Vehicle Model Recognition in Video

Suhan Lee, Jeonghwan Gwak and Moongu Jeon

School of Information and Mechatronics, Gwangju Institute of Science and Technology (GIST), Gwangju 500-712, Korea
{suhan, gawk, mgjeon}@gist.ac.kr

Abstract

Vehicle model recognition from images is one of challenging fields for supporting intelligence transport system. Existing methods only deal with vehicle in fixed view. However, in video, rotation is occurred and it decreases performance of recognition. To overcome this problem, rectification method about skew though rotated view is needed. We employ the robust license plate detection method using straight line template matching to get ROI. Also, we perform recognition process toward vehicle in rotated view whose angle is 0° to 15° after proposed image rectification from proposed rectification method.

Keywords: *Vehicle Model Recognition, Pose-invariant Object Recognition, License Plate Detection, Feature Matching*

1. Introduction

Vehicle model recognition has more significance in recent years as a result of increased demand for security and effectiveness. For example, this technology together with a license plate recognition system allows figuring out fake license plates, and thus it can be used to identify and track specific vehicles such as those of stolen and VIPs. Further, a tollgate system ensures proper toll rates are charged if it identifies a vehicle type through vehicle model recognition.

There are several researches about object recognition method that deal with character [1], tag [2], face [3] and so on. This paper considers vehicle model recognition methods in detail. First, manual method to obtain region of interest (ROI) for classification is implemented [4]. Although their combined descriptor is comparably good, they do not achieve good accuracy. The ROI is generally derived from front view of a vehicle. License plates have the same shape and size for each type whose information can be easily gathered. Therefore, license plate is the best standard to get ROI. Second, [5] presents an investigation into various feature extraction techniques to automatic vehicle model recognition through obtained ROI from a license plate. Even though this system showed relatively good recognition rates, they don't have accurate license plate detection system for getting ROI and do not consider rotation of a vehicle in picture.

The two main contributions of this work are as follows. First, we devise a new robust method for accurate total rectangular region for obtaining exact ROI. Since the objective of license plate detection methods [6] is generally getting characters in license plate, optimized method for getting ROI should be adopted. Second, we propose an image rectification method to increase recognition accuracy in a rotated view. Entrance or lane is broader than the width of a vehicle, and drivers generally don't drive their vehicles exactly in the forward direction. Further, cameras for vehicle model recognition may not be installed in the exactly front view to the vehicle's entering direction. ROIs for recognition are very different depending on rotation although they are caused from same model. Thus, rectification method that can transform rotated view into front view is valuable because it can increase accuracy in final result.

An organization of this paper is given as follows. In Chapter 2 and 3, we present our proposed system for pose invariant vehicle model recognition and overall procedure is illustrated in Figure 1. In Chapter 4 and 5, experimental results and conclusion are dealt with.

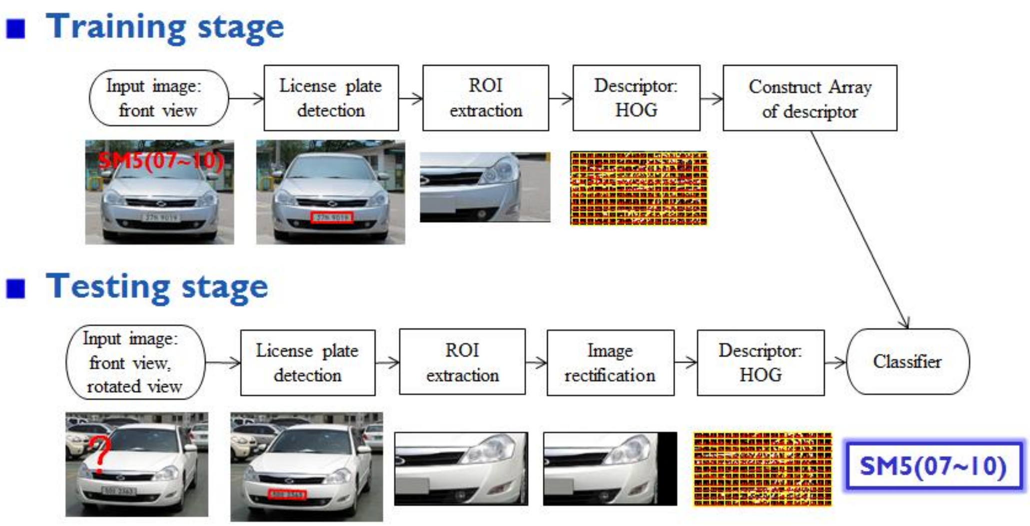


Figure 1. Overall Procedure

2. ROI Extraction using License Plate Detection

2.1. Detection of consecutive big numbers

Among existing license plate detection methods, we choose the method that is extracting coordinates of outer rectangle after detecting big characters in a license plate. This is because this method shows low false positive rate and can be utilized together with a license plate recognition system. In Korea, the big characters are always numbers. To search a region which may include a license plate, we detect only vertical edges in total region of an image. Given regions by vertical edge analysis, process for finding consecutive big numbers is described in Figure 2.

(a) Original image (b) Candidate regions (c) License plate (d) Found characters

Figure 2. Searching for Big Characters

After the original image in Figure 2(a) is transformed into a gray image, we use vertical edge filter. In [7], the authors found that the only use of vertical edge is good to represent a region including big numbers in a license plate. Regions including many vertical edges are only considered as candidate regions where a license plate is contained. In the regions, we use a thresholding method to find desired big numbers. In each candidate, we can find blobs of big numbers in the license plate by connected-component labeling with proper parameters such as maximal size, minimal size and aspect ratio and so on.

In each candidate region after thresholding, we can find blobs of big numbers in license plate by connected-component labeling with proper parameters such as maximal size, minimal size and aspect ratio and so on. In Korea, there are four consecutive big numbers so have form of arrangement as Figure 3. Thus, we can use rules of arrangement we design. Because the number of the rules doesn't affect operational time largely, we propose many rules in Table 1. Structure of license plates is different in other countries but similar method can be applied.

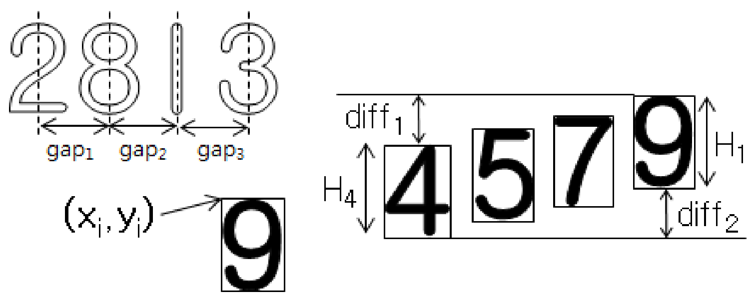


Figure 3. Form of Arrangement of Big Numbers

Table 1. Rules of Arrangement of Big numbers in i < j < k

$\alpha < H_i < \beta$	(1)
$0.1 < W_i / H_i < 0.9$	(2)
$0.4 \times H_1 < gap_i < 1.3 \times H_1$	(3)
$(gap_j+gap_k) \times 0.35 < gap_i < (gap_j+gap_k) \times 0.65$	(4)
$\lvert y_i-y_j \rvert < 0.3 \times H_1$	(5)
$H_j > H_1 \times 0.8$	(6)
$\lvert diff_1-diff_2 \rvert < 0.5 \times H_1$	(7)

2.2. Extraction of outer rectangle

Using the detected big characters, we can estimate total rectangular region of license plate. We use straight line template matching [8] to estimate total rectangular region of license plate. Shape of template is similar to two adjacent straight lines as illustrated in Figure 4. We assign negative value in the inside to find a start location of a rectangular region. This is because there are several parallel lines around sides of the license plate, and we should choose one line as a side for each side. Its search range and direction are based on the type of a license plate.

Searching is achieved in horizontal and vertical edge image without threshold. In template, gradient is determined by arrangement of big numbers and length is decided by its location and size. In license plates having long horizontal length, bolts in both sides interrupt matching calculation, and thus we eliminate the part of negative values in the template about regions including a bolt. For later process, skewed license plate in original image is rectified into exact rectangle to make a normalized image for classification.

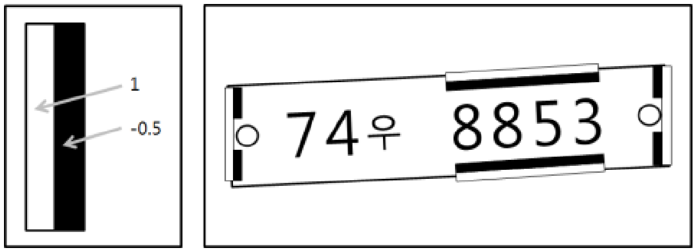


Figure 4. Straight Line Template Matching

2.3. Designation of region for classification

ROI is assigned from detected license plate as illustrated in Figure 5. Since their size and ratio of the horizontal to the vertical are different in each type of license plates, we consider this problem to obtain ROI which has the same size for all vehicles. We get real size of license plate with unit of millimeter (*mm*). Through this size, we make ROI that has 2,080 *mm* as width and 660 *mm* as height from multiplication with original lengths of each license plate. Both left and right sides as in Figure 3 will be used to match local features for rectification of a rotated view, and it is used to adopt one side image for classification.

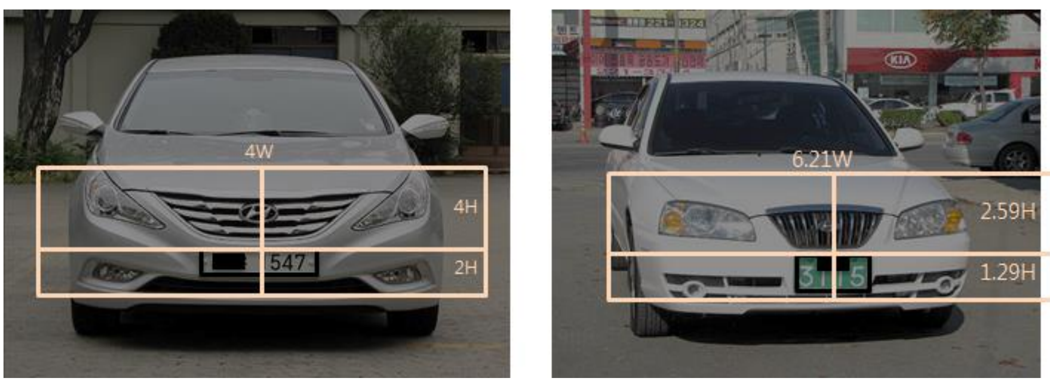


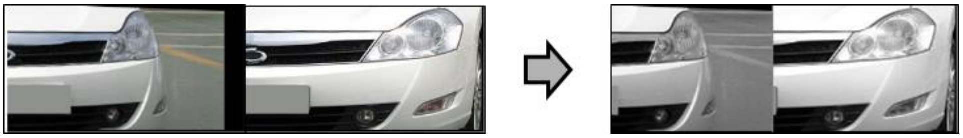
Figure 5. ROI from the Detected License Plate

3. Vehicle model recognition using symmetry

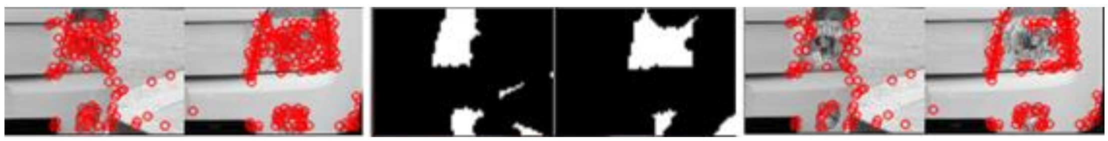
3.1. Adjustment of ROI

An image rectification method is required because ROI in the same model can be different according to the variation of angles because a vehicle pictured in rotated view has very different ROI. To correct designed ROI for increasing accuracy in recognition, [9] employs Hough transform and contour detection respectively. Our proposed method is based on feature matching [10], and its structure is slightly similar with stereo matching. Front part of most vehicles has property of symmetry, and this is the key property to construct a transform function.

Before extraction of features, we should exclude unnecessary regions for achieving better matching performance, and its steps are illustrated in Figure 6. First, the left side in ROI is flipped in *y*-axis (Figure 6(a)). Since left region of both ROIs which are close with a license plate and the blank area in the boundary of an image are unnecessary for matching, we eliminate them in the step of matching (Figure 6(b)). Regions of lamps have too irregular textures that decrease matching accuracy (Figure 6(c)). To eliminate the drawback, we estimate the regions of lamps (Figure 6(d)), and exclude features for the regions as follows: 1) Edge in both ROIs are extracted and transformed to binary image with a proper threshold (50 was used in this work). 2) Dilation and erosion operation is then performed. Finally, features without lamps are obtained (Figure 6(e)).



(a) Flip of left side in *y*-axis (b) Exclusion of the center region and gap

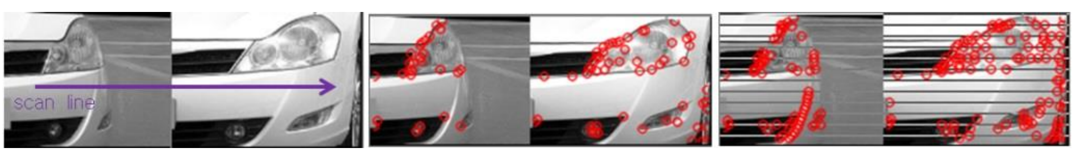


(c) Features with lamps (d) Estimated regions of lamps (e) Features without lamps

Figure 6. Exclusion of Unnecessary Regions

3.2. Local feature detection

FAST feature detector [11] has fast speed and also is known to result in good performance. The detector draws a circle composed of 16 pixels with center, *p*. If more than 9 contiguous pixels are brighter or darker than intensity of *p* above a proper threshold (35 was used in this work), we can assume that *p* is a corner point. FAST feature detector has a special property which is found in this work as illustrated in Figure 7. Preferentially, we draw horizontal lines whose intensity is 255 (which is the original intensity) with regular intervals that are 8 to 10. After that, FAST can detect some of points having a vertical edge as a feature. Since the scan direction is horizontality, points having a vertical edge can be easily distinguished. Thus, the usage of vertical edge as a local feature helps to achieve robust results through an increased number of matching.

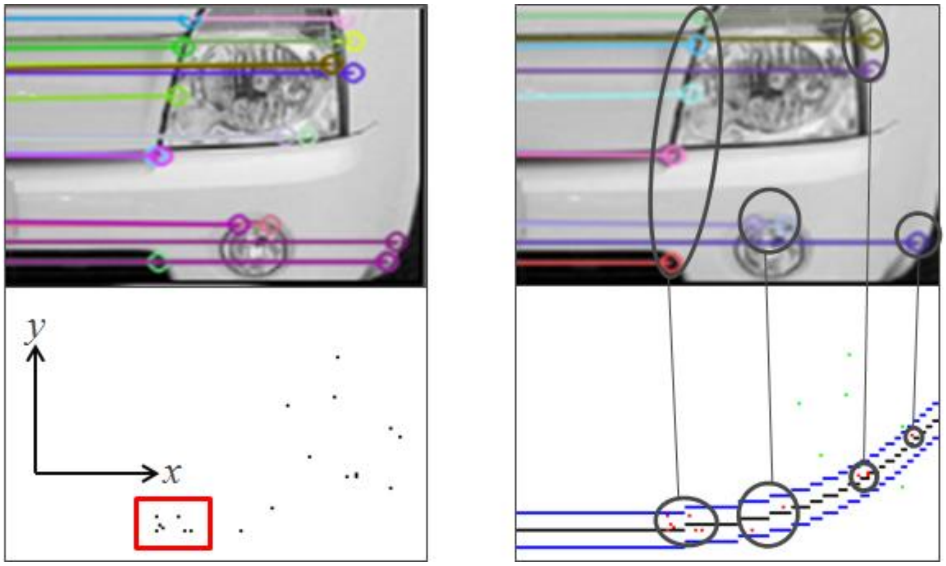


(a) The scan line (b) FAST without horizontal lines (c) FAST with horizontal lines

Figure 7. Exclusion of Unnecessary Features

3.3 Matching and transformation

Due to its efficiency in terms of processing time, we execute matching obtained local features in Section 3.2 through SURF descriptor [12]. Since we found that matching results of just using SURF have high error rates, two methods are employed to reduce the error rates. First, we use that a side image appears longer than another side. For all matching points, a side with a bigger x-value gets a higher score. The side with a higher score between the left and the right is selected in classification after image rectification. This is because the side appears long and thus contains more information. The matching points that contribute to increase the score of the selected side are only employed in the next stage. Second, the concept of random sample consensus (RANSAC) [13] is employed to remove remaining errors. A dataset to implement RANSAC is composed as follows. x-value of points is the same with x-value of the side image we opt in matching points and y-value of points is difference between x-values of a matching and then y-value is divided by two. y-value is similar with disparity in a stereo matching algorithm. An example of a dataset is illustrated in Figure 8.



(a) Before modified RANSAC (b) After modified RANSAC

Figure 8. Distribution of Data from Matching Points

We use modified RANSAC whose model is a curve that is dependent on the dataset as illustrated in Figure 8(b). Through experiment, we suppose that some polynomial curves are similar with the ideal curve. There are 10 polynomial curves heuristically designed using Eqs (1) and (2). *x*-range of function is 135 because we use only 75% of a width (180 in this work) due to exclusion of the left region. We can find best polynomial curves having more inliers than others among the 10 curves around height *h* which is average of *y* values in section that *x*-value is smaller than 60. The section is represented as a red box in Figure 8(a), and *y*-value of the left region which is excluded in matching is *h*.

$$y = m \times 4 \times 10^{-6} \times x^3 + h, \text{ where } m = 1 \text{ to } 5 \quad (1)$$
$$y = m \times 3 \times 10^{-8} \times x^4 + h, \text{ where } m = 1 \text{ to } 5 \quad (2)$$

As a result, we can get the transformation function from inliers. If the angle is big in a rotated view, the gradient of the function will be also big. A side image that appears long is translated in the left direction in *x*-axis as large as *y*-value of the function. In Figure 9, there are some views to evaluate our results through eye. ROI in rotated view after image rectification is more similar with ROI of same model in front view than before.

(a) ROI in a rotated view (b) ROI in a rectified image (c) ROI in a front view

Figure 9. Comparison of ROIs

3.4 Used visual descriptor and classifiers

In the proposed system, license plate detection was adopted for obtaining accurate ROI and image rectification to transform ROI in a rotated view into ROI in a front view. Histogram of oriented gradients (HOG) [14] was used as a visual descriptor. Since the HOG descriptor operates on localized cells with normalization, the method is invariant to geometric and photometric transformations.

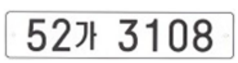
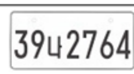
In the classification stage, the *k*-nearest neighbor (*k*-NN) and the support vector machine (SVM) [15] algorithms are used to classify vehicle models for input images. The *k*-NN is the simplest classification algorithm and this has different results depending on the value of *k*. SVM is a classification algorithms using a hyper-plane that has the largest distance to the nearest training data point of any class, and kernel functions are employed to solve non-linear problems. The *k*-NN is used as a baseline classification algorithm for the comparative study. SVM is specially adopted in the classification because of its good generalization performance.

4. Experimental Results

5.1. License plate detection

License plate detection about five types is experimented as represented in Table 2. Used image size is 640 x 480 and size of license plate should be larger than approximately 70 x 30. Average speed is 0.04s and total accuracy is about 98%.

Table 2. Result of License Plate Detection

type	No. of images	No. of success
	73	72
	148	147
	22	21
	9	8
	15	13
Total	267	261

5.2. Vehicle model recognition

Vehicle images of size 640×480 were taken in distance with a camera from 3 to 5 meters. We investigate vehicle models with high sales volume in South of Korea for the last ten years and select 40 vehicle models as represented in Table 3. For each model, 8 train images and 2 test images are employed. Thus, the number of training and testing data are 320 (=40×8) and 80 (=40×2), respectively. Angles of a rotated view range from 9° to 15°.

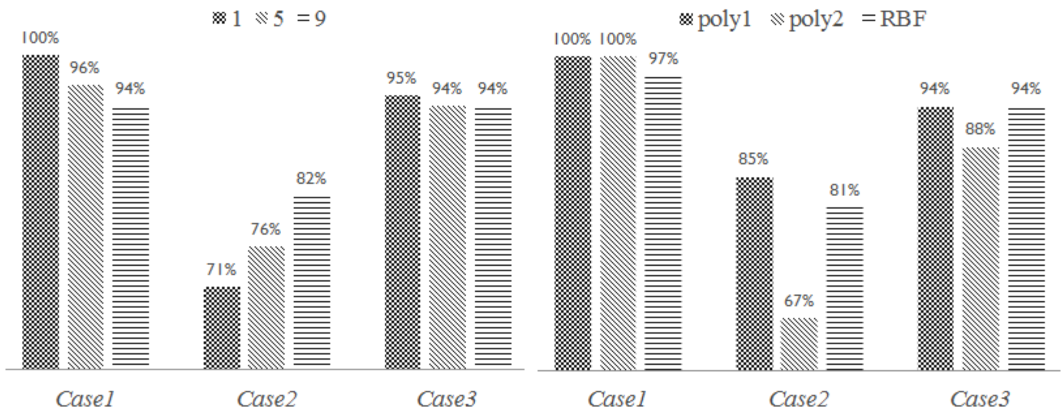
Table 3. Used Vehicle Models

No.	model name	model year	No.	model name	model year
1	Avante_XD	00~03	21	K5	10~
2	Avante_XD	03~06	22	K7	09~
3	Avante_HD	06~10	23	Morning	08~11
4	Avante_MD	10~	24	Morning	11~
5	EF_Sonata	01~04	25	Sportage	04~07
6	NF_Sonata	04~07	26	Sportage_R	10~
7	YF_Sonata	09~	27	Forte	08~12
8	Grandeur_XG	02~05	28	Sorento_R	09~12
9	Grandeur_TG	05~09	29	Bongo3	04~
10	L_Grandeur	09~11	30	Pride	05~11
11	Poter	86~04	31	SM3	05~11
12	Poter2	04~	32	SM3	11~
13	Starex	97~04	33	SM5	05~07
14	Starex	04~07	34	SM5	07~10
15	Grand_Starex	07~	35	SM5	10~12
16	Santafe_CM	05~08	36	SM7	08~11
17	i30	07~11	37	Lacetti_P	08~11
18	Tucson	04~09	38	Matiz	09~11
19	Tucson_IX	09~	39	Tosca	08~11
20	Genesis	08~	40	Korando	96~05

The experiments were carried out for three representative cases. *Case1* is for the images with original front views, and *Case2* is for those with rotated views without image rectification, and finally, *Case3* is for those with rotated views with image rectification. For the images whose license plate is detected, the recognition accuracy is calculated as follows:

$$\text{Accuracy (\%)} = \frac{\mathit{The\ Number\ of\ Success}}{\mathit{The\ Number\ of\ Test\ Images\ with\ a\ Detected\ License\ Plate}} \times 100 \quad (3)$$

Figure 10 shows the experimental results from classifiers which are *k*-NN with different k values and SVM with different kernels. In both front view and rotated view, 78 license plates are detected and in the 80 images, accuracy rate is gotten. *Case1* takes about 0.08s. In *Case2* and *Case3*, additional 0.06s is required due to the image rectification process. *k*-NN shows good result with the smallest *k* and SVM with linear and the RBF kernel show robust results. Finally, we can observe that image rectification increases accuracy rate (at least from 11 up to 24%) in comparison with *Case1* and *Case3*.



(a) *k*-NN with different *k* values (b) SVM with different kernels

Figure 10. Result of Vehicle Model Recognition

5. Conclusion

A novel framework for vehicle model recognition based on the robust license plate detection (Section 2) and the image rectification algorithm using symmetry (Section 3) is proposed in this work. Even for a small change of an angle, the recognition performance can be severely affected in terms of accuracy. To resolve this drawback, we propose an image rectification scheme.

In the experiments for vehicle model recognition (Section 4), our image rectification method showed that it assist to increase accuracy in case of pictured vehicles in rotated views. Since installed cameras generally deal with the range of angles, our image rectification method can be directly applicable to real-world applications. The authors acknowledge that the proposed system has some failures which are mainly attributed to similar ROI or wrong matching results by irregular illumination. The possible solutions are (i) combining other feature descriptors such as shape context of lamps for the failure due to similar ROI, and (ii) using several image restoration methods for the failure due to irregular illumination.

Acknowledgment

This research was supported by the GIST(Gwangju Institute of Science and Technology), Korea, under the Pratical Research and Development support program supervised by the GTI(GIST Technology Institute) and the Pioneer Research Center Program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (2012-0009460).

References

[1] B. Kessab, C. Daoui, B. Bouikhalene, M. Fakir and K. Moro, “Extraction Method of Handwritten Digit Recognition Tested on the MNIST Database”, International Journal of Advanced Science and Technology, vol. 50, **(2013),** pp. 99–110.

[2] B. Jang and C. Kim, “Improved Tag Identification Method for reducing the Process Time of Tag Recognition”, International Journal of Advanced Science and Technology, vol. 43, **(2012),** pp. 59-66.

[3] R. Adipranata, *et al.*, “Fast Method for Multiple Human Face Segmentation in Color Image”, International Journal of Advanced Science and Technology, vol. 3, **(2009),** pp. 19-32.

[4] B Zhang and C Zhao, “Classification of Vehicle Make by Combined Features and Random Subspace Ensemble”, ICIG, **(2011).**

[5] V. S. Petrovic, T. F. Cootes, "Analysis of Features for Rigid Structure Vehicle Type Recognition", BMVC, **(2004).**
[6] C. N. E. Anagnostopoulos, I. E. Anagnostopols, I. D. Psoroulos, V. Loumos and E. Kayafas "License plate recognition from still images and video sequences: A survey", IEEE Trans. Intell. Transp. Syst., vol. 9, no. 3, **(2008).**
[7] H. Bai and C. Liu, "A hybrid license plate extraction method based on edge statistics and morphology", in Proc. Int'l Conf. on Pattern Recognition, vol. 2, **(2004),** pp. 831-834.
[8] S. Lee, J. Gwak and M. Jeon, "Effective license plate detection algorithm using area reduction and template matching", KSPC, vol. 25, no. 1, (**2012).**
[9] M. Ramappa and S. Krishnamurthy, "Skew Detection, Correction and Segmentation of Handwritten Kannada Document", vol. 48, **(2012),** pp. 71–88.
[10] A. Baumberg, "Reliable Feature Matching across Widely Separated Views", Proc. Conf. Computer Vision and Pattern Recognition, **(2000).**
[11] D. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", IJCV, vol. 20, no. 2, **(2004),** pp. 91-110
[12] H. Bay, T. Tuytelaars and L. J. Van Gool, "SURF: Speeded Up Robust Features", ECCV, **(2006).**
[13] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography", Comm. ACM, vol. 24, no. 6, **(1981),** pp. 381-395.
[14] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection", Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol. 2, **(2005),** pp. 886-893.
[15] C. Cortes and V. Vapnik, "Support-vector network", Machine Learning, vol. 20, **(1995),** pp. 273-297.