Effective Algorithms and Methods for Automatic Number Plate Recognition

Amirgaliyev Beibut
ABY Applied Systems
Almaty, Kazakhstan

Kairanbay Magzhan
ABY Applied Systems
Almaty, Kazakhstan

Kenshimov Chingiz
ABY Applied Systems
Almaty, Kazakhstan

***Abstract*—nowadays, an automatic number plate recognition (ANPR) system is a key aspect in traffic congestion. This will help to minimize the different kind of violations on the road. Advanced systems for tracking and identifying stolen, unauthorized vehicles are based on automated number plate recognition technology. This paper's main objectives are to review other methods and propose our own algorithm. A short review is performed on the various methods of number plate recognition algorithms. Further explanations of the proposed algorithm are illustrated in graphical forms to show how the algorithm works. This paper concluded with tests and evaluation results.**

***Index Terms*—Automatic number plate recognition, plate area, segmentation, optical character recognition.**

I. Introduction

In the new global economy, traffic congestion became a central issue for majority of the developing countries. The number of cars are increasing rapidly; respectively the numbers of violations are increasing. Speeding, stealing the cars and other wide spectrum of violations on the road are general things in our daily lives. Shortages of parking places are reasons to entering unauthorized cars the private areas and spending a lot of time to find free places in parking lots.

Automated number plate recognition system is a key aspect in resolving all the problems listed before.

However, in order to create such system, it will be needed to construct such algorithm which will identify the location of number plate in the frame, extract the characters from it, and then recognize them.

In this paper we are going to review such algorithms, then propose a method to resolve this problem.

A. Related Work

There are various solutions of relevant problems. The main issues in number plate recognition are climate conditions, environmental interference, and accuracy of number plate localization. One of the methods of recognizing the number plate is utilizing the color characteristics and probability distribution of the license plate between the two lights [4]. Another popular method of number plate recognition algorithm is template matching [2]. The License Plate Detection algorithm based on template matching was designed and written for managing the parking lot system by identifying the unregistered cars from off-campus. At the same time vertical edges-based car license plate detection [3] are popular, too. However, others prefer to find the location number plate using horizontal and vertical projections of image [8]. The Genetic Algorithm [10] and Hough transform [6] can be applied to detect the license plate area. At the same time, the combination of edge statistics and mathematical morphology showed good results [11], [12] and they use block-based algorithm. Another algorithm [13], which is based on rows' distances counts the existent edges and if this number is more than some threshold value then number plate is recognized. Wavelet transform-based algorithm extract the important features to be used for number plate localization. The advantages of this algorithm, it will allow to find more than one number plates in the frame.

Some of the methods above are very complex and requires too much computation time, which is a bit difficult to use in real time applications. However, other approaches could be used only in specific countries with specific characteristics of number plate like background color, etc.

B. Research Objectives

The following list gives the objectives of this research paper:

- To review other algorithms which recognize the number plate of vehicles.
- To propose a method of solving the automated number plate recognition problem.
- To evaluate and test the proposed algorithm and presents the evaluation results.

II. Proposed Approach

The overall problem consists of three parts:

- Plate area detection
- Segmentation and extraction of characters from number plate
- Optical character recognition of extracted symbol etc.

A. Plate Area Detection

An image with number plate will be given as an input to the program and the number plate must be identified then cropped as output image to the next stage. In order to determine the number plate from whole image, the following sequence of actions must be performed to the image.

- Gray scale image - in this stage we need to read the image and convert it to gray scale format. Such conversion will not lead to loss of important data, at the

same token it will be more convenient to work with one channel in preference to three (red, green and blue).

- Blur - the noise is a main issue in our problem. In order to reduce them it is better to blur the image. There are different types of smoothing [5], The following cumulative error distribution graph (Fig. 1) shows the comparison among each of them. According to the Fig. 1, we can say that the homogeneous smoothing is the best one in comparison to others.
- Vertical edge detection - the number plate contains the characters. As we know, the characters contain mostly vertical edges in comparison to horizontal. That is why, one of the best approaches is to find vertical edges that are very close to each other [1]. The edge detection is basic and fundamental operations in computer vision field. There are different kind of edge detectors like Prewitt, Sobel, Canny and etc. Each of them is used for different cases and problems. We use Prewitt, Sobel and modified version of Sobel [9] to solve proposed problem. However, after investigation and testing we came to conclusion to use the modified version of Sobel, because it correctly identifies vertical edges and reduce the most of horizontal edges that impede finding the number plate. However, in order to use modified version of Sobel edge detection the gradient magnitude and gradient direction must be used. They are identified by the following formula in the Eq. 3:

$$|\nabla L| = \sqrt{L_x^2 + L_y^2} \qquad \theta = \operatorname{atan2}(L_y, L_x) \tag{1}$$

Where, L_x and L_y are derivatives of image. Using the value of θ it is possible to find only vertical edges. If the value of θ will be between 45 and 135 then we will get only vertical edges. The Fig. 2 illustrate the result and difference of each edge detector.

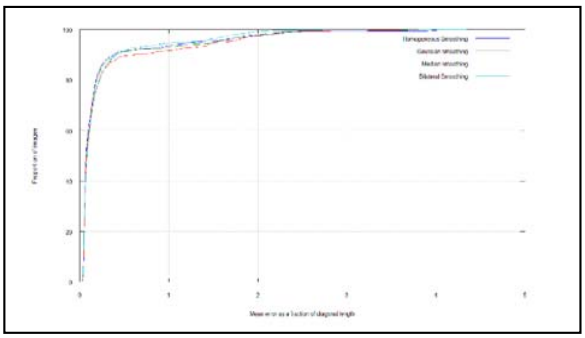

Fig. 1. Cumulative error distribution for smoothing

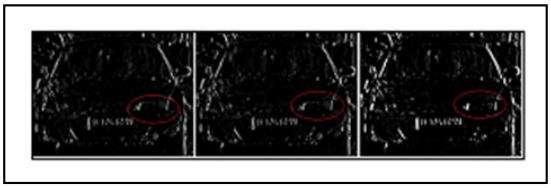

Fig. 2. The difference among vertical edge detectors

- Binary image contains only two colors such as black and white. There are wide spectrum of methods for threshold like Otsu, Niblack, Souvola, Wolf, Feng etc. [7]. Each of them used in special cases for different purposes. It is more convenient to work with binary image. After finding vertical edges, we will apply Otsu threshold to our current image.
- Close morphology used mostly to combine close elements together. Such as our goal is to find the area of number plate, we do not need much information about characters. That is why we apply close morphology, where all letters and digits combined. The Fig. 3 illustrates the result of morphology operation.
- Find contours. After applying the close morphology, we will find those contours that look like number plate, where the area and aspect ratio of contour must be taken into account. At the same time, the contours must be located horizontally like in the Fig. 3.
- Find correct candidate. In order to find the correct candidate among others, each number plate must be investigated. First of all, we will compute the feature vector for each candidate. In order to do that we will use Histogram of oriented Gradients. It will create histogram for each candidate based on gradients. After that we will have feature vector with specific length. Giving all of these vectors to the Support Vector Machine and setting it with specific settings we will be able to obtain the correct number plate among others. Another approach which is based on probabilistic theory was implemented. In this method, the contours were found first, then the contours which do not look like a character were omitted. Remained contours are investigated by identification whether they are located in one line or not. Then by extraction of each character from the number plate, identified the probability of each character. Summing up all of them, and taking into account these values, the best candidate was chosen. However, after testing, we came to the conclusion that the first algorithm works better in comparison to second one.

B. Segmentation

From number plate characters must be extracted. There are two basis algorithms for segmentation, where the first one is based on projection of image into X axis, however, the second one is based on finding of contours that looks like a character. According to investigation and testing, we came to conclusion that second algorithm works better than the first one. The Fig. 4 shows the result of this algorithm.

At the same token, before segmentation the image must be converted to binary format. In this case the Otsu method is not good solution. After applying different binary algorithms, we came to the conclusion that Nick, Niblack, Souvola algorithms works better in comparison to other algorithms.

C. Optical character recognition

Extracted character must be recognized. For recognition we used modified version of 1NN algorithm. The character was divided to small 49 subparts like in the Fig. 5. For each subpart the number of white pixels should be counted. The feature vector that contains 49 elements will identify each character. In the Fig. 6 feature vectors for A, B and C classes were illustrated. For each class average element based on feature vectors should be calculated, then for unknown element the distance to all average elements of each class must be calculated. Unknown element will be joined to those neighbor classes that are closest to that element

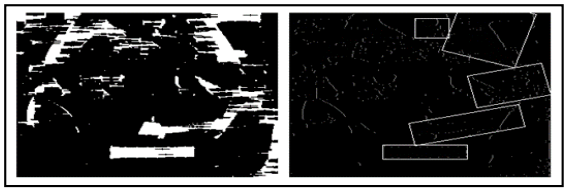

Fig. 3. Image after morphology and finding contours

Fig. 4. Segmentation

Fig. 5. The character that divided to small (7 x 7) subparts

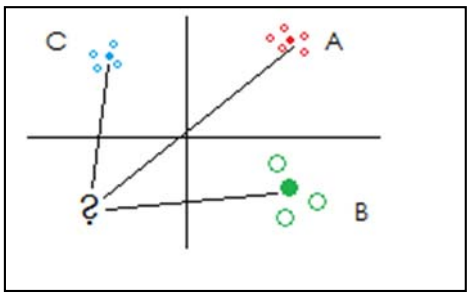

Fig. 6. 1NN for optical character recognition

III. Result and Discussion

The algorithm was tested with 1469 real car photos. The cars were taken from different sides and in different climate conditions.

A. Plate Area Detection

The tests were divided into five parts. Some subparts determines from which side of (front or rear) the photo was taken. Table 1 shows the result of plate area detection using these test cases:

TABLE I. The result of plate area detection

Front	Rear	Subpart 1	Subpart 2	Subpart 3
95.3 %	93.15 %	95.5 %	94.46 %	93.88 %

Taking the average value of results above, we will get the average performance of plate area detection algorithm: (95.3% + 93.15% + 95.5% + 94.46% + 93.88%) / 5 = 94.458%.

B. Segmentation

The number plates were grouped based on their formats to be segmented. Different kinds of formats of number plates have been used.

At the same token, there were used the number plates from other countries such Russia, Kyrgyzstan.

However, we encounter with the some company cars that have their own number plate formats. That is why we came to conclusion to decompose all test cases to subparts based on their formats. The Table 2 shows the result of character extraction for each group of number plate:

TABLE II. The result of segmentation

Sup part 1	Sub part 2	Sub part 3
203 out of 360	2553 out of 2870	5036 out of 7808
56.38%	77.83%	64.49%

Taking into account the value for each group, we found the total performance of segmentation algorithm (56.38% + 77.83% + 64.49%) / 3 = 66.23%. The Fig. 7 shows the extracted letters after segmentation.

Fig. 7. Extracted letters after segmentation

C. Optical Character Recognition

Some characters like “5 and S” and “O and 0” look similar. Taking into account this fact we test our OCR solution. Our

proposed algorithm correctly found 90 characters out of 100, or in other words, the accuracy of proposed algorithm is 90%.

IV. Conclusion and Feature Work

Main parts of number plate recognition system were successfully implemented. Our proposed solution works in general cases, where there are no limit for the distance from camera to the vehicle and weather conditions. However, for specific problems, when the distance from camera to the vehicle will be constant the performance of our system will increase.

For the future work, we need to improve the segmentation part and gather more data for training. As well as optical character recognition can be improved using other popular algorithms like Artificial Neural Network and Markov chain.

For future, we plan to develop an automatic number plate recognition system, which will have its own data base, user interface and authorize people cars by identification of number plate.

References

[1] Kuo-Ming, H., Ching-Tang, H. (2010). A Real-Time Mobile Vehicle License Plate Detection and Recognition. Retrieved from Tamkang Journal of Science and Engineering, Vol. 13, No. 4, pp. 433-442 (2010)

[2] Benjapa, R., Kittawee, K., Paruhat P., Thaweesak Y. (2012). License Plate Detection Based on Template Matching Algorithm. Retrieved from International Conference on Computer and Communication Technologies (ICCCT'2012) May 26-27, 2012 Phuket

[3] Beverly, S., Will, H., Peter, L., Patrico, R. Automatic Number Plate Recognition. Retrieved from CS 175 Fall '12.

[4] Ondrej, M. (2007), Algorithmic and Mathematical Principles of Automatic Number Plate Recognition Systems. Retrieved from B.Sc. Thesis, Bruno university of Technology.

[5] Yoshimori, S., Mitsukura, Y., Fukumi, M., and N. Akamatsu. (2003) "License plate detection using hereditary threshold determine method," Retrived from Lecture Notes in Artificial Intelligence, vol. 2773, V. Palade, R. J. Howlett, and L. C. Jain, Eds. New York: Springer-Verlag, pp. 585–593.

[6] Duan, T., Hong Du, T., Phuoc, T., Hoang, N. (2005). "Building an automatic vehicle license plate recognition system," Retrived from Proc. Int. Conf. Comput. Sci. RIVF, pp. 59–63

[7] B. Hongliang , L. Changping, "A hybrid license plate extraction method based on edge statistics and morphology" in Proc. ICPR, 2004.

[8] H. J. Lee, S. Y. Chen, S.Z. Wang, "Extraction and recognition of license plates of motorcycles and vehicles on highways" in Proc. ICPR, 2004, pp. 356-359

[9] A. Broumandia, M. Fathy, "Application of pattern recognition for license plate recognition", presented at the ICGST Int. Conf. Graphics, Vision and Image Processing (GVIP), Dec 2005.

[10] Bradski, G., Kaehler, A. (2008). Learning OpenCV. Retrieved from O' Relly Media Inc., 1005 Gravenstein Highway North, Sebastopol, CA 9547.

[11] Baggio, D., Emami, S., Escriva, D., Ievgen, K., Mahmood, N., Saragih, J. Shilkrot, R. (2012) "Mastering OpenCV with Practical Computer Vision Projects". Retrieved from Packt Publishing.

[12] Wenjing, J., Xiangjian, H., Huaifeng, Z., Qiang, W. (2007). "Combining Edge and Colour Information for Number Plate Detection" Retrieved from Proceeding of Image and Vision Computing, pp. 227-232.

[13] Khurram, K., Imran, S., Claudie, F., Nicole, V. "Comparison of Niblack inspired Binarization methods for ancient documents" Retrieved from http://www.ppgia.pucpr.br/~facon/ Binarizacao/NiblackComparison.pdf