Implementation of Road Traffic Signs Detection Based on Saliency Map Model

Woong-Jae Won, Minho Lee, Joon-Woo Son

Abstract—**In this paper, we proposed a new road traffic sign detection model based on human-like selective attention mechanism for implementing interactive workload manager system. Since the road traffic sign boards have dominant color contrast against backgrounds, we consider the color opponents and its edge information with center surround difference and normalization as a pre-processing, which is effective to intensify the sign board color characteristics as well as reduce background noise influence. After constructing the road traffic sign saliency map using the edge and color feature maps, the candidate road traffic sign regions are selected by local maximum energy searching with entropy maximization algorithm to find suitable size of the sign board areas. Computational experiment results show that the proposed model can successfully detect a road traffic sign board.**

I. INTRODUCTION

Recently, The National Highway Traffic Safety Adminstration (NHTSA) reported that the 25-35% of the traffic accidents or 1.2 million vehicle crashes per year in the United State are resulted from driver distraction and inattention [1, 2]. And, the Treat et. al. reported that the human error mainly cause the vehicle crash [3]. Furthermore, it's known that more increasing numbers of in-vehicle systems such as vehicle information and entertainment system (audio, navigation and mobile phone) dramatically increase car accidents as resulting by decreasing a driver's attention. Furthermore, Individuals aged 65 years and over represent the most rapidly growing segment of the driving population, and are keeping their licenses longer [4]. However, many older drivers face impaired visual functioning cataract and related visual impairment is highly prevalent [5,6].

Thus, the advanced driver assistance system and an interactive workload manager system have been received more attention from many intelligent vehicle research community to support and disburden the driver to significantly increase driving safety and comfort [7, 8, 9, 10]. For implementing these systems, the traffic sign detection and recognition technology have been important issues for research in computer vision research community [9, 10, 11].

The vision based traffic sign recognition problem has some beneficial characteristics such as unique design of a traffic sign board, which means the sign shape variations are small, and significantly color contrast against the backgrounds [12]. So, several color-based sign detection model and shape-based sign detection model can be easily introduced using these beneficial characteristics [12, 13, 14, 15]. Even through the conventional vision based models have a good performance for detecting the road traffic sign, there remain number of challenges for successful recognition of a road traffic sign board. First, weather and lighting condition are significantly varying in traffic environments, diminishing the advantage of the claimed object uniqueness. Additionally, as a camera installed in a vehicle is moving, additional image distortions, such as motion blur and abrupt contrast change, frequently occur. Moreover, the sign board installation and surface material can physically change over time, and are influenced by accidents and weather, which induces the rotation of sign board and degenerated color information.

Recently, some researchers have proposed saliency map models based on biologically motivated selective attention mechanism, which are imitating human-like early visual processing, to overcome those problems in complex environment [16, 17, 18]. Therefore, we also consider these models for constructing a road traffic sign saliency map (RTS_SM) as a preprocessing.

Since, the color information of a road traffic sign board mainly good contrast level against the visual environment, we consider the color opponents coding reflecting human visual characteristic [16, 17, 18, 19]. And, we adopt the center surround difference and normalization (CSD&N) algorithm with Gaussian pyramid processing to reduce noise influence in various scene and size of road traffic sign for input feature images, reinforce traffic sign area and inhibit non-traffic sign area [16, 17, 18]. And the road traffic candidate regions are simply selected by local maximum searching, and entropy maximization algorithm using the road traffic sign saliency map [17].

This paper is organized as follows; Section 2 describes the proposed traffic feature map extraction and candidate road traffic sign selection model. The experimental results will be

Manuscript received December 15, 2007. This work was supported by the Deagu Gyeongbuk institute of science and technology (DGIST) Basic Research Program of the MOST, and Components and Materials Technology Development Program, VISS (Vision-based Intelligent Steering System) Project.

Woong-Jae Won is Division of Advanced Industrial Science and Technology, Dagegu-Gyeongbuk Institute of Science and Technology, 75 Gongdanbuk2gil, Dalseo-Gu, Taegu 704-230, Korea (e-mail: wwj@dgist.ac.kr).

Minho. Lee is with School of Electrical Engineering and Computer Science, Kyungpook National University, 1370 Sankyuk-Dong, Puk-Gu, Taegu 702-701, Korea (e-mail: mholee@knu.ac.kr).

Joon-Woo Son is Division of Advanced Industrial Science and Technology, Dagegu-Gyeongbuk Institute of Science and Technology, 75 Gongdanbuk2gil, Dalseo-Gu, Taegu 704-230, Korea (e-mail: json@dgist.ac.kr).

followed in Section 3. Section 4 presents our conclusions and discussion.

II. Road Traffic Attention Model

A. Overview of the Proposed Road Traffic Detection Model

Fig.1 shows the proposed road traffic sign detection model. After extracting the red(R), green(G), blue(B) color feature from input color image, The normalized red(r), green(g), blue(b) and yellow(y) color features are extracted from the R, G, B [16, 17, 18, 19]. The red-green and blue-yellow (rg_by) color opponent feature information are extracted from r, g, b, y color information in order to reflect road traffic color contrast characteristics. Also, the edge of red-green and blue-yellow (rg_by) color opponent feature is considered as road shape information in order to reflect load sign shape characteristics. After we extract color opponent and edge information from r, g, b, and y, the color and edge feature maps are constructed by the Gaussian pyramid processing and CSD&N algorithm [16, 17, 18]. Then, the RTS_SM model is constructed through the weighted sum of edge and color feature maps.

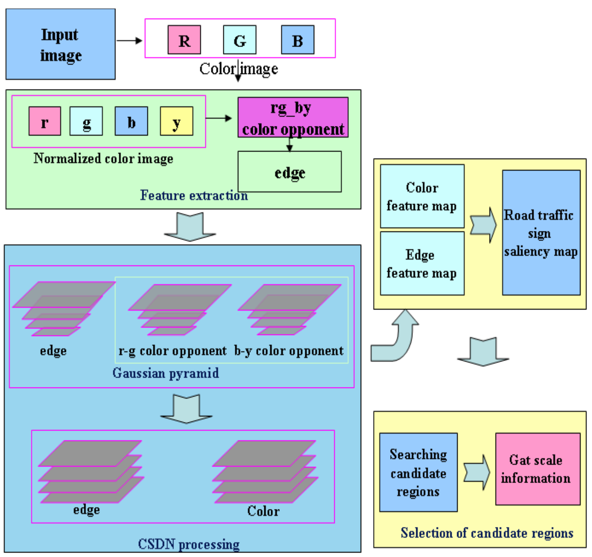

Fig. 1. The proposed Road Traffic Sign Attention Model

After constructing the RTS_SM, the candidate road traffic sign areas are selected by a simple local search for maximum energy. Then, we adapt an entropy maximization approach to select a proper scale of the road traffic sign regions in the RTS_SM [20].

B. Color and Edge feature Extraction

In order to reduce influence of luminance, the normalized color features are considered as following the human early visual processing mechanisms [16, 18, 19].

After extracting R, G, B, Y from the input image, the normalized r, g, b, and y color features are extracted by Eq (1).

$$
\begin{aligned}
r &= \begin{cases} R - \dfrac{G+B}{2}, & r > 0 \\ 0, & r \le 0 \end{cases} \\
g &= \begin{cases} G - \dfrac{R+B}{2}, & g > 0 \\ 0, & g \le 0 \end{cases} \\
b &= \begin{cases} B - \dfrac{R+G}{2}, & b > 0 \\ 0, & b \le 0 \end{cases} \\
y &= \begin{cases} \dfrac{R+G}{2} - B - \dfrac{|R-G|}{2}, & y > 0 \\ 0, & y \le 0 \end{cases}
\end{aligned}
\tag{1}
$$

Then, for reflecting the characteristics of road traffic sign boards such as color and edge contrast against environment, the rg_by color opponent feature and edge information (e) are obtained by Eqs. (2) and (3).

$$rg\_by = \| r - g | - | b - y \| \tag{2}$$

$$e = \sqrt{(rg\_by \cdot S_x)^2 + (rg\_by \cdot S_y)^2} \tag{3}$$

where S_x and S_y are sobel operators [21].

Fig. 2 shows the procedure of extracting the color and edge features. As shown Fig. 2, the edge of rg_by color opponent feature is dominant in the road traffic regions.

Fig. 2. The procedure of extracting edge and color features

C. Road Traffic Saliency Map Extraction

In order to reduce noise influence in various scenes, and intensify the road traffic sign regions, we implement the on-center and off-surround operation by the Gaussian pyramid images with different scales from 0 to nth level whereby each level is made by the sub-sampling of 2^n, thus it is able to construct 3 feature bases such as $E(\bullet)$, $RG(\bullet)$, and $BY(\bullet)$, [16, 18, 22]. Then, the center-surround features are constructed by the difference operation between the fine and coarse scales in the Gaussian pyramid images [16, 17, 18]. Consequently, the three center-surround feature bases such as, $E(c,s)$, $RG(c,s)$, and $BY(c,s)$ can be obtained by the following Eqs. (4) – (6):

$$E(c,s)=|e(c)-e(s)| \tag{4}$$

$$RG(c,s)=|r(c)-g(s)|-|g(c)-r(s)| \tag{5}$$

$$BY(c,s)=|b(c)-y(s)|-|y(c)-b(s)| \tag{6}$$

where the "c" is the finer scale pyramid image, the "s" is the coarse pyramid image, and "-" represents the interpolation to the finer scale and point-by-point subtraction. For real time operation, we reduce the input image size from 320 x 240 to 160 x 120. Therefore, the 7 different scales of Gaussian pyramid features are extracted from the edge, RG, and BY features. After obtaining the 7 levels of Gaussian pyramid features, we select five levels of Gaussian pyramid feature from 2nd level to 6th level of edge, RG and BY Gaussian pyramid features for the CSD&N algorithm. The selected Gaussian pyramid features are combined into 2 feature maps as shown in Eqs. (7) and (9) where $\overline{E}$, and $\overline{C}$ stand for edge, and color opponent feature maps, respectively. These edge and color feature maps are obtained through the across-scale addition "$\oplus$" [16,17, 18].

$$\overline{E}=\bigoplus_{c=2}^{3}\bigoplus_{s=c+2}^{3}N(e(c,s)) \tag{7}$$

$$\overline{C}=\bigoplus_{c=2}^{3}\bigoplus_{s=c+2}^{3}N(RG(c,s)+BY(c,s)) \tag{8}$$

Then, the RTS_SM is generated by equation (8).

$$TS(x,y)=(W_c\cdot\overline{C}(x,y)+W_E\cdot\overline{E}(x,y)) \tag{9}$$

where W_c and W_E represent the weight factor for color feature map and edge feature map, respectively.

As the shown in Fig. 3, the 5 layers of RG, BY color opponent and edge Gaussian pyramid features are selected from the 7 layer Gaussian pyramid features which are extracted from the r, g, b, y, and rg_by color opponent. Then, these pyramid features are constructing 4 layers of CSD images for color and edge feature maps through the CSD processing. The color and edge feature maps are obtained by normalizing each feature after summing the 4 layers CSD features. After we get the color and edge feature maps, the road traffic sign map is extracted through the weighted sum of edge and color feature map. The road traffic sign map effectively indicates more salient road traffic areas than another area by considering the characteristics of color contrast information in road traffic sign boards.

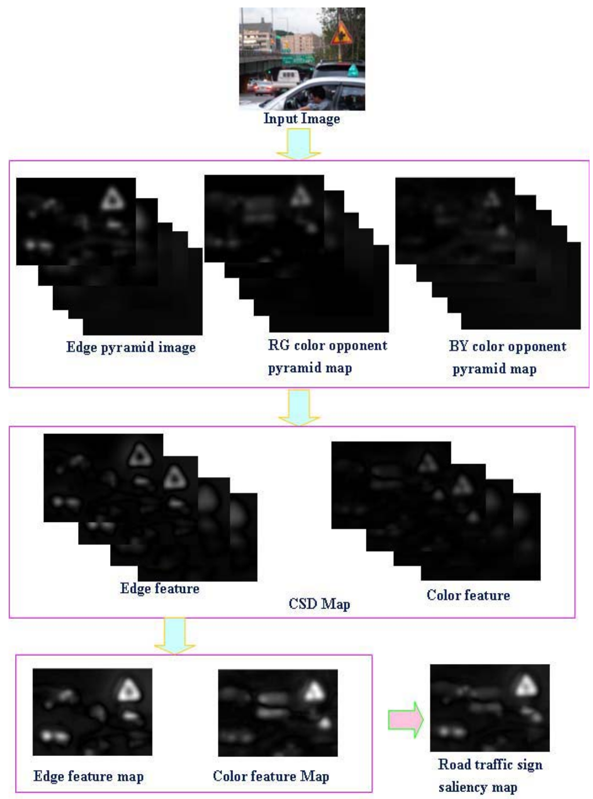

Fig. 3. The procedure of extracting the road traffic sign saliency map model

D. Selection of Road Traffic Regions

Since the RTS_SM intensifies the road traffic sing regions and diminishes complex backgrounds, we can simply select a candidate region for a road traffic sign through searching the local maximum energy with fixed window size by shifting the pixels the RTS_SM as shown Fig. 4.

Each candidate road traffic sign region, which is selected in the RTS_SM, may have different size. In the proposed model, an entropy maximization approach based on Kadir's approach was adapted to select a proper scale of road traffic sign region in the RTS_SM [20].

For each candidate region of a road traffic sign board, the proposed model chooses those scales at which the entropy is at a maximum, or has peaked, then the entropy value is weighted according to some measure of the self-dissimilarity in scale-space of RTS_SM. The most proper scale for each road traffic sign area centered at a location $\mathbf{x}$ is obtained by Eq. (10) with satisfied the Eq. (11) .

$$scale(\mathbf{x}) = \arg\max_{s}\{H_D(s,\mathbf{x})\times W_D(s,\mathbf{x})\}$$

(10)

$$s_{cd}\{s : H_D(s-1,sp) < H_D(s,sp) > H_D(s+1,sp)\} \quad (11)$$

where d is the set of all descriptor values, $H_D(s,\mathbf{x})$ is entropy defined by Eq. (12) and $W_D(s,\mathbf{x})$ is inter-scale measure defined by Eq. (13).

$$H_D(s,\mathbf{x}) \triangleq -\sum_{d\in D} p_{d,s,\mathbf{x}} \log_2 p_{d,s,\mathbf{x}} \quad (12)$$

$$W_D(s,\mathbf{x}) \triangleq \frac{s^2}{2s-1}\sum_{d\in D}\left|p_{d,s,\mathbf{x}} - p_{d,s-1,\mathbf{x}}\right| \quad (13)$$

where $p_{d,s,\mathbf{x}}$ is the probability mass function for scale s, position $\boldsymbol{x}$, and descriptor value d which takes on values in D. The probability mass function $p_{d,s,\mathbf{x}}$ is obtained from the histogram of the pixel values of the road traffic sign area, which is centered at the location $\mathbf{x}$ with size s in the RTS_SM.

As shown the Fig. 4, the proposed model can simply select the road traffic regions and get a proper scale of a salient area.

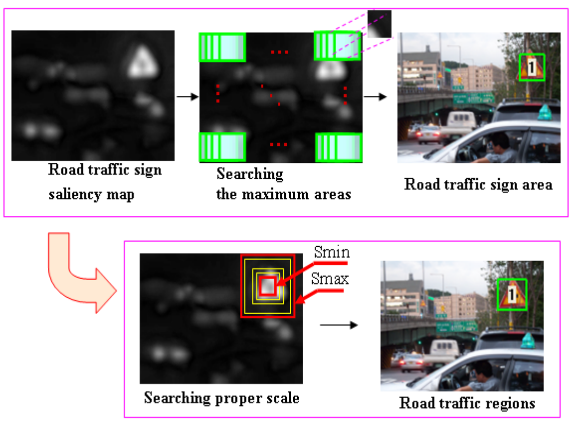

Fig. 4. Selection of candidate region for road traffic sign boards

III. Experiment Result

In our experiment, we took a video for urban, rural and highway road scenes while driving. Then, we captured 679 scenes including various shapes and sizes of the road traffic sign boards.

In order to reduce the computation load of the proposed model, we make 3x3 Gaussian filter mask for constructing the Gaussian pyramid features. We consider the reduced size of an input image as 160 x 120. And, we set W_c =0.5, and W_E =0.5 for the weighted sum of edge and color feature map to construct the RTS_SM, which will be optimally designed by a training process in our future research. Also, we set the fixed window size as 15x15 for local maximum searching to select the road traffic sign area, and S_{min} =2 and S_{max}=64 to search a proper scale for each selected road traffic sign region.

Fig. 5 shows the experimental results of the proposed road traffic sign detection model. In here, normalized r, g, b, y and edge of rg_by are considered as input features for making color and edge feature maps. Then, the RTS_SM consisted of color and edge feature maps. And, the candidate region of road traffic sign boards is selected through searching local maximum point with entropy maximization algorithm. As shown in Fig. 5, the proposed model presents larger salient values in road traffic sign area than another area. Moreover, we can get a proper scale of candidate regions in the RTS_SM, which is very useful for analysis of the contents in traffic sign boards.

Fig. 6 shows the performance of the proposed road traffic sign detection model in various situations. The accuracy of the selection rate for road traffic sign boards is 94.52% within 5th maximum areas for 679 road scenes which consist of urban road scenes, 201 rural scenes, and 168 highway scenes.

Fig. 7 shows some experiment results of proposed road traffic attention model in various environment and road

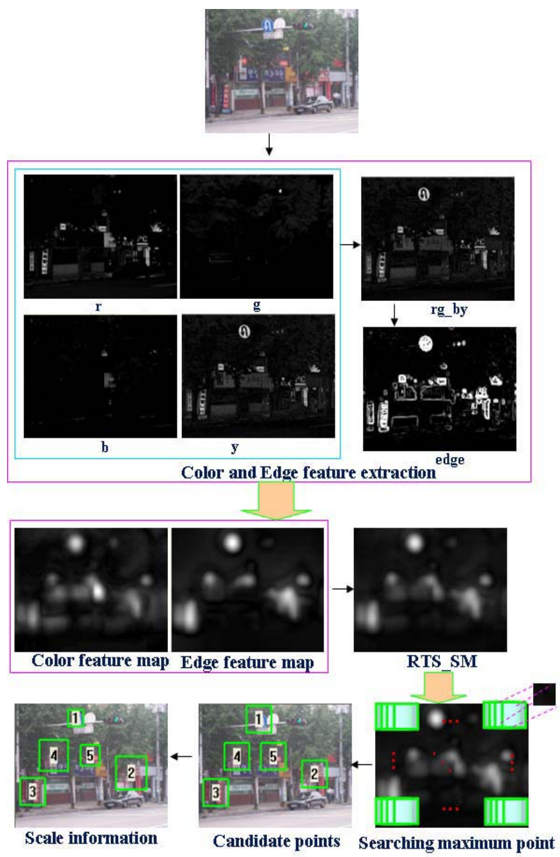

Fig. 5. Overall procedure in experimental results for the proposed road traffic sign detection model

traffic sign. As shown Figs. 6 and 7, the proposed model can robustly

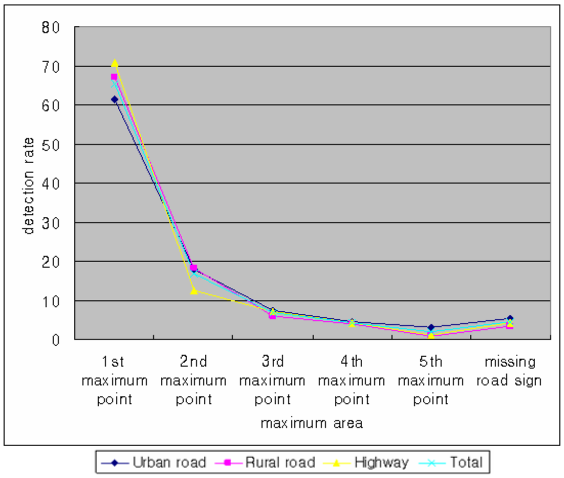

Fig. 6. The performance of the proposed road traffic sing selection model in various road condition.

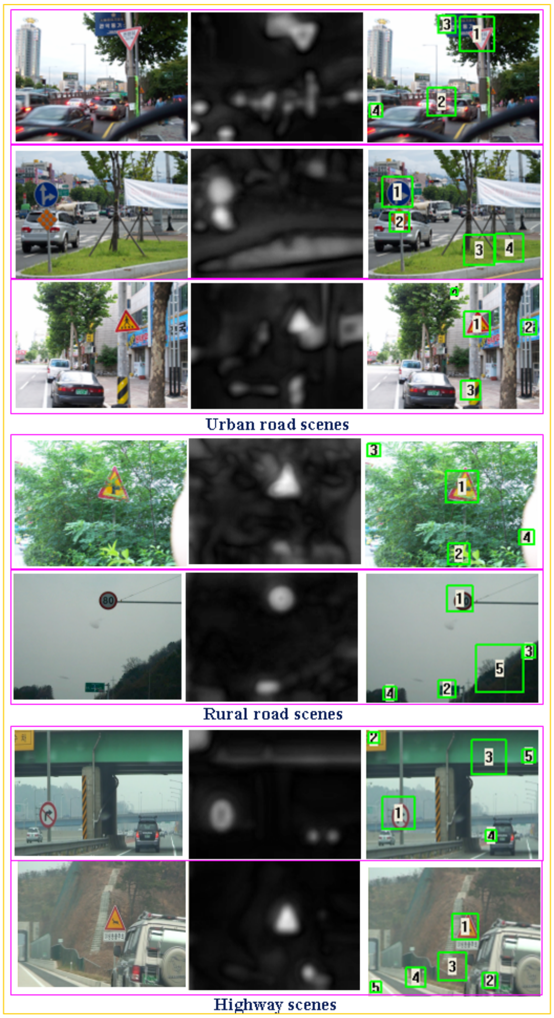

Fig. 7. Overall Experimental results for the proposed road traffic sign detection model in various situations.

detect the road traffic sign regions in complex visual environment with deformed road traffic sign boards by rotation and different scale.

IV. CONCLUSION

We proposed a new road traffic selection model for selection of the road traffic sign. In order to make more salient road traffic sign area than another areas, we consider traffic sign color contrast characteristic against environment with CSD&N algorithms. And, we can select the candidate regions of road traffic sign board through searching the local maximum energy with entropy maximization algorithm. In real environment, the proposed model successfully localizes a salient area and detects the road traffic sign regions.

As further works, we are considering for more plausible verifying process to check whether the localized area contains traffic sign boards, and also trying to compare the performance of our proposed method with that of another approaches. Also, we are trying to find an optimal weight factors in the feature maps by training process.

ACKNOWLEDGMENT

This work was supported by the Deagu Gyeongbuk institute of science and technology (DGIST) Basic Research Program of the MOST, and Components and Materials Technology Development Program, VISS (Vision-based Intelligent Steering System) Project.

REFERENCES

[1] L. R. Shelton, “Statement before the subcommittee on highways and transit, committee on transportation and infrastructure”, U. S house o representatives, May 9, 2001. Available: http://www.nhtsa.dot.gov/nhtsa/announce/

[2] J. C. Stutts, D. W. Reinfurt, L. S. Staplin,E. and A. Rodgman, “The role driverdistraction in traffic crashes,” ,Report Prepared for AAA Foundation for Traffic Safety, June 10, 2003.

[3] J. R. Treat, N.S. Tumbas, S. T. McDonald, D. Shinar, R.D. Hume, R.E. Mayer, R. L. Stansifer, and N. J. Castellan, “Tri-Level study of the Causes of Traffic Accidents: Final Report-Executive Summary,” Department of Transportation, National Highway Traffic Safety Administration, Washington D.C , Tech. Rep. DOT HS805 099, 1979.

[4] S. Lyman, s. A. Ferguson. E.R. Braver, and A. F. Williams, “Older driver involvements in police reported crashes and fatal crashes: trends and projections,” *Injury Prevention*, vol 8, no. 2, pp.116-120, 2002.

[5] E. Rochtchina. B. N. Mukesh, J. J. Wang, C. A. McCarty, H. R. Taylor, and P. Mitchell, “Projected prevalence of age-related cataract and cataract surgery in Australia for the years 2001 and 2021: pooled data from two population based surveys,” *Clinical and Experimental Ophthalmology*, vol. 31, no. 3, pp. 233-236, 2003.

[6] J. A. Jacko, I. I. Scott, F. Sainfort, et al, “Older Adults and Visual Impairment: What Do Exposure Times and Accuracy Tell Us About Performance Gains Associated with Multimodal Feedback? ,” *in Proc. the SIGCHI conference on Human factors in computing systems*, Florida, USA, 2003,vol. 5., no. 2 pp. 33-40.

[7] P. Green, “Driver Distraction, Telematics Design, and Workload Managers Safety Issues and Solution,” *in Proc. The 2004 International Congress on Transportation Electronics*, Warrendale, 2004, pp.165-180.

[8] SAVE_IT, Safety Vehicle(s) using adaptive Interface Technology (SAVE_IT) Program, Volpe National Trasnportation Systems Center. Available: http://www.volpe.dot.gov/hf/roadway/saveit/

[9] J. G. Walker, N. Barnes and K. Anstey, "Sign Detection and Driving Competency for Older Drivers with Impaired Vision," in *Proc. the 2006 Australasian Conference on Robotics & Automation (ACRA2006),* Auckland, New Zealand, December, 2006.

[10] L. Fletcher, G. Loy, N. Barnes, A. Zelinsky, "Correlating driver gaze with the road scene for driver assistance systems," *Robotics and Autonomous Systems*, vol. 52, pp. 71–84, 2005.

[11] G. Loy and N. M. Barnes, "Fast shape-based road sign detection for a driver assistance system," *in Proc. the IEEE/RSJ Intelligent Robots and systems,* pp.70-75, 2004.

[12] C. Bahlmann, Y. Zhu, M. Pellkofer, et all, "A System for Traffic Sign Detection, Tracking, and Recognition Using Color, Shape, and Motion Information," *in Proc. IEEE Intelligence Vehicles Symposium 2005(IV2005),* Las Vegas, NV, pp. 255-260, June 2005.

[13] H. Fleyeh and M. Dougherty, "Road and Traffic sign detection and recognition," *in Proc. the 10th EWGT Meeting and 16th Mini-EURO*, 2005.

[14] L. C. Gomez and O. Fuentes, "Color-Based Road Sign Detection and Tracking," *in Proc. Image Analysis and Recognition(ICIAR),* Montreal, CA, Agust 2007.

[15] M. Benallal and J. Meunier, "Real-Time Color Segmentation of road signs," *in Proc. Canadian Conference on Electrical and Computer Engineering (IEEE CCECE 2003)* ,pp.1823-1826, 2003.

[16] L, Itti, C. Koch, E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. Patt. Mach. Intel,* 1998, vol. 20, no. 11, pp. 1254–1259.

[17] W. J. Won and J. Y. Yeo, S. W. Ban, M. Lee, "Biologically Motivated Incremental Object Perception Based on Selective Attention", *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 21, no. 8, pp. 1293 -1305, 2007.

[18] S. J. Park, K. H An, and M. Lee, "Saliency map model with adaptive masking based on independent component analysis," *Neurocomputing*, vol.49, pp. 417–422, 2002

[19] E. B. Goldstein, *Sensation and perception 4th ed*, An international Thomson publishing company, USA, 2001.

[20] T. Kadir and M. Brady, "Scale, saliency and image description", *International Journal of Computer Vision*, vol. 45, pp. 83 -105, 2001.

[21] R. C. Gonzalez, R. E. Woods, *Digital Image Processing 2nd ed*, Prentice-Hall, 2001.

[22] J.M. Ogden, E. H. Adelson, J. R. Bergen, and P. J. Burt, "Pyramid Methods in Image Processing", *RCA Engineer,* vol. 30, no. 5, pp. 33–41, 1984.