Two-stage algorithm for license plate extraction

Marko Lukic, Milan Tuba and Raka Jovanovic

***Abstract*—**License plate recognition is an important practical problem with many applications in intelligent transport systems. The most important step is license plate extraction since recognition can be done by commercial character recognition systems. License plate extraction is a hard problem due to many variations in images including distance, angle, illumination, color of the plate, etc. In this paper we propose two-stage algorithm for license plate extraction. The first stage uses dynamically adjustable thresholding for determining potential license plate areas. The second stage uses vertical edge density, contrast information and entropy to extract license plate from other candidate regions. Our proposed algorithm was tested on standard benchmark images and it successfully extracted license plates.

Keywords—License plate extraction, dynamic thresholding, segmentation.

I. Introduction

LICENSE plate extraction (LPE) is the first step in license plate recognition (LPR) which plays an important role in intelligent transport systems (ITS). License plate recognition is widely used in various real-life situations such as parking lots control, traffic law enforcement, automatic toll collection, traffic road monitoring [1], and many others. The fact that every vehicle has a unique plate number and that by it both, the vehicle and the owner can be identified makes the license plate recognition very important task in todays life. License plate recognition can help in catching criminals, finding stolen vehicles and locating and ticketing traffic offenders. It is inevitable that in the future traffic tolls will be collected only automatically and it cannot be done without license plate recognition. All this reasons make the license plate extraction important and that is why this problem has become active research issue.

Growing number of vehicles has made this problem harder to solve due to traffic jams and the fact that it should work in all weather conditions during day or night makes this task difficult. Typical problems besides weather conditions are different angles of taking photos of vehicles, size of the plate that varies depending of the vehicle distance from camera, bad state of the plates, different resolutions of images including low ones, etc. Complex background can also be problematic as well as different types of illumination. For those license plate systems that are built for international use [2] several more problems occur: more diversity between plates including font of the characters, color, size and shape of the plates [3]. Also, different cameras are used for capturing images: night vision, infrared, etc.

License plate recognition usually consists of three parts: license plate extraction, character segmentation and character recognition [4], [5]. Input is the image with vehicle in it and the output is a string with license plate characters. Before the start of the search for the license plates in image, preprocessing is a necessary step. Depending on occasion, image transformation or correction must be done to make character segmentation and character recognition easier. The first step in license plate extraction is to locate the regions in the image that are candidates for license plate. Depending on the method used there probably will be more regions extracted. Next task is to discard regions that do not contain plates, which can be done by several methods. After extracting license plate from the image, character segmentation need to be done; it is a process of extracting characters from the plate and preparing them for recognition. An extracted license plate from the previous stage may have some problems, such as tilt and nonuniform brightness or dirt in some regions of the plate. Algorithms used for segmentation must overcome all this problems [3]. Image segmentation algorithms are used in many different fields so numerous methods were proposed. Those methods usually include multilevel image thresholding [6], [7]. Finally, the third step is to recognize characters and get license plate number as the output. Again, several problems may occur like different size and thickness of the characters, different fonts, dirt, tilt, etc. Because of all these reasons algorithms for recognition must be robust and efficient.

License plate localization is usually the main problem [8] in license plate recognition systems since commercial systems for printed [9] text recognition are available and some license plate recognition systems do not even need conversion of the plate characters into ASCII code. It is also the most important of all three steps because it has the strongest impact on the overall accuracy of entire system. An algorithm for locating and extracting plates must be efficient on all picture sizes and qualities [10], [11].

In this paper a two-stage method for locating license plates on still images is presented. An algorithm for extracting license plates is proposed and tested on different images sizes and quality. Quality of the proposed algorithm was measured using several metrics.

This paper is organized as follows. In Section II literature review of algorithms for extracting plates from still images is presented. In Section III our algorithm for license plate extraction is proposed. Section IV contains the results obtained by our proposed algorithm while Section V gives conclusions.

M. Tuba and M. Lukic are with the Graduate School of Computer Science, John Naisbitt University, Belgrade, Serbia.

R. Jovanovic is with the Qatar Environment and Energy Research Institute, Hamad Bin Khalifa University, Doha, Qatar.

This paper is supported by the Ministry of Education, Science and Technological Development of Republic of Serbia, Grant No. III-44006.

II. Literature Review

Optical character recognition is an old and very widely researched scientific field. Digit recognition is a subfield which can be divided to handwritten digit recognition [12], [13] and printed digit recognition such as digits from license plates [14].

License plate detection is already widely used all over the world and many different approaches have been successfully used [15], [16], but growing traffic jams and the growing need for the use of this system on the increasing number of roads make this problem still popular for researchers.

In paper [8] a morphological-based method is used. In the first step of their method vertical edges are extracted using Sobel mask. In the next step candidate regions are found by histogram analysis. Last stage uses morphological operators to exactly locate the license plate. Method was tested on 400 images of different quality with various types of illumination and the license plate detection rate of success was 83.50 percents.

Different method with 92.4 rate of success was presented in paper [17] using a wavelet transform based method for extracting important contrast features as guides to search for desired license plates.

In paper [18] an edge based multi-stage approach has been used. In their work authors have concentrated on localization of license plate regions from true color still snapshots captured from realistic situations. The technique is based on a novel multi-stage approach for analysis of vertical edge gradients from contrast stretched gray-scale images. The success rate of that method was 89.2 percents.

A very efficient method is presented in paper [19]. Authors used combination of few methods and their approach is divided into four sections. First one is vertical edge detection, second is edge statistical analysis, third is hierarchical-based license plate location and finally morphology-based license plate extraction. Success rate of their method is 99.6 percents.

As said before, poor image quality and complex backgrounds are problems that need to be solved. In paper [20] a robust plate extraction method for complex backgrounds is described. The method was divided into two steps. First step is searching candidate areas from the input image using gradient information, and second is determining the plate area among the candidates and adjusting the boundary of the area by introducing a plate template. In their experiments they used images from large underground parking place and had more than 90 percents rate of success.

Some researchers proposed a method for license plate recognition from video sequences. In paper [21] an efficient algorithm based on analysis of maximally stable extremal region (MSER) detection was presented. After a one-time detection of a plate it is robustly tracked through the sequence by applying a modified version of the MSER tracking framework. This method had promising results.

Some license plate recognition systems use more complex methods that are combinations of few methods in all three steps of license plate recognition. In paper [22] there are two major contributions. One contribution is a new binary method, the shadow removal method, which is based on the improved Bernsen algorithm combined with the Gaussian filter and the other one is a character recognition algorithm known as support vector machine (SVM) integration. They improved traditional methods of binarization which made their algorithm robust to the variance of illumination, view angle, position, size, and color of the license plates when working in a complex environment. The overall rate of success for the entire system was 93.54 percents.

In paper [23] a two-layer Markov network was proposed to formulate the joint segmentation and recognition problem in a 1-D case, where three kinds of nodes are defined. In paper [24] support vector machines were used for segmentation and for recognition of the characters from the plates, plate recognition system was used.

Several different methods were precisely described in paper [25] and compared to each other. It was concluded that license plate recognition may be further exploited in various ways such as vehicle model identification, under-vehicle surveillance, speed estimation, and intelligent traffic management and that it can be yet improved in future.

In [26] self-creating disk-cell-splitting (SCDCS) algorithm was proposed for training the radial wavelet neural network (RWNN) model. The proposed SCDCS combined with least-square based RWNN model was employed for the recognition of license plate characters. The proposed model performed better recognition compared to classical RWNN.

In [27] shadow aware license plate recognition system was proposed to recognize license plates. Proposed system achieved high recognition rate by applying shadow detection and removal, as well as rotation adjustment. Multilayer perceptron was a powerful tool to perform the recognition process.

As can be concluded from experimental results from above a lot of license plate recognition systems have high rate of success, but many of them are not so effective when it comes to recognizing plates on international roads especially with license plates from different countries. In paper [2] efficient method for recognizing multi-style license plate detection was presented.

Despite the success of license plate recognition systems in past decades there is still place for progress in the methods used and this is still an active field of research. In next section we propose our method for locating and extracting license plate from images.

III. Proposed method for license plate extraction

We propose a two-stage algorithm for license plate extraction. After preprocessing the image to remove noise, tilt and distortions, process of extracting plate from image can be started. It uses facts that the plates are rectangle, white (in Europe) regions with high contrast and with significant number of vertical edges. The first phase performs conversion to gray scale or retains RGB information, depending on the possible colors of the plates. For most regions this conversion is desirable.

The critical step in the first phase is to determine an appropriate threshold for binarization. Thresholding is the simplest method for image segmentation [28], [29], but adequate for this purpose. For license plate extraction bright regions need to be determined first, but if threshold is to low there will be too many candidate regions, while if it is to high maybe plate region will not be extracted at all, due to various illumination conditions, front lights, etc. The brightness distribution of various positions of a license plate image may vary due to the condition of the plate and the effect of the lighting environment. Because of these factors it is not possible to determine threshold that will always give good results so appropriate threshold has to be generated, not only for each image but also for different regions in one image. There are several methods to compute adequate value for threshold. The simplest way is to find the average brightness of the image:

$$\mu = \frac{\sum_{i=0}^{N}\sum_{j=0}^{M} p(i,j)}{N * M} \tag{1}$$

where $p(i,j)$ is intensity of pixel (i,j) and $N \times M$ is the image dimension, and to compute threshold somewhat higher.

More precise way to find threshold is by using the image brightness histogram. An image histogram acts as a graphical representation of the tonal distribution in a digital image. The histogram shows the number of pixels in the image (vertical axis) with a particular brightness value (horizontal axis) in a plot and this information that histogram contains is very useful for computing threshold.

Different regions in an image can have very different brightness (sky, directly illuminated regions, etc.). In order to overcome this problem we use dynamically adjusted thresholding where average image brightness is corrected by weighed addition of local brightness and the size of the region for local brightness is determined by that region's uniformity measured by variance. Such combination makes it possible to recognize license plate region with almost certainty, while minimizing inclusion of other regions.

The second stage tries to find license plate among other candidate regions. It uses weighted combination of contrast information and entropy. Since license plates contain dark symbols on bright background and many strong vertical edges it can be discriminated using variance of brightness as a contrast measure and entropy as additional factor. Variance of brightness is calculated as

$$\sigma^2 = \frac{1}{M * N}\sum_{i=1}^{N}\sum_{j=1}^{M}[p(i,j) - \mu]^2 \tag{2}$$

where μ is average brightness calculated by Eq. 1, $M \times N$ is dimension of the image and $p(i,j)$ is brightness of the pixel.

Entropy is defined by:

$$E = -\sum_{i=1}^{L}(p_i * \log_2(p_i)) \tag{3}$$

where L is the number of intensity levels, p_i represents intensity probability calculated from histogram.

By using mentioned features, variance of brightness and entropy, to discard extracted regions that do not contain plates it is highly possible that only one region will remain as candidate, one containing the plate.

For edge detection we use modified Sobel mask. Sobel mask is a discrete differentiation operator, computing an approximation of the gradient of the image intensity function. The Sobel operator is based on convolving the image with a small, separable, and integer-valued filter in the horizontal and vertical directions and it is quick and efficient. It contains two masks, the horizontal one and the vertical one. Horizontal mask is defined as following:

Fig. 1: Original image

Fig. 2: Binary image

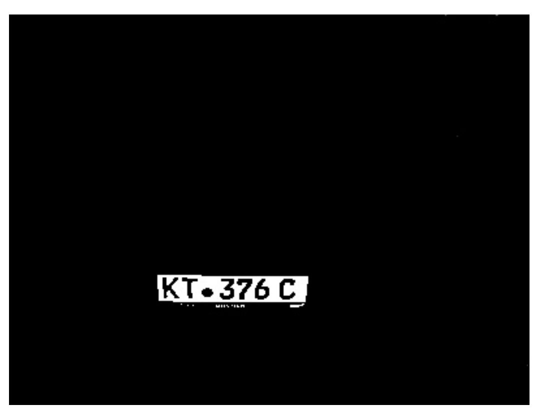

Fig. 3: Binary image after discarding false regions

Fig. 4: Original image

Fig. 5: Binary image

Fig. 6: Binary image after discarding false regions

$$T = \begin{vmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{vmatrix} \qquad (4)$$

The vertical mask is defined as the following:

$$T = \begin{vmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{vmatrix} \qquad (5)$$

We use only the vertical one due to the fact that only vertical edges need to be detected. From the total number of vertical edges pixels compared to the area of the candidate region, one feature for license plate recognition is determined. The other feature is entropy of the region, which is lower for more uniform regions and higher for regions with larger difference, together with variance of the region's brightness.

IV. Experimental results

For experimental results we have used standard benchmark images that can be found at on-line data base [30]. Due to the adaptive threshold and efficient method for discarding candidate regions that do not contain license plates our method proved to be efficient.

In Fig. 1 is the original image of the vehicle, while Fig. 2 shows image after turning into binary by using dynamic threshold computed as described in the previous section. As can be seen, after this first stage several bright regions remain, including the one containing license plate. In Fig. 3 is image after discarding regions with low number of vertical edges and lower entropy and brightness variance, leaving only license plate segmented.

In Fig. 4, 5 and 6 the same process can be seen on a different image with the same result.

We tested our proposed algorithm on 30 images and in all cases it was able to extract license plates.

V. Conclusion

Two-stage license plate recognition algorithm was proposed in this paper. Potential license plate areas are determined during the first phase by dynamically adjusted thresholding. This method proved to be efficient since it always marked license plates with addition of relatively few other areas. License plate was extracted at the second stage using vertical edge density, contrast measured by dispersion and entropy. Proposed method was tested on standard benchmark images and it successfully recognized license plates. Future research may include additional steps in the preprocessing stage to improve results for cases with license plates in more difficult positions.

Acknowledgment

This research is supported by the Ministry of Education, Science and Technological Development of Republic of Serbia, Grant no. III-44006.

References

[1] S. Kaur and S. Kaur, "An efficient approach for number plate extraction from vehicles image under image processing," *International Journal of Computer Science and Information Technologies*, vol. 5, no. 3, pp. 2954–2959, 2014.

[2] J. Jiao, Q. Ye, and Q. Huang, "A configurable method for multi-style license plate recognition," *Pattern Recognition*, vol. 42, no. 3, pp. 358–369, 2009.

[3] S. Du, M. Ibrahim, M. Shehata, and W. Badawy, "Automatic license plate recognition (ALPR): A state-of-the-art review," *IEEE Transactions on circuits and systems for video technology*, vol. 23, no. 2, pp. 311–325, 2013.

[4] S. Ozbay and E. Ercelebi, "Automatic vehicle identification by plate recognition," *World Academy of Science, Engineering and Technology*, vol. 9, no. 41, pp. 222–225, 2005.

[5] H.-J. Lee, S.-Y. Chen, and S.-Z. Wang, "Extraction and recognition of license plates of motorcycles and vehicles on highways," in *Proceedings of the 17th IEEE International Conference on Pattern Recognition (ICPR 2004)*, vol. 4, 2004, pp. 356–359.

[6] I. Brajevic and M. Tuba, "Cuckoo search and firefly algorithm applied to multilevel image thresholding," in *Cuckoo Search and Firefly Algorithm: Theory and Applications*, X.-S. Yang, Ed. Springer International Publishing, 2014, pp. 115–139.

[7] M. Tuba, "Multilevel image thresholding by nature-inspired algorithms-a short review," *The Computer Science Journal of Moldova*, vol. 22, no. 3, pp. 318–338, 2014.

[8] F. Faradji, A. H. Rezaie, and M. Ziaratban, "A morphological-based license plate location," in *2007 IEEE International Conference on Image Processing*, vol. 1, 2007, pp. I–57–I–60.

[9] J. Matas and K. Zimmermann, "Unconstrained licence plate and text localization and recognition," in *Proceedings 0f the 2005 IEEE Intelligent Transportation Systems*, 2005, pp. 225–230.

[10] Z. Yao and W. Yi, "License plate detection based on multistage information fusion," *Information Fusion*, vol. 18, pp. 78–85, 2014.

[11] M. Wafy and A. M. M. Madbouly, "Efficient method for vehicle license plate identification based on learning a morphological feature," *IET Intelligent Transport Systems*, vol. 10, no. 6, pp. 389–395, 2016.

[12] E. Tuba and N. Bacanin, "An algorithm for handwritten digit recognition using projection histograms and SVM classifier," in *23rd IEEE Telecommunications Forum Telfor (TELFOR)*. IEEE, 2015, pp. 464–467.

[13] E. Tuba, M. Tuba, and D. Simian, "Handwritten digit recognition by support vector machine optimized by bat algorithm," in *24th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, (WSCG 2016)*, 2016, pp. 369–376.

[14] L. Janowski, P. Kozłowski, R. Baran, P. Romaniak, A. Glowacz, and T. Rusc, "Quality assessment for a visual and automatic license plate recognition," *Multimedia Tools and Applications*, vol. 68, no. 1, pp. 23–40, 2014.

[15] C.-N. E. Anagnostopoulos, "License plate recognition: A brief tutorial," *IEEE Intelligent Transportation Systems Magazine*, vol. 6, no. 1, pp. 59–67, 2014.

[16] S. Du, M. Ibrahim, M. Shehata, and W. Badawy, "Automatic license plate recognition (ALPR): A state-of-the-art review," *IEEE Transactions on circuits and systems for video technology*, vol. 23, no. 2, pp. 311–325, 2013.

[17] C.-T. Hsieh, Y.-S. Juan, and K.-M. Hung, "Multiple license plate detection for complex background," in *19th IEEE International Conference on Advanced Information Networking and Applications (AINA'05)*, vol. 2, 2005, pp. 389–392.

[18] S. Saha, S. Basu, M. Nasipuri, and D. K. Basu, "License plate localization from vehicle images: An edge based multi-stage approach," *International Journal of Recent Trends in Engineering*, vol. 1, no. 1, pp. 284–289, 2009.

[19] B. Hongliang and L. Changping, "A hybrid license plate extraction method based on edge statistics and morphology," in *Proceedings of the 17th IEEE International Conference on Pattern Recognition (ICPR 2004)*, vol. 2, 2004, pp. 831–834.

[20] S. Kim, D. Kim, Y. Ryu, and G. Kim, "A robust license-plate extraction method under complex image conditions," in *Proceedings of the 16th IEEE International Conference on Pattern Recognition*, vol. 3, 2002, pp. 216–219.

[21] M. Donoser, C. Arth, and H. Bischof, "Detecting, tracking and recognizing license plates," in *Asian Conference on Computer Vision*. Springer, 2007, pp. 447–456.

[22] Y. Wen, Y. Lu, J. Yan, Z. Zhou, K. M. von Deneen, and P. Shi, "An algorithm for license plate recognition applied to intelligent transportation system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 3, pp. 830–845, 2011.

[23] X. Fan and G. Fan, "Graphical models for joint segmentation and recognition of license plate characters," *IEEE Signal Processing Letters*, vol. 16, no. 1, pp. 10–13, 2009.

[24] S.-Z. Wang and H.-J. Lee, "A cascade framework for a real-time statistical plate recognition system," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 2, pp. 267–282, 2007.

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

[26] R. Cheng, Y. Bai, H. Hu, and X. Tan, "Radial wavelet neural network with a novel self-creating disk-cell-splitting algorithm for license plate character recognition," *Entropy*, vol. 17, no. 6, pp. 38–57, 2015.

[27] S. A. El-said, "Shadow aware license plate recognition system," *Soft Computing*, vol. 19, no. 1, pp. 225–235, 2015.

[28] M. Tuba, N. Bacanin, and A. Alihodzic, "Multilevel image thresholding by fireworks algorithm," in *25th International Conference Radioelektronika (RADIOELEKTRONIKA)*, April 2015, pp. 326–330.

[29] A. Alihodzic and M. Tuba, "Improved bat algorithm applied to multilevel image thresholding," *The Scientific World Journal*, vol. 2014, p. 16, 2014.

[30] Sayman, "Sayman inc," 2016. [Online]. Available: http://www.saymannet.com/LPR-License-Plate-Recognition