Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Visual Campus Road Detection for an UGV using Fast Scene Segmentation and Rapid Vanishing Point Estimation

Min Zhu*, Yisha Liu, Yan Zhuang* and Huosheng Hu*****

** Research Center of Information and Control, Dalian University of Technology, Dalian, China, (e-mail: zhuang@dlut.edu.cn)*
*** School of Information Science and Technology, Dalian Maritime University, Dalian 116026, China. (e-mail: liuyisha@dlut.edu.cn)*
****School of Computer Science and Electronic Engineering, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, U.K.*

Abstract: Vision-based road detection plays a key role for Unmanned Ground Vehicles (UGVs) working in an unknown outdoor environment. The estimation of the vanishing point is a practical solution for general road detection using monocular vision, which however is not good enough for robust road detection in a campus environment due to the strong noises of texture orientations generated from roadside trees and buildings. In this paper, a novel system framework is proposed by combining the fast scene segmentation (FSS) and the rapid vanishing point detection. The proposed FSS algorithm can segment a single image into road and non-road regions based on the similarity analysis of color histogram, which can eliminate the inherent noises in the trees and buildings and improve the robustness of road detection effectively. Before voting for the vanishing point, we use Canny algorithm to extract the edges in the road region roughly segmented in FSS step. Since most of the strong texture orientations exist in the extracted edges, the computational complexity in the voting stage can be reduced significantly. Experimental results implemented on a real UGV platform show the validity and robustness of the proposed approach.

Keywords: Road detection; scene segmentation; vanishing point detection; unmanned ground vehicles.

1. INTRODUCTION

Encouraged by the exciting performances of Google driverless car and the teams in the DARPA Grand Challenge, it is clear that fully autonomous unmanned ground vehicles (UGVs) could be deployed in the highway in the near future. Vision systems are the most common sensors used in road detection for UGVs. Many vision-based systems have been developed to perform structured road detection recently. Compared with the road detection task in structured environments, unstructured road detection is a more challenging one for UGVs.

Many researchers focus on using supervised machine learning algorithms to segment the scene for road detection. The Adaboost-based region segmentation was proposed to detect the road areas of unstructured roads in Alon et al. (2006). Some other self-supervised learning algorithms (e.g. Fuzzy Support Vector Machines) were utilized in rural roads or campus roads, which need many types of road images to train an effective classifier in Zhou et al. (2010). Another kind of general road detection framework is based on road vanishing point estimation. Rasmussen (2004) proposed vanishing point detection based on texture orientations firstly for unstructured roads. A general road detection algorithm was proposed by Kong et al. (2010). To improve the texture-based vanishing point detection, Moghadam et al. (2012) proposed another local dominant orientation estimation algorithm using four Gabor filters with orientations {0º, 45º, 90º, 135º} and a robust voting scheme. In Miksik (2012), a rapid vanishing point estimation approach was implemented by expanding Gabor wavelets into a linear combination of Haar-like box functions. Moreover, super-pixels were used in the voting scheme to reduce the computational complexity. Many researchers focus on improving the speed of the vanishing point detection to meet the requirements of GUVs by adding a sky pixel removal technique in Moghadam et al. (2012). However, it is still not enough to meet the requirements of real-time and robust performance in real-world applications for UGVs.

In our work, an UGV operating in a campus environment, the texture orientations and features of roadside trees and buildings seriously affect the accuracy of vanishing point detection. A novel approach is proposed in this paper to solve the problem. Firstly, we combine the fast scene segmentation (FSS) algorithm with the rapid vanishing point detection approach to eliminate the inherent noises in the trees and buildings, so that a more robust performance of road detection can be obtained in the campus environment. Secondly, the vanishing point detection is based on the estimation of Gabor-based dominant orientations in the possible road region. Before voting for the vanishing point, Canny algorithm (Canny, 2009) is used to extract the edges in the possible road region, which can reduce the number of voters to ensure the real-time road detection for UGVs. Then we use the vanishing point constrained dominant border detection algorithm to delimit the two borders of the road.

The rest of the paper is organized as follows. Section 2 describes a novel fast scene segmentation approach using the similarity analysis of color histogram between the patches and the road model. Section 3 presents a rapid vanishing point detection approach based on Gabor and road border detection. The experimental results and data analysis are given in Section 4. Finally, a brief concluding remark is given in Section 5.

2. FAST SCENE SEGMENTATION

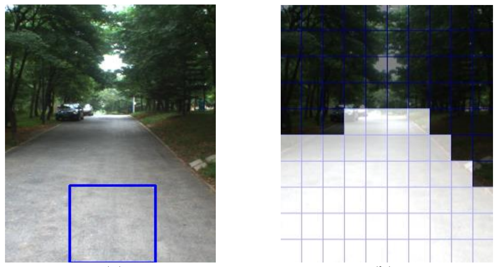
The vanishing point detection approach introduced in Kong et al. (2010) can do a good job for UGVs running in an open field environment, which however is not good in suburb or campus environments since the textures extracted from roadside trees and buildings will affect the robustness of the vanishing point detection. As shown in Fig. 1, the red point is the manually labelled ground truth vanishing point, while the blue one is the false detection result due to the strong noises inherent in the trees. In these cluttered outdoor scenes, the sky pixel removal technique proposed by Moghadam et al. (2012) cannot be adopted to reduce the computational complexity of vanishing point detection. In order to overcome this deficiency, an approach based on the similarity analysis of color histogram is utilized here to segment a single image into road and non-road roughly.

Fig. 1. The red point is a manually labelled ground truth vanishing point, while the blue one is a false detection result due to the noise interference.

First, we segment an image into some fixed-sized patches. We use fixed-sized patches rather than super-pixels because super-pixels cannot guarantee real-time road detection task in our work. Moreover, there is no need to delimit the accurate road region in this stage since the purpose of FSS is to eliminate the inherent noises in the roadside trees and buildings, so the fixed-sized patches are enough for FSS. The approach of the ultimate accurate road region delimitation will be given in the next section.

In our platform, images are normalized to the uniform size with a height of 320 and a width of 240. Its height is larger than the width so that the UGVs can obtain the vision information at a farther end of the road. According to the testing results conducted on our real UGV platform, a fixed-sized neighborhood of 32×24 pixels is adopted in our FSS and we segment an image into 100 fixed-sized patches. FSS utilizes the similarity analysis of color histogram between these patches and a priori road model. In our experiments, the width of our UGV is about 1/3 of the width of testing roads in our campus, so we choose a square region (96×96 pixels), whose side length is 2/5 of the width of testing images, as the priori road model.

As shown in Fig. 2(a), the limited middle region (the size of the region is 96×96) in front of the UGV is assumed as the priori road model in our work. We compute the color histograms over the priori road region and each patch in six color channels: RGB, HSV. The histogram of i th patch and the priori road region is defined as H_i and H_p, respectively. All distributions are quantized into 12 bins. At last, the Bhattacharyya distance, as a metric, is used to express the similarity between each patch and the priori road model.



(a) (b)

Fig. 2. (a) The region in the blue rectangle box (96×96) is assumed to be the priori road region in our work. (b) The white region is the possible road region segmented by FSS.

The Bhattacharyya distance between the i th patch and the priori road model is defined as follows (Bradski et al., 2008):

$$D_{\text{Bhattacharyya}}(H_i, H_p) = \sqrt{1 - \sum_k \frac{\sqrt{H_i(k) \cdot H_p(k)}}{\sqrt{\sum_k H_i(k) \sum_k H_p(k)}}} \quad (1)$$

For Bhattacharyya matching, $D = 0$ indicates a perfect match and $D = 1$ indicates a total mismatch. If the Bhattacharyya distance between the i th patch and the priori road model is smaller than a threshold λ, this patch is viewed to be the part of road region as follows:

$$\begin{cases} patch\ \ is\ \ road, & \text{if } D_p < \lambda \\ patch\ \ is\ non\text{-}road, & otherwise \end{cases} \quad (2)$$

In our experiment, λ =0.79 is chosen to result in the highest possible road region detection accuracy. The detailed illustration of the quantitative analysis of choosing λ is presented in Section 4.2. Fig. 2 (b) shows a sample result of FSS, and the patches classified as road are drawn with white ones.

3. CAMPUS ROAD DETECTION USING RAPID VANISHING POINT ESTIMATION

Fig. 3 shows our novel road detection framework combining FSS and the rapid vanishing point detection. As can be seen, we first use FSS to segment a single image into road and non-road regions approximately. FSS can adapt to the changing conditions of lighting and surface color without the cost to train a complex classifier. Then texture orientations are calculated by Gabor filter and the edges are extracted by Canny algorithm. Only the points at these edges in the possible road region have the right to vote for the vanishing point. Then, the vanishing point constrained dominant border method is used to detect the two borders of the road.

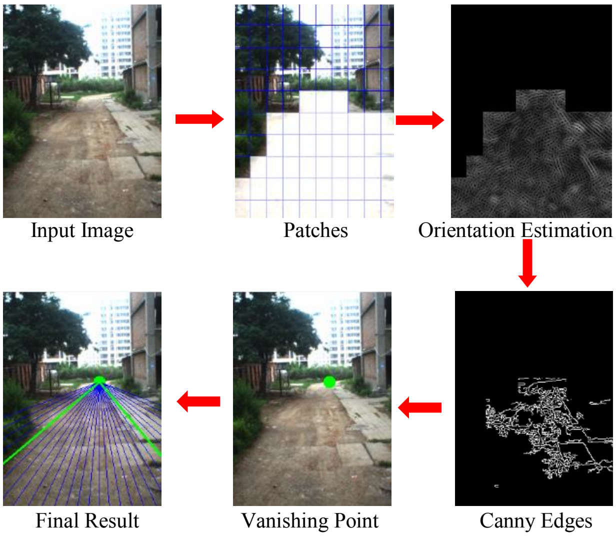


Fig. 3. The overview of proposed framework combining FSS and the rapid vanishing point detection.

3.1 The Texture Orientation Estimation

The texture orientation estimation based on Gabor filters is demonstrated to be accurate in Rasmussen (2004). The texture orientation $\theta(x,y)$ at pixel $p(x,y)$ can be computed by the convolution of a single image and a Gabor kernel as follows (Kong et al, 2010):

$$\varsigma_{\omega,\phi} = I \otimes \psi_{\omega,\phi} \tag{3}$$

where I is the gray image, and $\psi_{\omega,\phi}$ is the Gabor kernel with a scale ω (radial frequency) and an orientation ϕ.

The average of responses at the different scales is defined as the response at orientation θ. Then, the orientation $\theta(x,y)$ obtaining the maximum average response is chosen as the texture orientation at pixel $p(x,y)$ (the detailed algorithm see Kong et al., 2010). For our campus road environment, we choose 3 scales and 13 orientations to ensure the detection accuracy and computational efficiency.

3.2 Vanishing Point Estimation and Road Border Detection

The vanishing point can be voted by texture orientation $\theta(x,y)$. In our experiment, it is a realistic assumption that the pixels in the top 50% portion of the image can be regarded as the vanishing point candidates. The pixels below a vanishing point candidate V can vote for it. We use the locally adaptive soft-voting scheme proposed by Kong et al. (2010) which can overcome the situation that the higher vanishing point candidate will obtain more potential voting pixels, which may lead to false detection. In the soft-voting scheme, for each vanishing point candidate V, it defines a region R_V, a half-disk below V centered at V, whose radius is set to be $r_1 = 0.35\times\Upsilon$, where Υ is the length of the image diagonal. The pixels P in R_V can vote for V as follows:

$$Vote(P,V) = \begin{cases} \dfrac{1}{1+[\gamma d(P,V)]^2}, & if\, \gamma \le \dfrac{5}{1+2d(P,V)} \\ 0, & otherwise \end{cases} \tag{4}$$

where $\gamma = \angle((PV), \theta_P)$ is the angle between the direction of the line PV and θ_P, and θ_P is the texture orientation at pixel P.

For the campus road, trees and buildings will generate strong texture orientations that induce the false results in vanishing point detection. So we combine FSS and the vanishing point estimation. The pixels in the possible road region segmented by FSS are chosen as voters (see Fig. 4). Furthermore, before voting for the vanishing point we use Canny algorithm to extract the edges in the road region roughly segmented in FSS step. Since most of the strong texture orientations exist in the extracted edges, the computational complexity can be reduced significantly. Only the pixels in the possible road region and at these edges have the right to vote (see Fig. 5). Canny algorithm just costs a few milliseconds but this procedure can reduce about 80% of the computational time by reducing the number of voters.

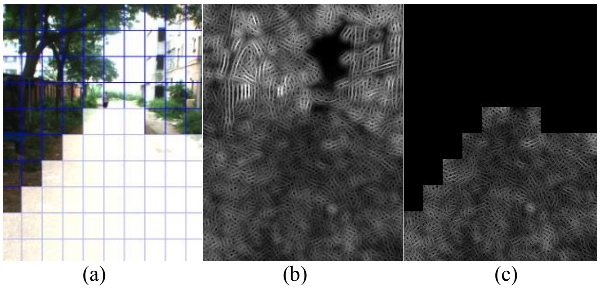


Fig. 4. (a) The result of FSS; (b) The result of Gabor filters on the whole image; (c) The result of Gabor filters on the possible road region segmented by FSS.

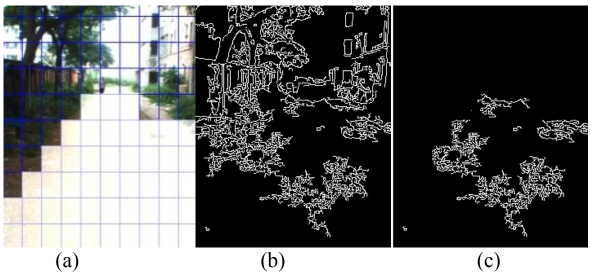


Fig. 5. (a) The result of FSS; (b) The result of Canny algorithm on the whole image; (c) The result of Canny algorithm on the possible road region segmented by FSS.

The candidate obtaining the most scores according to (4) is selected as the vanishing point. The experiment results show that our framework has higher detection accuracy than the one in Kong et al. (2010) for campus roads where usually have trees and buildings, as shown in Fig. 6.

Fig. 6. Comparison of vanishing point detection. The blue point is detected by Kong et al. while the green one is detected by our method, respectively.

Finally, we use the vanishing point constrained dominant border detection method in Kong et al. (2010) to find the two most dominant borders of the road. According to the two borders, we can then obtain the final road region.

3.3 Performance Evaluation

In order to evaluate the detection result, we propose a novel method to quantify the detection accuracy. Most of the researches just show the final detection result by marking the vanishing point and road borders on the image, but don't give a method to quantitatively express the road detection accuracy. Our performance evaluation method can do this well by analysing the true positive (TP), false positive (FP) and false negative (FN).

As shown in Fig. 7, we label two ground truth road borders manually, as the red lines, while two detected road borders are expressed by green lines. The intersection point of two manually labelled road borders is treated as ground truth vanishing point, as the red point in Fig. 7, while the green point is the detected one. Through the road borders we can obtain the final road region. The region "A+B" expresses the ground truth road region, and the region "A+C" expresses the detected road region. The number of pixels in region "A" is $num(p_A)$, the number of pixels in region "B" is $num(p_B)$ and the number of pixels in region "C" is $num(p_C)$. They represent TPs, FNs and FPs, respectively. "*Precision*" is to evaluate the road detection method quantitatively as follows:

$$Precision = num(p_A) / \{num(p_A) + num(p_C)\} \qquad (5)$$

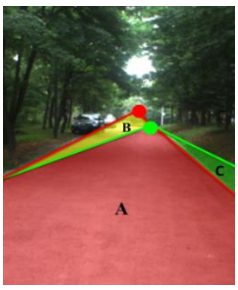


Fig. 7. The illustration of our performance evaluation method.

4. EXPERIMENTAL RESULTS

4.1 System Setup and Testing Image Dataset

All of the algorithms in this paper have been tested on our UGV platform, namely Smart-Cruiser that is a home-developed UGV equipped with multiple lasers and monocular cameras (see Fig. 8). Our road detection framework proposed in this paper only utilizes monocular vision information from the camera, FlyCapture Flea2, which is mounted on the top of our Smart-Cruiser UGV platform, as shown in Fig. 8.

Fig. 8. Smart-Cruiser, a home-developed UGV equipped with multiple lasers and monocular cameras. A FlyCapture Flea2 camera is mounted on the front top of our UGV platform.

1000 representative images used in our experiments are from the dataset that is acquired with the monocular camera in the real-world traffic situations at our campus in different seasons.

4.2 Experimental Results of Fast Scene Segmentation

With the Fast Scene Segmentation (FSS) approach proposed in Section 2, the road detection task can eliminate the inherent noises in the trees and buildings, and also improve the UGV's robustness in real-world applications. In order to select the proper threshold λ, a statistical analysis result is given in Fig. 9 using the ROC (Receiver Operating Characteristic) curve. In our experiment, we tune λ in (2) from 0 to 1 with an interval of 0.1 and test on our image dataset. The ROC curve of true positive rate and false positive rate is given in Fig. 9. In order to obtain the best segmentation results with the high true positive rate and low false positive rate, we set λ=0.79 in our experiments, and the corresponding true positive rate and false positive rate are 0.9483 and 0.8×10^{-2}, respectively (the red point in Fig. 9).

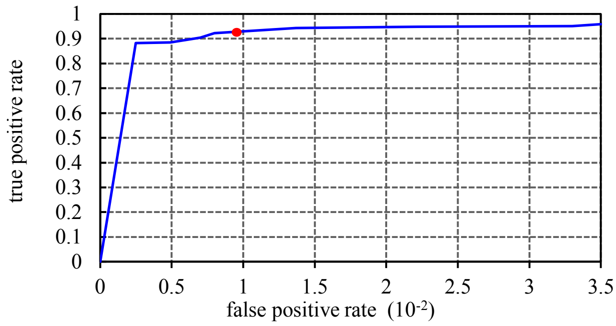


Fig. 9. The ROC curve for the fast scene segmentation (FSS) results with different thresholds.

A group of fast scene segmentation (FSS) results are given in Fig. 10 in which the detected road regions are marked with white patches. Although there may be some false negatives and false positives in these results, most of the road regions have been recognized approximately and the roadside trees and buildings have also been eliminated. The average time-cost in this stage is about 100*ms*.

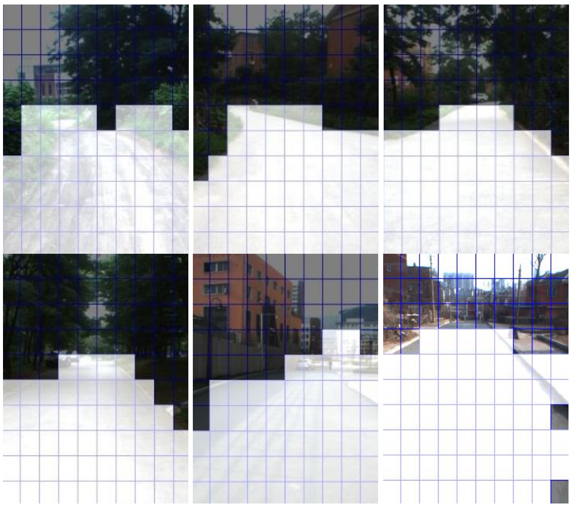


Fig. 10. Some experimental results of FSS.

4.3 Experimental Results of Road Detection

We compare our road detection method with Kong et al. (2010) using the image dataset acquired in our DUT campus. As shown in Fig. 11, the road detection results using our method (FSS+Canny) are delimited by green points and lines, while the detection results using Kong's method are delimited by blue ones.

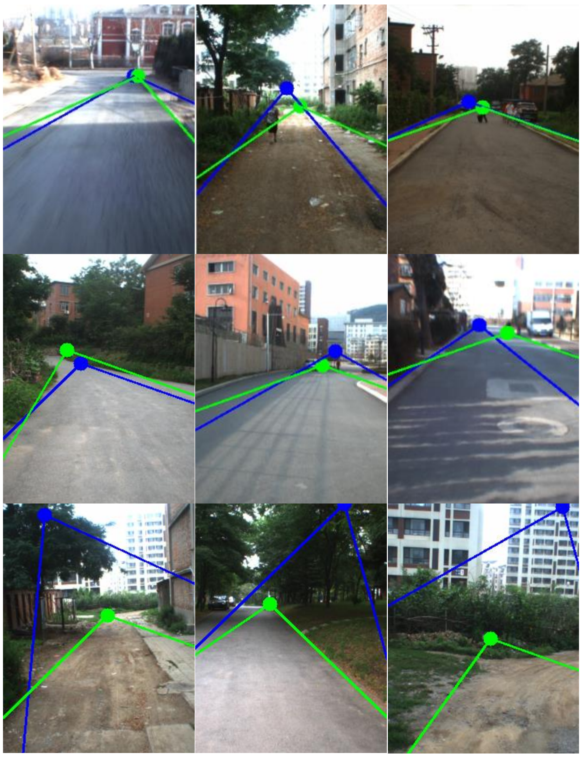


Fig. 11. Some experimental results of road detection.

The statistical analysis results shown in Fig. 12 provide a straightforward performance comparison between our method and Kong's one. As introduced in Section 3.3, Equation (5) is used to calculate the *"Precision"* to evaluate the real-world road detection performance in our experiments. We set the *"Precision"* from 0 to 1 with an interval of 0.1 and for each *"Precision"* interval calculate the corresponding percentage of the image number. As shown in Fig. 12, *"Precisions"* of our results are mainly concentrating on the interval between 0.7~0.97, while the ones of Kong's method are concentrating on the interval between 0.46~0.8.

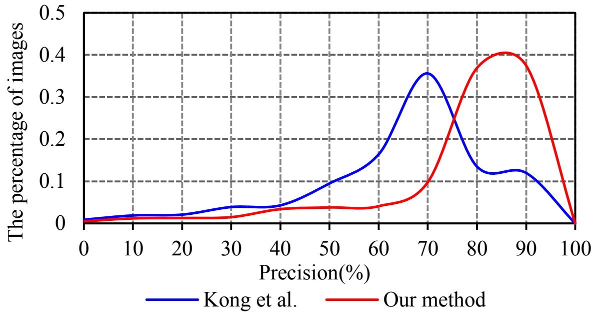


Fig. 12. In different precision intervals, the percentage of the number of images.

It should be noted that the method proposed in this paper is focused on the unstructured road detection problem in a campus environment. The road detection method based on vanishing point estimation proposed by Kong et al. (2010) can work well in general open field environments, but not in campus scenes since the inherent noises of strong texture orientations extracted from the roadside trees and buildings affect the system's accuracy and robustness significantly.

4.4 Timing

Adding Canny algorithm to vanishing point detection can reduce the number of the voters and increase up the computational efficiency. Our image dataset is divided into 25 groups and each group has 40 images according to the variations coming from shifting seasons, lighting conditions and variant scenes. Fig. 13 shows the average number of voters of each group by using different methods, and here we only show the result of 15 groups which are selected from 25 groups randomly. It can be seen that Canny algorithm may reduce about 80% the number of voters, but there are still certain noises from trees and buildings in the residual voters. Therefore, both FSS and Canny algorithm are adopted in our work to reduce the computational burden and improve the detection accuracy.

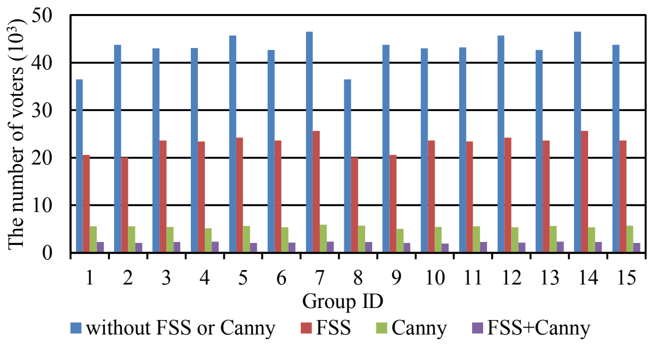


Fig. 13. The number of voters using different algorithms tested on 15 groups data randomly selected from 25 groups.

The runtime analysis is carried out on a PC with a Core2 Duo 2.0-GHz CPU, and the time-cost in every step of our method and Kong's one are shown in Fig. 14 and Fig. 15, respectively. To better illustrate the time-cost, we randomly select 18 groups from the testing dataset and calculate the average time-cost of every group. As shown in Fig. 14, we need about 100*ms* to implement fast scene segmentation (FSS), about 200*ms* for Gabor filter, about 110*ms* to perform vanishing point detection with Canny algorithm, and about 100*ms* for road border estimation. The total average time-cost of our method is about 510*ms*. It should be noted that the time-cost of Canny algorithm is not illustrated in Fig. 14 since it just costs a few milliseconds.

Fig. 15 shows the time-cost for the Kong's method tested on the same 18 groups data. As can be seen, the time-cost of Gabor filter is about 200*ms*, the time-cost of vanishing point detection is about 500*ms*, and the time-cost of road border estimation is about 100*ms*. Although there are only three steps in the Kong's method, its overall time-cost is about 800*ms*, which is a little longer than ours.

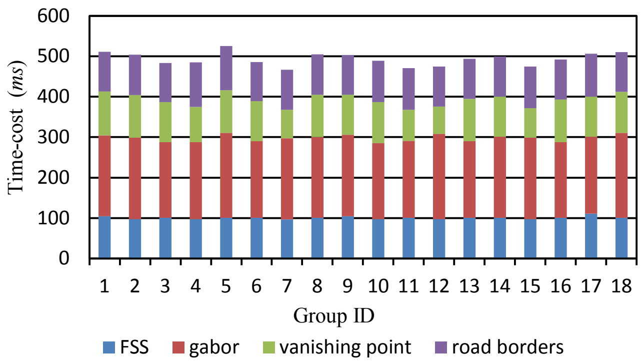


Fig. 14. The time-cost of our road detection method with fast scene segmentation and Canny algorithm (18 groups selected from the testing dataset randomly).

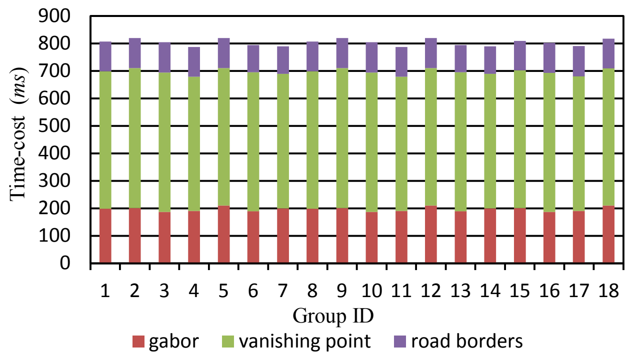


Fig. 15. The time-cost of Kong's method tested on the same 18 groups data shown in Figure 14.

5. CONCLUSION AND FUTURE WORK

In this paper, the novel fast scene segmentation (FSS) algorithm is proposed for visual campus road detection based on a monocular vision system. Compared with the traditional road detection method that is based on vanishing point detection, the proposed algorithm is combined with a rapid vanishing point detection algorithm to perform unstructured campus road detection.

In our work, the FSS algorithm is used to segment a single image into road and non-road regions approximately, which can eliminate the inherent noises in the roadside trees and buildings. A small region in front of the UGV is chosen as a road model by assuming that the small region in front of the UGV must be the road. Canny algorithm is utilized to extract the edges in the possible road region to reduce the computational complexity in the vanishing point voting stage. Experimental results show that our methods can improve the road detection accuracy, and reduce the computational complexity effectively.

In the future work, we plan to solve the problem of robust road detection with a sharp turning as shown in the last image in Fig. 11. Furthermore, other real-time detection algorithms will also be utilized to test the accuracy and robustness in real-world applications.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China (Grant Nos. 61305128) and the Training Program Foundation for the University Talents by the Liaoning Province (Grant Nos. LJQ2013008).

The authors would like to thank Dr. Kong Hui for providing the Matlab codes of his method in Kong et al. (2010).

REFERENCES

C. Rasmussen. Grouping dominant orientations for ill-structured road following. in Proc. IEEE Int. Conf. Computer Vision Pattern Recognition. volume 1, pages 470–477, 2004.

G. Bradski and A. Kaehler. Learning OpenCV: Computer Vision with the OpenCV Library. Cambridge, MA: O'Reilly Press, 2008.

H. Kong, J-Y. Audibert and J. Ponce. General road detection from a single image. IEEE Transactions on Image Processing. pages 2211-2220, 2010.

J. Canny. A computational approach to edge detection. IEEE Trans. Pattern Anal. Machine Intell. volume PAMI-8, pages 679-698, 2009.

O. Miksik. Rapid Vanishing Point Estimation for General Road Detection. IEEE International Conference on Robotics and Automation. Pages 4844-4849, 2012.

P. Moghadam, J. Starzyk, A. Wijesoma and, W. S. Wijesoma. Fast vanishing-point detection in unstructured environments. IEEE Transactions on Image Processing. pages 425-430, 2012.

P. Moghadam, and J. Dong. Road direction detection based on vanishing-point tracking. IEEE/RSJ International Conference on Intelligent Robots and Systems. pages 1553-1560, 2012.

S. Zhou and K. Iagnemma. Self-supervised learning method for unstructured road detection using fuzzy support vector machines. IEEE/RSJ International Conference on Intelligent Robots and Systems. pages 1183-1189, 2010.

Y. Alon, A. Ferencz, and A. Shashua. Off-road path following using region classification and geometric projection constraints. in Proc. IEEE Conf. Computer Vision Pattern Recognition. pages 689–696, 2006.