Estimating the Vanishing Point of a Sidewalk

Seung-Hyuk Jeon
Korea University of Science and Technology (UST)
Daejeon, South Korea
+82-42-860-4980
soulhyuk8128@etri.re.kr

Yun-koo Chung
Electronics and Telecommunications Research Institute (ETRI),
Daejeon, South Korea
+82-42-860-5032
ykchung@etri.re.kr

Ho-sub Yoon
Electronics and Telecommunications Research Institute (ETRI),
Daejeon, South Korea
+82-42-860-5233
yoonhs@etri.re.kr

ABSTRACT

In this paper, we propose an algorithm estimating the vanishing point (VP) of a sidewalk in a man-made environment from a single image. For finding the VP, the lines in an image are efficiently exploited as a clue because the projections of parallel lines in an image intersect at a VP. However, there are too many noises disturbing line detection in natural scene images, for example trees, pedestrian and shadows. Thus, we suggest a noise reduction technique called orientation consistency pass filtering (OCPF) for improving line detection performance. An edge orientation at the window center of OCPF is compared to its neighbor edge orientations for calculating orientation difference. The center pixel is removed if the difference of edge orientations greater than a threshold, and preserved otherwise. In addition, we suggest a novel vanishing point detection method using edge orientation voting (EOV), which predict VP position accurately. The lines filtered by OCPF can generate the VP candidates using bottom-up extended lines. The VP candidates receive supports from all edges below the currently inspected VP candidate. The most supported VP candidate is selected as dominant VP. This proposed method was implemented and tested on 600 sidewalk image database that has 640x320 resolutions. 74.3 % of the test sidewalk images are in range from 0 pixels to 20 pixels from manually marked VP.

Keywords

Sidewalk Recognition, Line Detection, Vanishing Point Estimation, Natural Scene Images.

1. INTRODUCTION

Our goal is to estimate the vanishing point (VP) of a sidewalk using an single image with unknown camera parameters to obtain navigate information. The images we would deal with are taken on the sidewalk. They have a sidewalk direction around image center. So, they are preserved perspective of parallel lines. Most sidewalk images have so complicated scenes we need to simplify as essential elements like lines and segmented region as possible. We use only lines as essential elements. Figure 1 shows that only the direction of green lines pass near the vanishing point in perspective. Other two color lines are roughly parallel to same color lines and have VPs at the infinite because they are roughly parallel to the image plane. A cue for the sidewalk direction is the VP at the end of the sidewalk, noted by the green dot (see Figure 1). For finding the vanishing point, the lines in an image are extracted and efficiently exploited. Our contribution is composed twofold. (1) a noise reduction technique considering variation of edge orientation called orientation consistency pass filtering (OCPF). It

Figure 1. An sidewalk image; Each color correspond to divided angle of lines in the image. Red and green , yellow are assigned if angles are 25°~ 165°, 80°~ 110°,165°~ 0°~ 25° respectively**.**

improves line detection performance. (2) A VP voting method called edge orientation voting (EOV) that is supported by edges satisfying orientation criteria.

2. RELATED WORK

Many researches on line detection and VP estimation have been studied over the past decades. Generally, VP estimation is based on lines clustering. So line detection is carried out well in advance.

One of well-known line detection methods is progressive probabilistic hough transform (PPHT) [Mat99]. Though it is improved version of hough transform, high edge density regions such as foliage of trees, are regarded as lines because without using edge direction information. The chain code of edges can be used sidewalk detection[Yu08] in terms of connections to inclined column. The drawback of chain code method is losing information horizontal and vertical edges because they are not connected diagonally. Horizontal lines supporting VP consist of horizontal edges. Depending on viewpoint, boundary of sidewalk have many vertical edges. It also tends to be generated in noise data such as tree branch due to considering edge connections only in image axis. Line segment detector (LSD)[von10] considers edge orientation along with edge magnitude for line detection. Its main algorithm is region growing based false detection control. The region growing is detecting lines well, but in some case, the OCPF is superb. Details are described in section 3. Line detection.

VP estimation is useful for many applications ranging from navigation assistance to understanding geometric context of a scene [Tar09]. Most VP estimation methods based on visual scenes are under the Manhattan world assumption [Cou03]. For finding the VP, a number of lines are clustered by its parallel principle and orthogonal principle [Mir11]. But, these methods are targeting the Manhattan world such as buildings have hexahedron shape and simple textures. Because we already know that the detection target is a sidewalk, it is possible to estimate VP without Manhattan assumption. In brief, we find maximum crossing point of extracted lines. Gabor texture orientation and semicircle voting region should be used for VP estimation of road [Hui09]. Clearly, it is good method for off-road, but there are so many occlusion of VP such as pedestrian or a case VP lie hidden over horizontal line in a sidewalk. Another VP estimating method is random sample and consensus (RANSAC) using distance from VP candidate to lines [Geo11]. We improve this method separating inliers and outliers. Details are described in section 4. Vanishing point estimation.

3. LINE DETECTION

The first step of a proposed line detection method is extracting edges from image using canny edge detector. We use similar method with [Geo11] for extracting canny edge detection (see Figure 2).

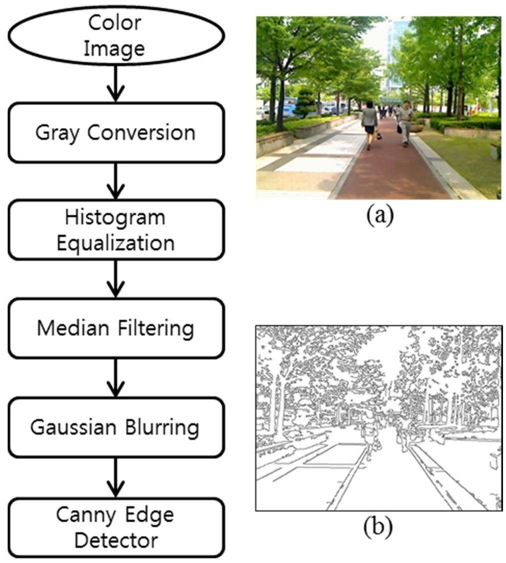


Figure 2. Image pre-processing and canny edge detection; (a) Original color image, (b) Canny edge image.

Although many noises are reduced by using above pre-processes and canny edge detection, much remains to be reduced. We need segmented lines those element pixels have similar directions in the image. Meanwhile, non-lines such as tree branch and pedestrians would be regarded as noise. According to our observation, the sidewalk have simple repeated pattern with tetragonal boundary along the length and width. In addition, most man-made structures such as buildings, boundary stones and bus stations are parallel to sidewalk. On the contrary, most non sidewalk objects have complex disorderly patterns. For this reason, if an edge has neighbors inconsistent edge orientation then erased else preserved. We designed OCPF described in Algorithm 1, Figure 3 and Figure4.

Input : I_m is the canny edge gradient magnitude image.
I_a is the canny edge gradient angle image.

Output: I_f is the OCPF image.

For x = 1, 2,…, N_x for x positions

For y = 1, 2,…, N_y for y positions

if $I_m(x, y) > 0$ then

$A_v(x, y) \leftarrow \frac{1}{n(W)} \sum_{W \ni (x_w, y_w)} |I_a(x, y) - I_a(x_w, y_w)|$

```
5      if A_v < Ø then
6          I_f(x, y) ← Line mark
7      else
8          I_f(x, y) ← Null
9      end
10   end
11 end
12 end
```

Algorithm 1. OCPF

N_x is the number of pixels along the x-axis. N_y is the number of pixels along the y-axis. The W is defined as A set $= \{(a,b) | (a,b) \in R_\theta,\ I_m(a,b) > 0,\ (a,b) \neq (x,y)\}$. $n(W)$ is the number of elements in W, R_θ is the rotated rectangle (see Figure 3). The 1 and 2 of Algorithm 1 iterate the computation of angle variation (A_v) along the x- and y-axis, respectively. The 3 of Algorithm 1 check if the position (x, y) is an edge in the I_m. The 4 of Algorithm 1 calculates the average of the acute angles between $I_a(x,y)$ and $I_a(x_w,y_w)$ of edges in R_θ except for a centroid position of R_θ. The 5 of Algorithm 1 determines whether a current center position is a line pixel or not by comparing to threshold angle value Ø .

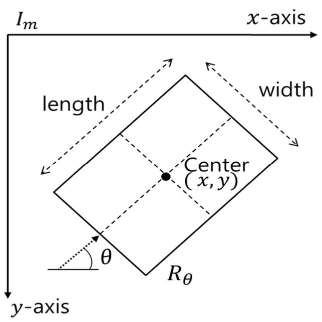


Figure 3. Rotated rectangle

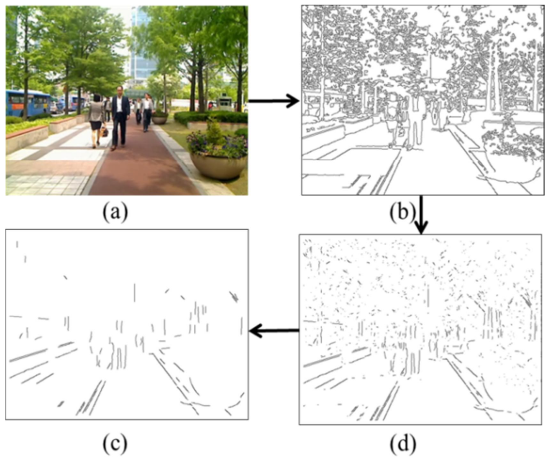


Figure 4. OCPF filtering procedure; (a) Original color image, (b) Canny edge image, (c) Filtered short length out image, (d) OCPF image.

After Algorithm 1, If the number of connected pixels is greater than a threshold value then we take the connected pixels as line (Figure 4. c). Rotated rectangle R_θ of Figure 3 is designed to adjust manually its length and its width in order to reduce impact of noise. The θ is automatically assigned from sobel edge of canny edge detector. Figure 4 shows the procedural result of OCPF. Algorithm 1 is correspond to the procedure b to d of Figure 4. We regard segments shorter than length of R_θ as noise in d of Figure 4 because length of R_θ represent the minimum length of lines. So we use the threshold T in proportion to the length of R_θ. A strong point of the OCPF is fewer responses to noise that has various orientation region compare to other line detection methods in natural scene image (see Figure 5). After performing PPHT and LSD, the shadows of the trees remain in the images.

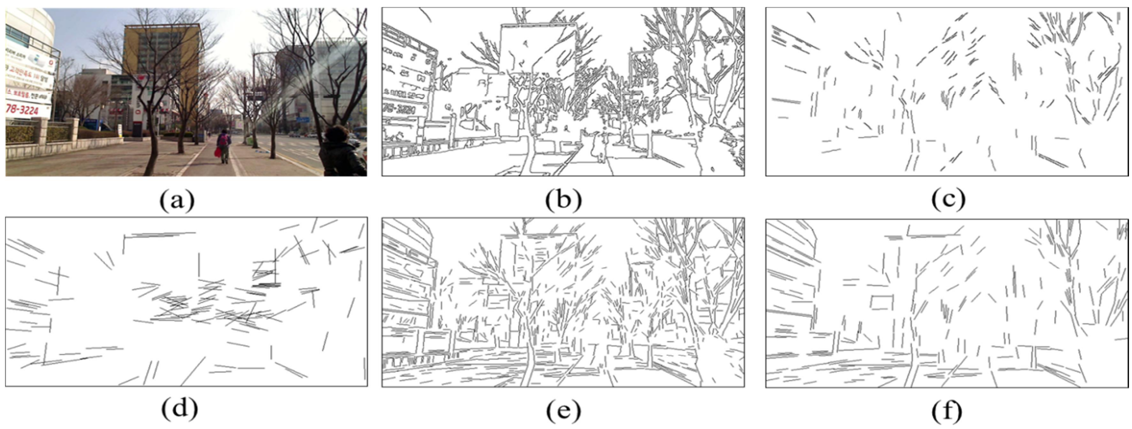


Figure 5. A comparison of line detection methods; (a) Original color image (b) Canny edge image (c) Short length filtered out OCPF image (d)PPHT image (e) LSD image (f) Filtered short length out LSD image.

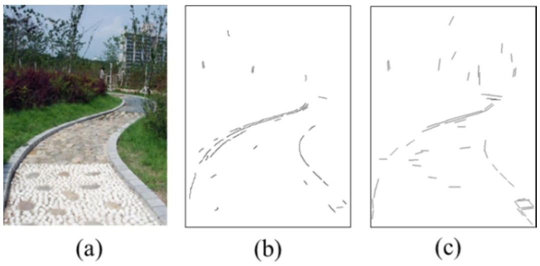
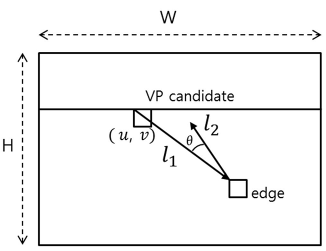


Figure 6. A comparison between OCPF and LSD; (a) Original image (b) OCPF image (c) LSD image

However, OCPF image not only reduces noise, but also have the most lines of the boundary of the sidewalk because OCPF cut noise lines into short lines to be erased. The other strong point remains curve as long line (see Figure 6). It can be used for the estimation of the shape of sidewalk.

4. VANISHING POINT ESTIMATION

For finding candidate VP that is an intersection point of two lines, we use bottom-up extended lines linearly approximated line segments are drawn from bottom of the line to top of the image, increasing pixel value 1 (c of Figure 8). The pixels on the bottom-up extended lines greater than 2 are regarded as VP candidate points because these points are an intersection of two more lines. We can choose one of the VP candidates as dominant VP using the function of Figure 7. It considers acute angle between l_1 (VP candidate to an edge) and l_2 (direction of edge), if it is smaller than threshold T then assign voting value y-position of an edge (inlier) else voting value 0 (outlier). This method applies all candidate VPs voted by edges positioned under v and accumulate the voting values into VP estimation dense image (I_v, e of Figure 8). Above mentioned method is called edge orientation voting (EOV).



$$I_v(u, v) = \sum_{x=0}^{W} \sum_{y=v+1}^{H} f(x,y)$$

$$where\ f(x,y) = \begin{cases} y, & \theta < T \\ 0, & \theta > T \end{cases}, \ \theta = I_a(x,y)$$

Figure 7. Edge orientation voting (EOV).

The y-position of an edge represents voting weight. The more close to bottom of image, the more voting weight because most images have sidewalk plan at the bottom of image. And then we choose maximum value as VP in I_v (a green dot in f of Figure 8). Figure 8 shows total procedure of VP estimation. The d of Figure 8 shows efficiency of OCPF again. The bottom-up extended line image reduces the time computing a point of intersection of two lines. As inspecting candidate VP, we can reduce time complexity for searching dominant VP. The white pixels of Figure 8 e are the candidate VPs.

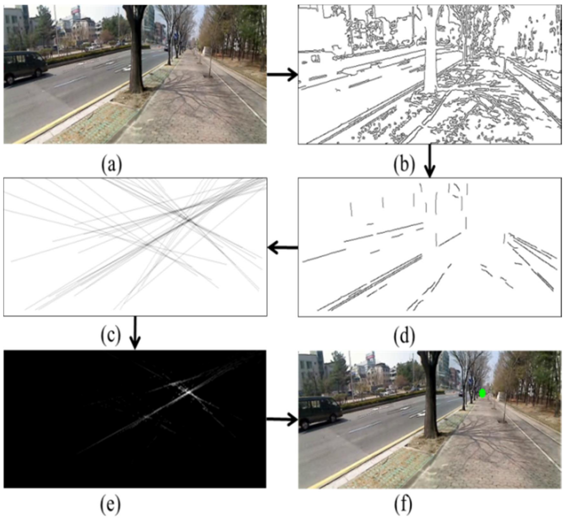


Figure 8. Vanishing point estimation; (a) Original image (b) Canny edge image (c) Bottom-up extended line image (d) OCPF image (e) VP estimation dense image (f) VP marked image(a green dot).

5. EXPERIMENT AND RESULT

VP estimation is tested on 600 sidewalk images taken in the Daejeon, South Korea in 12 different place each of 50 images along the sidewalk. These sidewalk images exhibit large variations in color, texture, illumination and amvient enviroment. All images normalized to the same size with height of 320 and width of 640 due to a wide view angle in order to imatate human. To assess VP localization, all the images are manually marked by those who are trained to know the VP concept. We compared the EOV method to other two methods. One of them is Locally adaptive soft voting(LASV) method [Hui09]. LASV method uses convolved responses of 5 scales and 36 orientation garbor kernels for estimating texture orientations. And for VP estimation, set half-disk voting region centered at overall VP candidates.

Another method is eyeDog [Geo11] which extracts lines usings canny edge detector and hough transform.

The intersection points of extracted lines are VP candidates. The VP candidates are selected randomoly and then calculating distances form it to lines. A VP that has most lines closer than a threshold is elected. This method is the random sample and consensus (RANSAC). The comparisions of VP estimation accuracy are shown in Figure 9.

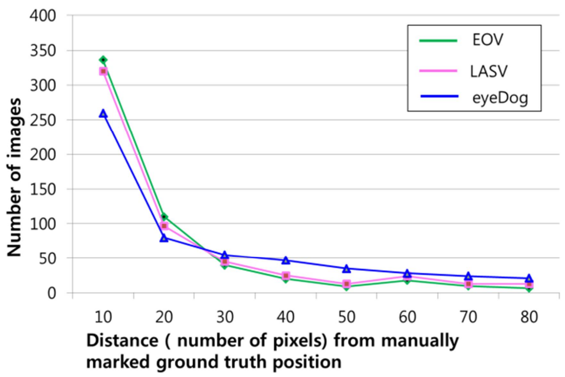


Figure 9. VP estimation accuracy.

EOV shows the best performance in our database. 56 %, 53% and 43% of the test sidewalk images are in range from 0 pixels to 10 pixels EOV, LASV and eyeDog respectively. 74.3 %, 70% and 50% of the test sidewalk images are in range from 0 pixels to 20 pixels EOV, LASV and eyeDog respectively. Figure 10 shows some VP estimation result of EOV. We can verify the robustness to noise such as shadows, tree branch and pedestrians.

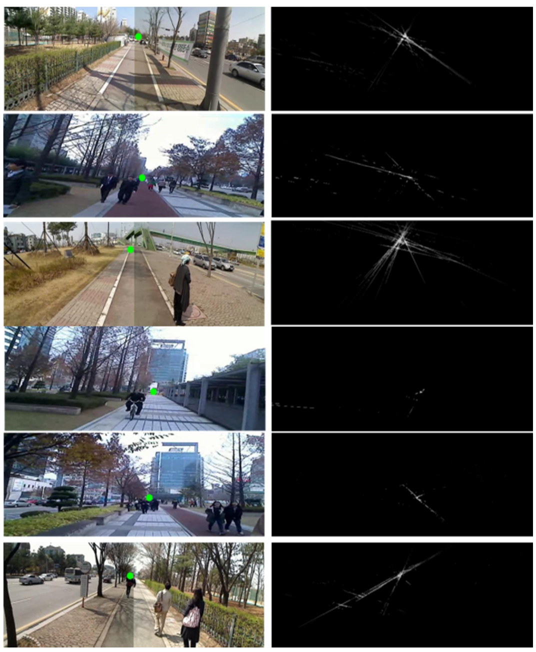


Figure 10. Some results of EOV in our database.

Figure 11 show that how the EOV works if there are different light condition 1~3 row of Figure 11 are in the evening, 4 row is in the morning, 5 row is in snowy weather and 6 row is a fog-shrouded sidewalk image. The column 1~3 of Figure 11 are VP estimation denoted by green dot, OCPF images and VP dense images respectively.

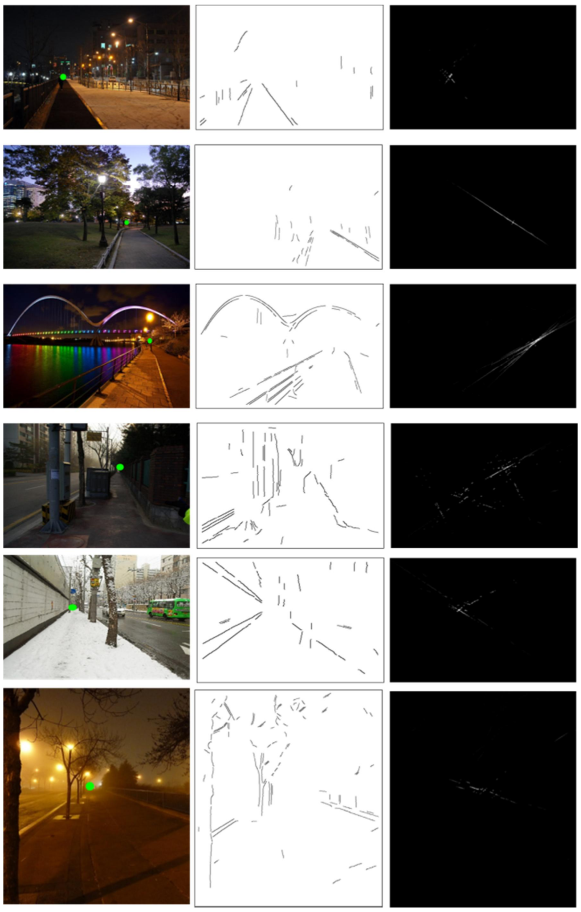
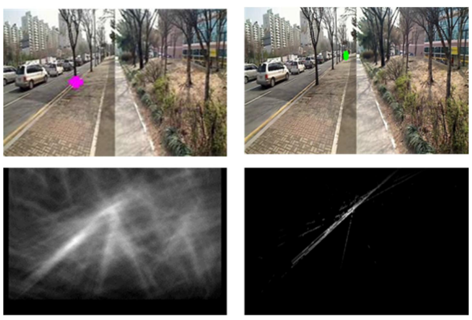


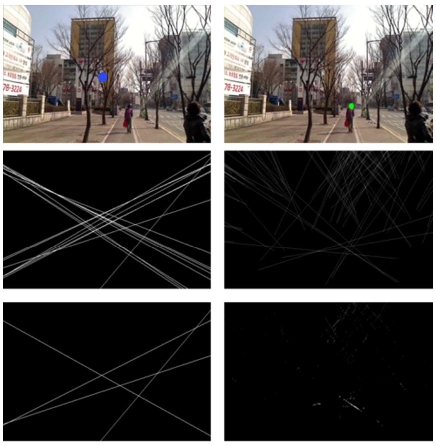
Figure 11. Some result of EOV in various light condition.



(a) (b)

Figure 12. A comparison of LASV (a) and EOV (b).

Figure 12 compares the VP estimation result using LASV and EOV. It reveals that more smaller region inspected in EOV and LASV`s false VP occurs in narrowly crossed two lines. But EOV is supported from different angle lines and detect right VP blurredly because OCPF reduce the impact of narrowly crossed two lines (erasing lines between two closed lines).



(a) (b)

Figure 13. A comparison of eyeDog (a) and EOV (b).

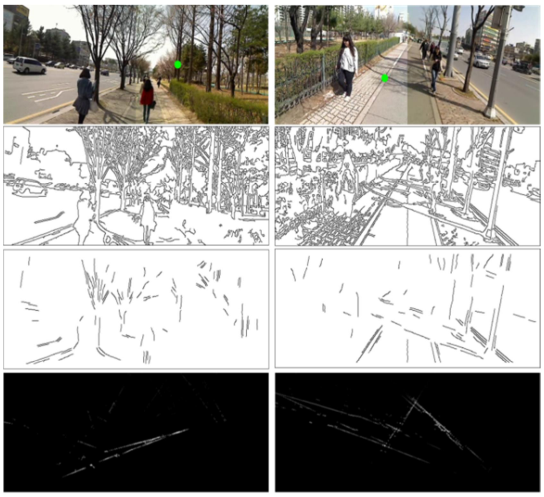
Figure 13 shows that the eyeDog method has wrong lines because the noises , tree branches, are not removed appropriately. On the contrary, EOV has an accurate VP but it has many tree branches nevertheless. Because EOV scheme concentrate on the lines close to the bottom of the image.

6. SUMMARY AND CONCLUSION

A novel framework for estimating VP of the sidewalk from one single image is proposed based on edge orientation using OCPF and EOV. The OCPF shows better results than PPHT and LSD in sidewalk images in terms of noise reduction and curve lines detection. The EOV shows better results than LASV and eyeDog in our database. Figure 13 show VP estimation failed case. Canny edge detector is not working well because of blurred edges (a). Another case is that the very dark shadow on the sidewalk goes wrong lines (b). These are future works.

7. ACKNOWLEDGMENTS

This research was supported by the Converging Research Center Program through the Ministry of Science, ICT and Future Planning, Korea (2013K000329)



(a) (b)

Figure 13. A VP estimation failed case.

8. REFERENCES

[Mat99] Matas, J. and Kittler, J., Progressive probabilistic hough transform for line detection, Computer Vision and Pattern Recognition, 1063-6919, 1999.

[Yu08] Yu, J. and Han, Y. and Hahn, S, An efficient extraction of on-road object and lane information using representation method, Signal Image Technology and Internet Based Systems, pp.327-332, 2008.

[von10] von Gioi, R.G. jakubowicz, J. Morel, J.-M. Randal, G, LSD: A fast line segment detector with a false detection control, Pattern Analysis and Machine Intelligence, pp.722-732 , 2010.

[Tar09] Tardif, J.P, Non-iterative approach for fast and accurate vanishing point detection. International Conference on Computer Vision, pp1250-1257, 2009.

[Cou03] Coughlan, J. and Yuille, A. Manhattan world: Orientation and outlier detection by bayesian inference. Neural Computation, 15(5):1063.1088, 2003.

[Mir11] Miraei, F.M., Vanishing Points Estimation and Line Classification in a Manhattan World, International Conference on Computer Vision, pp.2454-2461, 2011.

[Hui09] Hui Kong, Vanishing point detection for road detection, Computer Vision and Pattern Recognition, pp.96-103, 2009.

[Geo11] Georgios G., eyeDog : An Assistive-Guide Robot for the Visually Impaired , Proceedings of the 4th International Conference on Pervasive echnologies Related to Assistive Environments, 2011.