7] Minnesota Department of Transportation, Office of Traffic, Safety & Technology and SRF Consulting Group, Inc., "Evaluation of non-intrusive technologies for traffic detection," United States Department of Transportation, Federal Highway Administration, Minneapolis, 2010.

[8] S. H. Yu, J. W. Hsieh, Y. S. Chen and W. F. Hu, "An automatic traffic surveillance system for vehicle tracking and classification," in Scandinavian Conference on Image Analysis (SCIA2003), Göteborg, Sweden, 2003.

[9] R. P. Avery, Y. Wang and G. S. Rutherford, "Length-based vehicle classification using images from uncalibrated video cameras," University of Washington, 22 July 2004. [Online]. Available: http://faculty.washington.edu/yinhai/wangpublication_files/ITSC_05_VC.pdf. [Accessed 21 March 2015].

[10] Open Network Video Interface Forum Inc., "ONVIF™ core specification," August 2013. [Online]. Available: http://www.onvif.org/specs/core/ONVIF-Core-Specification-v240.pdf. [Accessed 20 June 2015].

[11] A. Ambardekar, M. Nicolescu, G. Bebis and M. Nicolescu, "Vehicle classification framework: a comparative study," EURASIP Journal on Image and Video Processing , 2014.

[12] H. Rabiu, "Vehicle detection and classification for cluttered urban intersection," International Journal of Computer Science, Engineering and Applications (IJCSEA), vol. 3, no. 1, pp. 37-47, 2013.

[13] A. R. M. Ribeiro and P. L. Correia, "Automatic vehicle counting based on surveillance video streams," Instituto de Telecomunicações, Lisbon, Portugal, 2012.

[14] N. Buch, J. Orwell and S. A. Velastin, "Detection and classification of vehicles for urban traffic scenes," in Visual Information Engineering, 2008. VIE 2008. 5th International Conference, Xian China, 2008.

[15] P. K. Mishra, M. Athiq, A. Nandoriya and S. Chaudhuri, "Video-based vehicle detection and classification in heterogeneous traffic conditions using a novel kernel classifier," IETE Journal of Research, vol. 59, no. 5, pp. 541-550, 2013.

[16] I. Vujović, M. Jurčević and I. Kuzmanić, "Traffic video surveillance in different weather conditions," Transactions on Maritime Science, vol. 3, no. 1, pp. 32-41, 2014.

[17] V. Kwigizile, M. Selekwa and R. Mussa, "Highway vehicle classification by probabilistic neural networks," [Online]. Available: https://www.aaai.org/Papers/FLAIRS/2004/Flairs04-114.pdf. [Accessed 10 May 2015].

[18] O. Sidla, L. Paletta, Y. Lypetskyy and C. Janner, "Vehicle recognition for highway lane survey," in 7th International IEEE Conference on Intelligent Transportation Systems, 2004.

[19] M. Pancharatnam and D. U. J. Sonnadara, "Vehicle counting and classification from a traffic scene," in Proceedings of the 26th National IT Conference, Colombo, Sri Lanka.

[20] G. Zhang, R. P. Avery and Y. Wang, "A Video-based vehicle detection and classification system for real-time traffic data collection using uncalibrated video cameras," Department of Civil and Environmental Engineering, University of Washington, Washington, 2006.

[21] E. Ince, "Measuring traffic flow and classifying vehicle types: a surveillance video based approach," Turkish Journal of Electrical Engineering and Computer Science, vol. 19, no. 4, 2011.

[22] R. C. Gonzalez and R. E. Woods, Digital image procesing, third edition, Upper Saddle River, New Jersey: Pearson Education Inc., 2008.

[23] G. Dougherty, Pattern Recognition and Classification, An Introduction, New York: Springer Science+Business Media, 2013.

[24] A. R. Webb and K. D. Copsey, Statistical pattern recognition, Third Edition, Chichester, West Sussex, UK: John Wiley & Sons, Ltd, 2011.

[25] R. O. Duda, P. E. Hart and D. G. Stork, Pattern Classification, Second Edition, John Wiley & Sons, 2000.