

An Adaptive Vehicle License Plate Detection at Higher Matching Degree

Raphael C. Prates¹, Guillermo Cámara-Chávez¹,
William Robson Schwartz², and David Menotti¹

¹ Computing Department, Federal University of Ouro Preto, Brazil

² Computer Science Department, Universidade Federal de Minas Gerais, Brazil

Abstract. In this paper, a novel approach for vehicle license plate detection that improves in both efficiency and quality over the common multiscale search method is proposed. The detection efficiency is improved by employing the result of a single scale sliding window search as a promising guess of the license plate location. The quality is assured by locally refining the initial detection in multiple scales. The main benefit of our method is that we have reached a more precise detection with the analysis of 20 times fewer detection windows with high reliability (96% recall and 70% precision). We also compared our method with an edge-based hybrid approach.

Keywords: License plate location, Sliding window, HoG feature, Multiscale search, Matching degree.

1 Introduction

Automatic License Plate Recognition (ALPR) is a field in which many researches have been working in recent years, reflecting the important role that it plays in numerous real-life applications such as automatic toll collection, traffic law enforcement, parking lot access control, and road traffic monitoring. ALPR systems usually consist of four main modules: image preprocessing, license plate detection/localization, character segmentation and character recognition [6]. Among these modules, license plate detection is the most challenging, since it has to deal with a large variation of plates (perspective, color, font, inclination, occlusion) and environments (illumination and background complexity). Besides, the errors in the license plate detection reflect in the entire ALPR system.

The main goal of a license plate detection module is to find the correct location of a vehicle license plate in an image. For achieving the ALPR system requirements, this task must be effective and in real-time. Since the license plate has a rectangular rigid format with a known aspect ratio (depending on the image capture perspective), it can be detected by analyzing a large set of possible rectangles in the image, an approach known as sliding window search.

Sliding window techniques can be applied with different types of descriptors. In [7], the authors used a sliding window technique to scan the outcome preprocessed vertical edge density image searching for license plate regions. Nonetheless, this technique is highly dependent on the license plate size. In [5], a sliding

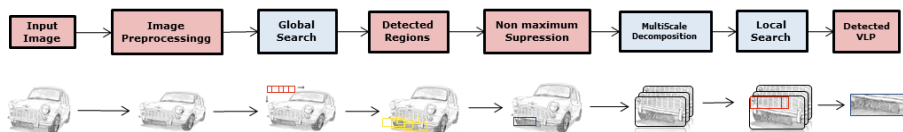


Fig. 1. Overall layout of the proposed *adaptive sliding window detection approach*. First, the image is converted to grayscale and descriptors are extracted. A sliding window search using a single scale is conducted aiming at reaching an initial approximation of license plate positions. After the application of a non-maximum suppression in the detected bounding boxes, a multiscale local search is employed on a region around the initial detection to reach better SVM responses assuring a better matching degree between the real license plate and the detection region.

window approach based on Haar-like features is used along with an AdaBoost classifier for license plate detection. Aiming at speeding up the localization process by early rejection of negative windows, a cascade of classifiers framework is associated with the Haar-like features and AdaBoost in a multiscale sliding window search in [18]. In order to obtain reliable results for more complex scenes/background, the Gentle Adaboost is applied [21,15]. However, this cascade still suffers with an unacceptable false positive rate. In [16], a discrete wavelet transform is employed and the vertical and horizontal edges are calculated from specific subbands. A sliding window scans the whole image eliminating undesired edges and detecting regions like license plates.

A common choice when dealing with controlled environments are the edge detection algorithms, specially the vertical edge information extracted using Sobel filter [20]. However, their effectiveness (precision/recall) drops when using complex images due their high sensitivity to noise [15]. In some cases, a better alternative is to apply block-based methods that can detect regions presenting high edges density, instead of focusing on specific edges [11]. Edge-based methods can be combined with morphological steps to eliminate undesirable edges, forming a hybrid approach [2,9,10,8].

Connected Component Analysis (CCA) has also been employed for license plate extraction aiming at measuring component area and aspect ratio. This information is used to find the best plate candidates. In [17], the authors reached 96.62% of correct extraction and 1.77% of false alarms in a four hour low resolution video. In [22], the license plate detection is formulated as visual matching of characters using SIFT features and employing Principal Visual Words representation. However, this method is not suitable for low resolution images and severe perspective distortion. Color is another feature used in license plate detection, but depends on stable light conditions employed in restrict conditions[6].

In this paper a novel coarse-to-fine sliding window search method using HoG descriptors and linear SVM classifier is employed for the license plate detection problem. We focus on better strategies to search for possible rectangle regions containing license plate, preventing the common pyramid of images. Initially, we perform a single scale sliding search that provides knowledge for a local

multiscale refinement. We can highlight our main contributions as the followings: (1) A multiscale sliding window detection algorithm using HoG descriptors is implemented and the detection effectiveness is analyzed; (2) A fast and novel method using an initial single scale detection followed by a multiscale local search on promising region is proposed; (3) Our method is compared with an edge-based hybrid approach [12] using two different scenarios/databases, presenting results that ensure a better performance with a more reliable detection in terms of effective region localization.

2 Proposed Method

Previous works have shown that the HoG can be successfully applied in different detection problems, reaching high recall rates (detection) with high precision rates (few false positives) [19,13]. Recent works have shown the applicability of HoG in the license plate detection problem, being used in the false positive elimination problem [21]. However, the common pyramid based approach to deal with different scales is very time consuming.

To tackle the problem at different scales in an acceptable computation time, we propose a two-stage approach. The first stage employs a sliding window approach using a single scale detection to achieve an initial approximation of the license plate locations. Due to wide range of scale variation, the results of this stage do not assure the necessary matching for the recognition system. For this reason, a local multiscale search is then conducted to achieve a higher response of the SVM classifier.

The stages performed are: (1) A sliding window search using a single scale is conducted, so that for each detection window in the image the HoG descriptors are extracted and classified by a SVM classifier. Then, a non-maximum suppression algorithm is applied; (2) For each resulting positive region, a local search method is conducted in order to reach better classification responses in a more fine and multiscale search. The final detection window is classified as the license plate region. Figure 1 shows schematically the steps employed in the detection approach. These steps are detailed in the next subsections focusing on the most important ones (in blue), since they compose the contributions of our work.

HoG Descriptor Overview: Histogram of oriented gradients (HoG) are an excellent choice for licence plate detection due to their capacity for object detection. These features were initially proposed for pedestrian detection [4], and also applications in other domains [19,13] can be found. They are based on the idea that an object can be well discriminated by its shape, which, in turn, can be represented by local HoGs.

In order to compute the HoG descriptor of a detection window, we firstly need to obtain the gradient magnitude and orientation for each pixel using a Sobel 1-D filter. Then, the image is divided into cells and oriented histogram are calculated for each cell. Each pixel votes in a bin determined by its orientation with a weight calculated using its magnitude. Blocks are then used to group

cells and apply local normalization. Different blocks are overlapped in such a way that the same cell contributes in more than one block with different values due the local normalization.

Sliding Window Search: Initially, the input image is converted from RGB to grayscales to reduce the influence of light conditions. Then, a sliding window of fixed size scans the image, from top-left to bottom-right, extracting HoG descriptors for each region. A SVM classifier is employed to predict the likelihood of a detection window be a license plate. Responses above a threshold are chosen as initial estimation of license plate positions. The stride between two adjacent windows is sparse and the whole process is conducted in a single scale to reduce the computation time. The assumption (corroborated by experimental results) is that even in a single scale with large stride, the HoG features are able to find fair approximations of the license plate locations.

Local Search: From the initial estimation supplied by the sliding window search, a local search at different scales and positions is performed on a region centralized in the initial window location. The stride is reduced in order to execute a fine search. The goal of this stage is to locate the highest SVM prediction on this region, assuming that a better prediction corresponds to a higher matching between predicted and real license plate.

The local search starts with six different scales to cover the spectrum of plate dimensions by searching for the scale where the highest likelihood is found. From the scale that obtains the highest SVM prediction of this initial local search, an iterative binary search is performed around the current scale until the improvement vanishes or reaches a minimum. It is expected that this approach finds excellent matches by analyzing quite less windows than in a multiscale search through the entire image (the accurate results obtained in the experiments confirm this claim).

3 Experiments

The experiments were performed on a Pentium Core i7, 2.10 GHz with 8 GB of RAM. Our implementation is on C++ using OpenCV 2.4.6 library and will be freely available in [3] after the paper approval. Our method is compared with the method described in [12], since they provided their source code, such that a fully fair comparison could be carried out.

Databases and Evaluation: For measuring the performance of the proposed method, two online available databases are considered: (1) A Greek database [1] with 345 images (640×480 pixels); and (2) A Brazilian database [12] with 377 images (800×600 pixels). Both databases contain challenging aspects. In the Brazilian database, pedestrians, fences, texts, and other vehicle parts are present in the background, while in the Greek database, the landscape in which the vehicle is inserted changes a lot. Figure 2 shows examples of the databases



Fig. 2. Localization examples in the Greek (first row) and Brazilian databases (second row) for sliding window and local search stages. Red and blue windows stand for ground-truth and detected windows, respectively. Note that in some situations the windows are so precisely detected that red windows are overlapped by blue ones.

considered in our tests. The databases were manually labeled with rectangle bounding boxes containing the license plate. The matching degree between a positive response of our SVM classifier and the ground truth is calculated using the Jaccard's coefficient between two windows d_1 and d_2 as $J(d_1, d_2) = (d_1 \cap d_2)/(d_1 \cup d_2)$. A protocol similar to the one used in [15], where the Jaccard coefficient is applied to compute the best match $m(r; R)$ for a rectangle r to a set of rectangles $r_0 \in R$, defined as

$$m(r; R) = \max \{J(r, r_0) | r_0 \in R\}. \quad (1)$$

Then, the recall and precision can be computed as $Recall = (\sum_{r \in T} m_t(r, E))/|T|$ and $Precision = (\sum_{r \in E} m_t(r, E))/|E|$, in which T and E represent the set of ground-truth and resulting rectangles. The term $m_t(a, B)$ is assigned to 1 when the best matching given by Eq. 1 is above a threshold value t . These evaluation metrics are used to assess the accuracy of the proposed approach.

In addition, the edge-based approach is setup using the training set, instead of the cross-validation as described in [12]. Finally, an integral histogram representation, presented in [14], is used for a fast HoG feature description computation and the HoG features are adjusted to find the best configuration for the license plate detection problem.

Experimental Setup: The parameters of our approach were adjusted using a training/validation set composed of 80% of the images, and the remaining 20% was used for testing. **Detection Window:** As the average aspect ratio of the license plates in our training set is 3, our canonical window are setup to this aspect. 9 and 3 pixels were horizontally and vertically added to each size of the license plate. A detection window of 108×36 pixels was employed. **HoG descriptor:** The HoG descriptor configuration applied is a 6×6 pixels cell grouped in blocks of 2×2 cells, generating a feature vector with 3,060 variables per detection window. L1-norm is employed to normalize the HoG histogram extracted from each block. **Sliding window approach:** The stride between two detection windows is 6 pixels in the sliding window search, while in the local search is 3

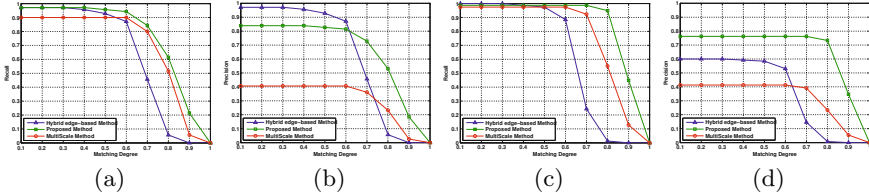


Fig. 3. Recall and Precision versus the Matching Degree of detection in the Greek ((a) and (b)) and Brazilian ((c) and (d)) databases. Curves in blue-triangle, green-squared, and red-circle represent the effectiveness of the Hybrid edge-based, Proposed, and Multiscale methods, respectively.

pixels. In the multiscale search we used 20 scales, which values were adjusted empirically. **SVM learning:** The SVM threshold is adjusted to reach 1 False Positive Per Image (FPPI) at maximum and the regularization parameter C is setup using cross-validation on the training set. For each database, a SVM classifier is separately trained. Negative samples were automatically extracted from the background (no intersection with the labeled real license plates). Initially, 20 negative examples for each image are sampled, but with one round bootstrapping (adding harder examples to train the classifier, bringing down the false positive rates), 1000 more negatives samples were added to the training set.

Results: Figure 2 shows examples of Brazilian testing database. Bounding boxes in red are employed to represent the ground truth used in our validation method, while the blue ones are the detected bounding boxes. Each example has two images, the one in the left represents the response of a single scale sliding window and the one in the right represents a final response of our local multiscale search. These images show that the initial detection is not sufficient to a recognition stage of a license plate and with the multiscale local search is possible to reach the necessary matching degree to perform a further analysis of the plate characters. Note that the top-left image in Figure 2 shows an example of detection with matching degree of 0.53 with a detection inadequate to an ALPR system, proving the need of studies in more reliable detections.

Figures 3(a) and (b) shows the results obtained using the Greek database comparing our proposed method with the multiscale traditional approach and the hybrid edge-based (HEB) method proposed in [12]. All methods evaluated performed well in the Greek database, achieving recall above 90% at a matching degree of 0.5. However, for higher matching degree, the results of HEB method decrease rapidly since it detects inaccurately the target regions. Besides, our method process only 4% of the total windows analyzed in a multiscale approach, and it surprisingly presents an average recall %5 better, proving its efficiency. We also highlight that our method can achieve a 30% precision higher than the other methods when high matching degrees are required (Figure 3(b)).

The recall vs matching degree curves obtained in the Brazilian database are shown in Figure 3(c). At matching degree of 0.8, the proposed method reaches a recall of 94.87% against 1.28 % and 71.80% of HEB and the multiscale methods,

Table 1. Efficiency comparison of sliding window methods

	Database		Database (secs)	
	Brazilian	Greek	Brazilian	Greek
Total wind. (#)	254,310	155,350	Baseline 54.5(± 1.43)	64.2(± 1.56)
Analyzed wind. (%)	5.7(± 0.43)	4.0(± 0.41)	Proposed 11.3(± 0.59)	7.7(± 0.72)

respectively. This better matching degree may allow a more accurate recognition of the characters, reflecting in the overall ALPR system performance. Also, we obtained a precision that is much superior to both other methods in all range of matching degrees. For instance, at 0.7 of matching degree, the precision of our method is 78% against 15% and 40% of the HEB and multiscale methods, respectively (Figure 3(d)).

The figures in Table 1 allow us to perform an efficiency analysis of the results reported in Figure 3. Note that for each image in the Brazilian and Greek databases, the full multiscale approach (our baseline) analyzes 254,310 and 155,350 windows, respectively. As we can observe, in both datasets, the proposed method analyzed 20 times fewer windows in average (5.7 and 4.0) achieving at least $5\times$ speed up. Besides, the efficiency gain the quality of obtained results (higher matching degrees) is still improved. Nonetheless, it is also important to highlight that due its simplicity, the run time achieved by the HEB method [12] is less than one second by image. This issue is the major disadvantage of our approach and it is discussed further as future works.

Also note that the speed up obtained by our method on the Brazilian database is smaller than Greek database. This can be explained by the lower precision on the Brazilian database compared to Greek database. That is, we have more windows for processing at the local multiscale search stage in the Brazilian database than in the Greek database.

4 Conclusions and Future Work

A novel license plate detection method using HoG descriptors in a two-stage approach was proposed in this paper. First, a single scale sliding window search is applied in order to reach a rough initial location of the license plate. This first step cannot assure an accurate detection. Thus, a refinement of the bounding box locations is applied aiming at reaching a higher matching degree between the detected bounding box and the real license plate location. The proposed method was evaluated in two publicly available databases and the results were compared to a hybrid edge-based method and with a traditional multiscale search (baseline), presenting higher recall and precision rates at higher matching degrees (*e.g.*, 0.7). Despite the high performance and the reduced number of windows analyzed, our implementation is still expensive for the real-time requirement of ALPR system, which can be significantly accelerated using, for instance GPU for HoG descriptor computation and SVM detection.

Acknowledgments The authors would like to thank the Brazilian National Research Council – CNPq (Grant #477457/2013-4) and the Minas Gerais Research Foundation - FAPEMIG (Grant APQ-01806-13).

References

1. Anagnostopoulos, C.N.E., et al.: License plate recognition from still images and video sequences: A survey. *IEEE TITS* 9(3), 377–391 (2008)
2. Bai, H., Liu, C.: A hybrid license plate extraction method based on edge statistics and morphology. In: *ICPR*, pp. 831–834 (2004)
3. de Carvalho Prates, R.F.: Vehicle license plate location (VLPL) algorithms, <https://github.com/pratesufop/vlpl>
4. Dalal, N., Triggs, B.: Histograms of Oriented Gradients for Human Detection. In: *CVPR*, vol. 1, pp. 886–893 (2005)
5. Dlagnekov, L.: License plate detection using adaboost. Tech. rep., Department of Computer Science and Engineering, University California of San Diego (2004)
6. Du, S., Ibrahim, M., Shehata, M.S., Badawy, W.M.: Automatic license plate recognition (ALPR): A state-of-the-art review. *IEEE TCSVT*, 311–325 (2013)
7. Du, Y., Shi, W., Liu, C.: Research on an efficient method of license plate location. In: *ICAPIE*, pp. 1990–1995 (2012)
8. Tan, J.L., Abu-Bakar, S.A.R., Mokji, M.M.: License Plate Localization Based on Edge-Geometrical Features Using Morphological Approach. In: *ICIP* (2013)
9. Lalimi, M.A., Ghofrani, S.: An efficient method for vehicle license plate detection in complex scenes. *Circuits and Systems* 2(4), 320–325 (2011)
10. Lalimi, M.A., Ghofrani, S., McLernon, D.: A vehicle license plate detection method using region and edge based methods. *COMPAG* 39(3), 834–845 (2013)
11. Lee, H.-J., Chen, S.-Y., Wang, S.-Z.: Extraction and recognition of license plates of motorcycles and vehicles on highways. In: *ICPR*, vol. 4, pp. 356–359 (2004)
12. Mendes, P.R., Neves, J.M.R., Tavares, A.I., Menotti, D.: Towards an automatic vehicle access control system: License plate location. In: *SMC* (2011)
13. Minetto, R., et al.: T-HOG: An effective gradient-based descriptor for single line text regions. *Pattern Recognition* 46(3), 1078–1090 (2013)
14. Porikli, F.: Integral histogram: A fast way to extract histograms in cartesian spaces. In: *CVPR*, pp. 829–836 (2005)
15. Wang, R., Sang, N., Huang, R., Wang, Y.: License plate detection using gradient information and cascade detectors. *Optik*, 1–10 (2013) (in press)
16. Wang, Y.-R., Lin, W.-H., Horng, S.-J.: A sliding window technique for efficient license plate localization based on discrete wavelet transform. *ESWA* 38(4) (2011)
17. Wu, H.-H.P., Chen, H.-H., Wu, R.-J., Shen, D.-F.: License plate extraction in low resolution video. In: *ICPR*, pp. 824–827 (2006)
18. Zhang, H., Jia, W., He, X., Wu, Q.: Learning-based license plate detection using global and local features. In: *ICPR*, vol. 2, pp. 1102–1105 (2006)
19. Zhang, W., Sun, J., Tang, X.: Cat head detection - how to effectively exploit shape and texture features. In: *ECCV*, vol. 4, pp. 802–816 (2008)
20. Zheng, D., Zhao, Y., Wang, J.: An efficient method of license plate location. *Pattern Recogn. Let.* 26(15), 2431–2438 (2005)
21. Zheng, K., Zhao, Y., Gu, J., Hu, Q.: License plate detection using haar-like features and histogram of oriented gradients. In: *Int. Symp. on Industrial Electronics* (2012)
22. Zhou, W., Lu, Y., Tian, Q.: Principal visual word discovery for automatic license plate detection 21, 4269–4279 (2012)