Journal of Universal Computer Science, vol. 15, no. 4 (2009), 786-804
submitted: 30/6/08, accepted: 27/2/09, appeared: 28/2/09 © J.UCS

Graph-based Approach for Robust Road Guidance Sign Recognition from Differently Exposed Images

Andrey Vavilin
(University of Ulsan, Ulsan, Korea
andy@islab.ulsan.ac.kr)

Kang-Hyun Jo
(University of Ulsan, Ulsan, Korea
jkh2009@islab.ulsan.ac.kr)

Abstract: In this paper we present an approach to detect traffic guidance signs and recognise the structure of junction information on them. The detection algorithm is based on using differently exposed images. These images are combined into one using tone mapping technique in order to minimize effects of bad environment conditions and low dynamic range of CCD-cameras. This technique allows robust sign detection in various lighting conditions. To localize sign candidates color segmentation is used. To minimize number of false detection filtering operations based on geometrical and color properties is applied. Recognition process is based on graph theory. Each sign candidate is decomposed into principal components and the region which represents junction structure is mapped into a graph. This graph is checked for possible mapping mistakes. Finally, the graph is analyzed in order to extract all possible paths of junction crossing. These paths must represent the real structure of the junction and correspond to the road law. The proposed method allows more effective detection in different lighting and environmental conditions such as insufficient or excessive lighting, rain, fog etc compared with conventional approaches.

Keywords: HDR, sign detection, sign recognition, graph theory
Categories: I.4.0, I.4.6, I.4.8, I.4.9

1 Introduction

Road signs provide drivers with information which is important for safety driving and effective navigation. One of the most important groups of signs is road guidance signs. These signs represent structure of junctions with additional information about directions. Correct recognition of these signs should provide driver with information about the path he has to follow on a junction in order to reach the destination point. Choosing the correct path is the primary question for safety driving and effective navigation. The main problem in recognition of road guidance signs is the complexity of the arrow region which represents the junction. This region could have very complex structure and contain different additional elements such as road numbers. Most of the previous works are limited by extraction directions instead of full recognition of junction structure [Azami, 96] [Lee, 03] [Kato, 02] [Vavilin, 06].

Road guidance signs are not as well structured as warning and restricting signs. They may consist of different elements composed in different styles. Hence the recognition system could not be based on any kind of matching. More detailed description of the sign model is given in Section 2 of this paper. In the proposed paper recognition is based on structural analysis of main sign components (arrow region, text labels).

Another problem is sign detection. One of the most common detection techniques is color segmentation in RGB or HSI color space. Color-based detection is reasonable, considering high-contrast colors used for the sign plate. However, this technique has some problems. Among the most challenging problems are bad lighting and weather conditions, huge number of sign-like objects in urban areas and limitations of dynamic range of CCD cameras (Figure 1). Problems of environment conditions and filtering out false-candidates were recently studied, while the problem of dynamic range was not mentioned in previous works.

Figure 1: Cases difficult for detection due to dynamic range limitations of camera

Dynamic range is the ratio between the highest and lowest luminance values in one scene which could be measured by sensor. Conventional CCD-cameras have a limitation of dynamic range in the level of 8 to 14 f-stops. However, the dynamic range of a real scene can reach values of 24 f-stops and higher. This may cause some image regions to be too bright (overexposed) or too dark (underexposed) to recover information. If a road sign will appear in these areas it could be undetected. One of traditional approaches for solving the dynamic range problem is using HDR imaging [Rovid, 07] [Ferwerda, 96] [Várkonyi-Kóczy, 08] and others. This method is based on using several images taken from same position but with different exposure values. These images are combined into a HDR image which contains details from all images. Then, a tone mapping operation is applied in order to compress the HDR image into an LDR image. The main disadvantage of this technique is the high computational cost. Some approaches required additional information about images such as shutter speed [Mertens 07], [Ferwerda, 96]. Other methods are based on combining parts of different exposures without additional information [Rovid, 07], [Vavilin, 08]. Processing time for one image could take more than one minute depending on image size and number of differently exposed images. It is thus impossible to use this technique in original approach for such an application as sign detection. However, the idea of using differently exposed images seemed to be the only way to solve the dynamic range problem. In the proposed paper the value of each pixel is computed as a superposition of pixel values from differently exposed images instead of recovering real value of luminosity with further tone mapping. It allows increasing brightness in overexposed images (and its decreasing in underexposed ones) in real-time. The price for this operation is lower contrast in recovered areas compared with traditional HDR.

Figure 2: Example of generating HDR image from multiple exposures. Image courtesy of J. Levenson

This paper is organized as follows: description of the used sign model and classification of Korean information signs is given in Section 2. Detection procedure, consisting of dynamic range improvement and color segmentation, is described in Section 3. Our recognition approach based on graph analysis is shown in Section 4. Sections 5 and 6 present experimental results and conclusions.

2 Sign model

Guidance signs in Korea have rectangular shape and could be one of three colors: blue, green or brown. Brown signs were not considered in this research as long as they contain information about places of interests without directions. Each sign can be disjointed into three components: arrow region, text regions and distance region optionally (Figure 3). The most important part of a guidance sign is the arrow region. It represents the structure of the next junction and shows the possible ways to cross it. Typically, it is the largest element of the sign. Guidance signs contain information about directions which is represented by text regions. The number of text regions is same as the number of arrowheads in arrow regions. Optionally, it can include distance region which shows distance to the crossroad. Most of the previous approaches are based on recovering possible direction from arrow regions without path recovering. These techniques can provide driver with information about direction but not the paths to cross the junction. Based on the complexity of arrow region, 6 classes of signs can be distinguished [Vavilin, 07]. Another classification is given in [Kato, 02].

Figure 3: Information sign model.

- One directional. The sign contains one or more arrow regions with simple structure. Each arrow region shows one direction only (Figure 4.a)
- Simple crossroad. The sign contains one arrow region which represents information about intersections of two or more roads or about road forks (Figure 4.b).
- Rotary. The sign represents information about roundabout intersection. Arrow region contains a circle-like part and at least two outgoing directions. To reach a destination the driver has to go counter-clockwise around rotary (Figure 4.c).
- Junction. The sign contains information about a multilevel junction with an intersection. The arrow region consists of a main road and branches which intersects the main road.
- Combined (rotary + junction). The arrow region contains both rotary and junction structures (Figure 4.e).
- Complex junction. The arrow region contains information about a junction with a complicated structure. It can contain several junctions or a main road can be intersected more than one time (Figure 4.f).

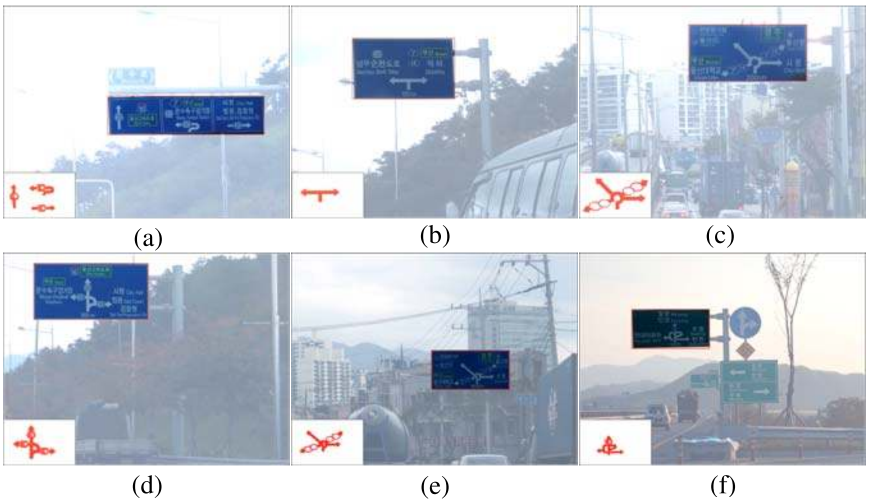


Figure 4: Guidance signs with different types of arrow region: (a) one directional; (b) simple crossroad; (c) rotary; (d) junction; (e) combined; (f) complex junction

3 Sign detection

In the proposed framework detection is based on color properties of a sign with further shape-based verification. Main strategy for detection can be formalized as "detect as many candidates as possible". Such strategy is efficient for urban areas where a number of sign-like objects is significant. However it makes an additional challenge for the filtering algorithm.

In order to increase the detection rate, the idea of HDR was used [Vavilin, 08]. For that purpose, 3 differently exposed images were used as an input. These images were combined into one using simple tone mapping operation. It allows recovering data from areas which could be lost because of the dynamic range problem and increases the probability of detection in bad weather conditions. The main idea of the detection method is shown in Figure 5.

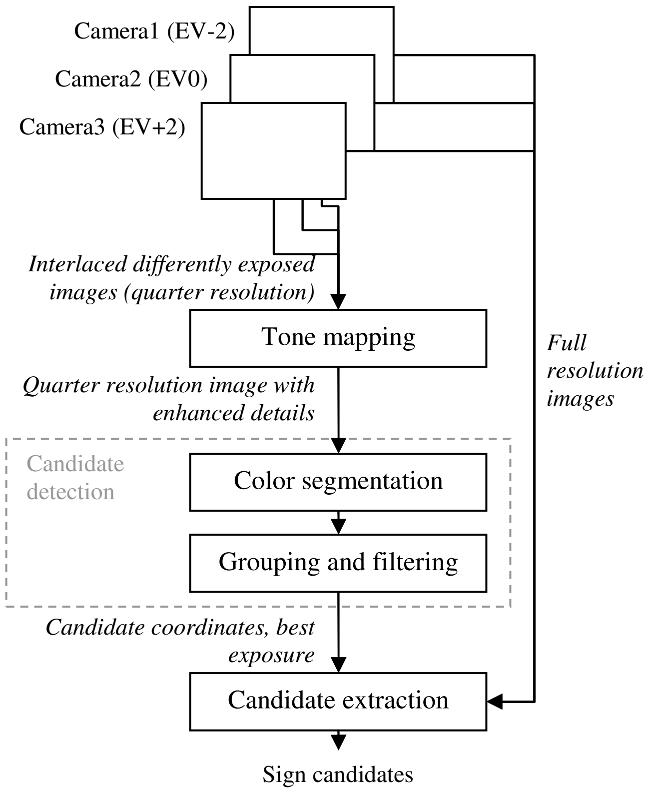


Figure 5: Sign detection scheme

The goal of the tone mapping operation is recovering details in overexposed and underexposed regions. We can define overexposed regions as regions with pixel intensity close to the highest feasible (255 in our case) and underexposed regions as regions with pixel intensity close to 0. Pixels whose intensity is close to the boundaries of feasible set have lower confidence comparing with pixels from the middle of this range. The idea of the proposed method is to combine underexposed pixels of normally exposed image I^0 with corresponding pixels of overexposed image I^{+2}. Similar, overexposed pixels are combined with corresponding pixels of underexposed image I^{-2}. This combination is a superposition of corresponding pixels with different exposures which could be formalized as follows:

$$I^C_{i,j} = \begin{cases} w\left(I^0_{i,j}\right)\cdot I^0_{i,j} + \left(1 - w\left(I^0_{i,j}\right)\right)\cdot I^{+2}_{i,j},\ if\ \ I^0_{i,j} < 127.5 \\ w\left(I^0_{i,j}\right)\cdot I^0_{i,j} + \left(1 - w\left(I^0_{i,j}\right)\right)\cdot I^{-2}_{i,j},\ if\ \ I^0_{i,j} > 127.5 \end{cases} \quad (1)$$

where I^C is combined image, I^{-2}, I^{+2} and I^0 are underexposed, overexposed and normally exposed images respectively and w is weighting coefficient, computed by a Gaussian-like function centered at 127.5:

$$w(x) = \exp\left(-4\frac{(x-127.5)^2}{127.5^2}\right) \quad (2)$$

In order to decrease the computational time, color segmentation and filtering is applied to interlaced image: each second row and column of a combined image is used for these processes. However, for the further processing full resolution candidates are used. Computation costs could also be decreased by using lookup tables for a weighting function instead of its permanent computing.

$$T = \{t_i \mid t_i = w(i), i = 0,1...255\} \quad (3)$$

Thus (1) can be written as:

$$I^C_{i,j} = \begin{cases} t_{I^0_{i,j}} \cdot \left(I^0_{i,j} - I^{+2}_{i,j}\right) + I^{+2}_{i,j},\ if\ \ I^0_{i,j} < 127.5 \\ t_{I^0_{i,j}} \cdot \left(I^0_{i,j} - I^{-2}_{i,j}\right) + I^{-2}_{i,j},\ if\ \ I^0_{i,j} > 127.5 \end{cases} \quad (4)$$

Examples of tone mapping and comparison with traditional approach are shown in Figures 6 and 7. Sign detection is based on color segmentation in RGB color space. Result of segmentation is a binary mask which obtained as follows:

$$B_{i,j} = \begin{cases} 1,\ if\ \dfrac{\max\left(I^{CG}_{i,j}, I^{CB}_{i,j}\right) - I^{CR}_{i,j}}{I^{CR}_{i,j}} > 0.4 \\ 0, otherwise \end{cases} \quad (5)$$

Next step of proposed algorithm is grouping connecting components. In this step we extract sign candidates from the binary masks obtained on the previous step. Two 8-connected pixels $B_{p,q}$ and $B_{m,n}$ of bitmap are connected if the distance in color space defined as $d(A,B) = \sqrt{\left(A^R - B^R\right)^2 + \left(A^G - B^G\right)^2 + \left(A^B - B^B\right)^2}$ is less than $\gamma = 0.3 \cdot \min\left(I_{p,q}, I_{m,n}\right)$.

(a) (b)

Figure 6: Tone mapping example

(a) (b) (c)

Figure 7: Tone mapping vs traditional approach: (a) input image; (b) candidates detected by proposed method; (c) candidates, detected by non-HDR approach

During this step main geometrical properties of a sign candidate such as area, proportions and bounding box are computed. These parameters are used for filtering operation to eliminate sign-like objects from candidate list. The filtering operation is based on both color and geometrical properties of signs. The most important sign-parameters are grouped in Table 1.

Filtering parameter	Description	Values
Minimal size	Filters candidates which are too small to be a road sign or too small for recognition.	100x50 pixels
Minimal area		1500 pixels
Minimal height to width ratio	Filters candidates which have proportions impossible for road signs. Height and width of bounding box is used in this step.	0.4
Maximal height to width ratio		1.8
Minimal area of sign candidate to area of its bounding box ratio	Filters candidates which are to thin or too thick. Information signs are rectangles; hence the area should be almost same to bounding box area.	0.6
Maximal area of sign candidate to area of its bounding box ratio		1
Rectangularity check	Information signs have rectangular shape. It can be checked by calculating difference between area of bounding box and area of sign candidate.	10% of candidate area
Number of white pixels inside the candidate	White color used for symbols inside the information signs. Usually area of these symbols is around 25-30% of total sign area.	10-50%

Table 1: Sign properties which were used for filtering

Next step is candidate extraction. A combined image is good for detection, as long as it contains more details compared with a normal image. It is generated from several images taken from different positions. Even when the cameras are located very close to each other there will still be a small difference in positions of images. Hence, the combined image will contain distortions. These distortions are not critical for the detection step, while it may cause serious problems in recognition process. This problem is illustrated in Fig.8. To overcome this problem the following strategy was used: first, we detect sign candidates using a combined image. Then we extract a corresponding region from the image which has the best local exposure in the region where the signs candidate is located. Suppose the k-th candidate has width kw, height kh and its topmost left point has coordinates (kx,ky). The best exposure for this candidate will be selected according to the following equations:

$$R(k) = \arg\left(\max_{j=\{-2,0,+2\}} D(j,k) \right) \qquad (6)$$

where $D(i,k)$ is level of details in the region, defined as

$$D(j,k) = \frac{1}{255} \sum_{x=kx}^{kx+kw-1} \sum_{y=ky}^{ky+kh-1} \max\left(\left| I^{j}_{x+1,y} - I^{j}_{x,y} \right|, \left| I^{j}_{x,y+1} - I^{j}_{x,y} \right| \right) \qquad (7)$$

Part of the image with best exposure selected using Equation (6) which corresponds to the sign candidate will be used in further operations instead of the combined image. An example of this process is shown in Figure 8. More details of this approach could be found in [Rovid, 07] and [Vavilin, 08].

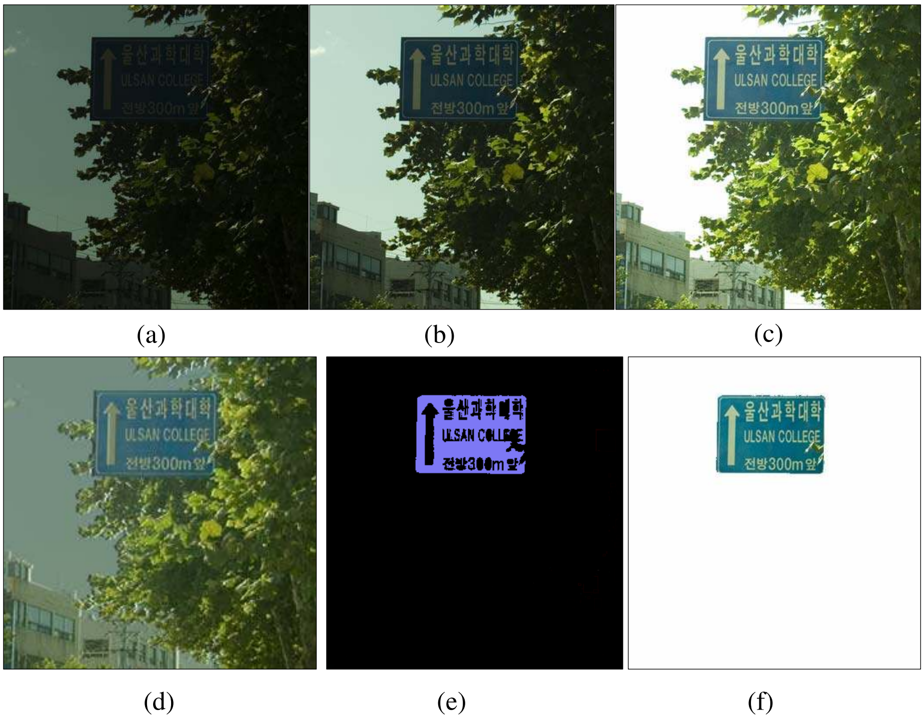


Figure 8: Selecting of best candidate exposure (a)~(c) crop of differently exposed input images, (d) image, combined using proposed algorithm, (e) detected candidate, (f) selected exposure

4 Sign recognition

The main difficulty in road guidance signs recognition is their complex structure. Unlike warning or restricting signs, the structure of guidance signs is not strictly defined. As far as guidance signs are used to provide drivers with information about crossroads, the recognition of guidance sign means an extraction of all possible moving directions with descriptions. Another type of information provided by a guidance sign is possible paths to pass the crossroad. This information is critical for the crossing junction with complex structure (Figure 4.(c)~(f)). Typically, a driver has very little time to choose the correct destination and crossing path. The proposed method provides the driver with information about possible directions with corresponding paths. Main steps of recognition algorithm are shown in Figure 9.

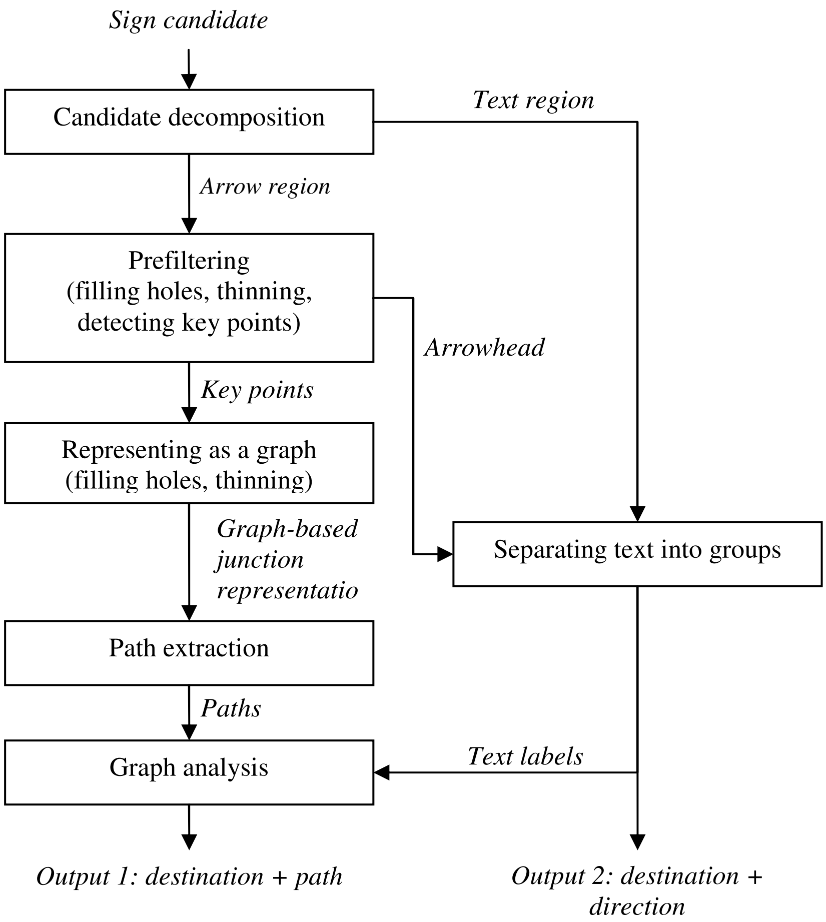
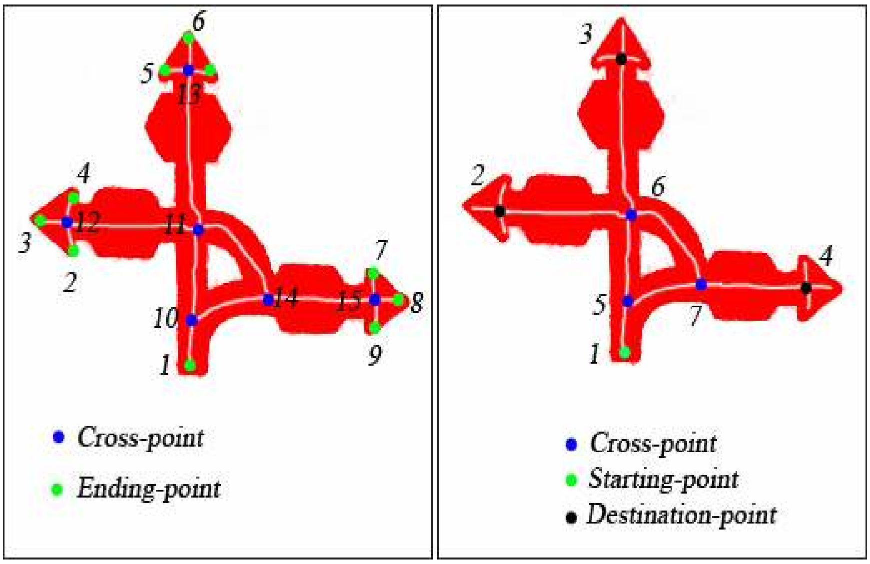


Figure 9: Recognition algorithm

After extracting signs candidates from the image each of them is analyzed in order to separate symbols from background. Then all symbols are classified into one of three principle components as described in Section 2. The most important problem in this step is the correct detection of an arrow region. Typically, an arrow region is the biggest symbol. However, in some cases a sign can include more than one arrow region. We choose all symbol clusters with an area greater than 90% of the maximum area. If more than 5 clusters were selected then symbol information was extracted incorrectly or sign contains no arrow region. After an extraction of the arrow region it should be transformed into a set of points. For that purpose a thinning operation is used. Points which are important for further processing are shown in Figure 10. These points correspond to junction entrances, exits and crossings of the road. Using information about positions of arrowheads we can separate text labels into several groups by the number of arrowheads. Each text group describes one of the possible directions. Then the arrow region is mapped into a graph with a number of vertices equal to the number of the key points. An analysis of this graph gives us information about possible paths. This process was described in [Vavilin, 07].



(a) (b)

Figure 10: Key points of arrow region

Using information about key-points (Figure 10.b) arrow is represented as two graphs. First graph (*C*) is a connectivity graph which shows connections between key points. Second graph (*A*) shows angles between key points. For the arrow region shown in Fig.10 graphs can be presented as follows:

$C =$

	1	2	3	4	5	6	7
1	0	0	0	0	1	0	0
2	0	0	0	0	0	1	0
3	0	0	0	0	0	1	0
4	0	0	0	0	0	0	1
5	1	0	0	0	0	1	1
6	0	1	1	0	1	0	1
7	0	0	0	1	1	1	0

$A =$

	1	2	3	4	5	6	7
1					90		
2						0	
3						270	
4							180
5	270					90	10
6		180	90		270		45
7				0	190	135	315

Then we extract paths from the destination points to the starting point using the following block-diagram.

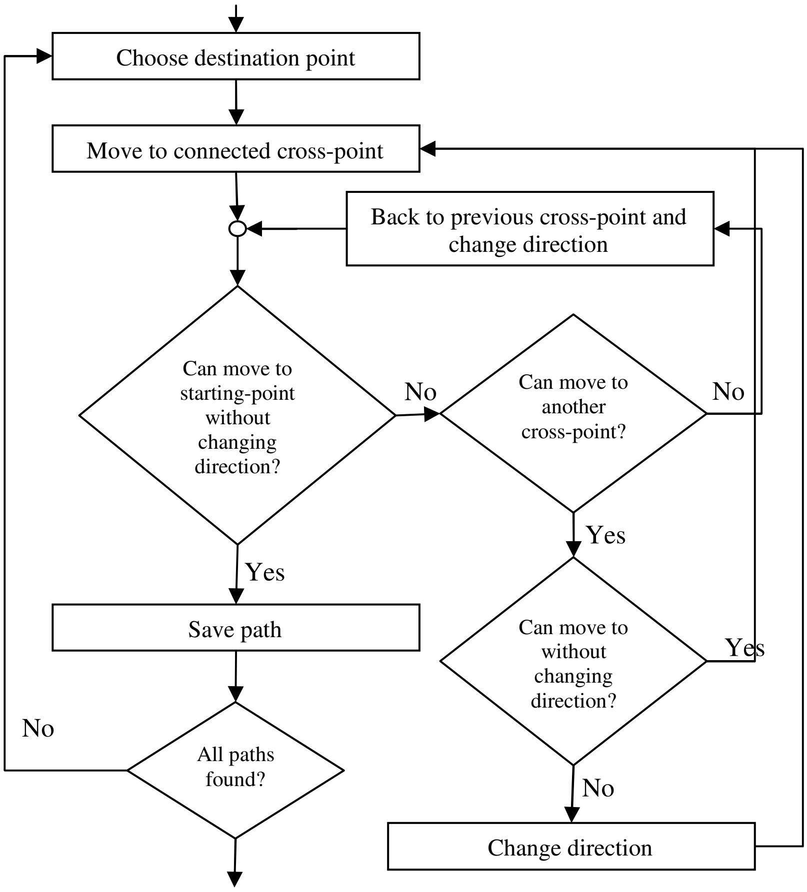


Figure 11: Path extraction algorithm

5 Experimental results

All experiments were done on Pentium-IV 3.2GHz with 1 GB RAM under Borland builder environment. Images of size 640x480 were used. Each scene represented by 3 images was taken with different exposures. Exposure value was changed with step of 2 f-stops. Thus, images were taken with EV-2, EV0 and EV+2. As reference methods for sign detection, we used the detection method proposed in previous works [Vavilin, 07] and the shape-based detection based on [Wu, 07]. Images with normal exposure (EV0) were used as input images for traditional detection method. Two types of tests were done.

First type was used to evaluate detection process. For these tests all images were separated into 3 groups by lighting conditions. The first group (Figure 12.a) contains images taken in good environmental conditions. These images are easy for detection, thus, the detection rate is 100% for both HDR and non-HDR approach and 96% for shape based detection (see Table 2). Images of second group (Figure 12.b) were taken in bad weather conditions such as rain or fog or during evenings. These conditions complicate the detection process; however, the detection rate is still higher than 90%. Nevertheless, the detection method based on single image shows worst results for this image group. The third image group (Figure 12.c) contains images of scenes with high dynamic range. These images are the most challenging case for non-HDR approaches. In addition, even if the sign was correctly detected, further processing may be difficult due to the low quality of the extracted candidate. This problem is illustrated by Figure 13. Detection results are shown in Table 2.

Image group	**Number of signs**	**Shape-based detection [Wu 07]**	**Color-based (non HDR)**	**Detected using HDR image**
1	25	24(96%)	25 (100%)	25 (100%)
2	20	19(95%)	18 (90%)	19 (95%)
3	15	10(66%)	6 (40%)	15 (100%)
Total	60	53(88%)	49 (81%)	59 (98%)

Table 2: Detection results.

A second type of tests was applied to the recognition stage. For these tests images were divided into 6 groups according to complexity of presented signs (see Section 2 for more details). In addition, 30 synthetic images with different types of arrow regions were added. The recognition process provides two types of results. First output contains pairs “direction-description”. The second type of results contains pairs “direction – path”. The first result is correct if the directions presented in the sign were correctly extracted and matched to the corresponding text description. The second result is correct if the path from the junction entrance to the destination corresponds to the correct junction crossing path according to the road law for all extracted paths. The results of recognition are summarized in Table 3. In this table columns entitled from 1 to 6 show the complexity of the arrow region according to the proposed classification. Total values were computed without synthetic images.

Examples of successful recognition for the main types of signs are shown in Figure 14. Figure 15 shows result of failed recognition caused by an incorrect extraction of an arrow region due to the small size of the candidate (100x51). This size is a minimal acceptable size according to filtering parameters; however, it was too small for correct recognition.

(a)

(b)

(c)

Figure 12: Examples of input images of first (a) second (b) and third (c) groups (see text for details)

(a) (b)

Figure 13: Sign candidates extracted from third image group using traditional (a) and proposed (b) methods.

Figure 14: Recognition examples. Output consists of extracted candidate (a), direction label (b), arrowheads(c), and extracted paths (d)

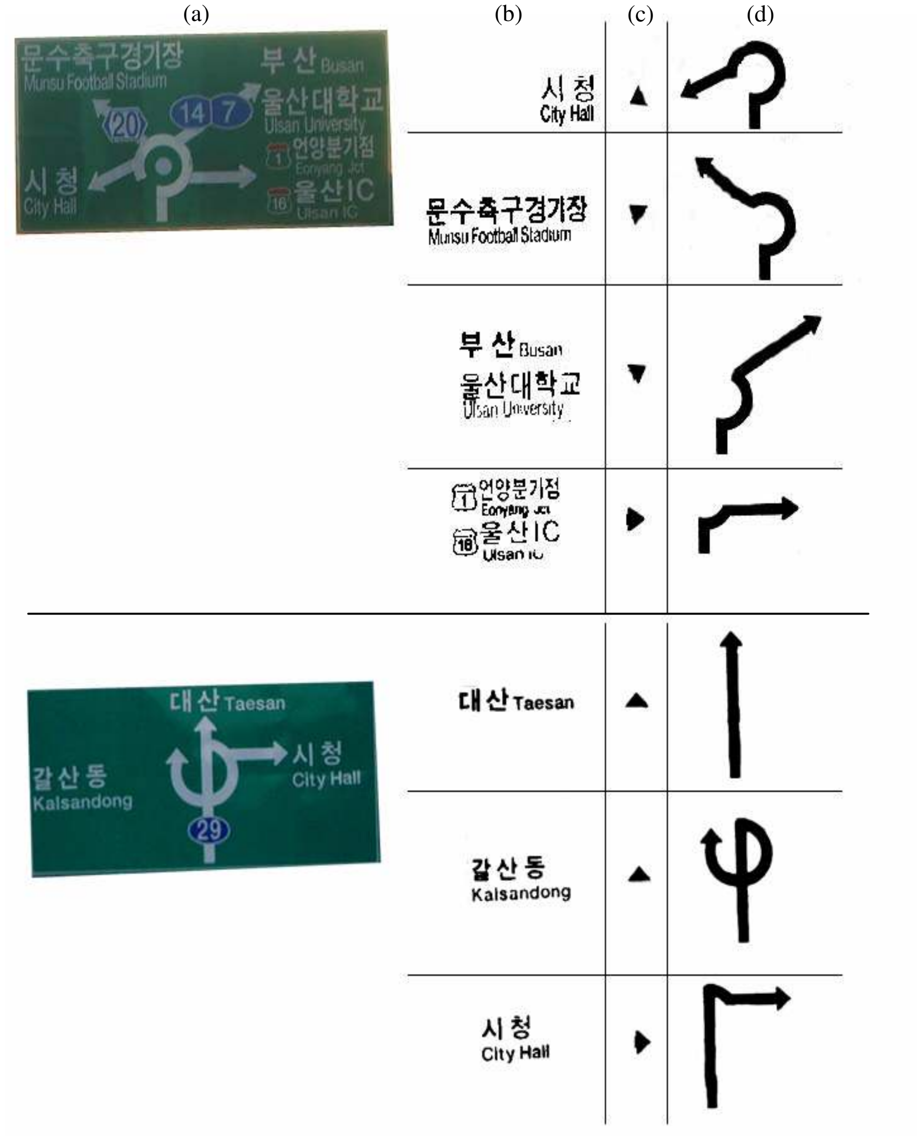


Figure 14 (continued): Recognition examples. Output consists of extracted candidate (a), direction label (b), arrowheads(c), and extracted paths (d)

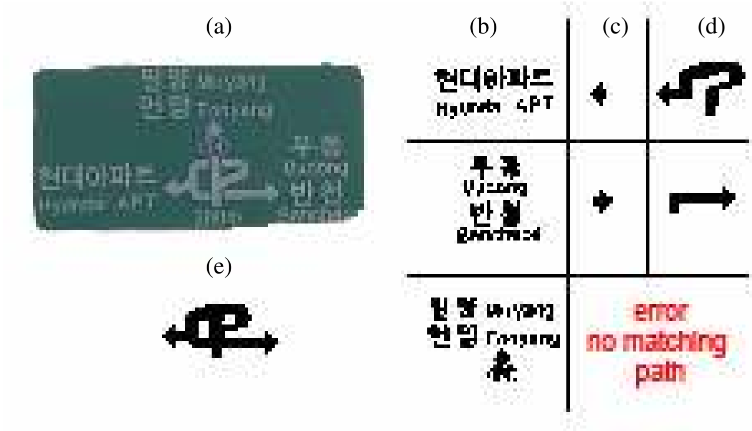


Figure 15: Failed recognition example. Output consists of extracted candidate (a), direction label (b), arrowheads(c), and extracted paths (d)

Image group	Synthetic	1	2	3	4	5	6	Total
Number of signs	30	10	10	10	10	10	10	60
Direction extraction	30	10	10	10	10	9	9	58 (96%)
Path extraction	30	10	10	10	9	8	9	56 (93%)

Table 3: Recognition results

Average computational time for 50 test images is presented in Table 4.

Stage	Average time (sec)	Accumulated time (sec)
Tone mapping	0,029	0.029
Color segmentation	0.028	0.057
Filtering and candidate extraction	0.018	0.075
Detecting arrow region and converting it into graph	0.030	0.105
Recognition and result verification	0.010	0.115
Path extraction	0.020	0.135

Table 4: Average computational time

6 Conclusions

Experimental results indicate that the proposed algorithm allows more effective detection of information sign compared to the traditional color-based approach. Overall detection rates for the proposed method, color-based detection from single image and shape-based detection are 98% and 81% and 88% respectively. The proposed method is especially effective for the scenes where the dynamic range of a CCD camera is not enough. The detection rate of our method is 100% against 40% for the traditional approach. Shape-based detection provides better results compared with single image detection (66%), but detected signs are difficult to recognize due to their low intensity. Another problem of the shape-based detection is a high false-positive detection rate. Recognition method shows a 96% rate for the "destination-direction" extraction and 93% for the path extraction. These values can be increased by using better thinning technique and a more precise symbol extraction method.

Average computational time for an image with one sign is less than 0.2 sec. This enables the proposed techniques to be used in a real-time application. This time may be improved by several optimizations such as using lookup tables instead of recalculating equation (2) and grouping tone mapping with color segmentation.

In future works, we would like to improve the performance of the algorithm by adding text recognition stage and recovering geometrical aberrations such as skew, rotation, etc. Text recognition enables additional result verification for label segmentation. This verification is based on the idea that Korean and English labels must be similar for one direction.

Acknowledgements

The authors would like to thank to Ulsan Metropolitan City and MEST and MKE of Korean Government which partly supported this research through the NARC and post BK21 project at University of Ulsan.

References

[Azami, 96] Azami, S, Katahara, S., Aoki,M.: Route Guidance Sign Identification Using 2-D Structural Description. Proceedings of the IEEE Intelligent Vehicles Symposium, pp. 153–158, 1996.

[Azami, 96]Azami, S., Aoki, M.: Route Guidance Sign Recognition. Proceedings of the IEEE Intelligent Vehicles Symposium, pp. 338–343, 1995.

[Ferwerda, 96] Ferwerda, J., Pattanaik. S.N., Shirley, P., Greenberg, D.P.: A Model of Visual Adaptation for Realistic Image Synthesis, Proceedings of SIGGRAPH 1996, pp. 249-258. 1996.

[Kato, 02] Kato, T., Kobayasi, A., Hase, H., Yoneda, M.: An Experimental Consideration for Road Guide Sign Understanding. In: ITS. Intelligent Transportation System,. IEEE, pp. 268–273, Singapore, 2002.

[Lee, 03] Lee, J-H., Jo, K-H.: Traffic sign recognition by division of characters and symbols. In: The 7th Korea-Russia International Symposium, vol.2, pp. 324-328, 2003

[Lim, 08] Lim, H-C., Vavilin, A., and Kang-Hyun Jo, K-H.: “An Efficient Method of Road Sign Recognition”, CASS2008, Seoul, Korea, pp. 277-280 2008.

[Mertens, 07] Tom Mertens Jan Kautz Frank Van Reeth, “Exposure Fusion”, Proceedings of 15th Pacific Conference on Computer Graphics and Applications, Maui, Hawaii • October 29-30 & November 1-2, 2007.

[Rovid, 07] Rovid, A.; Varkonyi-Koczy, A.R.; Hashimoto, T.; Balogh, S.; Shimodaira, Y.: Gradient Based Synthesized Multiple Exposure Time HDR Image. Instrumentation and Measurement Technology Conference Proceedings, 2007 IEEE, pp.:1-6, 2007.

[Várkonyi-Kóczy, 08] Annamária R. Várkonyi-Kóczy, András Rövid, and Takeshi Hashimoto: “Gradient-Based Synthesized Multiple Exposure Time Color HDR Image”, IEEE Transactions on Instrumentation and Measurement, vol.57 No8, Aug 2008.

[Vavilin, 06] Vaviloin, A., Jo, K-H.: Road Guidance Sign Recognition in Urban Areas by Structure, Proceedings of the 1st International Forum On Strategic technology, pp293-296, Korea, 2006.

[Vavilin, 07] Vavilin, A., and Jo, K-H.: Automated Junction Structure Recognition from Road Guidance Signs. In: ICIC 2007, LNCS, pp. Qingdao, China, 2007.

[Vavilin, 08] Vavilin, A., Lim, H-C., Jo, K-H.: Robust Road Sign Detection Using Differently Exposed Images. In: 14th Korea-Japan Joint Workshop on Frontiers of Computer Vision, pp.445-450, Beppu, Oita, Japan, 2008.

[Vavilin, 08] Vavilin A. and Jo, K-H.: “Recursive HDR Image Generation from Differently Exposed Images”, Graphcon 2008, Moscow, Russia 2008.

[Wu, 07] Gang Wu; Weijie Liu; Xiaohui Xie; Qiang Wei.: A Shape Detection Method Based on the Radial Symmetry Nature and Direction-Discriminated Voting. In: International Conference on Image Processing 2007, vol. 6, pp. 169-172 2007.