Context-Adaptive Approach for Vehicle Detection Under Varying Lighting Conditions

David Acunzo[+]   Ying Zhu[+]   Binglong Xie[+]   Gregory Baratoff[++]

[+]Siemens Corporate Research Inc.   [++]Siemens VDO

Princeton, NJ, USA   Lindau, Germany

Abstract— **This paper presents a vision-based vehicle detection method, taking into account the lighting context of the images. The adaptability of a vehicle detection system to lighting conditions is an important characteristic on which little research has been carried out. The scheme presented here categorizes the scenes according to their lighting conditions and switches between specialized classifiers for different scene contexts. In our implementation, four categories of lighting conditions have been identified using a clustering algorithm in the space of image histograms: Daylight, Low Light, Night, and Saturation. Classifiers trained with AdaBoost are used for both Daylight and Low Light categories, and a tail-light detector is used for the Night category. No detection is made for the Saturation case. Experiments have shown a considerate improvement in the detection performance when using the proposed context-adaptive scheme compared to a single vehicle detector for all lighting conditions.**

I. INTRODUCTION

Extensive research has been carried out recently for driver assistance systems involving on-board vision sensors. The main motivations for this research are the increasing need for safer roads, the decreasing cost of visual sensors and the improved computing power offered by modern technologies. Related applications include lane departure warning, traffic sign recognition, pedestrian and vehicle detection systems.

To realize these functionalities, challenging problems need to be addressed. Since the sensor is on-board and seeing outdoors scenes, such a system needs to be robust enough to deal with random and drastic changes of the environment. Numerous previous studies related to vehicle detection systems focused on the robustness against the large variance of vehicles' appearance, while assuming fairly constant lighting conditions (see [8] for a review). However, equally important is the issue of being able to deal with drastic changes of the lighting conditions in an outdoor environment. To our knowledge, fewer works have addressed this challenging problem. In [2] and [4], the authors proposed a system that switches between day and night. A system described in [1] deals with reduced visibility conditions with a stereo sensor and color detection. In [7], the author presented a vehicle tracking scheme for both daytime and night time. However, a system able to perform smooth transition from daylight to night has yet to be developed.

We focus our work on the detection of preceding vehicles driving in the same direction as the host car. The data acquisition apparatus used for our study consists of a single monocular CMOS camera mounted inside the host vehicle and capturing image sequences of road scenes ahead. In this paper, we propose a vehicle detection method to deal with changes of lighting conditions from bright day to night, taking into account transitional contexts such as dawn, dusk, and other low light conditions. In our previous work [11][10], we have noticed a drop of performance when the lighting condition changes (*e.g.* during dawn and dusk), which motivated us to focus our effort on the work being presented here. We propose a novel detection method called *context-adaptive detection*. The two key ideas of our method are **i)** Automatic context categorization of the input frames based on the histogram of pixel intensities and **ii)** Context-adaptive detection using specialized classifiers to deal with each context.

The paper is organized as follows. In Section II, we give an overview of the method proposed. The concept of lighting context is described in Section III, and the specialized detectors for each context are presented in Section IV. In Section V, experimental results are presented to validate the proposed method. We conclude our work in Section VI.

II. OVERVIEW OF THE METHOD

The goal is to build a robust autonomous vehicle detection system that can deal with various lighting conditions, which changes the appearance of vehicles drastically. Figure 1 shows images captured under various lighting conditions. We remark that during daytime, the most salient features are the objects' contours and textures. At night, such texture information is lost, but tail-lights are easily perceived. During the transition time of dawn and dusk, both information may be present, but with variable saliency and contrast. These changes need to be taken into account by the detector in order to achieve satisfactory performance in each case.

In our work, we deal with an uncontrolled environmental condition, *i.e.* the ambient light, which introduces additional variation in vehicle appearance besides the variation already existing among different types of vehicles. The conventional approach of distinguishing vehicles from backgrounds with a binary classifier [8] would be insufficient to handle large changes of lighting conditions. Similar problems have been addressed in computer vision research in the area of multi-view object detection, where pose changes introduce large variation in object appearance, and it has been shown that categorizing the object appearance according to their shapes and combining individual specialized binary classifiers substantially improves the performance of an object detector [6]. Following a similar idea, we categorize captured image
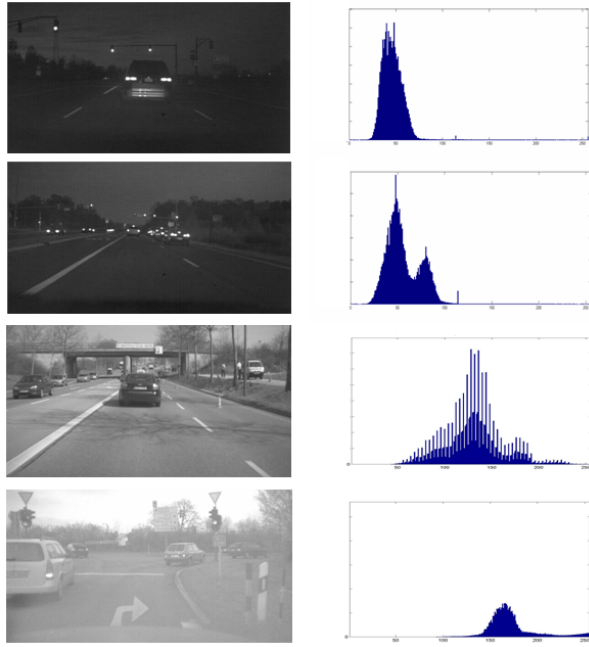
Fig. 1. Example scenes taken under different lighting conditions. The lighting aspect of the images is reflected by the shape of their histogram.

Fig. 2. Principle of the context-adaptive detection method

frames into sub-classes according to the lighting condition and build a dedicated detector for each category.

A simple way to deal with vehicle detection during day and night time would be to use a specialized detector for each of the two cases: a daytime detector focusing on texture information and a night time detector utilizing tail-light information. As proposed in [4], the system would switch from one detector to the other according to the mean value of the pixels intensities. In such a system, one can expect a drop of performance during transition time, when the main features of vehicle appearance is a mix of tail-lights and textures.

In order to build a scheme able to switch smoothly from day time to night time, two problems need to be addressed: **i)** Define a context variable to characterize the lighting condition; **ii)** Build a dedicated classifier for each context. Figure 2 shows the principles of our method. Off-line tasks refer to preliminary stages necessary to implement a context-adaptive system based on our method. We first acquire image samples reflecting various lighting conditions. Clustering is then performed based on the histograms of these images. The clustering scheme identifies different contexts, which enables the learning of dedicated classifiers on vehicle and non-vehicle examples acquired from the corresponding context. A context classification scheme is also performed on-line by the context switch engine in order to switch to the appropriate detector dedicated to the context of the input image.

III. CONTEXT CATEGORIZATION

A. Lighting Context

We introduce the concept of *lighting context* to describe the condition of environment lighting, which is reflected
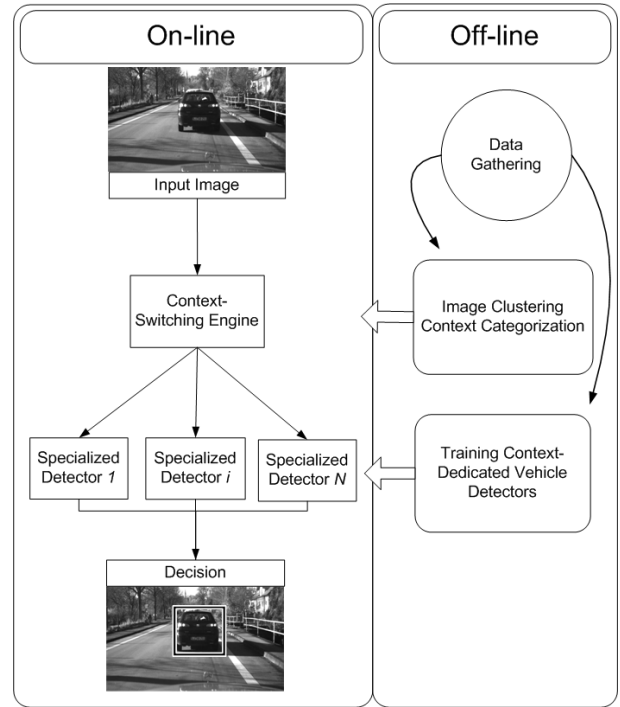
in the measurable image intensities. In addition to ambient lights, image intensities also depend on camera parameters such as exposure time, camera gain, etc.. Since vehicle detection will be performed in the image domain, it is more tangible to define the lighting context from a space that integrates all of the above imaging parameters. The histogram, being the distribution of the pixel values of an image, reflects its overall intensity level and is considered to be a viable indicator of the lighting context. A number of traffic scenes are shown in Figure 1 along with their histograms. We have observed a strong correlation between the distribution of the pixel values and the lighting conditions, which suggests that the histogram is a good indicator to determine the lighting context.

It is worth to point out that not all image pixels are relevant to describe the context in which our target objects are present. With the settings of our data acquisition system, the upper third of the frames shows large variations because of various objects that can occlude the sky, such as trees, mountains or buildings. These variations make it more complex to infer the lighting context. In comparison, the lower two-thirds of the image, covering the road surface and the vehicle appearance, show more consistent and stable distribution of the pixel intensities. We will consequently define the lighting context of an image from the histogram of its lower two-thirds. In the following discussion, we describe how context categories are defined. In the first step, i.e. image clustering, image samples are partitioned into several clusters. In the second step, i.e. context categorization, each cluster is assigned a context category.

Algorithm 1 K-Mean Clustering Algorithm
> **Given:** N patterns to be clustered
> Choose K cluster seeds, to coincide with K patterns chosen to represent a variety of lighting conditions
> **repeat**
>> Assign each pattern to the closest cluster center
>> **for all** k clusters **do**
>>> Recompute the cluster centers using the current cluster membership
>>
>> **end for**
>
> **until** the convergence criteria is met
> **return** the K centers from last iteration

B. Image Clustering

Image samples are first grouped into a number of clusters, where images with similar histograms are categorized into a same cluster. The criterion retained for clustering is the similarity between histograms of image lower parts. Substantial work has been carried out on data clustering, and many algorithms and similarity measures have been developed for that purpose [3]. The K-mean algorithm (Algorithm 1) is used in our work for image clustering. In the K-mean algorithm, the number k of clusters is chosen *a priori*, and the final clustering depends on the initial conditions. Since we want to group images according to their context, it is an advantage to keep some control over the clustering process through the choice of the initial seeds, to guarantee that the clustering results are relevant to our purpose. The output of the K-mean algorithm is the K cluster centroids obtained when the convergence criterion is reached, and each image sample is assigned a *cluster label*.

In our implementation, we choose to use the following distance measure derived from the Bhattacharyya coefficient,

$$D_B(H_1, H_2) = 1 - \sum_{j=1}^{N_b} \sqrt{H_{1,j} \times H_{2,j}} \qquad (1)$$

where $H_{i,j}$ denotes the jth bin of the histogram H_i and N_b denotes the number of bins. The distance measure is bounded, i.e. $0 \leq D_B(H_1, H_2) \leq 1$. Note $D_B(H_1, H_2) = 1$ when there is no overlapping between H_1 and H_2, and $D_B(H_1, H_2) = 0$ when H_1 and H_2 are identical. Other alternatives can be considered, such as the Euclidean or Wasserstein distances. In practice, the number of clusters K is defined to achieve a good tradeoff between complexity and within-class variation. Figure 3 shows the initial clustering results with $K = 13$ clusters in our implementation.

C. Context Categorization

After the image clustering step, we obtain K clusters, with K large enough to achieve low within-class variation. Images belonging to the same cluster are close to each other in the histogram space. To detect vehicles under varied lighting conditions, one possibility is to build a specific detector for each image cluster, but this may lead to an overly complex system. To prevent this, we merge the K clusters into C categories, with $K > C$. Each category contains a few adjacent image clusters with relatively similar lighting conditions. We call this procedure *context categorization*, where a cluster label is mapped into a *context label*. In this procedure, we have the control over the number of categories C, and how to group the clusters into categories. This allows us to keep a balance between within-class variance and algorithm complexity. The final category would define the context label. In our implementation, we have defined four categories ($C = 4$): Night, Low Light (LL), Daylight (DL) and Saturation. Figure 3 shows a representation of the 13 clusters grouped into the 4 categories, with some representative examples from each category. In practice, we allow overlaps between different context categories by sharing clusters on the boundary to guarantee smooth transition between different categories.

D. Context Switch Engine

Through image clustering and context categorization, we obtain two functions. The first function f maps an image frame I to a cluster label $l_{cluster}$ by identifying the nearest centroid among the K clusters. Denote the histograms of the cluster centroids obtained by the K-mean algorithm as $\{H_1, \cdots, H_K\}$ and the histogram of an image I as h_I. f is written as

$$f(h_I) = argmin_{l=1,\cdots,K} D_B(h_I, H_l) \qquad (2)$$

The second function g maps a cluster label $l_{cluster}$ to a context label $l_{context}$,

$$l_{context} = g(l_{cluster}) \qquad (3)$$

A context switch engine is used on the fly to assign a context label to each incoming frame I_t. The context switch engine first decides on the cluster label of the nearest centroid. The final context label is assigned to the image through context categorization.

$$l_{context,t} = g(f(h_{I_t})) \qquad (4)$$

Different treatments are given to frames categorized into different contexts. An adaptive approach with specialized detectors trained by a learning algorithm is adopted to detect

vehicles in the Daylight and the Low Light conditions. A night time detector can be used to detect vehicle lights in the Night condition. When images are completely saturated due to direct sunlight, no detection is performed. The computational overhead introduced by the context switch engine is negligible compared to the context specific detectors.

IV. SPECIALIZED DETECTORS

A. Vehicle Detectors for Daylight and Low Light conditions

1) Vehicle/Non-Vehicle Classification via Boosting: We use the AdaBoost algorithm, presented in [5], to train vehicle detectors for Daylight and Low Light conditions with appearance cues, though many other alternatives can be considered, such as SVM or Neural Networks. AdaBoost learns a strong classifier composed of a number of weak classifiers $\{h_i\}$,

$$H(I) = \sum_{i=1}^{N} \alpha_i h_i(I) \qquad (5)$$

where I denotes an image sample being classified, N denotes the number of boosting rounds performed, h_i denotes the $i - th$ weak classifier ($h_i(I) \in \{-1, +1\}$), and α_i denotes its weight. The final decision $sign[H(I)]$ on an image patch depends on the weighted votes from the weak classifiers. The classification label assigned to the image I is $sign[H(I)]$. AdaBoost is an iterative learning algorithm where training samples are re-weighted and a series of new weak classifiers are learned. In each iteration, a weak classifier that best discriminates the weighted samples is chosen. The weight α_i of the resulting weak classifier is determined by the classification error. The sample distribution is then modified by increasing the weights of the misclassified samples and decreasing the weights of samples classified correctly. This allows weak classifiers focusing on the previously misclassified examples to be selected.

To train vehicle detectors with AdaBoost algorithm, we have designed a set of image filters to characterize salient features of vehicle appearance. For the experiments described in Section V, we have included the Haar wavelet-like features presented in [9]. Under Daylight condition, vehicle appearance is fully visible, and salient features include edges, textures and contours. When the ambient light is decreasing, a part of this information is lost, but tail-lights become a more salient feature. We consequently also include image features that represent tail-lights for the Low Light condition.

2) Training the Classifiers: The classifiers are trained over a large number of vehicle and non-vehicle images from each category. To stabilize the classifier performance, we performed several rounds of bootstraps to extend the set of training data. This procedure consists of extending
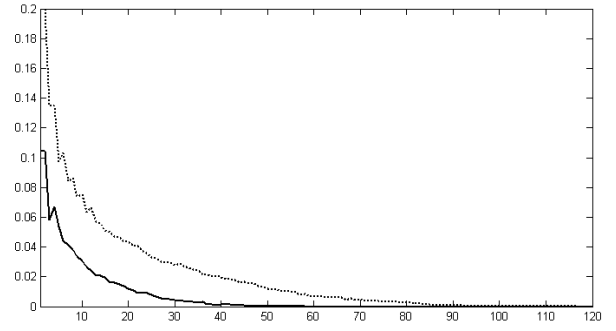

Fig. 4. Comparison of the training error of the Low Light detector (dashed line) and the Daylight detector (solid line) throughout the boosting rounds. The Low Light detector is more difficult to train.

the non-vehicle dataset by running the classifier on testing images and adding the false alarm examples into the training data. Figure 4 shows the performance of classifiers on the corresponding training data. The Daylight classifier shown here was trained over 1972 positive and 6214 negative samples. The Low Light classifier was trained over 3026 positive and 6862 negative samples. Figure 4 shows that the Low Light classifier is more difficult to train. Indeed, the Low Light classifier covers cases from late afternoons to early evenings, with vehicle tail-lights turned on and off, which results in higher within-class variation compared to the Daylight classifier.

B. Vehicle Detector for Night

As Figure 3 illustrates, in the context category labeled as Night condition, the texture information about vehicle appearance is completely lost. Instead, vehicle lights become the only stable and salient feature for detecting and tracking vehicles. A specific vehicle detector has been developed for the Night condition, which detects and tracks vehicle tail lights. Since the focus of this paper is on context recognition and adaptation, detailed discussions are not included here on the specific detection approach developed for the Night condition.

V. EXPERIMENTS

A. Experimental framework

1) Image Clustering and Context Categorization: We performed image clustering on a set of 8046 images from our database. We binned the pixel values of the lower part of the image into 51 bins. 34 iterations were performed by the K-mean algorithm before the convergence criterion was reached. As described in section III-B, each of the 13 clusters was then categorized into one of four categories: Night, Low
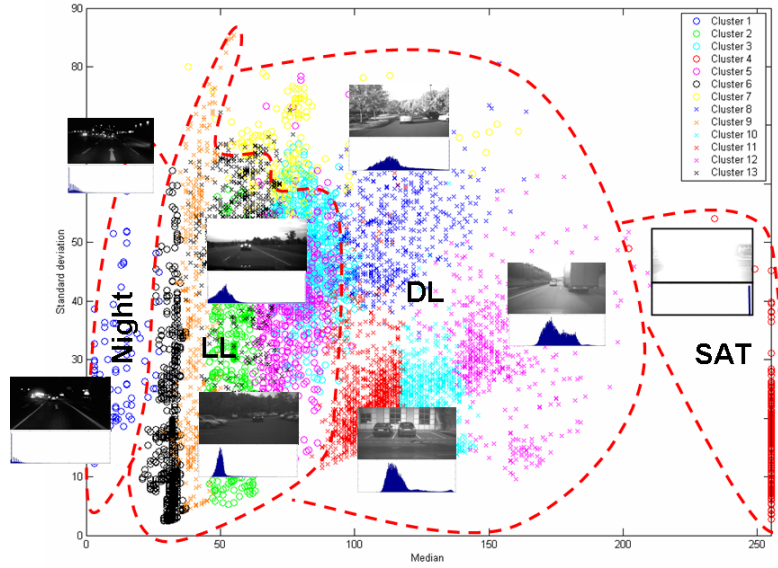
Fig. 3. Results of the image clustering and context categorization process. All 8,046 training sample are represented in a 2 dimensional space: the x-axis and the y-axis represent respectively the median value and the standard deviation of the histograms. The samples are partitioned into $k = 13$ clusters (shown in different colors and symbols) which are then merged into $C = 4$ categories (separated by red dashed curves): Night, Low Light (LL), Daylight (DL) and Saturation. Representative example images along with their histograms are shown for each category.

Light, Daylight and Saturation. The results of the image clustering and context categorization is shown in Figure 3.

2) Validation Experiments: Once the categories of context have been defined, a set of vehicle and non-vehicle examples were extracted from sequences of images from each category. For the validation experiments described here, Low Light and Daylight samples were used to train and test vehicle classifiers.

A k-fold cross-validation, with $k = 5$, has been performed on the set of vehicle and non-vehicle examples from Low Light and Daylight frames. A total of 3775 vehicle examples and 9718 non-vehicle examples in the Low Light condition, and 2465 vehicle examples and 9646 non-vehicle samples in the Daylight condition were used in the validation. As Figure 5 illustrates, the validation process is described as follows. For each round of the cross-validation, we trained three classifiers:

- A Low Light classifier trained with the set of Low Light training samples
- A Daylight classifier trained with the set of Daylight training samples
- A non-specialized classifier trained with the whole set of training samples

In each classifier trained with the AdaBoost algorithm, 250 features were used. To evaluate the performance of the context-adaptive classification, we compare the error rate of the non-specialized classifier, with the error rate of the context-adaptive detector over all the testing data. The error rate of the context-adaptive detector is calculated by accumulating the error of the two specialized classifiers in their respective context.

B. Results

In Figure 6 and Table I, we compare the performance of the different detectors on the testing data. We observe a considerable improvement when using the context information of the image. The more difficult category of the two is the Low Light one. This is consistent with the training results shown in Figure 4. The classification performance increases for both categories when using the dedicated classifiers, which is particularly valuable for the Low Light case, for which the non-adaptive classifier shows poor performances.

When running the detection algorithm on images of actual road scenes, the detector scans image frames and performs classification at each image location to determine if a vehicle appears. Since each frame contains a limited number of vehicles, and most of the scanning areas being tested are non-vehicle patches. Consequently, even a small drop of the false alarm rate will decrease significantly the the number of false alarms. In our experiment, for a 100% true detection rate, the non-adaptive classifier shows a minimum false alarm
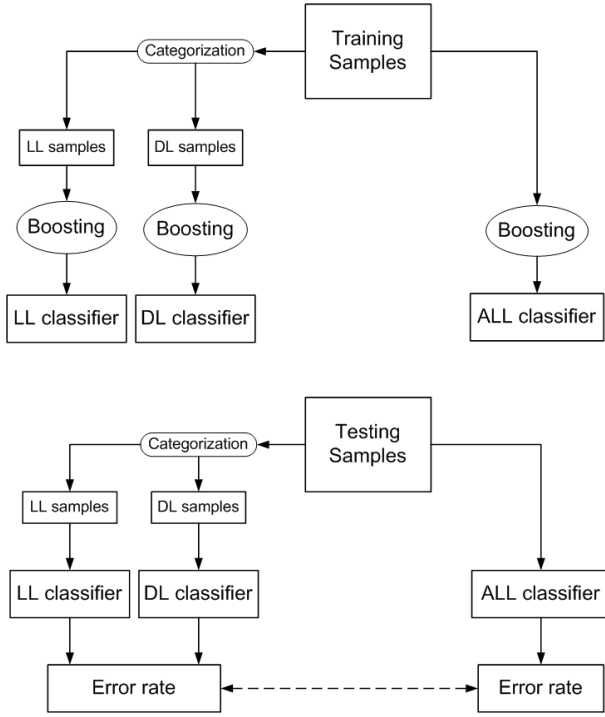
Fig. 5. The Low Light (LL) and Daylight (DL) classifiers, together with the context categorization process, form the context-adaptive classifier. Its performance is compared with a non-specialized classifier (named ALL), trained and tested with the same sample sets.
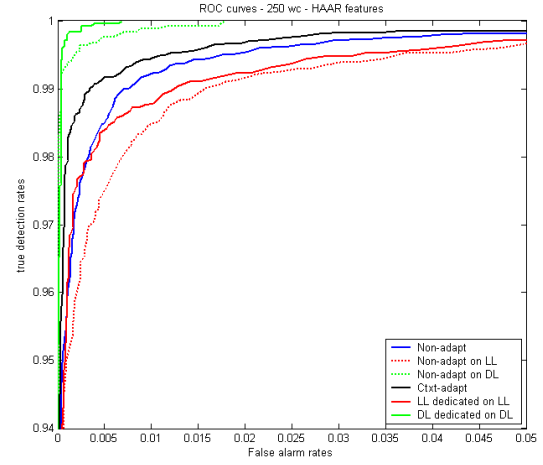

Fig. 6. ROC curves comparing the performance of the context-adaptive classifier (black line) *vs* the non-adaptive classifier (blue line) on the testing datasets. The performance of the non-adaptive classifier on each context category is displayed in dotted lines (dotted red for LL, dotted green for DL). The performances of the dedicated classifiers is displayed in solid lines (solid red for LL, solid green for DL). Each ROC curve was generated by averaging the error rates obtained in 5-folder cross-validation. Quantitative results are also given in Table I.

rate of 0.096, which drops to 0.084 when using the context adaptive classifiers.

Other experiments have also been run, where we used a different number of weak classifiers and included different kinds of features. The results obtained in those experiments are similar to the ones shown here, which confirms the expected performance improvement introduced by context adaptation. The performance improvement shown here is consequently not specific to the parameters used in the classifiers.

C. Test on Road Scene Images

Experiments on images captured of actual road scenes were also conducted. To test the context-adaptive detector, a set of 20 images were randomly extracted from videos captured on the road at different time of the day, including 10 images under the Low Light condition and 10 images under the Daylight condition. In order to detect vehicles of different sizes with the same classifiers, we resized each image multiple times. A scanning window was used to test various image locations to determine if there was any vehicle present. A total of 74853 image patches were tested in

True	Average Number of False Alarms					
Det.	ALL		LL		DL	
Rates	non-adapt.	adapt.	non-adapt.	adapt.	non-adapt.	adapt.
1.000	311.5	274.4	368.4	331.8	27.4	5.4
0.999	196.8	166.0	224.0	199.2	13.4	1.6
0.998	119.8	84.6	158.0	140.8	7.2	1.2
0.997	89.8	58.0	85.6	77.8	5.4	0.8
0.996	70.2	47.6	76.0	66.4	3.2	0.8
0.995	53.2	34.4	63.4	48.6	2.6	0.8
0.990	23.0	10.8	27.4	20.8	0.6	0.6
0.985	15.2	5.8	16.6	9.2	0.4	0.6
0.980	10.8	3.6	12.4	6.2	0.2	0.6
0.970	5.6	2.2	5.8	2.8	0.2	0.4

	Average Number of Test Samples		
	LL	DL	ALL
vehicles	754.8	492.4	1247.2
non-vehicles	1713.6	1544.2	3257.8

TABLE I

SET OF VALUES FOR THE ROC CURVES OF THE DIFFERENT TRAINED CLASSIFIERS. THE NUMBER OF SAMPLES AND FALSE ALARMS IS THE AVERAGE OVER THE DIFFERENT k-FOLD ROUNDS.

Threshold	Low Light Fa/Fr	Low Light Td	Daylight Fa/Fr	Daylight Td	Combined Fa/Fr	Combined Td
-0.04	1.30	1.00	1.80	1.00	1.55	1.00
-0.02	0.40	1.00	1.00	1.00	0.70	1.00
0.00	0.10	1.00	0.50	1.00	0.30	1.00
0.02	0.10	0.94	0.20	1.00	0.15	0.97
0.04	0.10	0.89	0.00	1.00	0.05	0.94
0.06	0.00	0.78	0.00	0.94	0.00	0.86

TABLE II

NUMBER OF FALSE ALARMS PER FRAME (FA/FR) VS TRUE DETECTION RATE (TD) USING DIFFERENT THRESHOLD VALUES.

Fig. 7.   Examples of detection results. True detections are shown in green and false alarms in red. Pictures on the left are categorized as Low Light, those on the right as Daylight.

the Low Light case and 27682 image patches tested in the Daylight case. The test images contain 18 vehicles under Low Light and 17 under Daylight. Most of the image patches being tested are non-vehicles.

Table II displays the receiver operating characteristics (ROC), where a set of true detection rates and the corresponding number of false alarms per frame were calculated by varying the threshold value on the the classifier response. Zero miss and zero false alarm was achieved by the Daylight detector when the threshold was set to 0.04. In the Low Light case, when the threshold value was set to 0, the detector produced one false alarm and no miss. Context switch was automatically done using the context categorization procedure. Consistent with the previous experiment, the detector performs better under the Daylight condition than the Low Light case. Some detection results are displayed in Figure 7.

VI. CONCLUSION

In this work, we proposed the concept of context-adaptive vehicle detection. In particular, we addressed the effect of environmental lighting and explored the possibility of applying adaptive detectors to deal with varying lighting conditions. A context switch engine was introduced to categorize the lighting condition, and specialized detectors were learned for each category. Promising results were obtained in preliminary testing in terms of improvement in the detection accuracy. Further research needs to be done to extend the current work in two directions. First, we need to utilize available information obtained in the camera control process to better characterize lighting conditions. Second, to further improve the detection accuracy, more studies need to be carried out on efficient features that are specific to different lighting conditions.

REFERENCES

[1] I. Cabani, G. Toulminet, and A. Bensrhair. Color-based detection of vehicle lights. In *IEEE Intelligent Vehicle Symposium 2005*, pages 278–283, June 2005.
[2] R. Cucchiara, M. Piccardi, and P. Mello. Image analysis and rule-based reasoning for a traffic monitoring system. In *IEEE/IEEJ/JSAI ITS '99*, pages 758–763, October 1999.
[3] A.K. Jain, M.N. Murty, and P.J. Flynn. Data clustering: A review. *Computing Surveys*, 31(3), September 1999.
[4] S. Kim, S.-Y. Oh, J. Kang, Y. Ryu, K. Kim, S.-C. Park, and K. Park. Front and rear vehicle detection and tracking in the day and night times using vision and sonar sensor fusion. In *IEEE International Conference on Intelligent Robots and Systems*, pages 2173–2178, 2005.
[5] R.E. Schapire and Y. Singer. Improved boosting using confidence-rated predictions. *Machine Learning*, 37(3):297–336, 1999.
[6] Y. Shan, F. Han, H.S. Sawhney, and R. Kumar. Learning exemplar-based categorization for the detection of multi-view multi-pose objects. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 1431–1438, June 2006.
[7] K.-T. Song and C.-C. Yang. Front vehicle tracking using scene analysis. In *IEEE International Conference on Mechatronics and Automation 2005*, volume 3, pages 1323–1328, July 2005.
[8] Z. Sun, G. Bebis, and R. Miller. On-road vehicle detection: A review. *Transactions on Pattern Analysis and Machine Intelligence*, 28(5), May 2006.
[9] P. Viola and M.J. Jones. Rapid object detection using a boosted cascade of simpe features. In *IEEE CVPR 2001*, pages 511–518, December 2001.
[10] Y. Zhu, D. Comaniciu, M. Pellkofer, and T. Koehler. An integrated framework of vision-based vehicle detection with knowledge fusion. In *IEEE Intelligent Vehicles Symposium*, pages 199–204, June 2005.
[11] Y. Zhu, D. Comaniciu, M. Pellkofer, and T. Koehler. Reliable detection of overtaking vehicles using robust information fusion. *Transactions on Intelligent Transportation Systems*, 7(4):401–414, December 2006.