Vertical Edge Based Algorithm for Korean License Plate Extraction and Recognition

Mei Yu*, Yong Deak Kim* *Regular Members*

ABSTRACT

Vehicle license plate recognition identifies vehicle as a unique, and have many applications in traffic monitoring field. In this paper, a vertical edge based algorithm to extract license plate within input gray-scale image is proposed. A size-and-shape filter based on seed-filling algorithm is applied to remove the edges that are impossible to be the vertical edges of license plate. Then the remaining edges are matched with each other according to some restricted conditions so as to locate license plate in input image. After license plate is extracted, normalized and segmented, the characters on it are recognized by template matching method. Experimental results show that the proposed algorithm can deal with license plates in normal shape effectively, as well as the license plates that are out of shape due to the angle of view.

Ⅰ. 서 론

Vehicle license plate recognition have many applications in traffic monitoring field, such as automatic tolling, parking control, finding stolen cars, gate control for secured area, enforcement purposes, for example, speed violation, entry to illegal traveling zones, and so on. Different from vehicle classification, license plate recognition identify vehicle as a unique.

Some methods to extract license plate in color image have been proposed, because license plates usually have appointed colors in their backgrounds and characters. In study [1], each pixel of an input image is converted into one of four color groups (green, red, white and the others) by neural networks, and then histogram of white, red and green are calculated to extract the license plate. It is claimed that the recognition rate is 91.25% with 80 test images. However, it is difficult to extract license plate when the vehicle has similar color to the background of the license plate.

Many methods have also been presented to extract vehicle license plates in gray-scale images, because color image processing usually takes more computational efforts than gray-scale image processing. Since the characters on a license plate usually make a clear "signature" which corresponds to strong gray level variations at somehow "regular" intervals, histogram based methods are used to extract license plate in input image [2]. However, other characters on vehicles will disturb the license plate extraction. Additionally, it is difficult for histogram based methods to extract license plate accurately if the plate is out of shape due to the angle of view, especially for Korean license plate which consists of characters of two rows. Hough transform for extracting edges is another method to extract vehicle license plate [3, 4], but it requires huge memory and computational efforts. Study [5] proposed a license plate recognition method based on histogram and template matching, and it is claimed that the method can correctly recognize 42 license plates within 50 images.

In this paper, a vertical edge based algorithm

* Division of Electronics Engineering, Ajou University
논문번호 : 99397-0928, 접수일자 : 1999년 9월 28일

for license plate extraction is proposed for gray-scale images, after that, character segmentation and recognition are also introduced. Fig.1 gives the flow of the proposed algorithm, in which the part within the dotted frame is the step of license plate extraction. Experimental results show that the proposed algorithm can recognize license plates in normal shape, as well as license plates that are out of shape due to the angel of view.

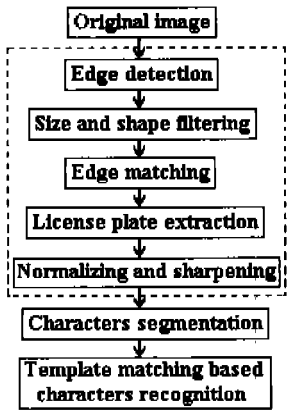

Fig. 1 Flow of license plate recognition.

Ⅱ. Vertical Edge Based Algorithm for License Plate Extraction

License plate extraction is the key step within license plate recognition system, and influences the accuracy of the system significantly. Even though license plates have appointed colors in their backgrounds and characters, the gray levels of license plates in images vary greatly due to different light conditions under which the images are taken. Moreover, the vehicle may have similar gray levels to the background of license plate. Thus, it is not easy to segment license plate by directly using information on plates' gray levels. By contrast, license plates usually have relatively clear outline, which can be used to separate license plates from vehicles. Based on the point of view, a vertical edge based method is proposed to extract license plate in gray-scale image in this paper.

2.1 Edge detection

Most of vehicles usually have more horizontal lines than vertical line. If the two vertical edges of a license plate is detected correctly, the four corners of the license plate can then be located, so that the license plate can be extracted accurately from the input image. Therefore, as an alternative, only the vertical edges of input image are employed to extract the license plate. For a given gray-scale image $\{G_{i,j}; 1\le i\le K, 1\le j\le L\}$, its vertical edge image is obtained by

$$D_{i,j}=\frac{1}{3}\sum_{y=j-1}^{j+1}|G_{i-1,y}-G_{i+1,y}|, \tag{1}$$

$$E_{i,j}=\begin{cases}1, & \text{if } D_{i,j}\ge T,\\ 0, & otherwise\end{cases} \quad 1\le i\le K,\ 1\le j\le L, \tag{2}$$

where $D_{i,j}$ represents vertical edge intensity function of the input image at pixel (i, j), while $\{E_{i,j}\}$ denotes the corresponding binary vertical edge image.

T in Eq.(2) is a threshold, and defined by

$$T=\begin{cases}T_D, & \text{if } T_D\ge T_0,\\ T_0, & otherwise\end{cases} \quad T_D=D_m+D_\sigma, \tag{3}$$

where D_m and D_σ denote the mean and standard deviation of $\{D_{i,j}; 1\le i\le K, 1\le j\le L\}$, respectively. T_0 (empirically, $T_0=26$) is the lower limit of the threshold. Fig.2(a) shows some input images, where the license plates in the last two images are out of shape due to the angle of view. The detected vertical edge images of Fig.2(a) are given in Fig.2(b).

2.2 Size-and-shape filtering

After the vertical edge image has been obtained, it is filtered so as to remove the edges that are impossible to be the vertical edge of license plate. Before filtering, morphological operation (dilation) is applied as a pre-processing. In the vertical edge image, an edge area is defined as a group of white pixels that are eight connected neighbors with each other. For each edge area, its size and shape are checked by

(a) input images (b) vertical edge images (c) filtered results (d) extracted plates (e) binary plates

Fig. 2 Examples of license plates extraction.

using seed-filling algorithm [6], since the algorithm visits all pixels within a connected area very quickly. The steps of size-and-shape filtering is as follows

1) For each white pixel $E_{i,j}$ in a vertical edge image, if the pixel has not been checked, then search the connected white edge area by using pixel $E_{i,j}$ as the seed.
2) Check whether the edge area forms a vertical beeline whose slope is within a predefined interval. If the edge area is smaller than a predefined square, or does not form a beeline, then fill the edge area with black.
3) Mark the whole pixels of this edge area as the checked pixels and continue to do the three steps until all white pixels in the vertical edge image have been checked.

During the size-and-shape filtering, the top and bottom coordinates of each remaining edge area are recorded for post-processing. Since the vertical edges of license plate may be cut off in the vertical edge image, edge areas are also recorded as one edge area if they have similar slope and ones top is quite close to the others bottom. The results of the size-and-shape filtering of Fig.2(b) are shown in Fig.2(c). It is seen that the filter has removed much noise in Fig.2(b).

2.3 Edge matching and license plate extraction

The ratio of width to height of Korean license plate is about 2:1, it can be used to judge whether two edge areas are the pair of vertical edges of license plate. Consider that the real ratio in the image may have some departure due to the angle of view, the possible range of the ratio is adopted from 1.4:1 to 3.3:1 in this paper. It is assumed that the license plate does not lean quite a lot in input image, therefore, the vertical coordinates of the two vertical edges of license plate should have small differences. It is restricted that the difference should be within half of the plate's height. Additionally, the two vertical edges of license plate will have similar height within the vertical edge image, which is also a restricted condition of the edge matching. In this paper, the height ratio of plate's two vertical edges is from 0.8 to 1.2. Let L be an edge area, X and Y be its corresponding horizontal and vertical coordinates, *Top*, *Bottom* and *Middle* represent its top, bottom and middle, respectively. N be the total number of edge areas in the vertical edge image. Then the edge matching is described in Fig.3.

Here, edge areas in vertical edge image are

for i=1 **to** N-1 **do**

{ $Y_{Middle}(L_i)=\frac{Y_{Bottom}(L_i)-Y_{Top}(L_i)}{2}$;

for j=i+1 **to** N **do**

if $\frac{Height(L_i)}{Height(L_j)}\geq 0.8$ **and** $\frac{Height(L_i)}{Height(L_j)}\leq 1.2$ **then** /* *Do* L_i *and* L_j *have similar height?* */

/* *Do* L_i *and* L_j *have similar vertical coordinates?* */

if $Y_{Middle}(L_i)>Y_{Top}(L_j)$ **and** $Y_{Middle}(L_i)<Y_{Bottom}(L_j)$ **then**

{ $ratio=\frac{|X_{Top}(L_i)-X_{Top}(L_j)|+|X_{Bottom}(L_i)-X_{Bottom}(L_j)|}{Height(L_i)+Height(L_j)}$;

/* *Is the ratio of width to height satisfies the condition?* */

if *ratio*>1.4 **and** *ratio*<3.3 **then**

{ Extract the region according to the pair of edge areas L_i and L_j,

and check whether it is license plate region during the plate segmentation and recognition;

if the extracted region is license plate region **then exit;**

}

}

}

Fig. 3 Edge areas matching algorithm.

compared with each other according to the above restricted conditions. If a pair of edge areas satisfies the conditions, it is regarded as the possible vertical edges of license plate, otherwise, another pair is checked. After the possible vertical edges of license plate have been found, the region is extracted and continue to be checked whether it is real license plate region during plate segmentation and recognition. For example, after license plate is segmented, the percentage of character regions (white pixels) on license plate is about from 10% to 40%. That is, if the percentage of character regions in the extracted region is lower than 10% or higher than 40%, it can not be license plate region. Thus, another pair of edge areas that satisfy the above restricted conditions of edge matching is searched. When license plate has been extracted, it is normalized and sharpened into a 200×100 gray-scale image, as shown in Fig.2(d).

Ⅲ. Character Segmentation and Recognition

In these steps, the normalized plate $\{P_{i,j};1\leq i\leq 200, 1\leq j\leq 100\}$ is first segmented into a binary image $\{B_{i,j}\}$. In Korea, the most common vehicle license plates are private plate and business plate. Their backgrounds are green and yellow respectively, while the corresponding characters are white and dark blue. The segmentations for them are a little bit different from each other.

For private plates, consider that the characters have higher gray levels than the background, the segmentation is represented as

$$B_{i,j}=\begin{cases}1, & \text{if } P_{i,j}\geq T_p,\\ 0, & otherwise\end{cases} \quad 1\leq i\leq W,\ 1\leq j\leq H, \tag{4}$$

where W and H denote width and height of the segmentation region, and the threshold T_p is defined by $T_p=P_m+P_\sigma$, in which the two terms are

$$P_m=\frac{1}{W\times H}\sum_{i=1}^{W}\sum_{j=1}^{H}|P_{i,j}|, \tag{5}$$

$$P_\sigma=\frac{1}{W\times H}\sum_{i=1}^{W}\sum_{j=1}^{H}|P_{i,j}-P_m|. \tag{6}$$

For business plates, the threshold T_p is defined by $T_p=P_m-P_\sigma$, and the pixels whose gray levels

Table. 1 License plate extraction and character recognition rate.

	License Plate Extraction	Character Recognition for Extracted License Plates				License Plate Recognition
		Region Name	Class Code	Usage Code	Serial Number	
Correct Recognition	699/710	698/699	1196/1210	691/699	2785/2796	670/710
Recognition Rate	98.45%	99.86%	98.84%	98.86%	99.61%	94.37%

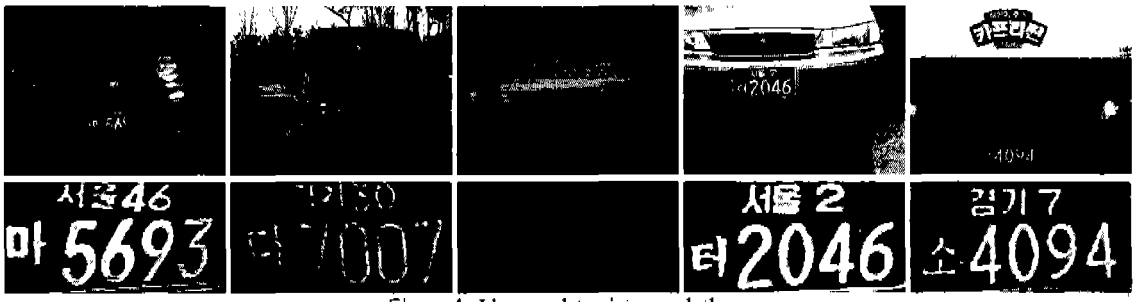

Fig. 4 License plates in normal shape.

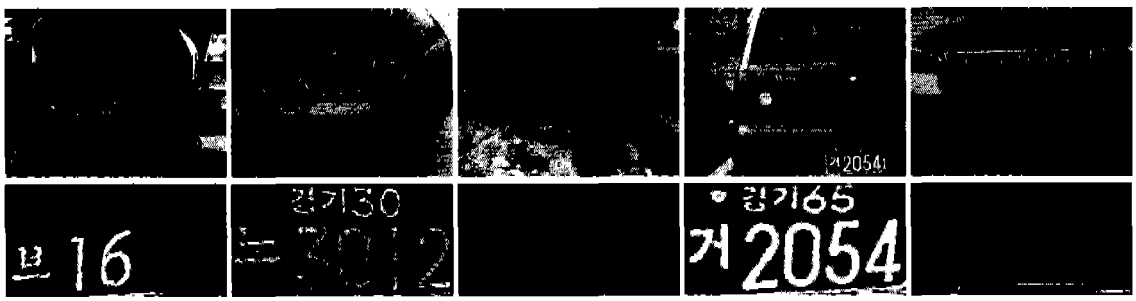

Fig. 5 Leaned license plates.

Fig. 6 Vehicle with similar color to plate.

are larger than the threshold are set to be 0, while the others are 1, because in this case the dark blue characters have lower gray levels than the yellow background.

Since the gray levels of different part of license plate may not be uniform because of the light condition, license plate is separated into three or four parts when it is segmented, that is, the part of characters in the up row (region name and class code), the part of Korean character in the down row (usage code), and the parts of four numbers in the down row (serial number). For each part, the threshold T_p is calculated separately according to Eq.(5) and Eq.(6). When the normalized license plate has been segmented into binary image, size-and-shape filter is employed to remove noise on the plate. Fig2.(e) shows the binary license plates $\{B_{i,j}\}$ obtained from Fig.2(d).

Then characters are segmented by using the horizontal and vertical histograms of the binary license plate image, combined with the knowledge of standard characters position on the license plate. Korean characters are normalized into 20×20 binary image, while the numbers are 8×16.

Since characters on license plates have the same font, character recognition algorithm based on template matching is employed because of its relatively lower computational efforts compared with neural networks. Template matching is also more tolerant to noise than structural analysis method.

There are totally 16 region names in our obtained Korean license plate images, which are 서울, 부산, 대전, 대구, 광주, 인천, 경북, 경남, 전북, 전남, 충북, 충남, 경기, 강원, 울산, and 제주. Korean characters of the region name is separated into two sets, one for the first characters, denoted by set S_1, another for the second characters,

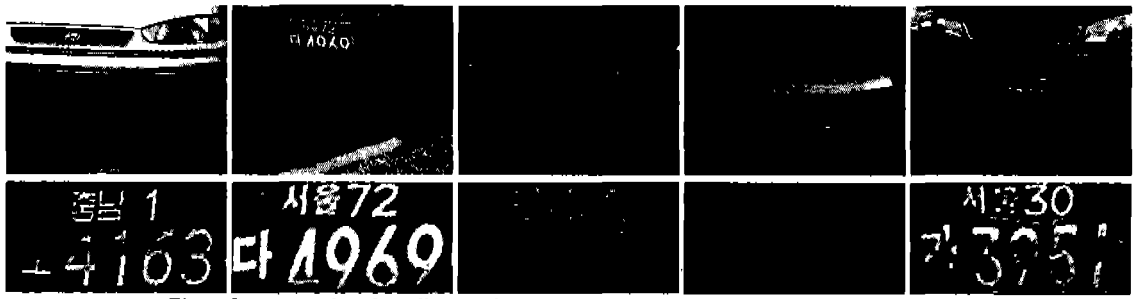

Fig. 7 Damaged or bent license plates.

Fig. 8 Dirty license plates.

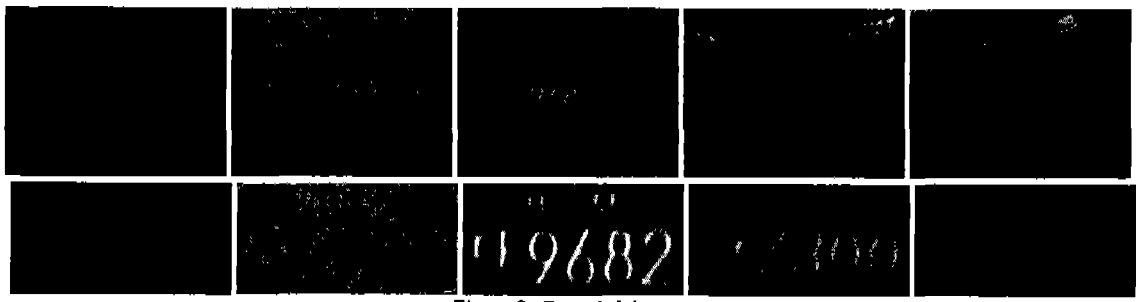

Fig. 9 Degraded images.

denoted by set S_2. Similarly, we can also define the other two sets S_3 (for usage code) and S_4 (for 0-9 ten numbers). Each unknown character is matched to the templates in the corresponding set, and the character is regarded as the one whose template has the smallest distance to the unknown character. When recognizing characters of region name, some rules are used to increase the recognition accuracy. For example, when 경기 is being recognized, the character 기 is not easy to be confused with other characters in set S_2, but 경 is similar to 강, 전 or 광. Since the reliability of 기 is higher than 경, the region name is still assumed to be 경기 if the second character is recognized as 기 and the first character is regarded as 강, 전 or 광.

Ⅳ. Experimental Results

Experiments have been implemented to test the efficiency of the above vertical edge based algorithm to recognize license plate in input gray-scale image, the size of which is 493×373. The test images are taken under various situations, such as sunshiny, raining, snowing, cloudy days. The license plates include both old and new ones.

To test the performance of the proposed algorithm under different situations, the experiments are implemented in the following six aspects: (1) license plates in normal shapes, (2) license plates that are out of shape or leaned due to the angle of view, (3) license plates which have similar color to vehicle bodies, (4) damaged or bent license plates, (5) dirty license plates, (6) degraded images (including under or over exposed images, and blurred images). Besides the experimental results in Fig.2, Figs.4-9 give more test images and their extracted license plates.

It is seen that the proposed algorithm can extract license plates in normal shape as well as skewed license plates, as shown in Fig.2 and Figs.4-5. The proposed algorithm is good at extracting license plates that have similar color to the vehicles. Fig.6 is such kind of examples, in which the first one is a green *Matiz* with green private plate, while the other is a yellow taxi with yellow business plate. The proposed algorithm also shows its efficiency in recognizing damaged or dirty license plates, as shown in

Fig.7 and Fig.8. Finally, Fig.9 gives some degraded test images, in which the first two images are under or over exposed, while the license plates in the last three images are blurred or noisy. The proposed algorithm succeeds in extracting license plates from these degraded images.

Table 1 gives the rate of license plate extraction and characters recognition of the proposed algorithm. Here, license plate recognition is thought as a failure so long as one character of license plate is incorrectly recognized. The second row in the table shows the number of correctly recognized license plates (or characters) and the total number of license plates (or characters). It is shown that the proposed algorithm has correctly extracted 699 license plates from total 710 test images. The 3-6 columns in the table give the character recognition rates of the 699 extracted license plates. For the total 710 input images, the number of correctly recognized license plates is 670, and the recognition rate is about 94.37%, as shown in the last column of Table 1. The average processing time of these 710 images is about 0.28 second per image, which include license plate extraction and character recognition, when implemented on PC 586(300MHz) in Borland C++ 5.02 language.

V. Conclusion

In this paper, a vertical edge based algorithm for extracting Korean license plate is proposed. To detect the two vertical edges of license plate, the obtained vertical edge image is filtered by a size-and-shape filter so as to remove the pixels that do not belong to the vertical edges of the license plate. Then the remaining edges are matched according to the knowledge of the license plate. Template matching is applied to recognize the characters on license plate. Experimental results show that the proposed algorithm can recognize license plates in normal shape as well as plates that are out of shapes due to the angle of view. In future study, we will improve the robustness of license plate recognition by combining the proposed algorithm with character-like area searching method, so as to locate license plate when its outline is unclear or corrupted by other things.

References

[1] E. R. Lee, P. K. Kim, and H. J. Kim, "Automatic recognition of a car license plate using color image processing", *Proc. of Ist International Conference on Image Processing*, vol. 2, pp. 301-305, USA, Nov. 1994.

[2] T. K. Lim, J. T. Ong, P.S.C. Lee, and T. Y. Lee, "Image segmentation and character recognition of vehicle licence plates for electronic road pricing", *Third International Conf. on Automation, Robotics and Computer Vision*, pp. 1663-1667, Singapore, Nov. 1994.

[3] B. T. Chun and H. S. Yoon, "A method to extract vehicle number plates by applying signal processing techniques", *Journal of the Korean Institute of Telematics and Electronics*, vol. 30B, no. 7, pp. 92-100, July, 1993.

[4] G. M. Kim, "The automatic recognition of the plate of vehicle using the correlation coefficient and hough transform", *Journal of control, automation and system engineering*, vol. 3, no. 5, pp. 511-519, 1997.

[5] D. U. Cho and Y. H. Cho, "Implementation of pre-processing independent of environment and recognition of car plate using histogram and template matching", *The journal of the Korean institute of communication sciences*, vol. 23, no. 1, pp. 94-100, 1998.

[6] Smith A. R., "Tint fill", *Computer Graphic*, vol.13, no.2, pp.276-283,1979.

Mei Yu received her B.S. and M.S. degrees in Computer Science from Hangzhou Institute of Electronics Engineering, China, in 1990 and 1993, respectively. She has been with Hangzhou Institute of Electronics Engineering as an assistant professor since 1996. She will receive her Ph.D. degree from Ajou University, Korea, in 2000. Her current research interests include image processing, computer vision, and intelligent transportation system.

Yong Deak Kim received his B.S., M.S. and Ph.D. degrees from the Department of Electronics Engineering at Yonsei University, in 1971, 1973 and 1976 respectively. From 1973 to 1974, he was a research member at the Dept. of Elect. Ecole Superieur D'electricite in France. From 1979 to 1980, he was a visiting professor at the Dept. of E.E. at Standford University. Since 1974, he has been a professor in the Dept. of E.E. at Ajou University in Korea. His current research interests include ATM Network, multimedia system, and intelligent transportation system.