Main Peak Finding for Signboard Recognition under Adverse Conditions

Thi Thi Zin†, Sung Shik Koh††, and Hiromitsu Hama†††

†Graduate School of Engineering, Osaka City University, Osaka, 558-8585 Japan
††Optics & Digital System Division, Samsung Techwin Co., Ltd., Seongnam-city, 462-703 Korea
†††Faculty of Engineering, Osaka City University, Osaka, 558-8585 Japan

Summary
This paper presents a new framework for signboard recognition of signboards under adverse conditions. The framework is composed of two main stages newly proposed here: the main peak finding algorithm based on HSV color space to remove regions of occlusion and reflection, and the template matching to recognize objects. In the experiments, we used 300 images taken under a great variety of adverse conditions including occlusion, reflection, specular highlights, and so on. According to the experimental results, the proposed system could give 95% recognition rate for images taken under such bad conditions. The results are shown to demonstrate the robustness and effectiveness of the system. Moreover, the proposed method is applicable for vision-based car and pedestrian navigation assistance to provide up-to-date visual information in accordance with users' demands, for example, for finding shops, billboards, landmarks and so on.
Key words:
Signboard, Uniform color region, Occlusion and reflection, Main peak finding, Template matching

1. Introduction

Until now, object recognition in real world environments is one of challenging tasks, which is widely taken up in many applications. Large amounts of visual information are embedded in natural scenes and signboards are good examples that have rich information contents. And automatic recognition of signboards is essential for automated driving and for driver assistance system. Most recognition systems typically involve the extraction of target features such as parts [1], components [2], and fragments [3] at the first step and their combination at the next step. H. Wersing and E. Korner [4] proposed the method using hierarchical processing for achieving invariant recognition. Others [5]-[7] proposed systems based on support vector machines, and eigen-space for object recognition, and evolutionary optimization approaches.

The illumination conditions, cluttered backgrounds and surface geometrics of objects in a scene may cause significant problems of occlusion, shadowing, inter-reflection and specularities which make it difficult to recognize objects. The recognition systems must be robust to the deterioration of signboard's images due to aging, the time of day, and also partial Occlusions and Reflections (O/R). Generally, 'O/R region' is used here in the wide meaning of the region not being included in any Uniform Color Region (UCR). Furthermore, a variety of weather conditions such as cloudy, foggy, dusky, rainy, and snowy should be taken into consideration. In this paper we particularly concern such bad conditions. The overview flowchart of our proposed framework is shown in Fig.1. In general, the complete system consists of two main stages: segmentation and recognition. This paper focuses on the latter one and we implement the proposed system as follows: (i) O/R removal by main peak finding method and (ii) recognition by template matching. All of the signboard images used here are already segmented and normalized into 64x64 beforehand.

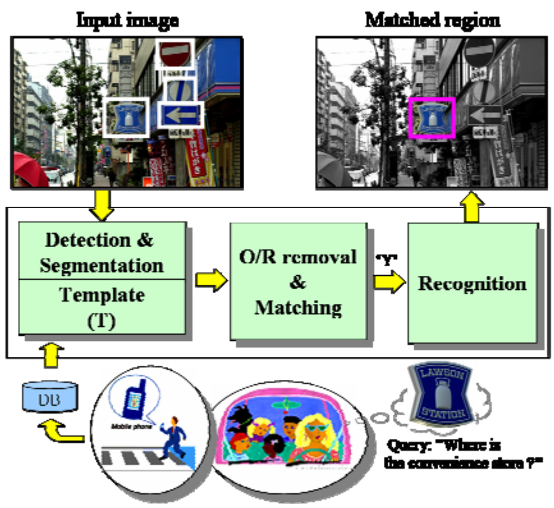

Fig. 1 Overview flowchart for the application.

Manuscript received December 8 2006.
Manuscript revised December 25 2006.

The illumination problems were treated in [8] and [9], but their systems were not enough for overall outdoor scenes. There were a method [10] for removing shadows from images and another one [11] for detecting specular highlights by a truncated least square approach. Estimation of illumination color was addressed in relation to color constancy [12]-[14], which attempted to remove the effect of illumination color from an image. Others [15]-[17] considered the recognition systems under disturbing environments and adverse conditions. In [15], the occlusion information acquired from the first step was used to decide the matching image by neural networks. H-M. Yang et al [16] treated the partially occluded signs in the training process by adding Gaussian noise. The partial occlusion was considered in [18] using a ring partitioned method.

In spite of every effort, a number of challenges remain for a robust recognition. To overcome these difficulties, we propose the robust system that can recognize signboards under various adverse conditions. Special attention is devoted to the robustness and flexibility of the recognition system for outdoor scene images taken under O/R. By the proposed framework, almost all images tested here are correctly recognized, even if they are under bad illumination conditions, slightly titled, deteriorated, and partially occluded. Moreover, our proposed method can work well when O/R region is nearly one half size of the object region.

This paper is organized as follows. In Section II, we present the recognition process with step by step. The experimental results are described in Section III. Finally, Section IV concludes this paper.

2. Recognition Process

Under outdoor scenes, real images may include some O/R regions. In such a case, we often get the incorrect recognition result. For example, the images of outdoor objects taken at nighttime are suffered from partial intensity changes due to the influence of street lights, headlights of cars and so on. Moreover, the occlusion may often occur due to electric poles, trees and so on. To overcome these difficult conditions, the proposed recognition system is developed based on two main stages:

[*Stage*1]: main peak finding for O/R removal, and

[*Stage*2]: template matching for recognition.

The first stage includes: 1) using color uniformity on templates, 2) finding the main peak, and 3) removing O/R regions. The second one is the template matching for robust recognition. The process flowchart is shown in Fig.2.

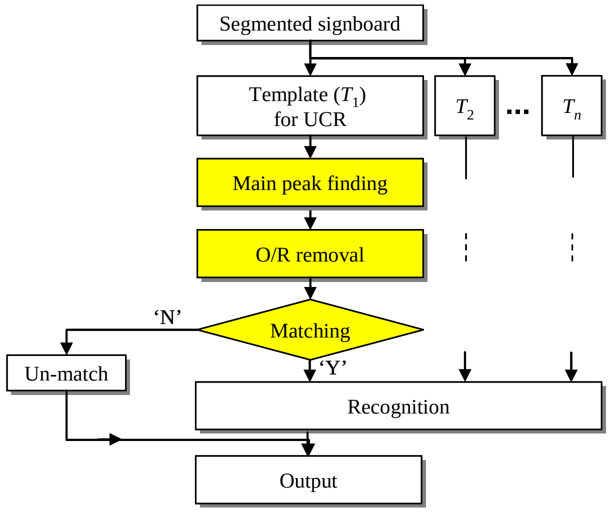

Fig. 2 Process flowchart.

2.1 Planar Object with UCRs

In outdoor scenes, signboards are good examples of planar objects. To use the color uniformity of planar objects composed of UCRs, the j^{th} template T_j is got by thinning a little the j^{th} UCR U_j. So, each object O is decomposed into templates and their border region BR, as shown in Fig.3. In this figure, there are four templates and the rest region is their border one. The black regions mean the backgrounds.

$$O=\bigcup_j U_j, \quad U_j \cap U_k = \phi, \quad \text{for } j \neq k,$$
$$T_j = \text{thinning}(U_j), \tag{1}$$
$$O=\left(\bigcup_j T_j\right) \cup BR, \quad |O| >> |BR| \text{ and } T_j \cap BR = \phi, \text{ for } \forall j,$$

where ϕ is an empty set, and $|O|$ and $|BR|$ mean the total numbers of pixels in O and BR, respectively.

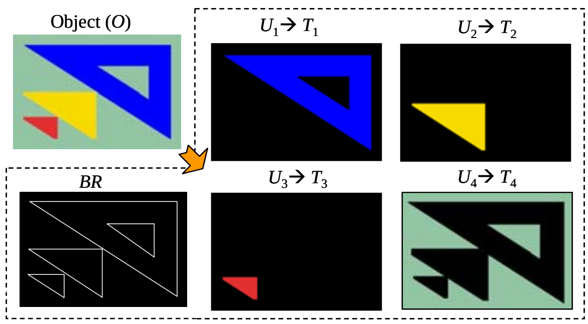

Fig. 3 Object with UCRs: templates (T_j, $j = 1,\ldots,4$) and border region (BR).

2.2 Main Peak Finding

Here, we propose the main peak finding algorithm based on each component on HSV. A hue-based space such as HSV was confirmed to be superior to RGB for our purpose through some pre-experiments. Hue histograms of objects without and with occlusion are illustrated in Fig.4. The main peak finding algorithm is composed of five steps.

[*Step*1] Make templates for the target object, and then obtain $\vec{M}$ of each pixel $P(i, j)$ in T.

$$\vec{M}(i, j) = \left(M_I(i, j), M_H(i, j), M_S(i, j)\right), \tag{2}$$

where M_I, M_H, and M_S mean the values of Intensity, Hue, and Saturation, respectively. We take Hue histograms for explanation, and the corresponding parameters to Intensity and Saturation can be got in the same way. The original range of each component is from 0 to 255.

[*Step*2] Smooth the histogram $h(k)$ of M_H by moving average as expressed in (3).

$$m_i = \frac{1}{(2p+1)}\left[\sum_{k=i-p}^{i+p} h(k)\right], \tag{3}$$

where $h(k)$ is the histogram value at level k, and m_i is the moving average of order $(2p+1)$.

[*Step*3] Find the maximum of m_i.

$$m_{\max} = \max\{m_i\} = m_{P_1} \text{ , at } i = P_1 .$$

[*Step*4] Calculate σ by (4).

$$\sigma_i = \left[\frac{1}{n}\sum_{k=i-q}^{i+q} (j-i)^2 h(k)\right]^{(1/2)}, \tag{4}$$

where, $n = \sum_{k=i-q}^{i+q} h(k).$

[*Step*5] Search the smallest value (σ_{min}) within the distance d from P_1.

$$\sigma_{min} = \min\{\sigma_i\} = \sigma_{P_2}, \quad \text{at } i = P_2, \quad \text{and} \quad P_1 - d \le P_2 \le P_1 + d. \tag{5}$$

From T, we extract the region which is within the distance $2\delta_2$ from P_2. δ_1 and δ_2 are defined in (6).

$$\delta_1 = \sigma_{P_1}, \quad \delta_2 = \sigma_{P_2} = \sigma_{\min}. \tag{6}$$

There is a reason for adopting P_2 instead of P_1. Since P_2 and P_1 are equal when the smoothed histogram m_i is symmetric, let us consider non-symmetrical case. In Fig. 5, the regions extracted by P_1 and P_2 are compared, and m_i and σ_i are expressed by a solid line and a dotted line on the left, respectively. The painted regions on the right side mean the regions which are within $2\delta_i$ from P_i (i=1, 2). From this figure, we should notice that the extracted region using P_2 consists of more uniform color pixels than that using P_1.

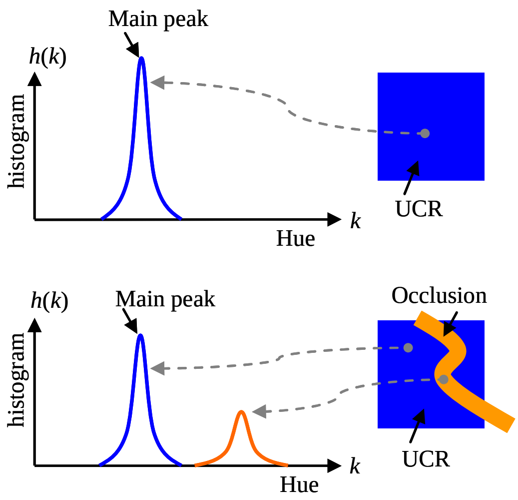

Fig. 4 Hue histograms of objects without and with occlusion.

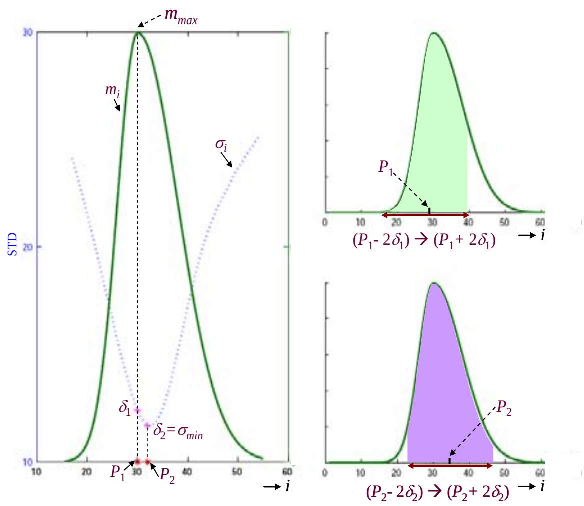

Fig. 5 Advantage of P_2 over P_1.

2.3 O/R Removal

The optimum values of parameters used here are decided from pre-experiments as p=8, q=16 and d=5 in (3)-(5). But it is possible to adopt other values in response to the situation. For the main peak finding algorithm, Intensity and Hue are mainly used. In HSV color representation, Hue has the greatest discrimination power among the components. Although Hue is the most useful attribute, it becomes meaningless when the Saturation is very low. For this reason, we use Saturation instead of Hue in the special case of "P_1<30 and δ_2<8 in M_S". Such a case is called 'achromatic case'. Otherwise, it is called 'chromatic case'.

In order to find O/R regions from an input image, the following procedures are performed for each template region.

[*Step*1] Find the main peak over the template region T.

[*Step*2] Extract X_1 which is the region of pixels whose values of one component of HSV are between P_2-2δ_2 and P_2+2δ_2. Here $\overline{X_1}$ is the conjugate region of X_1. In the same way, X_2 and $\overline{X_2}$ are obtained from another component of HSV.

[*Step*3] Compute the common region $V = X_1 \cap X_2$ and the conjugate region $\overline{V} = \overline{X_1} \cup \overline{X_2}$.

For explanation, two examples of O/R removal are shown in Fig.6. These outdoor scene images are influenced by specular highlight and occlusion, respectively. $I^{(k)}$ and $T^{(k)}$ mean an input image and a template of the k^{th} object. In Fig.6(a), UCRs are extracted using Intensity and Hue by the main peak finding algorithm. Hue histogram repeats (0-255) to show the main peak clearly. Another example using Intensity and Saturation is shown in Fig.6(b). In such a case, Saturation is very low, so we had not better use Hue component. Then we use Saturation instead of Hue in achromatic case.

2.4 Template Matching

For shape recognition, we employ temple matching based on color uniformity of the target signboard. Therefore, each template (T) is made of each UCR and its region is a little thinner than the original UCR as expressed in Fig.3. We decide whether or not the input signboard matches with the templates by using the following parameters of A and B defined by:

$$A_k = \frac{|V_k|}{|T_k|},\ B_k = \frac{|V_c \cap (S_c - T_k)|}{|T_k|}, \tag{7}$$

where, |.| is the number of pixels, T_k is the k^{th} template and $S_c = \bigcup_k T_k$. V_k and V_c are the extracted UCR with respect to T_k and S_c, respectively. If $A_k \geq Th_A$ and $B_k \leq Th_B$, we decide T_k is a 'match' template with the input image, otherwise, an 'un-match' template.

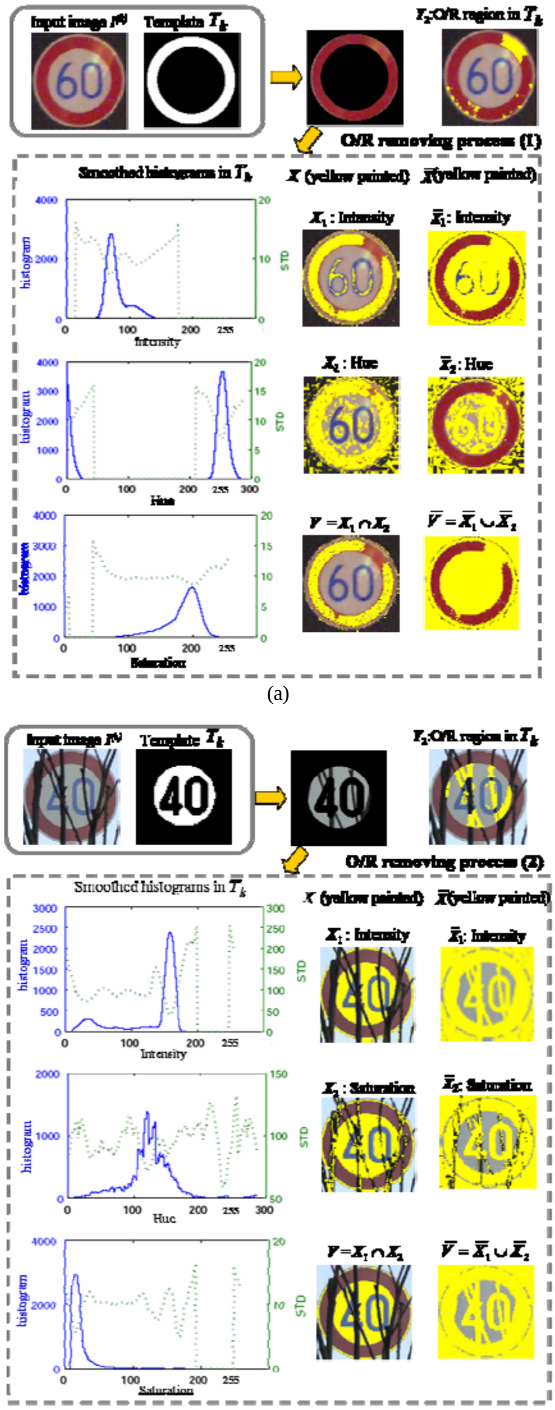

Fig. 6 Examples of O/R removal in real images: (a) reflection removal using Intensity and Hue (chromatic case), (b) occlusion removal using Intensity and Saturation (achromatic case).

3. Experimental results

A large amount of visual information is embedded in natural scenes. The proposed framework provides a simple algorithm therefore the real time implementation can be achieved. The basic ideas and techniques of this proposed framework are also potentially applied to recognition of any planar objects which are composed of UCRs. Some examples of signboards used in the experiments are shown in Fig.7. We tested various kinds of signboards under a great variety of bad illumination conditions and the proposed framework could give the satisfied recognition rate even though the input images are much influenced by O/R.

To evaluate the performance of the proposed system, we made experiments using 300 images of signboard images taken by various types of digital cameras from outdoor scenes under a great variety of illumination changes and cluttered backgrounds. All of the images were segmented and normalized into 64*64 beforehand. Among 300 images, we used 30 images for deciding parameters. Among 270 images, 250 images are learned object classes and 20 images are not belonging to any learned object classes. Almost all of cases are successfully recognized by our proposed framework, but a few failure cases are left and they are shown in the last row in Fig.7. We decide ‘match’ or ‘un-match’ of the input signboard to the target templates by using the values of A and B, described in (7). If ‘A’ and ‘B’ are satisfied their predefined thresholds, the template T is decided to be a ‘match’ template. If there is only one ‘match’ template, then this case is called ‘Single Matching’(SM). If more than one, it is called ‘Multiple Matching’ (MM). In adverse cases, for example, the images with very low resolution or very big occluded regions, MM cases sometimes happen. The algorithm was programmed by Matlab 7.0.4. From the experiment, the execution time takes from 0.15 sec to 0.27 sec for recognizing a signboard. The overall results using various types of signboards as shown in Fig.7 are summarized in Table 1.

Comparing with other approaches, our proposed framework is very strong to recognize planar objects with UCRs even under O/R. By using the proposed method, we got the 98% (without O/R: 100%, with O/R: 95%) recognition rate and 100% rejection rate even though under O/R regions. George K. Siogkas et al [19] considered the road sign recognition system in adverse conditions. They treated four categories of road signs, circular red, triangular red, circular blue and rectangular blue. The classification rule was based on the normalized cross-correlation of the cropped road signs with the reference templates and their system classified correctly 216 out of 266 signs (81.2%). In [18], although they took into account the degree of occlusion in the matching process, but not that of reflection. They proposed the fuzzy histogram matching with the ring partition method and used four gray scale ranges for red, blue and white, and other colors to extract the occlusion. Their system used only circular road signs and the matching rate was 93.9%. For the case of partial occlusion by a red traffic light, their system faced with some errors.

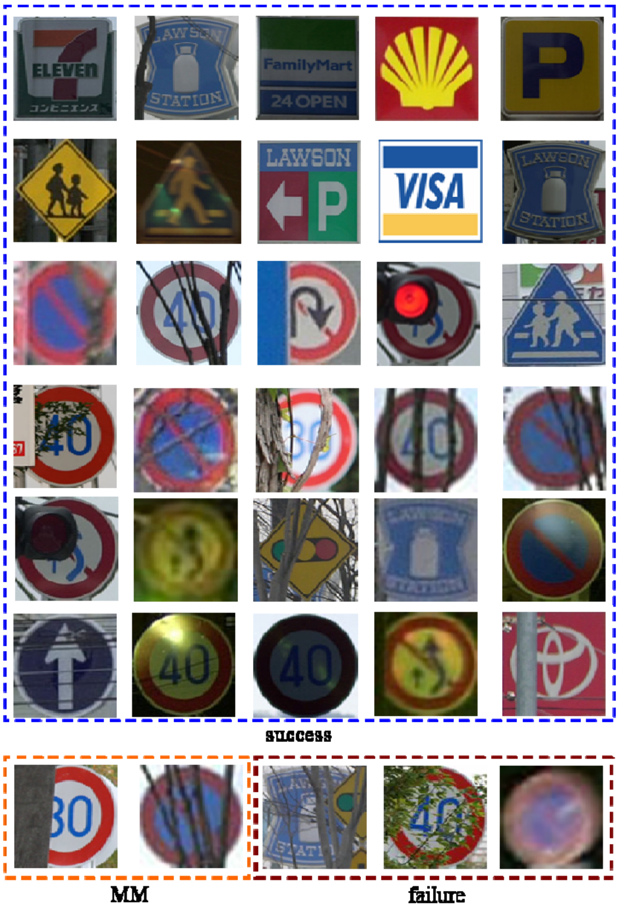

Fig. 7 Images tested in the experiments.

Table 1: The overall results

object class	Conditions	Images	Recognition			Rejection
			SM	MM	failure	
learned	without O/R	150	150	-	-	-
	with O/R	100	95	2	3	-
not yet learned	without O/R	10	-	-	-	10
	with O/R	10	-	-	-	10

4. Conclusion

The difficulties of image recognition in outdoor scenes have been the influence by partial reflection, specular highlight and occlusion which may cause recognition error. To overcome O/R influence, we have newly developed the main peak finding algorithm. This proposed system could dramatically improve the robustness and effectiveness for signboard recognition. Experimental results have shown the feasibility of the approach. For experiments, we used 300 images taken under a great variety of illumination conditions. Sometimes O/R regions are nearly one half of the object region due to occlusion, specular highlights and so on. Our proposed framework worked well under not only O/R influences but also under various bad conditions such as rain, fog, dusk-twilight, and nighttime. In addition, it can be easily extended to recognition in the case when the O/R regions are larger than the remained uniform regions. If we use two or more separated regions of templates instead of only one completed template, MM and failure cases will be solved easily.

The basic ideas and techniques of this proposed system can be potentially applied to segmentation of objects. Furthermore, our study intends further application to vision-based car navigation for providing up-to-date information to drivers' demands. Moreover, this system is applicable to other navigation services for customers, for example, finding information of shops, commercial goods, and landmarks to the pedestrians even underground and indoors.

References

[1] M. Weber, M. Welling, and P. Perona, "Unsupervised learning of models for recognition," in *Proc. ECCV*, Dublin, Ireland, vol. 1, pp. 18-32, 2000.

[2] B. Heisele, T. Serre, M. Pontil, T. Vetter, and T. Poggio, "Categorization by learning and combining object parts," in *NIPS*, Vancouver, 2001.

[3] S. Ullman, M. Vidal-Naquet, and E. Sali, "Visual features of intermediate complexity and their use in classification," *Nature Neuroscience*, vol. 5, no. 7, pp. 682–687, 2002.

[4] H. Wersing and E. Korner, "Learning optimized features for hierarchical models of invariant recognition," *Neural Comput.*, vol.15, no. 7, pp. 1559-1588, 2003.

[5] S.K.Nayar, S.A. Nene, and H. Murase, "Real-time 100 object recognition system," in *Proc. IEEE Int. Conf. Robot. Automat.*, Apr. 1996, pp. 231-235.

[6] D. Roobaert and M. V. Hulle, "View-based 3D object recognition with support vector machines," in *Proc. IEEE Int. Workshop on Neural Netw. for Signal Process.*, 1999, pp. 77-84.

[7] G. Schneider, H. Wersing, B. Sendhoff, and E. Korner, "Evolutionary optimization of a hierarchical object recognition model," *IEEE Trans. Syst., Man, Cybern. B*, Cybern., vol. 35, no. 3, pp. 426- 437, 2005.

[8] Mehran Yazdi and Andre Zaccarin, "Semantic object segmentation of 3D scene using color and shape compatibility," in *Proc. 6th World Multiconf. on Syst. Cybern. and Informat.*, Orlando, FL, USA, July 2002.

[9] Ping Tan, Long Quan, and Heung-Yeung Shum, "Highlight removal by illumination-constrained in-painting," in *Proc. of 9th IEEE Intl. Conf. on Comput. Vision*, 2003, pp. 164-169.

[10] G. D. Finlayson, S. D. Hordley, and M. S. Drew, "Removing Shadows from Images," in *Proc. ECCV*, 2002, pp.823-836.

[11] Jae Byung Park and Avinash C. Kak, "A Truncated Least Squares Approach to the Detection of Specular Highlights in Color Images," *Proc. of the 2003 IEEE Int. Conf. on Robot. Automat.*, Taipei, Taiwan, 2003.

[12] D. H. Brainard and W. T. Freeman, "Bayesian color constancy," *J. of the Optical Soc. of America A*, vol. 14, pp. 1393-1411, 1997.

[13] G. D. Finlayson, S. D. Hordley, and P. M. Hubel, "Color by correlation: A simple, unifying framework for color constancy," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, no. 11, pp. 1209-1221, 2001.

[14] G. Sapiro, "Color and illuminant voting," *IEEE Trans.Pattern Anal. Machine Intell.*, vol.21, no. 11, pp. 1210-1215, Nov. 1999.

[15] A.de la Escalera, J.M. Armingol, and M. Mata, "Traffic Sign Recognition and Analysis for Intelligent Vehicles," *Image & Vision Computing*, vol. 21, no. 3, pp. 247-258, Mar. 2003.

[16] H-M. Yang, C-L. Liu, and S-M. Huang, "Traffic Sign Recognition in disturbing Environments", *Proc. of the 14th Intl. Symp. on Methodologies for Intelligent Systems*, Maebashi, Japan, pp. 28-31, Oct. 2003.

[17] D.M.Gavrila, "Multi-feature Hierarchical Template Matching Using Distance Transforms", *Proc. IEEE Int. Conference on Pattern Recognition*, Brisbane, Auatralia, 1998.

[18] A. Soetedjo and K. Yamada, "Traffic Sign Classification Using Ring Partitioned Method," 2nd Joint Intl. Conf. on Soft Computing and Intelligent Systems and 5th Intl. Symp on Advanced Intelligent Systems (SCIS & ISIS 2004), Yokohama, 2004.

[19] George K. Siogkas and Evangelos S. Dermatas, "Detection, Tracking and Classification of Road Signs in Adverse Conditions," *Proc. of IEEE Mediterranean Electro -technical Conf.*, Benalmadena, Spain, pp.537-540, May 2006.

Thi Thi Zin received the B.Sc. and M.I.Sc. degrees from Yangon University and University of Computer Studies, Yangon, Myanmar in 1995 and 1999, respectively. She received the M.E. degree from Osaka City University in 2004. She is a doctoral candidate in Osaka City University, Japan. Her research interests are in ITS, Moving Image Processing, Mathematical Morphology and so on. She is a student member of IEEE and IEICE.

Sung-Shik KOH received his B.S., M.S., and Ph.D. degrees from Chosun University, Gwangju, Korea, in 1994, 1996, and 2002, respectively, all in electronic engineering. From 2004 to 2006, he has been a postdoctoral research scientist at Osaka City University, Japan. He is currently a senior research engineer at Samsung Techwin Co., Ltd., Korea. His research interests include digital image watermarking, 3D world reconstruction, ITS, and DSC AF.

Hiromitsu HAMA received the Dr. Eng. degrees in Electrical Engineering from Osaka University, Japan. He is currently a professor of Graduate School of Engineering at Osaka City University. His research interests include ITS (Intelligent Transportation Systems), 3D world reconstruction, computer vision, picture processing/understanding. He is a member of IEEE, SPIE, IEICE, IIITE, IPSJ, ASJ, J-FACE, and ISCIE.