Region-Based License Plate Detection

Wenjing Jia *, Huaifeng Zhang and Xiangjian He

Faculty of Information Technology, University of Technology, Sydney
PO Box 123, Broadway, NSW, 2007, Australia

Abstract

Automatic license plate recognition (ALPR) is one of the most important aspects of applying computer techniques towards intelligent transportation systems. In order to recognize a license plate efficiently, however, the location of the license plate, in most cases, must be detected in the first place. Due to this reason, detecting the accurate location of a license plate from a vehicle image is considered to be the most crucial step of an ALPR system, which greatly affects the recognition rate and speed of the whole system. In this paper, a region-based license plate detection method is proposed. In this method, firstly, mean shift is used to filter and segment a color vehicle image in order to get candidate regions. These candidate regions are then analyzed and classified in order to decide whether a candidate region contains a license plate. Unlike other existing license plate detection methods, the proposed method focuses on regions, which demonstrates to be more robust to interference characters and more accurate when compared with other methods.

Key words: License Plate Detection, Region, Mean-Shift Segmentation, Features, Mahalanobis Classifier

1 Introduction

In recent years, many researches on Intelligent Transportation Systems (ITS) have been reported. As one form of ITS technology, Automatic License Plate Recognition (ALPR) not only recognizes and counts vehicles, but distinguishes each as unique by recognizing the characters in the license plates. In the approach, a camera captures the vehicle images and a computer processes the captured images and recognizes the information on the license plate by applying various image processing and optical pattern recognition techniques.

* Corresponding Author.
Email address: wejia@it.uts.edu.au (Wenjing Jia).

Prior to the character recognition, the license plates must be located from the background vehicle images. This task is considered as the most crucial step in the ALPR system, which influences the overall accuracy and processing speed of the whole system significantly. Since there are problems such as poor image quality, image perspective distortion, other disturbance characters or reflection on vehicle surface, and the color similarity between the license plate and background vehicle body, the license plate is often difficult to be located accurately and efficiently. Even though many researches have been done in this area, this research is, however, far from a solved problem.

Methods have been proposed in order to solve the problem. In [1], [2], [3] and [4], the presence of abundant edges, especially vertical edges, in license plate regions due to the presence of characters is used to generate the candidate regions for classification. Combined with some prior geometrical properties of license plates, the algorithms can obtain good performance even when dealing with some deficient license plates [3][4]. Another kind of popular methods focuses on detecting the frames which hold the border lines of license plates, where the Hough Transform is widely used [5][6][7]. Besides the algorithms based on gray-level image processing, color information of license plates also plays an important role in license plates localization, where the unique color or color combination between the license plates and vehicle bodies are considered as the key feature to locate the license plates. In order to decide the exact color at a certain pixel, neural network classifier [8][9][10], genetic algorithm [11], etc. are widely used.

Currently, most researchers prefer a hybrid detection algorithm, where multiple features are involved in order to make the algorithm more robust. The algorithm proposed in this paper is also a hybrid algorithm. However, unlike the recently proposed hybrid method described in [3] and [4] (we call them *edge-based* method for simplicity hereafter), the proposed algorithm can be viewed as a *region-based* method. We firstly apply a mean shift procedure in spatial-range domain to segment the vehicle images to directly produce the candidate regions that may include license plate regions; multiple features are then used for classifying. Comparatively, in edge-based methods, candidate regions are actually generated artificially from features of characters (e.g., edges) within license plate regions. By doing so, candidate regions do not necessarily match the real license plates region, which in turn cannot guarantee the accurate detection result.

The remaining parts of the paper are organized in the following order. The proposed algorithm is described firstly, where candidate region producing, feature extraction, and Mahalanobis classification are introduced in Sect. 2, Sect. 3, and Sect. 4 respectively. Then, experimental results are given and analyzed in Sect. 5, where the proposed algorithm is compared with edge-based methods. Sect. 6 gives the conclusion.

2 Candidate Regions

In the area of applying computer for image processing, an image in 2-dimensional space can be typically digitized as a 2-dimensional lattice of r-dimensional vectors, where r is 1 for gray-level images and 3 for color images to represent the value at the point determined by the 2-dimensional lattice. The space of the lattice is known as the *spatial domain* while the gray level or the color value is known as the *range domain*. After a proper normalization, the space and range vectors can be concatenated to obtain a spatial-range domain of dimension $d = r + 2$. In the proposed method, a procedure named as *mean shift* [12] is applied for the data points in the joint spatial-range domain for image segmentation in order to produce candidate regions.

Mean shift algorithm is a nonparametric statistical method for seeking the main modes of a point sample distribution, in order words, seeking where the density function takes local maxima. The mean shift estimate of the gradient of a density function and the associated iterative procedure of mode seeking have been developed by Fukunaga and Hostetler [13] and largely forgotten until Cheng's paper [14] rekindled interest in it. Recently, Comaniciu proposed a practical method employing mean shift in the joint spatial-range domain of color images for discontinuity preserving filtering and image segmentation [12].

Using mean shift filtering in the joint spatial-range domain, each data point is associated to a point of convergence which represents the local mode of the density in the d-dimensional space. The output of the mean shift filter for a pixel is defined as the range value of the convergence point. The procedure continues until all pixels in the image are scanned. Since this process takes into account simultaneously both the spatial and range information of the image, it can achieve a high quality, discontinuity preserving spatial filtering. For the task of image segmentation, the convergence points sufficiently close in the joint spatial-range domain are further fused properly in order to obtain the homogeneous regions in the image. We give the procedure of mean shift image segmentation in the joint spatial-range domain as follows.

The segmentation operation is implemented on the mean shift filtered images. The input original image is firstly normalized with a uniform kernel function, where the bandwidth in spatial domain and range domain are denoted as h_s and h_r respectively.

Let $\{\vec{x}_j\}_{j=1,2,\ldots,n}$ and $\{\vec{z}_j\}_{j=1,2,\ldots,n}$ be the d-dimensional original image points and filtered image points in the spatial-range domain respectively. The subscripts s and r will be used to denote the spatial and range components of a vector respectively. The original image is assumed to be normalized using a uniform filter with the bandwidths in spatial domain and range domain of h_s

(a) (b)

Fig. 1. A vehicle image and its segmentation results using mean shift. (a) Original image with a dimension of 324 × 243. (b) Segmented image with $(h_s, h_r, M) = (7, 6.5, 200)$.

and h_r respectively. We use $\{L_j\}_{j=1,2,...,n}$ to denote a set of labels for different regions after segmentation.

(1) For each $j = 1, 2, ..., n$, run the mean shift filtering procedure for $\vec{x}_j$ and store the convergence point in $\vec{z}_j$, as:
 (a) Initialize $k = 1$ and $\vec{y}_k = \vec{x}_j$.
 (b) Compute $\vec{y}_{k+1} = \frac{1}{n_k} \sum_{\vec{x}_i \in S_1(\vec{y}_k)} \vec{x}_i$, $k \leftarrow k + 1$ until convergence, where n_k is the number of points in a window $S_1(\vec{y}_k)$ with unit radius and center at $\vec{y}_k$.
 (c) Assign $\vec{z}_j = (\vec{x}^s_j, \vec{y}^r_{conv})$ to specify that the filtered data at the spatial location of $\vec{x}_j$ will have the range components of the point of convergence $\vec{y}_{conv}$.
(2) Delineate the clusters $\{C_p\}_{p=1,2,...,m}$ in the joint domain by grouping together all $\vec{z}_j$ which are closer than h_s in the spatial domain and h_r in the range domain.
(3) For each $j = 1, 2, ..., n$, assign $L_j = \{p | \vec{z}_j \in C_p\}$.
(4) Eliminate spatial regions that are smaller than M pixels.

Thus, using mean shift procedure in the joint spatial-range domain, the input color vehicle images are segmented into many regions. Fig. 1(b) gives such a segmentation result when being applied in a vehicle image, where the pixels in a region share the same color and different regions are represented by different colors. The parameters used in segmentation are set experimentally. The selection criteria is related to the proportions of the size of license plates to that of vehicle images.

The resulted regions are hereafter called *candidate region* for simplification in this paper. Each candidate region will be further analyzed to decide if it really contains a license plate based on its feature analysis.

3 Feature Extraction

After the candidate regions are obtained by applying mean shift segmentation, features of each region are to be extracted in order to correctly differentiate the license plate regions from others. Various features have been utilized for this purpose. Such features include the size of the region, the width (or the height) of the region, the orientation of the characters, edge intensity, and position of the region. Some features, however, can only deal with images captured under very specific environment. In this experiment, three features are defined and extracted in order to decide if a candidate region contains a license plate, namely, rectangularity, aspect ratio and edge density. Even though these features are not scale-invariant, luminance-invariant, rotation-invariant, but they are insensitive to many environment changes.

3.1 Rectangularity

The license plate takes a rectangle shape with a predetermined length to width ratio in each kind of vehicles. Due to the view angle and uneven or curvy road surface, however, when taken into image, the license plate in a vehicle image is usually not a standard rectangular any longer. Under limited distortion, however, license plates in vehicle images can still be viewed approximately as rectangle shape with a certain aspect ratio. This is the most important shape feature of license plates.

Rectangularity is a measurement that represents how well an object fits its minimum enclosing rectangle (MER) [15](page 492), which is defined as the ratio of the area of the object A_O and the area of the object's MER A_{MER}. The *area* here is defined as the total number of pixels in the region.

One quick and straightforward method is to rotate the object region as a rigid body to get its MER. With this method, the object is rotated as a rigid body through a range in steps of $\Delta\theta$. After each incremental rotation, a horizontally oriented enclosing rectangle (ER) is fit to the boundary. The rotating angle at which the ER takes the minimum value is picked out. The size and dimension of the ER at this angle can be taken to be the A_{MER} and the dimension of the region.

3.2 Aspect Ratio

The *aspect ratio* is defined as the ratio of the width to the height of the region's MER. Since the MER of the object region can be computed via rotating the region in previous section, the dimension of the object's MER can be taken as the width and the height of the region.

3.3 Edge Density

Applying the above two features to filter the segmented regions, a lot of non-license-plate regions can be removed. However, there are still many candidate regions left which take similar rectangularity and aspect ratio features as the license plate regions do, such as often the head lights. Considering that the license plate regions generally take higher local variance in its pixels' values due to the presence of characters, an important feature to describe license plate region is local variance, which is quantized using the edge density.

The *edge density* is measured in a region R by averaging the intensities of all edge pixels within the region, as illustrated in (1):

$$D_R = \frac{1}{N_R} \sum_{m,n \in R} E(m,n) \tag{1}$$

where $E(i,j)$ represents the edge magnitude at location (i,j), and N_R is the number of pixels in region R. It is observed that most of the vehicles usually have more horizontal edges than vertical edges [16]. While in the region of license plates, or other character regions, there always contain abundant vertical edges due to the presence of characters. In order to reduce the complexity of algorithm, the vertical edges E_V only are used to compute the edge density. In this paper, edge information is acquired by applying a 3×3 Sobel edge detector in "vertical" direction only.

4 Mahalanobis Classification

When features have been extracted, people may have different methods to use these features. The most straightforward method is to set a threshold to each feature so as to make a positive or negative decision directly from each feature. By this way, some regions will be killed once there is one feature that does not look like a license plate. As this result, some researchers defined a so-called *possibility value* [1] or *gross confidence value* [17] for each candidate region to measure the possibility that the region contains a number plate. By this way, only one threshold is needed to make a decision. The possibility value is usually computed as the weighted sum of the extracted features. Obviously, the selection of the weighting coefficients affects the classification rate critically. However, in practice, this is generally decided empirically, like in [1], where

Table 1
The statistical mean data of the three features for three classes of regions

Classes	**Rectangularity**	**Aspect Ratio**	**Edge Density**
Plates-1	0.93	2.63	159.46
Plates-2	0.93	3.78	212.13
Non-plates	0.72	3.42	24.28

these parameters can be tuned for a class of images. The second method is to generalize a feature map based on a certain key feature. Then a proper binarization algorithm will be applied to convert the feature map into a binary-level image. Morphological operations, such as opening operation, and some specific merging rules are needed to joint small broken blocks into larger ones in order to get the final plate regions [3]. The last method is to input the features to a classifier, for example a neural network [2], in order to make the decision.

In this paper, we use a compromised method to take use of the features in order to make the decision. For each candidate region, above three defined features are extracted to compose the feature vector $\mathbf{x}$. A minimum-distance-based classifier is then applied to make the decision, where the Mahalanobis distance [18](page 85) is used as the distance measurement. We generate the models in the first place based on training using sample images, then the Mahalanobis distance between a feature vector and each model will be computed. An input feature vector x is classified by measuring the Mahalanobis distance from x to each of the centroid and attributing x to the models in which the Mahalanobis distance is minimum. In this way, no threshold is needed to set at all, while the training itself is not so hard to control like some complex classifiers.

In order to use the Mahalanobis distance as a minimum distance classifier, the involved features of each model must have a normal distribution. A study has therefore been made to define the distribution of probabilities of the chosen features. According to our observation, in our experiments there exist two classes of license plates, which suffer from a distinct difference in their aspect ratios. So, plus the non-plate region totally three classes of regions are defined, i.e., *plates-1*, *plates-2*, and *non-plates*. Statistics shows that the probability distributions of the three features of the two classes of plate regions approximately take normal distribution, which mean values for three classes of regions are shown in Table 1. It can be seen that two classes of plate regions share similar rectangularity features, but they are differentiated from each other in aspect ratio and edge density.

In this table, in order to reduce unnecessary computation, regions which size are smaller than 0.5% or larger than 5% of image size, which rectangularity are smaller than 0.5, or which aspect ratio are less than 1 or larger than 5 are removed from target candidates before classification. When the distance between the camera and vehicles is fixed or approximately fixed, the proportions of the size of license plates to that of vehicle images are basically determined. Thus, parameters can be chosen based on experiments. However, once assigned, these parameters need not be adjusted for other vehicles.

Feature vector's mean vectors $\{\mathbf{m}_k\}_{k=1,2,3}$ and covariance matrix $\{\mathbf{C}_k\}_{k=1,2,3}$ of three classes of regions have been computed using a training data set. In our experiments, 10% of sample data are randomly selected for training to compute the mean vector and covariance matrix. The others are used for testing. During testing, we measure the Mahalanobis distance [18](page 85) from the feature vector $\mathbf{x}$ to the mean vector $\mathbf{m}_x$ of the mode $\{\mathbf{w}_k\}_{k=1,2,3}$ of three classes of regions, which is defined in (2) and assign $\mathbf{x}$ to the class of the nearest mode.

$$d_k^2 = \|\mathbf{x} - \mathbf{m}_k\|^2 = (\mathbf{x} - \mathbf{m}_k)'\mathbf{C}_k^{-1}(\mathbf{x} - \mathbf{m}_k)k = 1, 2, 3 \tag{2}$$

In this experiments, as long as $d_1 < d_3$ or $d_2 < d_3$, the region is classified as a license plate region.

5 Experimental Results

In order to evaluate the performance of the proposed algorithm, we tested it on 57 various vehicle images captured from a highway intersection. The proposed algorithm can accurately locate license plates for those vehicle images where the background color of license plate is different from that of the vehicle body where the license plate adheres to. Moreover, several special cases are shown in order to compare its performance with that of other methods.

5.1 Robustness

This detection algorithm shows robustness to interference characters that may cause detection errors. Fig. 2 shows one case where there are interference characters ("ISUZU" at the top part of the picture) on the truck body. Since the interference characters have similar feature to those in the license plate, when local edge is used as the key feature in order to generate the candidate regions, it will cause misclassification since the region also has similar dimension feature and local variance feature as the plate. (Semantically, a license plate that contains only five letters is acceptable in Australia). Instead, our algorithm will succeed thanks to the mean shift segmentation (see Fig. 2(c)).

(a)

(b)

(c)

(d)

Fig. 2. (a) Interference characters ("ISUZU") in vehicle body and the license plate detection result. (b) Interference characters (two green letter P's) that are next to the license plate and the license plate detection results (c) and (d).

Fig. 2(b) shows another case that often happens in Australian vehicles, where there are interfering characters (two green letter P's) next to the license plate (the green letter "P" stands for "provisional", which is used to indicate a stage of driver licenses). Using edge-based methods, the candidate region will unavoidably become larger in horizontal direction since they care only the edge information, which results in that the dimension feature of the candidate does not meet the prior knowledge of license plates, and this in turn causes the detection failure. Comparably, our proposed algorithm has no difficulty at all when dealing with this situation (see Fig. 2(d)) thanks, again, to the mean shift segmentation in the joint spatial-range domain, where color information has been considered as well.

5.2 Accuracy

Since the candidate regions used in the algorithm are obtained directly from the segmentation results, which is decided by the spatial-range value of pixels, the detected regions are more close to the real area of the license plates, when compared with the methods where the candidate regions are actually generated from certain features of characters. In order to show the accuracy of detected regions using the proposed algorithm, we compare our experimental results with those obtained using the edge-based methods [3][4]. For those cases where both algorithms can find the correct position of license plates, experimental results show that the detected regions using our algorithm are

(a) (b) (c)

Fig. 3. The comparison between the detected results obtained using the edge-based methods (b) and the results obtained using our method (c) on a vehicle image (a).

much more close to the real license plates. Such an example is given in Fig. 3.

We use a matching rate to measure the matching degree between detected regions and the real ones. The matching rate is defined as $R = 1-(R_{FP}+R_{FN})$, where $R_{FP} = N_{FP}/N_T$ and $R_{FN} = N_{FN}/N_T$ are the rate of false positives (FP) and the rate of false negatives (FN). A *false positive*, also called *false alarm*, exists when a detection reports incorrectly that it has found a license plate pixel where none exists in reality. A *false negative*, also called a *miss*, exists when a detection reports incorrectly that a pixel does not belong to a license plate when, in fact, it does. We use the total number of false positives, N_{FP}, the total number of false negatives, N_{FN}, and the total number of pixels of real license plate regions, N_T, to calculate the matching rate. The real license plate regions are obtained manually with the help of lasso tools from the software Adobe Photoshop.

Under this definition, the detected region can only match 60.8% (26.0% false alarms and 13.2% miss) of the real plate region in Fig 3(b), while in our case (see Fig 3(c)), the matching rate is as high as 98.7%. Although the shapes of regions obtained using edge-based methods could be further modified by continually adjusting the parameters used in morphological operations to make them more close to a rectangle shape, candidate regions generated from the features of characters do not necessarily exactly match the real license plate region. This mismatch may in turn result in recognition error. A statistical analysis of the detection accuracy using the above definition shows that an average matching rate of 97.6% can be obtained with our method.

6 Conclusion

A region-based method is proposed for accurate license plate detection. The algorithm firstly generates candidate regions using mean shift segmentation, then decides if a region of interest really contains a license plate based on the feature analysis of candidate regions. The proposed algorithm shows great robustness when detecting license plate from vehicles which have a different color with that of the license plate where it adheres to. The proposed algorithm is demonstrated to be robust when dealing with vehicles where interference characters exist thanks to the its using mean shift segmentation in spatial-range domain to generate candidate regions. Experimental results also show more accurate detection results than the edge-based method as the candidate regions in this algorithm are obtained directly from color images rather than generated from features.

References

[1] S.-Z. Wang, H.-J. Lee, Detection and recognition of license plate characters with different appearances, in: Proceedings of 2003 IEEE Intelligent Transportation Systems Conference, Vol. 2, 2003, pp. 979–984.

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

[3] H. Bai, J. Zhu, C. Liu, A fast license plate extraction method on complex background, in: Proceedings of 2003 IEEE Intelligent Transportation Systems Conference, Vol. 2, 2003, pp. 985–987.

[4] H. Bai, C. Liu, A hybrid license plate extraction method based on edge statistics and morphology, in: Proceedings of the 17th International Conference on Pattern Recognition, Vol. 2, 2004, pp. 831–834.

[5] V. Kamat, S. Ganesan, An efficient implementation of the hough transform for detecting vehicle license plates using dsp's, in: Proceedings of 1995 Real-Time Technology and Applications Symposium, 1995, pp. 58–59.

[6] S. Gendy, C. L. Smith, S. Lachowicz, Automatic car registration plate recognition using fast hough transform, in: Proceedings of The Institute ofElectrical and Electronics Engineers 31st Annual 1997 International Carnahan Conference on Security Technology, 1997, pp. 209–218.

[7] Y. Yanamura, M. Goto, D. Nishiyama, M. Soga, H. Nakatani, H. Saji, Extraction and tracking of the license plate using hough transform and voted block matching, in: Proceedings of IEEE Intelligent Vehicles Symposium, 2003, pp. 243–246.

[8] W. Wei, M. Wang, Z. Huang, An automatic method of location for number-plate using color features, in: Proceedings of 2001 International Conference on Image Processing, Vol. 1, 2001, pp. 782–785.

[9] K. I. Kim, K. Jung, J. H. Kim, Color texture-based object detection: An application to license plate localization, in: Proceedings of the First International Workshop on Pattern Recognition with Support Vector Machines, 2002, pp. 293–309.

[10] S. H. Park, K. I. Kim, K. Jung, H. J. Kim, Locating car license plates using neural networks, Electronics Letters 35 (1999) 1475–1477.

[11] S. K. Kim, D. W. Kim, H. J. Kim, A recognition of vehicle license plate using a genetic algorithm based segmentation, in: Proceedings of 1996 International Conference on Image Processing, Vol. 2, 1996, pp. 661–664.

[12] D. Comaniciu, P. Meer, Mean shift: a robust approach toward feature space analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (2002) 603–619.

[13] K. Fukunaga, L. Hostetler, The estimation of the gradient of a density function, with applications in pattern recognition, IEEE Transactions on Information Theory 21 (1) (1975) 32–40.

[14] Y. Cheng, Mean shift, mode seeking, and clustering, IEEE Transactions on Pattern Analysis and Machine Intelligence 17 (1995) 790 – 799.

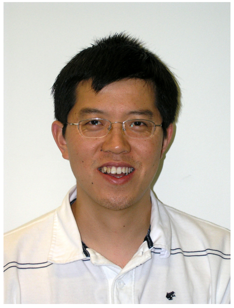
[15] K. R. Castleman (Ed.), Digital Image Processing, Prentice Hall, Toronto, 1996.

[16] M. Yu, Y. D. Kim, An approach to korean license plate recognition based on vertical edge matching, in: Proceedings of 2000 IEEE International Conference on Systems, Man, and Cybernetics, Vol. 4, 2000, pp. 2975–2980.

[17] D.-S. Gao, J. Zhou, Car license plates detection from complex scene, in: Proceedings of the 5th International Conference on Signal Processing, 2000, pp. 1409–1414.

[18] J. M. de Sá, Pattern recognition: concepts, methods, and applications, Springer, New York, 2001.

Wenjing Jia received her Bachelor degree from Changchun Institute of Posts and Telecommunications, China in 1999 and her Master degree from Fuzhou University, China in 2002. She was an associate lecture in Fuzhou University from 1999 to 2003. Currently, she is a PhD candidate in Faculty of IT at University of Technology, Sydney. Her current research interests include image processing, pattern recognition, object detection and recognition, and color image/video retrieval.



Huaifeng Zhang received his BS and MS degree from Dalian University of Technology and Xi'an Jiaotong University, China, in 1997 and 2000 respectively. In 2004, he received his PhD degree from Chinese Academy of Sciences, and then joined University of Technology, Sydney, Australia. His research interests include object detection and recognition, moving object tracking, 3D reconstruction and registration.

Xiangjian He, Associate Professor, received his BSc degree from Xiamen (Amoy) University, and MSc degree from Fuzhou University, China in 1982 and 1985 respectively. He received his PhD degree in the area of Computer Vision from the University of Technology, Sydney (UTS) in 1999. He is the Deputy Director of the UTS university-level Computer Vision Research Group, and a Senior Member of IEEE. He has been carrying out research in the areas of computer vision and image processing based on the Spiral Architecture since 1996. He is the chair of the 3rd International Conference on Information Technology and Applications (ICITA2005), and the chief editor of the conference proceedings published by IEEE CS Society. He is an editor of HIT Transaction on ECCN.