Development of automatic number plate recognition system

B Amirgaliyev[1*], M Kairanbay[1], Ch Kenshimov[1], D Yedilkhan[2]

[1] *ABY Applied Systems*

[2] *International Information Technologies University*

Received 1 March 2014, www.tsi.lv

Abstract

Today, the automatic number plate recognition (ANPR) system is a key aspect in traffic congestion. This will help minimizing the different kind of violations in the road. Advanced systems for tracking and fixing stolen, unauthorized vehicles are based on automated number plate recognition. This paper's main objectives is to review other methods and develop, at the same token evaluate our proposed approach. A very short review is performed on the various methods of number plate recognition systems. Further explanations of the proposed algorithm is illustrated in graphical forms to show how algorithm works. The paper is going to be concluded with test and evaluation results.

Keywords: ANPR, Plate area, Segmentation, OCR

1 Introduction

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

Automated number plate recognition system is a key aspect in resolving all the problems listed before. Adding other features to the system, we can identify and track the vehicle, fix the time and coordinates of appearance and disappearance of the cars.

Automated number plate recognition system consists of three main parts like number plate localization, number plate segmentation and optical character recognition.

2 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 colour 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 region.

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 can be used only in specific countries with specific characteristics of number plate like background colour etc.

3 Research objectives

The following list gives the objectives of this research paper:

1) To solve and develop the automated number plate recognition system;
2) To evaluate, test developed system and presents the evaluation results.

4 Proposed approach

The whole problem consists of three parts:

1) Plate area detection;
2) Segmentation and extraction of characters from number plate;
3) Optical character recognition of extracted symbols.

4.1 PLATE AREA DETECTION

The image with number plate will be given as an input to the program and the number plate must be identified then

[*] ***Corresponding author*** e-mail: amirgaliyev@gmail.com

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 image.

4.1.1 Grey scale image

In this stage, we need to read the image and convert it to grey 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).

4.1.2 Blur

The noise is a main issue in our problem. In order to reduce them it is better blur the image. There are different types of smoothing such as homogeneous, Gaussian, median as well as bilateral [5]. The following cumulative error distribution graph shows the comparison among each of them.

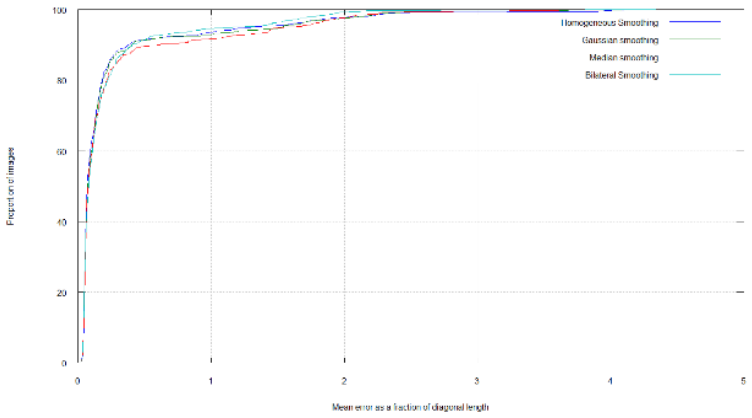

FIGURE 1 Cumulative error distribution for smoothing

According to the figure above, we can say that the homogeneous smoothing is the best one in comparison to others.

4.1.3 Vertical edge detection

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 too close to each other [1]. The edge detection is basic and fundamental operations in computer vision field. There are exist different kind of edge detectors like Prewitt, Sobel, Canny and etc. Each of them is used in different cases and problems. In our problem we use Prewitt, Sobel and modified version of Sobel [9]. 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 to find the number plate. Vertical edge detection can be implemented using convolution operation with specific matrix. For Prewitt and Sobel, the following matrices will be used.

$$\mathbf{G_x} = \begin{bmatrix} -1 & 0 & +1 \\ -1 & 0 & +1 \\ -1 & 0 & +1 \end{bmatrix} * \mathbf{A} \qquad \mathbf{G_y} = \begin{bmatrix} +1 & +1 & +1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} * \mathbf{A}$$

FIGURE 2 Prewitt edge detector for vertical and horizontal edges

$$\mathbf{G}_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * \mathbf{A} \qquad \mathbf{G}_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * \mathbf{A}$$

FIGURE 3 Sobel edge detector for vertical and horizontal edges

where G_x, G_y derivative of image in X and Y directions respectively, A – is an input image. 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:

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

FIGURE 4 Gradient magnitude and direction

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 Figure 5 illustrate the result and difference of each edge detector.

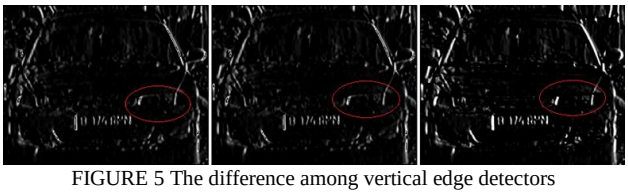

FIGURE 5 The difference among vertical edge detectors

4.1.4 Binary Image

Binary image contains only two colours such as black and white. There are wide spectrum of methods for threshold like Otsu, Niblack, Souvola, Wolf, Feng and 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.

4.1.5 Close morphology

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 together. The Figure 6 illustrates the result of morphology operation.

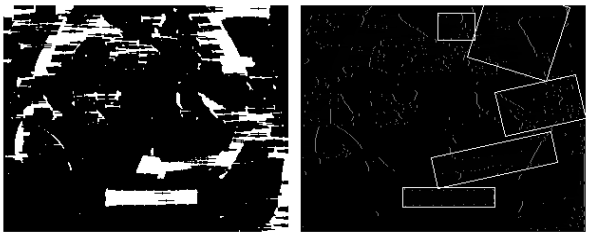

FIGURE 6 Image after morphology and finding contours

4.1.6 Find contours

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 Figure 6.

4.1.7 Find correct candidate

Find correct candidate. In order to find the correct candidate among others, each number plate must be investigated. First of all, characters from number plate must be extracted. In order to do that, we need to take a threshold of image using nick method for determining the texture inside of each contour (Figure 7). After extraction of characters, each symbol will be recognized and appropriately, recognition probability will be returned for each character. Summing up all these probabilities for each symbol in number plate and taking into account the value that identifies how much the contours located in one line, we will get set of probabilities for each candidate. When we already have set of probabilities for each candidate, we must to count the number of probabilities that are more than 0.9 (0.9 got from practical observation and testing). The candidate with maximum number of such probabilities will be considered as number plate.

FIGURE 7 Find contours that look like character

4.2 SEGMENTATION AND EXTRACTION OF CHARACTERS FROM NUMBER PLATE

From number plate characters must be extracted. There are two basis algorithms for segmentation, where first one is based on projection of image into X axis, however, the second one based on finding of contours that look like character. According to investigation and testing, we came to conclusion that second algorithms works better than first one. The Figure 8 shows the result of this algorithm.

4.3 OPTICAL CHARACTER RECOGNITION (OCR)

Extracted character must be recognized. For recognition, modified version of 1NN algorithm was used. The character was divided to small 49 subparts like in the Figure 9. For each subpart, the number of white pixels should be counted. The feature vector that contains 49 elements will identify each character. In the Figure 10 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 neighbour classes that are closest to that element.

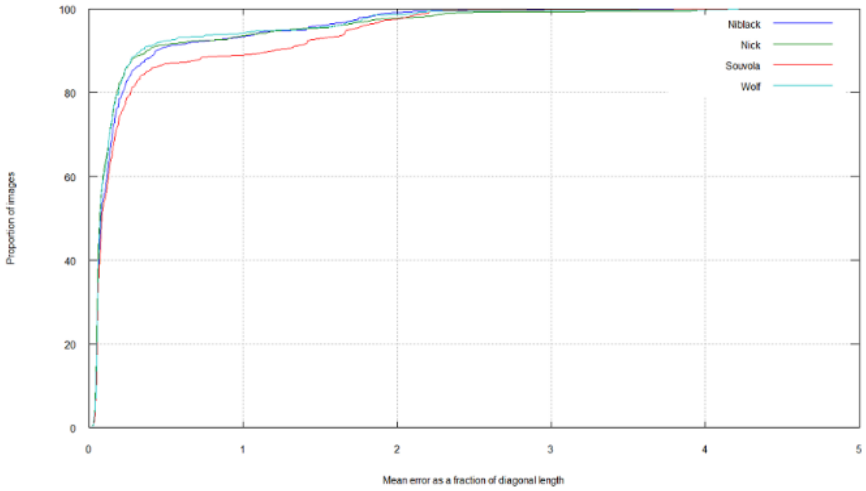

FIGURE 8 Threshold methods for identifying the texture in number plate

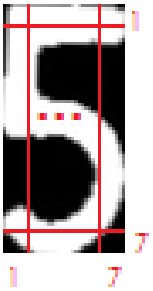

FIGURE 9 The character that divided to small (7 x 7) subparts

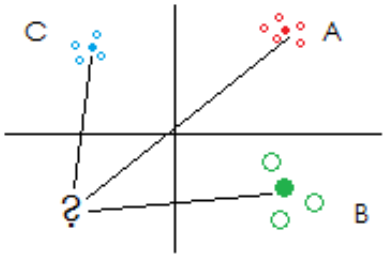

FIGURE 10 1NN for optical character recognition

The Figure 11 shows the result of whole algorithm.

FIGURE 11 The result of whole algorithm

5 Result and discussion

The program was tested with 1469 real car photos. The cars were taken from different sides and in different climate conditions. The Figure 12 illustrates the test cases that were used in testing stage.

FIGURE 12 Test cases

5.1 PLATE AREA DETECTION

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

TABLE 1 The result of plate area detection Points sizes and types styles

Front	Rear	Subpart	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 whole performance of plate area detection system: (95.3% + 93.15% + 95.5% + 94.46% + 93.88%) / 5 = 94.458%.

5.2 SEGMENTATION

The number plates were grouped based on their formats in order to segment. There were used different kinds of formats of number plate. The most popular are:

- (KZ) DDD LLL | DD (8 character);
- (KZ) DDD LL | DD (7 character);
- L DDD LLL (7 character);
- L DDD LL (6 character),

where D – is digit and L – is letter. The number plates above are new and old types of number plate in Kazakhstan, where (KZ) is a prefix and last two digits identify the region of Kazakhstan. As well as, there were used the number plates from other countries such as Russia, Kyrgyzstan etc. The following are their formats:

- L DDDD L (6 character)
- LLLL DDDD (8 character)
- LLL DDD (6 character)

At the same token, we encounter with the some company cars that have their own number plate formats like L DDDDDD (7 character). Looking to the number plate formats above, we came to conclusion to decompose all number plates based on number of characters. In other words, we have three groups, where in first group - 6 characters (60 number plates), second group - 7 characters (976 number plates) and last group contains the number plates with 8 characters (410 number plates).

TABLE 2 The result of segmentation

Types of number plate	6 character	7 character	8 character
Number of extracted characters	203 out of (60 * 6)	2553 out of (410 * 7)	5036 out of (976 * 8)
Percentage of extracted characters	56.38%	77.83%	64.49%

Taking into account the value for each group, we will find the total performance of segmentation algorithm (56.38% + 77.83% + 64.49%) / 3 = 66.23%. The Figure 13 shows the extracted letters after segmentation.

FIGURE 13 Extracted letters after segmentation

5.3 OPTICAL CHARACTER RECOGNITION

Some characters like “5 and S” and “O and 0” are looking 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 performance of our solution is 90%.

6 Conclusion and future work

The main parts of number plate recognition system was successfully implemented. Our proposed solution works for general cases, where there is no limit for the distance from camera to the vehicle and climate 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.

Acknowledgments

This research study was funded by the World Bank, Ministry of Education and Science, Kazakhstan. This paper would not have been possible without the assistance, support and patience of our co-workers. Finally, I would like to thank our parents for giving birth to us and supporting us throughout our life.

References

[1] Baggio D, Emami S, Escriva D, Ievgen K, Mahmood N, Saragih J, Shilkrot R 2012 *Mastering OpenCV with Practical Computer Vision Projects* Packt Publishing
[2] Kroto H W, Fischer J E, Cox D E 1993 *The Fullerenes* Pergamon:Oxford
[3] Benjapa R, Kittawee K, Paruhat P, Thaweesak Y 2012 License Plate Detection Based on Template Matching Algorithm *International Conference on Computer and Communication Technologies ICCCT'2012 May 26-27* Phuket
[4] Beverly S, Will H, Peter L, Patrico R 2012 Automatic Number Plate Recognition *Retrieved from CS 175 Fall '12*
[5] Kuo-Ming H, Ching-Tang H 2010 A Real-Time Mobile Vehicle License Plate Detection and Recognition *Tamkang Journal of Science and Engineering* **13**(4) 433-42
[6] Bradski G, Kaehler A 2008 *Learning OpenCV* O’ Relly Media Inc: 1005 Gravenstein Highway North, Sebastopol CA 9547
[7] Duan T., Hong Du T, Phuoc T, Hoang N 2005 Building an automatic vehicle license plate recognition system *Proc. Int. Conf. Comput. Sci. RIVF* 59–63
[8] Khurram K, Imran S, Claudie F, Nicole V 2005 *Comparison of Niblack inspired Binarization methods for ancient documents* http://www.ppgia.pucpr.br/~facon/Binarizacao/NiblackComparison.pdf
[9] Ondrej M 2007 *Algorithmic and Mathematical Principles of Automatic Number Plate Recognition Systems* BSc Thesis Bruno university of Technology
[10] Wenjing, J, Xiangjian H, Huaifeng Z, Qiang W 2007 Combining Edge and Colour Information for Number Plate Detection *Proceeding of Image and Vision Computing* New Zealand 227-32
[11] Yoshimori S, Mitsukura Y, Fukumi M, and Akamatsu N 2003 License plate detection using hereditary threshold determine method *Lecture Notes in Artificial Intelligence* **2773** ed V Palade, R J Howlett and L C Jain New York: Springer-Verlag 585–93

Authors

Beibut Amirgaliyev, born in November 27, 1986, Kazakhstan, Almaty

Current position, grades: Scientific director
University studies: PhD
Scientific interests: AI, Computer vision
Publications: 17
Experience: 7 years

Magzhan Kairanbay, born in October 26, 1991, Kazakhstan, Almaty

Current position, grades: Junior research fellow in Computer Vision field
University studies: Master, International IT University, Kazakhstan
Scientific interest: Computer Vision, Machine Learning, Artificial Intelligence
Publications: A Review And Evaluation of Shortest Path Algorithms, International Journal of Scientific & Technology Research Volume 2, Issue 6, June 2013
Experience: Junior research fellow in Computer Vision field in ABY Applied Systems**,** Development and analysis wide spectrum of projects related to Computer vision, Machine Learning.

Chingiz Kenshimov, born in May 20, 1987, Kazakhstan, Taraz

Current position, grades: Project manager
University studies: Master
Scientific interests: Machine Learning, AI, Computer Vision
Experience: 7 years

Didar Yedilkhan, born in May 30, 1990, Kazakhstan, Almaty

Current position, grades: PhD student
University studies: AI
Scientific interests: AI, Machine Learning
Publications: 3
Experience: 3 years