License Plate Character Segmentation using Partial Least Squares

Gabriel Gonçalves, William Schwartz
Department of Computer Science
Universidade Federal de Minas Gerais, UFMG
Belo Horizonte, Brazil
{gabrielrg, william}@dcc.ufmg.br

David Menotti
Computing Department
Universidade Federal de Ouro Preto, UFOP
Ouro Preto, Brazil
menottid@gmail.com

***Abstract*—A very important research topic nowadays is the Automatic License Plate Recognition (ALPC). This task consists in locating and identifying an on-track vehicle automatically. This task can be divided into the following subtasks: vehicle detection, license plate detection, characters segmentation and character recognition. This work proposes a new technique to perform character segmentation, which is considered solved in the literature, but in practice is a bottleneck for achieving a robust ALPC system. Our approach is a learning-based technique that uses a regression method known as Partial Least Squares to find the best points where the segmentation should be done between the characters. We perform experiments using a dataset composed of 2,000 license plates and three baselines to compare them with the results obtained by the proposed approach. In addition, we evaluate the usage of the PLS with five feature descriptors and our results show that our method is able to achieve a result up to 46.5% of accuracy, evaluated by the Jaccard measure.**

***Keywords*-Automatic license plate recognition; character segmentation; partial least squares.**

I. INTRODUCTION

Automatic License Plate Recognition (ALPR) is the name used to designate the tasks aiming at automatically locating and identifying on-track vehicles. It is a very important research topic, once it can be used in various problems of real life such as traffic speed control, toll collection, among others. Unfortunately, this problem is only solved for controlled traffic environments. Because of that, ALPR still has been theme of many research over the years. According to Du et al. [1], there are four main subtasks that compose an ALPR system: (1) vehicle detection in a video frame; (2) license plate detection; (3) character segmentation; and (4) character recognition.

A bottleneck of an ALPR is the License Plate Character Segmentation (LPCS), since an almost perfect or outstanding character segmentation is extremely necessary to achieve satisfactory results in the step of Optical Character Recognition (OCR). In this work, we propose a new technique to perform LPCS efficiently using a regression technique.

Our method, a learning-based technique, first estimates a Partial Least Squares regression model [2] in which lower responses are given in locations between characters. Then, on the testing phase, the regression method evaluates in a sliding window fashion all possible regions of a license plate and attribute a score to each one of them according to the association that has been learned in the first phase. Among the several horizontal lines of responses, the one that has the higher standard deviation is chosen to be analyzed further. In that way, a 1D function with peaks and valleys is selected, where the latter indicates the separation between two characters. Then, a dynamic programming algorithm is employed to estimate the valley points that better separate the characters. We perform experiments to validate our method using a dataset of Brazilian license plates.

The main contribution of this work can be pointed as the new technique to perform LPCS using Partial Least Squares. In addition, we consider a straightforward technique using only the geometrical information, i.e., the size of the license plate and its characters, as baseline to compare our proposed segmentation approach. Our experiments shows that our method achieve results up to 0.465 using the Jaccard coefficient measure.

The remainder of this paper is organized as follows. In Section II, some works correlated to ours are reviewed. In Section III, the proposed approach to perform segmentation of license plate characters is presented. Then, in Section IV, the experiments conducted to evaluate the proposed approach and the achieved results are described. Finally, in Section V, the conclusions and perspectives for future works are presented.

II. RELATED WORKS

In this section, we present a brief review of the literature that describes some works related to the main focus of this work, the license plate character segmentation (LPCS) problem.

Character segmentation is a task that has been theme of many researches in the past years. There are several fields in which it can be applied, including handwritten character recognition [3] and automatic license plate recognition systems [4].

In the case of ALPR, there are approaches employing different techniques to perform LPCS. According to Du et al. [1] there are five main technique categories: using pixel connectivity or Connected Component Analysis (CCA), pixel projection, prior knowledge of the characters location, using characters contours and approaches based on a combination of these techniques. Since our proposed technique is a learning-based technique, which is not very commonly in literature, it does not fit in these categories.

The dataset used in experiments has a high influence in the final results of the segmentation. In [5], the authors

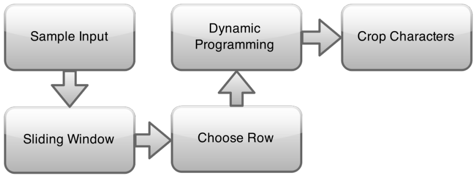

Fig. 1. There are five steps to perform the segmentation.

Fig. 2. Examples of the crop centroids on a license plate that feeds the PLS in the learning phase. The centroids are marked by the red points.

implemented a technique to segment the characters using pixel connectivity and evaluated it in databases manually and automatically cropped. The results showed that there is a large amount of variation in the accuracy among these two types of databases. Since in datasets with automatically cropped license plates, the amount of noise is higher, works that perform experiments on it have to develop a strategy to deal with it. For instance, in [6], the authors have to consider the problem of touching characters that occurs due to the perspective projection on the image captured by the camera.

There are straightforward LPCS techniques that achieve acceptable results in some cases. For instance, in [7], the authors performs an image binarization using an adaptive iterative technique and then segment the characters using the information of pixels connectivity. Soumya et al. [8] propose a technique to segment the characters by counting the number of the black pixels of each binarized license plate.

Works that proposes techniques to improve the quality of the LCPS are common in literature. For instance, Wang et al. [6] use two segmentation techniques in sequence to improve the results and Chuang et al. [9] employ super-resolution.

Our proposed technique is different from most approaches that exists in literature. To the best of our knowledge, there is not works in literature that employ regression by Partial Least Squares to perform segmentation of characters in license plates. Furthermore, learning-based approaches are rarely utilized to perform any kind of character segmentation. More details of the proposed approach are given in the next section.

III. Proposed Approach

The proposed approach is based on a regression technique using Partial Least Squares (PLS) to predict the segmentation point between characters of the license plate. Given a large supervised set of windows indicating their distance to the segmentation point, the approach learns a regression model. The evaluation/segmentation step is composed of the following steps: 1) Given a license plate image, a set of sliding window regions with stride of only one pixel for several scales is generated; 2) For each row of windows, the regression model is applied resulting in a 1-D function in which low values correspond to segmentation points for two consecutive characters, while high values are expected to represent intra character regions or even non-character locations; and 3) Given the best 1-D function, a dynamic programming algorithm is employed to find the combination of the seven regions that best segments the license plate characters. These steps are illustrated in Figure 1.

On the learning phase estimates a PLS regression model. In other words, the PLS learns to predict how far the considered window is from the segmentation point. The dependent variable of the regression model is the minimum distance of the window crop centroid to one centroid of the license plate. The crop centroid is defined by the points that the license plate must be cropped as illustrated in Figure 2. The independent variable is composed of features extracted by some descriptor for each window. To do this, we initially use some license plate from our database to be used as learning examples. Then, from this set, we generate a second and larger set of samples by a sliding window over each license plate of the initial set. The windows of the second set have the same height of one license plate character in the desired scale and the width 50% larger than height.

On the testing/segmentation phase, we generate all possible sliding windows (with stride of one pixel on both vertical and horizontal directions) within the license plate image, assigning to them the response predicted by the PLS regression model. This way, each column in the image presents multiple values (one for each horizontal row of the image). Then, we opt to choose a single horizontal row of sliding windows to segment the text region. For such, we compute the standard deviation of the values estimated by PLS of each horizontal row of slide windows as a score. The one that obtains the highest score is given as the horizontal region to be further analyzed represented by its 1-D function of predicted values. Note that by using only the slide windows of this horizontal row, each vertical column of pixels of the image is represented by a single value. Figure 3 illustrates the resulting values of this chosen 1-D function related to the centroid of the corresponding sliding window.

Once the 1-D function has been estimated and chosen, one needs to determine where are the vertical points to crop/segment the seven characters. On this phase, we defined a combined weight function, i.e.,

$$H(x) = \sum_{k=1}^{8} r_k(x), \tag{1}$$

in which $r_i(x)$ is the regression response of the i-th region of the license plate, i.e., the PLS prediction of that region/window. Since the points to crop the characters are provided in the learning phase as the lowest values, we want to find the combination of eight points, that are almost equidistant (there is a tolerance of 10%), to minimize the function $H(x)$. Dynamic programming is employed to reduce the computational complexity of the search. Finally, when the eight valleys are found, we have the seven character delimited by the eight crops, as illustrated in Figure 4.

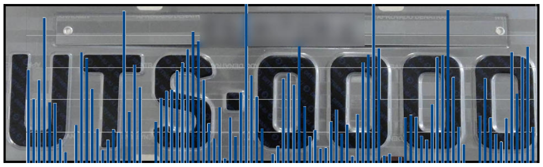

Fig. 3. The license plate has a single result for each column of the image highlighted by blue bars.

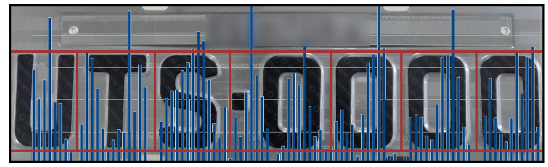

Fig. 4. The characters are cropped on the vertical points delimited by a search algorithm using the dynamic programming paradigm.

IV. Experimental Evaluation

In this section, we present a set of preliminary experiments performed to validate the proposed approach. We evaluate it using a dataset recorded in UFMG Campus This dataset contains 5,628 characters in the test set from 804 images of Brazilian license plates and 5,523 characters to learn (the training set) from 789 images of license plates. Besides, we utilize the Jaccard coefficient computed b etween the bounding boxes detected and the ground-truth of each character to measure the approach effectiveness. We compared our method to other three approaches.

The first approach performs segmentation using pixels vertically and horizontally projected. The image is binarized and the segmentation points are decided based on the amount of white pixels on both projections. This technique was combined to the pre-processing technique, SL*L, which is proposed by Nomura et al. [10]. The SL*L uses morphological operations on the image in order to remove some noises that could disrupt the segmentation.

The second approach used in our experiments was by Shapiro & Gluhchev [7]. It utilizes an adaptive technique for image thresholding followed by a process of Connected Component Labeling (CCL). After that, they utilizes a greedy algorithm to select the components that have size and shape closer to the ones of the characters and have a specific space, in pixels, between them. The seven best components are assumed to represent the characters bounding boxes.

The last approach used is very straightforward because it uses only the shape information of the real license plate shape, ignoring all possible distortions and noises that could exist on the image. Since the Brazilian license plate has seven characters and a hyphen between the third and the fourth ones, separating letters from digits, we just consider the hyphen as a character and divide the license plate into eight vertical parts. Afterwards, the horizontally crop is performed removing 15% of the license plate on the top and 5% on the bottom in order to crop the exact portion with the characters.

We evaluate our proposed method using five types of feature descriptors. The first one is the feature obtained from a pixel vectorization, where the pixels are directly used as features to the PLS. We call this descriptor here as Pixels as Features (PaF). In this case, we horizontally project all pixels of the binarized sliding-window in one row and we feed this projected row to the Partial Least Squares regression as the features of this image region. The pixels are normalized on the interval of 0 to 1 inclusive, where the projection with the largest value in the region becomes 1, and the lowest pixel becomes 0. The second, third and fourth features type are the ones given by the following descriptors: SIFT [11], SURF [12] and ORB [13] where the points to describe were determined by a grid of size 3x3. The last features type is extracted using Histogram of Oriented Gradient [14].

In Table I, one can see the results achieved by our approach. It is possible to note that the features given by the HOG descriptors adapted better to attend the needs of the method. The results illustrated on this table are the average of the jaccard achieved in all characters on the dataset, although, in the Figure 5, it is possible to see the number of character corrected segmented as a function of the Jaccard coefficient used as threshold. We determined that the character is well-segmented when its jaccard value is above 0.7. Based on that, it is possible to see that SIFT and SURF descriptors achieved better results on these thresholds.

The segmentation results obtained for each approach utilizing the Jaccard coefficient are illustrated in Table II. As one can see, none of the evaluated methods achieved satisfactory segmentation results. The techniques utilizing pixel projection and connected component labeling were capable of achieving much better results than our technique. We believe that this result obtained by our method can be explained by the features used. Besides the fact that using Histogram of Oriented Gradients we can improve our results slightly, they are still not good enough.

We also analyzed the characters that were satisfactory segmented as a function of the Jaccard coefficient. These results are illustrated in Figure 6. Considering 0.7 as a Jaccard threshold to obtain a satisfactory segmentation, our approach is not accurate enough to be employed in a reliable ALPR system. Our accuracy using 0.7 as Jaccard threshold is around 12%

TABLE I
Proposed Approach Results: The average jaccard results for all tested descriptors.

Approach	Jaccard Result
PaF	0.287
SIFT	0.384
SURF	0.385
ORB	0.313
HOG	0.465

TABLE II
Experiments Results: The average jaccard results for all implemented approaches.

Approach	Jaccard Result
Pixel Projecting [10]	0.601
Connected Component Labeling [7]	0.452
Prior Knowledge-Based	0.398
Our Approach with Hog	0.465

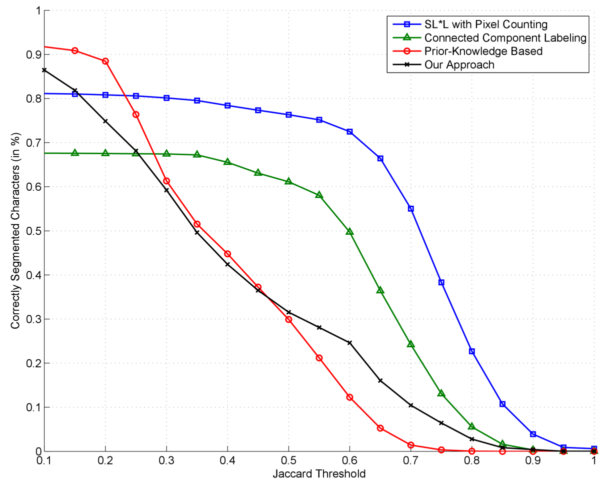

Fig. 6. Proportion of characters well-segmented as a function of the Jaccard coefficient.

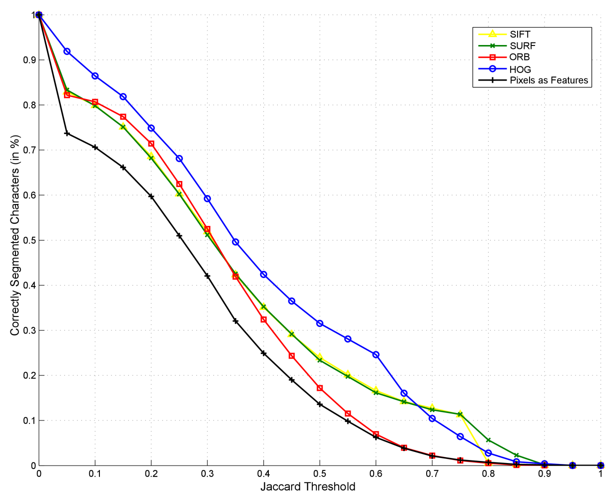

Fig. 5. Proportion of characters well-segmented as a function of the Jaccard coefficient using different descriptors in our approach.

with HOG. The baseline using Pixel Projecting was capable of achieving the best results, around 55%, of well-segmented characters. The baselines using connected component labeling and using Prior-Knowledge were able to segment only 25% and 2% of the characters,

V. Conclusions and Future Works

In this paper, we proposed an approach to perform segmentation of License Plate Character. It is a learning-based method that utilizes regression by Partial Least Squares to determine where are the points that the license plate must be cropped. We evaluated the use of five types of features in our method aiming to discover the best descriptor for this task. The features used were extracted using HOG, SIFT, SURF, ORB and the pixels vectorization.

The experimental results demonstrated that our method is not capable to achieve satisfactory results yet. The method was capable of achieving, according to the Jaccard coefficient, only 0.465 using HOG descriptor. In addition, using 0.7 the threshold for the Jaccard coefficient to determine whether the characters were segment properly, our method was capable of achieving 4% of well-segmented characters. Our method showed better effectiveness than two technique: using the prior knowledge of its shape; and using connected components labeling. In addition, the approach that use SL*L as pre-processing technique could achieve results up to 0.136 points higher than the method proposed.

For future works, we plan to improve the dynamic programming technique to determine the crop points since we believe that we could perform a fine-tuning its parameters. In addition, we propose to evaluate the segmentation methods using other metrics once the Jaccard coefficient is not very well suitable to evtter aluate LPCS techniques since it does not consider the location of the character within the windows.

Acknowledgment

The authors would like to thank the Minas Gerais Research Research Foundation – FAPEMIG (Grant APQ-00567-14) and the Coordination for the Improvement of Higher Education Personnel – CAPES (DeepEyes Project).

References

[1] S. Du, M. Ibrahim, M. Shehata, and W. Badawy, "Automatic license plate recognition (ALPR): A state-of-the-art review," *Circuits and Systems for Video Technology, IEEE Trans. on*, vol. 23, no. 2, pp. 311–325, 2013.

[2] H. Wold, "Partial least squares," *Encyclopedia of statistical sciences*, 1985.

[3] M. Vyas and K. Verma, "A comprehensive survey of handwritten character segmentation," in *Advanced Communication Control and Computing Technologies (ICACCCT), 2014 International Conference on*. IEEE, 2014, pp. 1462–1465.

[4] C.-N. Anagnostopoulos, I. E. Anagnostopoulos, I. D. Psoroulas, V. Loumos, and E. Kayafas, "License plate recognition from still images and video sequences: A survey," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 9, no. 3, pp. 377–391, 2008.

[5] L. Araújo, S. Pio, and D. Menotti, "Segmenting and recognizing license plate characters," in *Workshop of Undegraduate Works - Graphics, Patterns and Images, SIBGRAPI Conf. on*, 2013, pp. 251–270.

[6] R. Wang, G. Wang, J. Liu, and J. Tian, "A novel approach for segmentation of touching characters on the license plate," in *2012 International Conf. on Graphic and Image Processing*. International Society for Optics and Photonics, 2013, pp. 876 847–876 847.

[7] V. Shapiro and G. Gluhchev, "Multinational license plate recognition system: Segmentation and classification," in *Pattern Recognition, (ICPR) International Conf. on*, vol. 4. IEEE, 2004, pp. 352–355.

[8] K. R. Soumya, A. Babu, and L. Therattil, "License plate detection and character recognition using contour analysis," *Int. Journal of Advanced Trends in Computer Science and Engineering*, vol. 3, no. 1, pp. 15–18, 2014.

[9] C.-H. Chuang, L.-W. Tsai, M.-S. Deng, J.-W. Hsieh, and K.-C. Fan, "Vehicle licence plate recognition using super-resolution technique," in *Advanced Video and Signal Based Surveillance (AVSS), IEEE Int. Conf. on*, 2014, pp. 411–416.

[10] S. Nomura, K. Yamanaka, T. Shiose, H. Kawakami, and O. Katai, "Morphological preprocessing method to thresholding degraded word images," *Pattern Recognition Letters*, vol. 30, no. 8, pp. 729–744, 2009.

[11] P. C. Ng and S. Henikoff, "Sift: Predicting amino acid changes that affect protein function," *Nucleic acids research*, vol. 31, no. 13, pp. 3812–3814, 2003.

[12] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *Computer vision–ECCV 2006*. Springer, 2006, pp. 404–417.

[13] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: an efficient alternative to sift or surf," in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 2564–2571.

[14] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, IEEE Conf. on*, vol. 1, 2005, pp. 886–893.