License Plate Verification Method For Automatic License Plate Recognition Systems

B. Y. Amirgaliyev
ABY Applied Systems
Almaty, Kazakhstan
founders@abyas.kz

K. K. Kuatov
ABY Applied Systems
Almaty, Kazakhstan
kanat@abyas.kz

Z. Y. Baibatyr
ABY Applied Systems
Almaty, Kazakhstan
zhalgas@abyas.kz

C. A. Kenshimov
ABY Applied Systems
Almaty, Kazakhstan
cka@abyas.kz

M. Z. Kairanbay
ABY Applied Systems
Almaty, Kazakhstan
magzhan@abyas.kz

A. K. Jantassov
ABY Applied Systems
Almaty, Kazakhstan
abylay@abyas.kz

***Abstract*—Automatic license plate recognition (ALPR) is the technology for identification of license plate numbers from a video stream or an image without human interaction. ALPR has many applications such as registration or accounting (toll roads, checkpoints, and parking lots), search and tracking (recovery of stolen vehicles, capture of offenders, and other regulations of traffic laws), etc. There are many approaches used in contemporary ALPR systems that use different image processing techniques and machine learning algorithms for recognition. These systems display promising results of recognition on data with high resolution and quality. However, these systems encounter difficulties of recognition due to partial occlusions, not uniform illumination, dirty plates, and unwanted elements. Moreover, the types of license plates vary by different countries and sometimes even by regions. Therefore, many ALPR systems require specific approaches to attain high recognition rates. ALPR systems must also be able to make all the processing in real time because in most cases recognition time should not have influence on external processes. In this work, we consider different approaches of ALPR and provide our approach of license plate verification method which demonstrates very promising results.**

***Keywords*—Automatic license plate recognition, segmentation, license plate verification.**

I. INTRODUCTION

There are many approaches of recognition used in contemporary ALPR systems. Most of them consist of three main steps: license plate localization, character segmentation and optical character recognition (OCR). Each of these steps must have high performance rates in order to increase reliability and accuracy of the whole system.

License plate localization is the first step which is used to extract regions where license plates could be. In [1] and [2], vertical edge features are considered to be effective for license plate detection. These algorithms estimate density of vertical edges and search for most dense regions in the image. In [3], color edge detector is used because license plates can consist of only multiple colors. That is why pixels with such colors can be projected in both directions and the regions with high density are considered to be license plate candidates.

After license plate candidates are found they should be analyzed by some verification method. This step is not mandatory but in most cases it reduces number of candidates and improves the speed of ALPR system. In [4], heuristic analysis and priority selection algorithm is used for the license plate verification. This algorithm includes several different heuristics: the height of the band, number of peaks in vertical projection, a value of area under the graph near peaks, the aspect ratio of the candidate plate. After these heuristics are evaluated they are combined together calculating their weighted sum.

II. PROPOSED METHOD OF LICENSE PLATE VERIFICATION

In this work, we propose the license plate verification algorithm which is connected with the segmentation part. Character segmentation can provide us with necessary information about the contours of license plate characters. This is achieved by making connected component analysis (CCA) on the images. The contours are used to locate and extract each character in the plate. In this step, we faced several difficulties concerning precision and fullness of contour extraction such as:

- Some plates are covered with the dirt;
- Some parts of a plate are occluded or erased;
- Camera is set with inappropriate angle;
- Some extra elements around license plate.

For now, we leave these difficulties as future challenges because their solution requires more detailed study and more data.

Before sending all patches with characters to the recognition step we should verify that this candidate is really a license plate of a vehicle. To do that we use information about sizes and coordinates of the patches. In this work, we use four types of verification together. Each of these verifications is very simple but using their combination gives a strong verification method. These weak verifiers are:

- Overlapping of characters verifier;

- Angle between fitting line to the centers of segments and horizontal axis of the patch verifier;
- Deviation of the segments from the fit line verifier;
- Number of symbols verifier;

The first verifier uses the contours of blobs to estimate how far they are located. If contours are located too close to each other so that they intersect it means that these blobs are less likely to be characters of a license plate. Commonly there is a gap between characters of license plates of most countries. That is why bounding boxes of characters of license plates should not overlap each other. Rotation of a license plate can cause undesired overlapping of bounding boxes. Therefore, detected license plate candidates must be normalized using affine transformations in case of noticeable rotation. If verified candidate is not a license plate there is a good chance that this verifier will throw it off. The formula of estimating this verifier is:

$$V_1 = \left(\sum_{i \neq j} \left(1 - 2 \times \frac{|S_i \cap S_j|}{\min(|S_i|,|S_j|)}\right)\right) / N \qquad (1)$$

where V_1 is the value of the verifier, S_i and S_j are two segments and N is the number of comparisons of all segments. All the verifiers are calculated so that their values are between zero and one. To estimate the value of the first verifier intersection of each two neighboring segments is found and divided by minimum of these segments. This is done to increase effect of overlaps. It was empirically identified that value of overlap can also be increased by multiplying it by some coefficient. In this work, this coefficient is equal to two. Fig. 1 shows examples of segmentation without intersection (a) and with intersecting segments (b).

a b

Figure 1. a) Segmentation without intersections, b) Segmentation with intersecting segments

The second verifier works by estimating angle between a fitting line to the centers of segments and a horizontal axis of the patch. The horizontal axis of the patch is a simple imaginary straight line on the patch parallel to horizontal direction. The fitting line can be found using least squares method. This method minimizes the sum of the squares of the deviations of all center points. Since license plate characters are basically located in a horizontal straight line, the fitting line should not diverge from the horizontal axis to significant degree. The value of this verifier is calculated using the following formula:

$$V_2 = \frac{90 - |\alpha_f|}{90} \qquad (2)$$

where V_2 is the value of the verifier, α_f is the estimated angle. Fig. 2 demonstrates example of finding angle between the fitting line and the horizontal axis.

Figure 2. Angle between fitting line and horizontal axis

The third verifier is a deviation verifier which operates relying on the fitting curve found in the previous verifier. We claim that the sum of the distances between the centers of segments and fitting line must be small value. During the tests, it was detected that fitting line can sometimes have proper angle even if segments are located in different vertical positions. So this verifier is used to get rid of false candidates with noticeable deviation of segments from the fitting line. Here as a deviation value, we calculated vertical distance from points to the fitting line. To normalize distance maximum deviation must be initialized. The value of maximum deviation can be estimated using the height of the patch. Formula of this verifier is:

$$V_3 = 1 - \frac{\sum_{i=0}(D_i)}{(n \times D_{max})} \qquad (3)$$

where V_3 is the value of the verifier, D_i is the value of a deviation from the fitting curve, D_{max} is the value of maximum possible deviation and n is the quantity of segments. Fig. 3 shows deviations from fitting line.

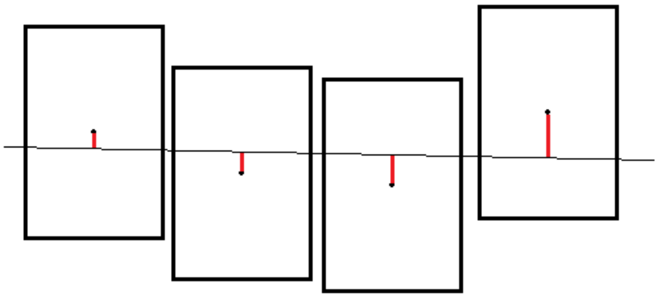

Figure 3. Deviations of contours are colored in red

The last and the simplest verifier checks the quantity of allocated segments. Each country has multiple license plate formats with defined number of characters. This information allows strong license plate verification method. So, if the number of characters does not suit to one of the defined formats the value of this verifier will decrease drastically. Its formula is as follows:

$$V_4 = 1 - \frac{|n - n_{avg}|}{n} \qquad (4)$$

where V_4 is the value of the verifier, n is the quantity of allocated segments and n_{avg} is the average quantity of characters on license plates.

After we estimated values of all four weak verifiers their arithmetic mean is calculated. Since the values of all four verifiers normalized arithmetic mean can be found using the following formula:

$$V = \frac{V_1 + V_2 + V_3 + V_4}{4} \quad (5)$$

where V is the resultant value of all verifiers. This value will be used as a resultant value of the candidate when it is compared to a predefined threshold. The exact value of the threshold should be found by doing a test on a big dataset. In this work, the threshold for the resultant verification value was empirically determined to be equal to 0.9. When this threshold value was used very low false positive and false negative rates were detected. Since ALPR systems require high recognition rates, lower threshold value can be used to decrease amount of false negatives almost to zero. However, the amount of false positive candidates will increase too.

This verification method was changed to check license plates with two rows of characters. If there are few characters in each row the second verifier is not used because of aforementioned difficulties in recognition process. The overlapping of characters verifier and number of symbols verifier are used on all character segments of an image. The third verifier is calculated independently on each row of a license plate. Instead of two fitting lines for each two rows, static horizontal lines are used to estimate the deviation values. Mean of the deviations of two rows is found for the further calculations. The resultant value of verifiers is found by the similar formula as in (5) but without second verifier.

III. Conclusion

This work introduces novel method of license plate verification for ALPR systems. Four weak verification approaches were developed and combined into one strong method. This method was implemented and tested on the set of 1470 images. Since verification of plate candidates was used after detection part, those candidates were checked on correspondence of their aspect ratio to its real value. Therefore, most part of the false candidates were discarded before the proposed verification method. This method found 25% of the remained false positive candidates without any false negatives. The accuracy rate of the method is suitable for ALPR systems but it is still possible to add other verifiers for the improvement of the algorithm.

References

[1] M. Sarfraz, M. J. Ahmed, and S. A. Ghazi, “Saudi Arabian license plate recognition system,” in Proc. Int. Conf. Geom. Model. Graph., 2003, pp. 36–41.

[2] D. Zheng, Y. Zhao, and J. Wang, “An efficient method of license plate location,” Pattern Recognit. Lett., vol. 26, no. 15, pp. 2431–2438, 2005.

[3] E. R. Lee, P. K. Kim, and H. J. Kim, “Automatic recognition of a car license plate using color image processing,” in Proc. IEEE Int. Conf. Image Process., vol. 2. Nov. 1994, pp. 301–305.

[4] O. Martinsky, “Algorithmic and mathematical principles of automatic number plate recognition systems,” B.Sc. thesis, department of intelligent systems, faculty of information technology, Brno University of Technology (2007).