

Mean Shift for Accurate License Plate Localization

Wenjing Jia, Huaifeng Zhang, Xiangjian He and Massimo Piccardi, *Member, IEEE*

Abstract—This paper presents a region-based algorithm for accurate license plate localization, where mean shift is utilized to filter and segment color vehicle images into candidate regions. Three features are extracted in order to decide whether a candidate region represents a real license plate, namely, rectangularity, aspect ratio, and edge density. Then, the Mahalanobis classifier is used with respect to above three features to classify license plate regions and non-license plate regions. Experimental results show that the proposed algorithm produces high robustness and accuracy.

I. INTRODUCTION

LICENSE plate localization, also called license plate detection, is used to locate license plates from vehicle images accurately and efficiently. This task is considered to be the most crucial step in automatic license plate recognition (ALPR) systems. Since there are problems such as poor image quality due to ambient lighting conditions, image perspective distortion, interference characters, etc., license plates in real environment are usually difficult to be located accurately and efficiently.

Many algorithms have been proposed in order to solve the problem. In [1]-[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]-[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]-[10], genetic algorithm [11], etc. are widely used.

Manuscript received March 1, 2005. This work was sponsored by the Australian Research Council under Large ARC-Discovery Grant DP0451666.

Wenjing Jia is with the Faculty of Information Technology, University of Technology, Sydney, PO Box 123, Broadway, NSW, 2007, Australia (wejia@it.uts.edu.au)

Huaifeng Zhang is with the Faculty of Information Technology, University of Technology, Sydney, PO Box 123, Broadway, NSW, 2007, Australia (hfzhang@it.uts.edu.au)

Xiangjian He is with the Faculty of Information Technology, University of Technology, Sydney, PO Box 123, Broadway, NSW, 2007, Australia (sean@it.uts.edu.au)

Massimo Piccardi is with the Faculty of Information Technology, University of Technology, Sydney, PO Box 123, Broadway, NSW, 2007, Australia (massimo@it.uts.edu.au)

Currently, most researchers prefer hybrid methods, where multiple features are combined 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 to segment the vehicle images to directly obtain the candidate regions that may include license plate regions. Multiple features are then used for classifying rather than generating candidate regions, while in the edge-based methods candidate regions are actually generated artificially from features of characters (e.g., edges) within the real license plate regions via proper morphological operations. By doing so, candidate regions do not necessarily match the real license plates region, which in turn cannot guarantee the accuracy of detection result.

The remaining parts of the paper are organized in the following order. The proposed algorithm is described firstly, where mean shift segmentation is introduced in Section II, feature extraction is introduced in Section III, and Mahalanobis classification is introduced in Section IV. Then, experimental results are shown and analyzed in Section V, where the performance of the proposed algorithm is analyzed and the results are compared with those obtained by the edge-based methods from the viewpoint of robustness of algorithm and accuracy of detection results. Section VI is the conclusion part and also shows the future research work.

II. MEAN SHIFT IMAGE SEGMENTATION

A. Mean Shift

Mean shift algorithm is a nonparametric statistical method for seeking the main modes of a point sample distribution. 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 in [12] and largely forgotten until Cheng's paper [13] rekindled interest in it. In [14], Comaniciu proposed a practical method employing mean shift in the joint spatial-range domain of color images for discontinuity preserving filtering and image segmentation [14].

Given n points data set in the d -dimensional space R^d , the multivariate kernel density estimator with kernel $K(x)$ and window radius (bandwidth) h is given by:

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (1)$$

The modes of this density, where density $f(x)$ takes local maxima, are located among the zeros of the gradient $\Delta f(x) = 0$.

We denote

$$g(x) = -k'(x) \quad (2)$$

where $k(x)$ is the profile of the kernel function $K(x)$. If $g(x)$ is some profile of a kernel function $G(x)$, G is called the *shadow* of kernel K [13].

It can be deduced that searching the modes of the density can be done by searching the convergent points of the mean shift without estimating the density [14]:

$$m_{h,G}(x) = \frac{\sum_{i=1}^n x_i g(\|\frac{x-x_i}{h}\|^2)}{\sum_{i=1}^n g(\|\frac{x-x_i}{h}\|^2)} - x \quad (3)$$

i.e., finding the points where the difference between the sample mean, using the kernel G , and x , the center of the kernel (window), converges to zero.

We denote the sequence of successive locations of the sample mean with shadow kernel G as $\{y_j\}_{j=1,2,\dots}$, where

$$y_{j+1} = \frac{\sum_{i=1}^n x_i g(\|\frac{x-x_i}{h}\|^2)}{\sum_{i=1}^n g(\|\frac{x-x_i}{h}\|^2)} \quad (4)$$

is the sample mean at $x = y_j$ computed with kernel G and y_1 is the center of the initial position of the kernel.

B. Mean Shift Segmentation

In this paper, the mean shift procedure is applied for the data points in the joint *spatial-range domain* [14]. Each data point becomes associated to a point of convergence which represents the local mode of the density in the d -dimensional space. The process takes into account simultaneously both the spatial and range information. The output of the mean shift filter for an image pixel is defined as the range information carried by the point of convergence. This process achieves a high quality, discontinuity preserving spatial filtering. For the segmentation task, the convergence points sufficiently close in the joint domain are fused to obtain the homogeneous regions in the image.

An image is typically represented as a 2-dimensional lattice of r -dimensional vectors (pixels), where r is 1 in the gray-level case and 3 for color images. The space of the lattice is known as the *spatial domain* while the gray level or the color is represented in the *range domain*. However, after a proper normalization, the location and range vectors can be concatenated to obtain a spatial-range domain of dimension $d = r + 2$.

The mean shift segmentation in the spatial-range domain is implemented on the mean shift filtered images. The input original image is 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 $\{x_j\}_{j=1,2,\dots,n}$ be the original image points, $\{z_j\}_{j=1,2,\dots,n}$ be the points of convergence, and $\{L_j\}_{j=1,2,\dots,n}$ be a set of labels for different regions.

1) For each $j = 1, 2, \dots, n$, run the mean shift filtering procedure for x_j and store the convergence point in z_j , as:

- Initialize $k = 1$ and $y_1 = x_j$.
- Compute according to (4), $k \leftarrow k + 1$ until convergence.
- Assign $z_j = (x_j^s, y_{conv}^r)$ to specify that the filtered data at the spatial location of x_j will have the range components of the point of convergence y_{conv} .

2) Delineate in the joint domain the clusters $\{C_p\}_{p=1,2,\dots,m}$ by grouping together all z_j which are closer than h_s in the spatial domain and h_r in the range domain, i.e., concatenate the basins of attraction of the corresponding convergence points.

3) For each $j = 1, 2, \dots, n$, assign $L_j = \{p | z_j \in C_p\}$.

4) Optional: eliminate spatial regions smaller than M pixels.

The superscripts s and r denote the spatial and range components of a vector respectively.

III. FEATURE EXTRACTION

After the candidate regions are obtained by applying mean shift segmentation, features of license plates are to be extracted in order to correctly differentiate the license plate regions from others.

According to our statistical analysis, compared with other non-license plate regions, license plate regions take a unique feature combination, such as they have rectangle shape, they have determined aspect ratio, and they have evenly distributed higher color variance. In this experiment, three features are defined and extracted in order to detect the license plate area from segmented candidate areas, namely, rectangularity, aspect ratio and edge density.

A. Rectangularity

Under limited perspective distortion, license plates can 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:

$$R = \frac{A_O}{A_{MER}} \quad (5)$$

where A_O is the area of the object and A_{MER} is the area of the object's MER. The *area* here is defined as the number of pixels of the object.

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

TABLE I
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

rotating angle at which the ER goes through 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.

In this project, considering that the tilted angles of vehicles caused by uneven or curvy road surface vary in a very limited range, the range of rotating angle is set as $[-15^\circ, 15^\circ]$ with $\Delta\theta = 1^\circ$ in order to reduce unnecessary computation.

B. Aspect Ratio

The *aspect ratio* is defined as the ratio of the width to the height of the region's MER:

$$R = \frac{W}{H} \quad (6)$$

Since the MER of the object region can be easily computed via rotating the region within a certain range, the dimension of the object's MER can be taken as the width and the height of the region.

C. 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 contrast due to the presence of characters, an important feature to describe license plate region is to quantify its local variance.

The *edge density* is measured in a region R by summing all edge pixels within the region:

$$D_R = \frac{1}{N_R} \sum_{m,n \in R} E(m,n) \quad (7)$$

where $E(i,j)$ represents the edge magnitude at pixel (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 are detected. In this paper, edge information is acquired by applying a 3×3 Sobel edge detector in "vertical" direction only.

IV. MAHALANOBIS CLASSIFICATION

Before the feature extraction, segmented regions are firstly filtered roughly using the number of pixels in each region, where too small or too big regions are removed in order to reduce unnecessary computation. The resulted regions are called region of interest (ROI) in this paper. For each ROI, above three features are extracted, i.e., rectangularity, aspect ratio, and edge density, to compose the feature vector \mathbf{x} as:

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \text{rectangularity} \\ \text{aspect ratio} \\ \text{edge density} \end{pmatrix} \quad (8)$$

In the project, according to our observation, there exist two classes of license plates, which suffer from a distinct difference in their aspect ratios. So, totally three classes of regions are defined, i.e., *plates-1*, *plates-2*, and *non-plates regions*. 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 feature in regard of rectangularity, but they are differentiated from each other in aspect ratio and edge density.

In this table, in order to reduce unnecessary computation, regions:

- 1) which size are smaller than 0.5% or larger than 5% of image size;
- 2) which rectangularity are smaller than 0.5; or
- 3) which aspect ratio are less than 1 or larger than 5.

are removed from target candidates before classification. All of the parameters here mentioned are set experimentally and they should not be set strictly. 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.

In order to correctly classify the regions, 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 sampled data are used for training to compute the mean vector and covariance matrix. The others are used for testing. During testing, we measure the Mahalanobis distance [17](page 85) from the feature vector \mathbf{x} to the templates $\{\mathbf{w}_k\}_{k=1,2,3}$ of three modes as shown in 9 and assign \mathbf{x} to the class of the nearest mode based on Mahalanobis distance:

$$d_k^2 = \|\mathbf{x} - \mathbf{m}_k\|^2 = (\mathbf{x} - \mathbf{m}_k)' \mathbf{C}_k^{-1} (\mathbf{x} - \mathbf{m}_k) \quad k = 1, 2, 3 \quad (9)$$

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

V. EXPERIMENTS

A. Experiments set-up

The proposed algorithm has been tested in various vehicle images captured from a highway intersection. The

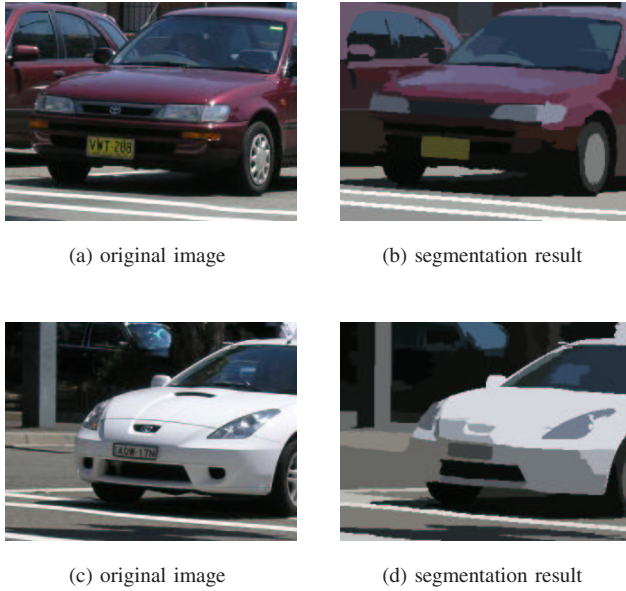


Fig. 1. Mean shift segmentation results.

target vehicle images are 324×243 color images. Several experimental results are shown in this paper to compare the performance with other methods. In order to do the classification, we employed the mean shift segmentation algorithm with $(h_r, h_s, M) = (5, 6.5, 400)$ to segment input vehicle images into candidate regions first. Parameters h_r , h_s , and M are set experimentally. The selection criteria of M are related to the proportions of the size of license plates to that of vehicle images. Regions that are smaller than 400 pixels, which are approximately 0.5% of total image size, will be eliminated.

Fig. 1 gives mean shift segmentation results on two vehicle images, where different regions are represented by different colors.

Segmented regions are then filtered roughly using the number of pixels (area) of the region to remove those very small and very large regions. For the remained regions (ROIs), three features are extracted and then input to the Mahalanobis classifier for classification to make final decisions.

B. Performance analysis

In order to evaluate the performance of the proposed algorithm, we tested it on 57 various vehicle images. 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 the localization results obtained using our method with those obtained using other methods.

1) *Robustness*: This algorithm shows robustness to interference characters that may result in detection errors. Fig. 2(a) shows one case where there are interference characters (“ISUZU” at the top part of the picture) on the truck

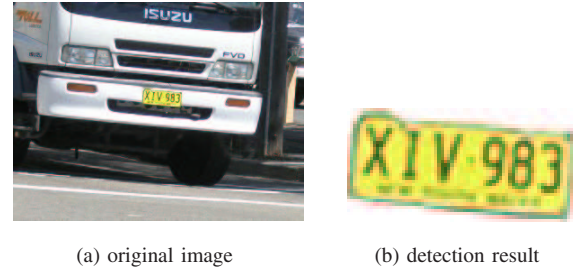


Fig. 2. Interference characters (“ISUZU”) in vehicle body and the license plate detection result.

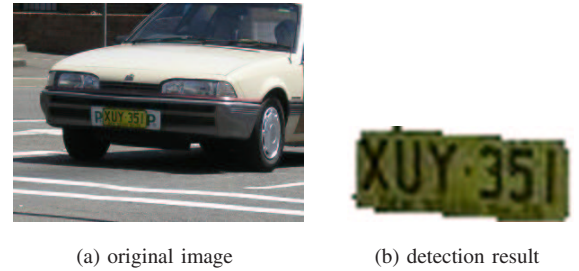


Fig. 3. Interference characters (two green letter P’s) that are next to the license plate and the license plate detection result.

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 like the plate. (Semantically, a license plate that contains only five letters is also acceptable in Australia). Instead, our algorithm will succeed thanks to the mean shift segmentation (see Fig. 2(b)). Fig. 3(a) 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, which results in that the dimension feature of the candidate does not meet the prior knowledge of license plate, and this in turn causes the localization of license plates being failed. Comparably, our proposed algorithm has no difficulty at all when dealing with these situations (see Fig. 3(b)) thanks, again, to the mean shift segmentation.

2) *Accuracy*: In order to show the accuracy of detection 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 are able to find the position of license plates, experimental results show that detected regions using our algorithm are much more close to the real license plates (see bottom pictures in Fig. 4).

In this paper, a matching rate is used to measure the matching degree between detected plates and the real ones,



Fig. 4. Comparison between the location results obtained using the edge-based methods (the middle row) and the results obtained using our method (the bottom row) on two vehicle images (the top row).

which is defined as:

$$R = 1 - (R_{FP} + R_{FN}) \quad (10)$$

where

$$R_{FP} = \frac{N_{FP}}{N_T} \quad (11)$$

and

$$R_{FN} = \frac{N_{FN}}{N_T} \quad (12)$$

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. In this experiment, N_{FP} is the total number of false positives, N_{FN} is the total number of false negatives, and N_T is the total number of pixels of real license plate regions which are obtained manually with the help of the software Adobe Photoshop.

The detected regions can only match 60.8% (26.0% false alarms and 13.2% miss) in Fig. 4c and 67.4% (6.7% false alarms and 25.9% miss) of the real plate regions, while in our case (see Fig. 4e and Fig. 4f), these matching rates are 98.7% for Fig. 4e and 97.5% for Fig. 4f respectively. Although the shapes of regions obtained using edge-based methods could be modified further to make them more close to a rectangle via continually adjusting the parameters used in morphological operations, candidate regions generated from the features of characters do not necessarily completely match the real license plate region. This mismatch may in

turn result in recognition mistakes. A statistical analysis of the detection accuracy using the definition in (10) shows that an average mismatching rate of 2.4% can be obtained with our method.

VI. CONCLUSIONS AND FUTURE WORKS

A. Conclusions

A robust and accurate license plate localization algorithm has been proposed. This algorithm employs mean shift segmentation in order to segment the regions of interest. Feature vectors are extracted from the segmented regions. A Mahalanobis classifier is then applied in order to correctly classify each region of interest into the correct class. The proposed algorithm shows great robustness when detecting the color license plates from vehicles which have a different color with that of license plate where the license plates adhere to. The proposed algorithm proved to be robust when dealing with cases where interference characters exist. Comparison has been done to show the better accuracy of detection results.

B. Future Works

Currently, the proposed algorithm needs 6 seconds to process a 324×243 color images tested on a PC with one Pentium 4 1.60GHz CPU, among which mean shift segmentation takes more than 5 seconds. So, future research work may consider a roughly scanning first on the input image in order to avoid unnecessary mean shift segmentation on those relatively uniform regions which definitely can not contain any characters. Then, applying the proposed algorithm to accurately localize the license plate.

VII. ACKNOWLEDGEMENT

The authors would like to thank the anonymous reviewers for their valuable comments.

REFERENCES

- [1] S.-Z. Wang and H.-J. Lee, "Detection and recognition of license plate characters with different appearances", *Intelligent Transportation Systems, 2003. Proceedings. 2003 IEEE*, vol. 2, pp. 979-984, 2003.
- [2] S. Kim, D. Kim, Y. Ryu, and G. Kim, "A robust license-plate extraction method under complex image conditions", *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, vol. 3, pp. 216-219 vol.3, 2002.
- [3] H. Bai, J. Zhu, and C. Liu, "A fast license plate extraction method on complex background", *Intelligent Transportation Systems, 2003. Proceedings. 2003 IEEE*, vol. 2, pp. 985-987, 2003.
- [4] H. Bai and C. Liu, "A hybrid license plate extraction method based on edge statistics and morphology", *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, vol. 2, pp. 831-834, 2004.
- [5] V. Kamat and S. Ganesan, "An efficient implementation of the Hough transform for detecting vehicle license plates using DSP'S", *Real-Time Technology and Applications Symposium, 1995. Proceedings*, pp. 58-59, 1995.
- [6] S. Gendy, C. L. Smith, and S. Lachowicz, "Automatic car registration plate recognition using fast Hough transform", *Security Technology, 1997. Proceedings. The Institute of Electrical and Electronics Engineers 31st Annual 1997 International Carnahan Conference on*, pp. 209-218, 1997.
- [7] Y. Yanamura, M. Goto, D. Nishiyama, M. Soga, H. Nakatani, and H. Saji, "Extraction and tracking of the license plate using hough transform and voted block matching", *Intelligent Vehicles Symposium, 2003. Proceedings. IEEE*, pp. 243-246, 2003.

- [8] W. Wei, M. Wang, and Z. Huang, "An automatic method of location for number-plate using color features", *Image Processing, 2001. Proceedings. 2001 International Conference on*, vol. 1, pp. 782-785 vol.1, 2001.
- [9] K. I. Kim, K. Jung, and J. H. Kim, "Color Texture-Based Object Detection: An Application to License Plate Localization", *Proceedings of the First International Workshop on Pattern Recognition with Support Vector Machines*, pp. 293-309, 2002.
- [10] S. H. Park, K. I. Kim, K. Jung, and H. J. Kim, "Locating car license plates using neural networks", *Electronics Letters*, vol. 35, pp. 1475-1477, 1999.
- [11] S. K. Kim, D. W. Kim, and H. J. Kim, "A recognition of vehicle license plate using a genetic algorithm based segmentation", *Image Processing, 1996. Proceedings., International Conference on*, vol. 1, pp. 661-664 vol.2, 1996.
- [12] Fukunaga, K. and L. Hostetler, "The estimation of the gradient of a density function, with applications in pattern recognition", *Information Theory, IEEE Transactions on*, 1975. 21(1): p. 32-40.
- [13] Y. Cheng, "Mean shift, mode seeking, and clustering", *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 17, pp. 790 - 799, 1995.
- [14] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis", *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, pp. 603-619, 2002.
- [15] K. R. Castleman, *Digital Image Processing*, Toronto: Prentice Hall, 1996.
- [16] M. Yu and Y. D. Kim, "An approach to Korean license plate recognition based on vertical edge matching", *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*, vol. 4, pp. 2975-2980 vol.4, 2000.
- [17] J.P. Marques de S., *Pattern recognition: concepts, methods, and applications*, New York: Springer, 2001.