

Long Range Traversable Region Detection Based on Superpixels Clustering for Mobile Robots

Huimin Lu^{1,2,*}, Lixing Jiang² and Andreas Zell²

Abstract— Traversable region detection is important for autonomous visual navigation of mobile robots. Only short range traversable regions can be detected using traditional methods based on stereo vision because of the limited image resolution and baseline of stereo vision. In this paper, we propose a novel method to detect long range traversable regions without using any supervised or self-supervised learning process. Superpixels are clustered using an improved spectral clustering algorithm to segment the image effectively, and after integrating short range traversable region detection based on u-v-disparity, the traversable region can be extended to long range naturally. The experimental results show that the proposed method works well in different outdoor/field environments, and the detecting range can be improved greatly in comparison with traditional methods. Furthermore, the proposed superpixels clustering algorithm can also be applied in other robot vision tasks like road detection and object recognition.

I. INTRODUCTION

To realize autonomous navigation, mobile robots should be able to detect traversable regions which we define as the regions that do not contain geometric obstacles in the environments. In comparison with laser scanners and ultrasonic sensors, stereo vision can provide richer information including color, texture and depth, and thus has become increasingly popular in this research topic. Stereo vision can measure the ranges to objects by calculating disparities between stereo images. After acquiring the disparities, traversable regions or obstacles can be detected robustly and efficiently using a series of approaches based on u-v-disparity.

V-disparity was originally introduced by Labayrade, et al. [1] aiming at detecting obstacles, where v is the ordinate of a pixel in the (u, v) image coordinate system. By accumulating pixels with the same disparity value d in each row, a v -disparity image (d, v) can be built. In the v -disparity image, perpendicular obstacles can be mapped to vertical lines whose pixel intensity represents the width of the obstacles, and the traversable region modelled as a succession of planes can be projected as slanted line segment, also named as ground correlation line in [2]. Hu, et al. [3] proposed the concept of u-v-disparity. Similar to a v -disparity image, a u -disparity image is built by accumulating pixels with the same disparity value d in each column, where perpendicular

obstacles can be mapped to be horizontal lines whose pixel intensity represents the height of the obstacles. So line detection algorithms like the Hough transform can be used to detect traversable regions and obstacles for mobile robots.

Although the u-v-disparity methods mentioned above work, only short range traversable regions can be detected, generally less than 10 meters [4]–[6]. At long ranges, the range/disparity data become too sparse or noisy to evaluate the traversability because of the limitations on resolution and baseline of stereo vision. The typical results can be found in Fig. 1. In this case, optimal path planning may not be realized, and mobile robots may drive into long cul-de-sacs, so the efficiency of visual navigation will be reduced greatly. To deal with this problem, the most famous work is using self-supervised online learning or near-to-far learning, to generalize short-range classification from stereo-based geometry to long-range classification from imagery [5]–[8]. The general steps are as follows: detecting a traversable region locally with stereo data, selecting features from the classified terrain, then learning a two-class classifier for traversable and non-traversable regions using supervised learning algorithms like support vector machine (SVM), and finally applying the classifier to classify long-range data.

However, humans can perceive the environment and find traversable regions by analysing the scene images naturally without a learning process. So our question is: is it possible to detect long range traversable regions just by image processing without any supervised or self-supervised learning process?

During the past decade, superpixels have become increasingly popular in computer vision applications, because they capture redundancy in the image and reduce the complexity of subsequent image processing tasks [9] [10]. Especially some recent implementations of the superpixel algorithm [11] can be run in real-time, which makes them possible to be used in robotics application. Good superpixel algorithms can adhere well to image boundaries, which benefits image analysis like semantic image segmentation. In this paper, we propose to cluster superpixels with an improved spectral clustering algorithm to segment the image into different regions with semantic meaning, and then extend the short range traversable region detected by a u-v-disparity based method to long range naturally, as shown in Fig. 1.

The next sections are organized as follows: the framework of our long range traversable region detection method will be proposed in section II; short range traversable region detection based on u-v-disparity will be introduced briefly in section III; in section IV, we propose to cluster superpixels

*This work was supported by National Science Foundation of China (No.61403409) and China Scholarship Council (No.201306115022).

¹College of Mechatronics and Automation, National University of Defense Technology, Changsha, China, 410073. lhmnew@nudt.edu.cn

²Department of Cognitive Systems, University of Tübingen, Tübingen, Germany, 72076. huimin.lu@wsii.uni-tuebingen.de, {lixing.jiang, andreas.zell}@uni-tuebingen.de

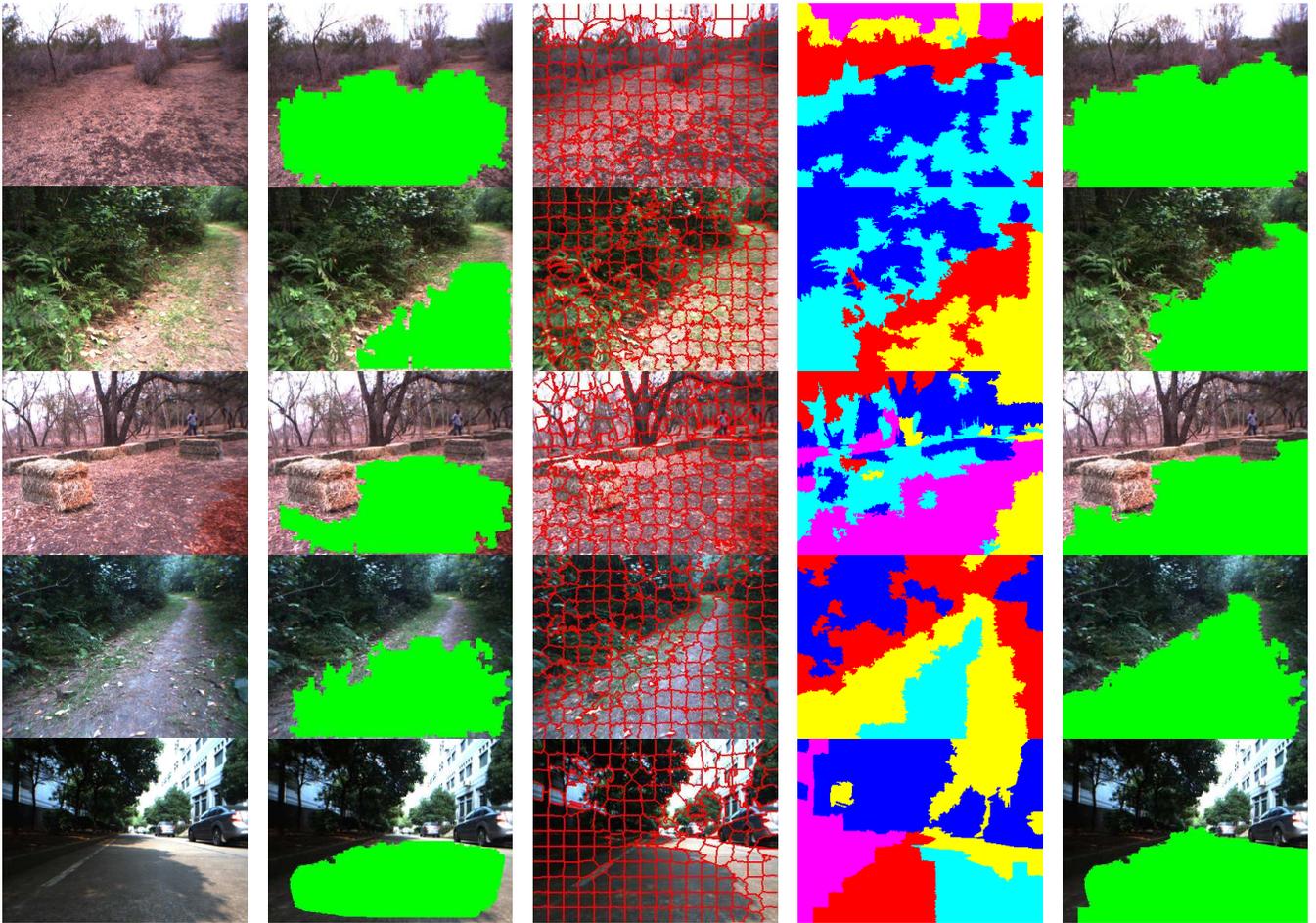


Fig. 1. (1st column) The original left images from stereo vision. (2nd column) Short range traversable region detection results (green) using the u-v-disparity method. (3rd column) The generated superpixels using the Preemptive SLIC algorithm. (4th column) The superpixels clustering results using the improved spectral clustering algorithm. (5th column) The final long range traversable region detection results (green) using the proposed method. The images in the 1st~4th rows are from the LAGR dataset [12], and the image in the 5th row is acquired on the NUDT campus by the NuBot rescue robot.

with an improved spectral clustering algorithm, and then extend the detected short range traversable region in section III to long range; experimental results with different datasets are presented in section V, and the real-time performance is also discussed; section VI concludes this paper.

II. LONG RANGE TRAVERSABLE REGION DETECTION FRAMEWORK

In this section, we propose the framework of our long range traversable region detection method, as shown in Fig. 2. After the left and right image are acquired by stereo vision, we calculate the disparity information using stereo matching, and detect short range traversable regions in the left image based on u-v-disparity. Then we use a fast version of the SLIC superpixels algorithm, namely Preemptive SLIC [11], to segment the left image into superpixels. After extracting color and texture features from each superpixel, we cluster the superpixels to segment the image into several regions with semantic meaning using an improved spectral clustering algorithm. By integrating short range detection results, the traversable region can be extended to long range naturally

without any supervised or self-supervised learning process. The proposed method is introduced in detail in the following sections.

III. SHORT RANGE TRAVERSABLE REGION DETECTION BASED ON U-V-DISPARITY

In the first author's earlier research [13], an efficient algorithm was proposed to detect traversable region based on u-v-disparity which is similar as the work in [14]. The algorithm is divided into the following steps:

- Generating the disparity image: after acquiring the left image and right image from stereo vision, we generate the disparity image firstly by applying ELAS (Efficient LARGE-scale Stereo Matching) [15], which performs quite well and fast even in relatively low-textured images.
- Calculating the obstacle disparity image and the non-obstacle disparity image: considering that traversable regions can be projected to be slanted line segments, called ground correlation lines, in the v-disparity image,

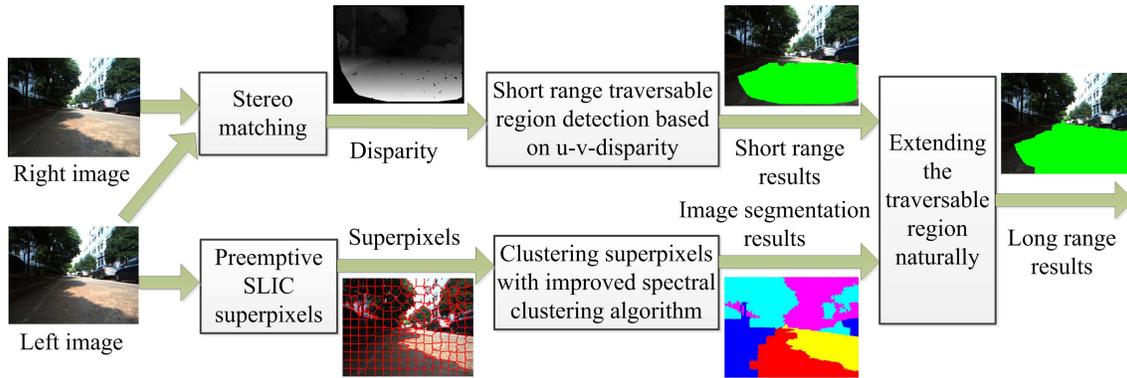


Fig. 2. The framework of the proposed long range traversable region detection method.

and the affection from non-traversable regions in the v -disparity image would make it difficult to detect correct ground correlation lines, we calculate the obstacle disparity image by binarizing the u -disparity image and performing a morphological close operation, and then generate the non-obstacle disparity image by subtracting the obstacle disparity image from the disparity image. The final v -disparity image can be calculated from the non-obstacle disparity image. During this process, the affection from non-traversable regions mentioned above can be removed as much as possible.

- Detecting the ground correlation line and projecting back: the ground correlation line can be detected easily by using the Hough transform, and traversable regions can be detected by projecting back all pixels in the v -disparity image that contribute to the ground correlation line on the left image.

The typical stereo images and traversable region detection results are shown in Fig. 3. More details about this algorithm can be found in [13], where experimental results show that the algorithm works well in various environments. However, because of the limitation of imaging resolution and baseline of stereo vision, only short range traversable regions can be detected, which reduces the efficiency of visual navigation, especially for long range path planning.

IV. LONG RANGE TRAVERSABLE REGION DETECTION BASED ON SUPERPIXELS CLUSTERING

Superpixels have become a good pre-processing step to reduce the computation cost of subsequent processing for many vision applications. In this paper, we propose to cluster superpixels according to the color and texture information from each superpixel, so the image can be segmented into different regions with semantic meaning, which will be used to extend the detected short range traversable region into long range.

A. Generating superpixels using Preemptive SLIC

Simple linear iterative clustering (SLIC) [9] is a good algorithm to generate superpixels that adhere to image boundaries with computational and memory efficiency. However, the

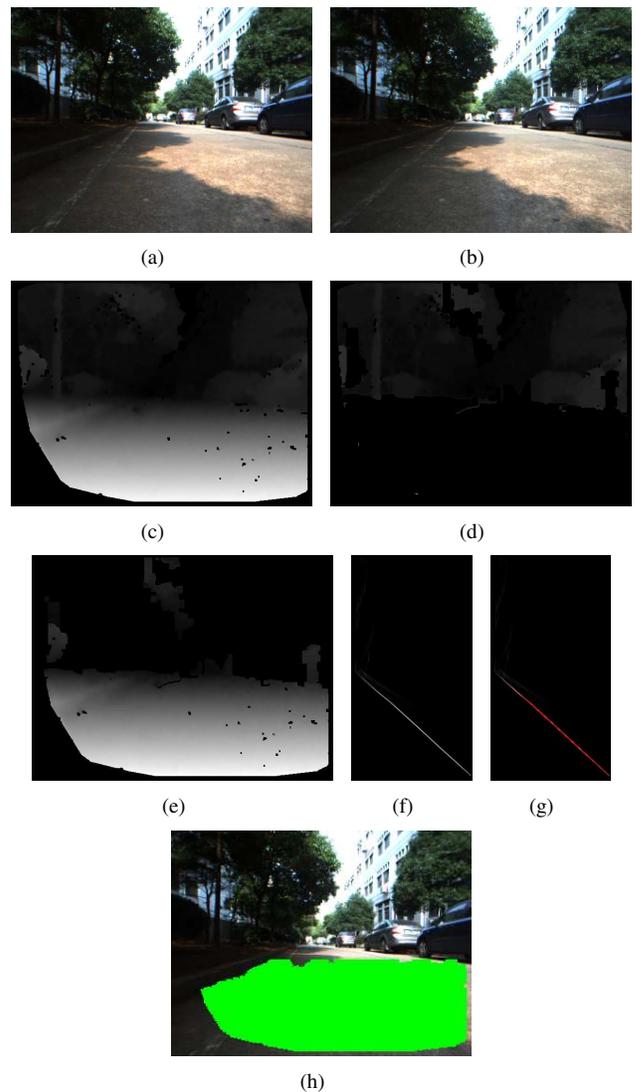


Fig. 3. (a)(b) The left and right image acquired on the NUDT campus by the NuBot rescue robot. (c) The disparity image. (d) The obstacle disparity image. (e) The non-obstacle disparity image. (f) The final v -disparity image. (g) The detected ground correlation line (red). (h) The detected traversable region (green).

computation cost is still too high on a standard desktop CPU for robot vision applications. In [11], Neubert, et al. present an adaption of SLIC, Preemptive SLIC, that preserves the high segmentation quality level of the original implementation and can be run in real-time. So we use Preemptive SLIC to generate superpixels on the left image from stereo vision, as shown in Fig. 1 and Fig. 4. In this paper, about 200 superpixels are generated from each image.

B. Extracting color and texture features from superpixels

After obtaining the superpixels, some feature should be extracted before clustering them. In this paper, we consider color and texture features that compliment each other. For the color feature, we first transform the image into HSV space, which is more robust to varying illumination than RGB space, and then extract a color histogram from each superpixel. In constructing the color histogram, the H, S, and V channel are discretized into 9, 8, and 6 bins respectively, which were determined experimentally. The dimension of the color histogram feature is 432.

Because the superpixels are not regular and fixed in shape, we calculate a texture value for each pixel, and then construct a histogram within each superpixel as the texture feature similar as that in the color histogram. We compared several methods with very low computation complexity to compute the texture value, including local binary patterns (LBP) [16], center-symmetric local binary patterns (CS-LBP) [17], local ternary patterns (LTP) [18], Weber's law descriptor (WLD) [19], and local homogeneity [20], and we found WLD performs best in our case. We compute the WLD value I_{wld} on the pixel x_c as follows:

$$I_{wld} = \arctan\left[\sum_{i=0}^7 \frac{(x_i - x_c) * 3}{x_c}\right] \quad (1)$$

where $x_i (i = 0, \dots, 7)$ are the eight neighbours of x_c . After normalizing I_{wld} into $[0, 255]$, we can construct a histogram within each superpixel as the texture feature. The dimension of the texture feature is 256. The WLD values computed from Fig. 3(a) are shown in Fig. 4(b). Finally, we acquire the feature set $\{c_1, \dots, c_n\}$ and $\{t_1, \dots, t_n\}$ from each image, where c_i and t_i are the color and texture feature respectively, and n is the number of superpixels.

C. Clustering superpixels using improved spectral clustering

As one of most popular modern clustering algorithms, spectral clustering often performs better than traditional algorithms like K-means [21]. Given a dataset $\{s_1, \dots, s_n\}$, the original spectral clustering can be described as follows:

- Construct the affinity matrix $A \in R^{n \times n}$, where $A_{i,j} = \exp(-d(s_i, s_j)^2 / 2\sigma^2)$, when $i \neq j$, and $A_{i,j} = 0$, when $i = j$. $d(s_i, s_j)$ is the distance between s_i and s_j .
- Compute the degree matrix D , where $D_{i,i} = \sum_{j=1}^n A_{i,j}$, and the normalized affinity matrix $L = D^{-1/2} A D^{-1/2}$.
- Assuming $1 = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k$ be the k largest eigenvalues of L , where k is the number of expected

clusters, construct the matrix $V = [v^1, v^2, \dots, v^k] \in R^{n \times k}$, where v^1, v^2, \dots, v^k are the corresponding eigenvectors.

- Form the matrix Y from V by normalizing each of V 's rows to have unit length, where $Y_{i,j} = V_{i,j} / (\sum_{j=1}^k V_{i,j}^2)^{1/2}$.
- Treat Y_i , each row of Y , as a point in R^k , and cluster $\{Y_1, \dots, Y_n\}$ into k clusters via the K-means algorithm to obtain the final clustering of the original dataset.

In our case, $s_i = \{c_i, t_i\}$. We calculate $d(c_i, c_j)$ and $d(t_i, t_j)$ by matching the feature histograms. We compared Correlation, Chi-Square, Intersection, Bhattacharyya, Euclidean distance and Earth Mover's Distance (EMD), and found that the Correlation distance is best for our application. Taking $d(c_i, c_j)$ as an example, it is calculated as follows:

$$1.0 - \left(\frac{\sum_{l=1}^L (c_i(l) - \bar{c}_i)(c_j(l) - \bar{c}_j)}{\sqrt{\sum_{l=1}^L (c_i(l) - \bar{c}_i)^2 \sum_{l=1}^L (c_j(l) - \bar{c}_j)^2}} + 1.0 \right) / 2.0 \quad (2)$$

where $\bar{c}_i = \sum_{l=1}^L c_i(l) / L$, and L is the dimension of the histogram. The smaller $d(c_i, c_j)$ means smaller distance between c_i and c_j .

Then we combine the color feature and the texture feature with adaptive weights. When comparing the mean value of all $d(c_i, c_j)$, namely \bar{d}_c , and the mean value of all $d(t_i, t_j)$, namely \bar{d}_t , we assume that a larger value means better discrimination between the superpixels using the corresponding feature, so we calculate $d(s_i, s_j)$ as follows:

$$d(s_i, s_j) = d(c_i, c_j) * \frac{\bar{d}_c}{\bar{d}_c + \bar{d}_t} + d(t_i, t_j) * \frac{\bar{d}_t}{\bar{d}_c + \bar{d}_t} \quad (3)$$

We also modify the original spectral clustering algorithm as follows:

- Because σ is a kind of scale parameter needed to be set when constructing the affinity matrix A , we calculate it autonomously by $\max_{i=1}^n (\min_{j=1}^n d(s_i, s_j))$, where $i \neq j$.
- We decide the number of clusters k autonomously according to the eigenvalues. For $1 = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$, we set $k = 1 + \arg(\max |\lambda_i - \lambda_{i+1}|_{\lambda_i \neq 1})$.
- The K-means algorithm used in the final step of spectral clustering converges quickly to a local optimum, so the initial setting of cluster centers will affect the clustering results, and it is not good to choose initial cluster centers randomly. We adopt the following method to choose initial centers by maximizing the distances between them: choosing the first center from $\{Y_1, \dots, Y_n\}$ randomly, and then choosing the other $k - 1$ centers one by one from $\{Y_1, \dots, Y_n\}$ with the largest distance to the already chosen centers. Good clustering results can be achieved using this method, as shown in Fig.1 and Fig. 4.

D. Extending the traversable region into long range

Using the improved spectral clustering algorithm, we can segment the image into several regions with sort of semantic meaning effectively, like ground with soil, ground with

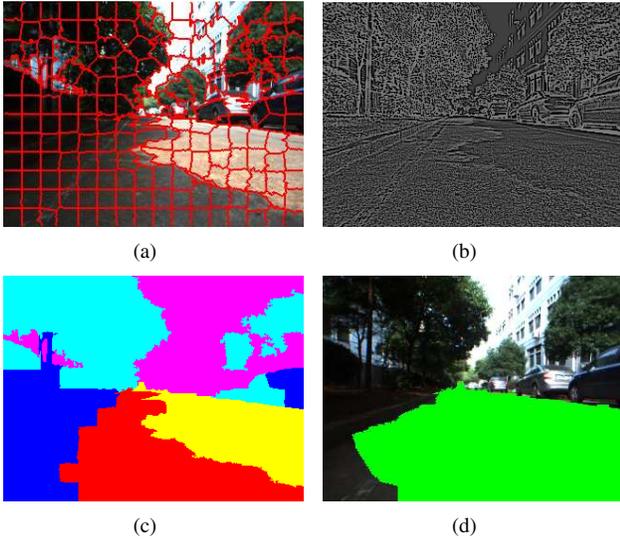


Fig. 4. The processing results on Fig. 3(a). (a) The generated superpixels using Preemptive SLIC. (b) The computed WLD values. (c) The superpixels clustering results using the improved spectral clustering algorithm. (d) The final long range traversable region detection results (green).

grass, ground with shadows, etc. Then we can extend the detected short range traversable region naturally using a simple strategy: for each clustered region, if over 50% of its area are within the detected short range traversable region, the whole clustered region is considered to be traversable. The union of the short range traversable region and these traversable clustered regions make up the final long range traversable region, as shown in Fig.1 and Fig. 4.

V. EXPERIMENTAL RESULTS AND DISCUSSION

A. Experimental setup

We use two challenging datasets to test our proposed long range traversable region detection method. The first one is a publicly available dataset from the LAGR project including six image sets [12], which were acquired by the robot equipped with a Bumblebee2 stereo vision system developed by Point Grey on different field environments with different terrain appearances and under different lighting conditions. There is a sequence of 100 images in each image set. The manually labelled results are provided as ground truth, where each pixel is classified to be one of three classes: obstacle, ground plane (traversable region), and unknown. The second dataset was acquired on the NUDT campus by the NuBot rescue robot developed by the first author’s research group in NUDT, as shown in Fig. 5. The robot is also equipped with a Bumblebee2 stereo vision system. The typical images from these two datasets are shown in Fig. 1 and Fig. 3.

B. The experimental results

Because ground truth is provided in the LAGR dataset, we can evaluate the performance of our method with quantitative results. Three criteria are used: precision, recall, and root-mean-square error (RMSE). They are defined as follows:

$$precision = \frac{TP}{TP + FP} \quad (4)$$



Fig. 5. The NuBot rescue robot equipped with a Bumblebee2 stereo vision system.

$$recall = \frac{TP}{TP + FN} \quad (5)$$

where TP is the number of pixels which are detected as traversable correctly, $TP + FP$ is the number of pixels which are detected as traversable, and $TP + FN$ is the number of pixels which belong to ground truth traversable region.

$$RMSE = \sqrt{\sum_{i=1}^N (\hat{y}_i - y_i)^2 / N} \quad (6)$$

where \hat{y}_i and y_i belong to $\{0,1\}$, N is the number of pixels in the image, and smaller RMSE means better performance. $\hat{y}_i = 1$ means that the corresponding pixel is detected as traversable, and 0 means non-traversable. $y_i = 1$ means that the corresponding pixel is traversable, and 0 means non-traversable according to ground truth. In this paper, we consider the obstacle class and unknown class as non-traversable.

The performance of our method is shown in Table I when testing with each image set of the LAGR dataset. The typical results are demonstrated in Fig. 1. We also test our method using our own dataset, and the typical results are shown in Fig. 1 and Fig. 4. More results on these two datasets can be found on our submitted accompanying video. The detection rate of traversable regions can be improved greatly with only a small reduction on the detection precision using our proposed method in comparison with the traditional method based on u-v-disparity. The traversable region can be extended to long range effectively. Furthermore, because disparity information can only be generated within the part of the left image with corresponding points in the right image, parts of short range traversable region are missing when using the method based on u-v-disparity. Most of these missing traversable regions can also be detected using our method. In comparison with the near-to-far learning method [7] using the same LAGR dataset and the same performance criteria RMSE, the performance of our method is comparable to the baseline near-to-far learning method (mean RMSE = 0.2692 and 0.273 respectively), although no supervised or self-supervised learning process is needed in our method.

We test the performance of our method when only using the color feature or the texture feature. As shown in Table

TABLE I

THE STATISTICS ON THE PERFORMANCE OF OUR PROPOSED METHOD IN COMPARISON WITH SHORT RANGE DETECTION RESULTS BASED ON U-V-DISPARITY. THE PERFORMANCE OF OUR METHOD WHEN ONLY USING THE COLOR FEATURE OR THE TEXTURE FEATURE, AND USING THE ORIGINAL SPECTRAL CLUSTERING BY CHOOSING THE INITIAL CLUSTER CENTERS RANDOMLY IN THE K-MEANS STEP IS ALSO PRESENTED.

		DS1A	DS1B	DS2A	DS2B	DS3A	DS3B	mean value
precision	the method based on u-v-disparity	0.9919	0.9772	0.9537	0.9525	0.9815	0.9924	0.9749
	our proposed method	0.9706	0.9293	0.9186	0.9213	0.9535	0.9779	0.9452
	only using the color feature	0.9742	0.9251	0.9163	0.9150	0.9556	0.9787	0.9441
	only using the texture feature	0.9732	0.9202	0.8438	0.8505	0.9472	0.9525	0.9146
	choosing initial cluster centers randomly	0.9722	0.9307	0.9208	0.9222	0.9504	0.9789	0.9459
recall	the method based on u-v-disparity	0.8212	0.8154	0.8382	0.8154	0.7616	0.5283	0.7634
	our proposed method	0.9170	0.9495	0.9400	0.9116	0.8884	0.8003	0.9011
	only using the color feature	0.9178	0.9488	0.9344	0.9117	0.8707	0.7697	0.8922
	only using the texture feature	0.9030	0.9041	0.9347	0.9184	0.8769	0.7275	0.8775
	choosing initial cluster centers randomly	0.9147	0.9472	0.9342	0.9076	0.8888	0.7959	0.8981
RMSE	the method based on u-v-disparity	0.3277	0.3092	0.3210	0.3380	0.3282	0.4820	0.3510
	our proposed method	0.2521	0.2347	0.2662	0.2856	0.2545	0.3224	0.2692
	only using the color feature	0.2459	0.2391	0.2743	0.2927	0.2680	0.3419	0.2770
	only using the texture feature	0.2632	0.2791	0.3450	0.3511	0.2693	0.3854	0.3155
	choosing initial cluster centers randomly	0.2528	0.2356	0.2697	0.2885	0.2571	0.3256	0.2715

I, the experimental results verify that the performance can be improved by using these two features together to cluster superpixels. We also compare the performance when using the improved spectral clustering algorithm or the original version by choosing the initial cluster centers randomly in the K-means step, as shown in Table I. Better performance can be achieved by using the improved spectral clustering algorithm. It is worth mentioning that even small improvements in the performance criteria are very valuable, because the traversable regions at long range only occupy small part of the image.

C. The real-time performance

We test the real-time performance of our method using 150 frames of stereo images from our own dataset. The needed computation time to perform each step is shown in Fig. 6. The dimension of the image is 640*480, and the computer is equipped with a 2.4GHz i7 CPU and 4GB memory. We find that the mean computation time is about 191ms, and only 5 frames of stereo images can be processed to detect long range traversable region. However, most of the time, about 104ms, is consumed by generating the disparity using stereo matching, so when using an RGB-D sensor which can output the disparity already, or using FPGA hardware to perform stereo matching, long range traversable regions can be detected at a frame rate of 11 fps. It is acceptable for visual navigation of mobile robots, because in our case, traversable regions are detected at long range, which does not need to be done for each frame of stereo images.

D. Discussions

Besides traversable region detection, our superpixels clustering algorithm can be used as a preprocessing step directly in many other robot vision tasks, like object recognition for service robots and road detection for unmanned ground vehicles, as shown in Fig. 7.

Although experimental results show that our method works well in different outdoor/field environments, the clustering results should be improved in some cases. As shown in

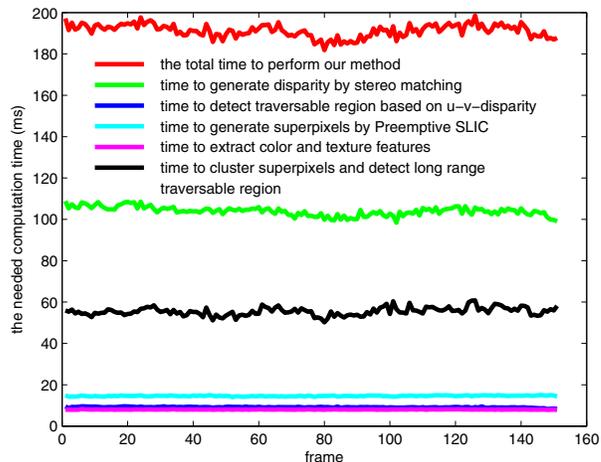


Fig. 6. The needed computation time to detect long range traversable region using the proposed method.

Fig. 8(a)(b), the clustering results are not always optimal or sometimes even wrong. There are also many situations which are difficult for our method to deal with. As shown in Fig. 8(c)(d), the image appearance is affected greatly by strong sunshine and shadows, so the clustering results are right but still not good for traversable region detection. Better features should be extracted from superpixels to improve our method and make it more robust to extreme environmental conditions.

VI. CONCLUSIONS

In this paper, we focus on detecting long range traversable regions without using any supervised or self-supervised learning algorithm. We proposed to cluster superpixels using an improved spectral clustering algorithm, so the image can be segmented into several regions with sort of semantic meaning effectively. After integrating short range traversable regions detected using a u-v-disparity based method, the

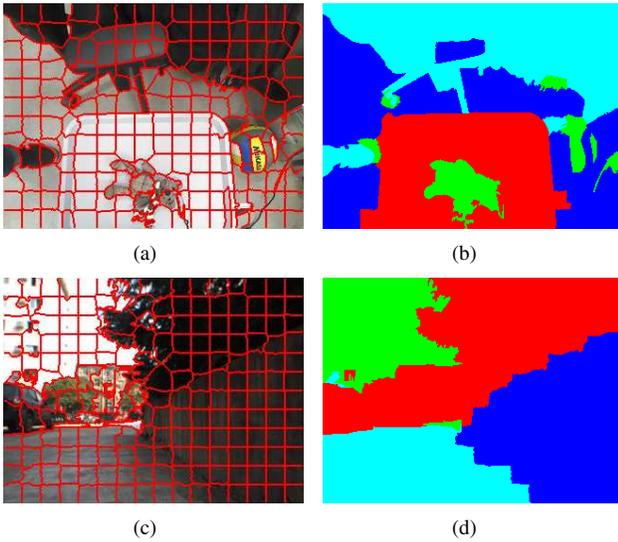


Fig. 7. (a)(b) The superpixels clustering results for object recognition. (c)(d) The superpixels clustering results for road detection. The toy and road regions can be segmented successfully.

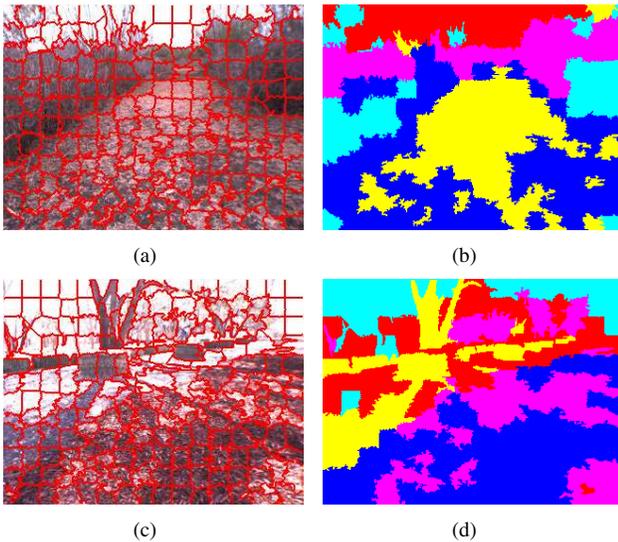


Fig. 8. (a)(b) The superpixels clustering results are not always optimal or sometimes even wrong. (c)(d) The superpixels clustering results are right but still not good for traversable region detection because of the great affection from strong sunshine and shadows.

traversable region can be extended into long range naturally. The experimental results show that our method works well in different outdoor/field environments, and long range traversable regions can be detected efficiently for visual navigation of mobile robots.

ACKNOWLEDGEMENT

We would like to thank Peer Neubert and Peter Protzel for their release of the Pre-emptive SLIC source code on <https://www.tu-chemnitz.de/etit/proaut/forschung/cv/segmentation.html.en>, and Simon Butler for his release of spectral clustering source code on <https://github.com/pthimon/clustering>.

REFERENCES

- [1] R. Labayrade, D. Aubert, and J.-P. Tarel, "Real time obstacle detection in stereovision on non flat road geometry through "v-disparity" representation," in *2002 IEEE Intelligent Vehicle Symposium*, June 2002, pp. 646–651.
- [2] A. Broggi, C. Caraffi, R. Fedriga, and P. Grisleri, "Obstacle detection with stereo vision for off-road vehicle navigation," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*, June 2005, pp. 65–65.
- [3] Z. Hu and K. Uchimura, "U-v-disparity: an efficient algorithm for stereovision based scene analysis," in *2005 IEEE Intelligent Vehicles Symposium*, June 2005, pp. 48–54.
- [4] A. Howard, M. Turmon, L. Matthies, B. Tang, A. Angelova, and E. Mjolsness, "Towards learned traversability for robot navigation: From underfoot to the far field," *Journal of Field Robotics*, vol. 23, no. 11-12, pp. 1005–1017, 2006.
- [5] M. Bajracharya, A. Howard, L. H. Matthies, B. Tang, and M. Turmon, "Autonomous off-road navigation with end-to-end learning for the lagr program," *Journal of Field Robotics*, vol. 26, no. 1, pp. 3–25, 2009.
- [6] M. Bajracharya, B. Tang, A. Howard, M. Turmon, and L. Matthies, "Learning long-range terrain classification for autonomous navigation," in *2008 IEEE International Conference on Robotics and Automation*, May 2008, pp. 4018–4024.
- [7] M. J. Procopio, J. Mulligan, and G. Grudic, "Learning terrain segmentation with classifier ensembles for autonomous robot navigation in unstructured environments," *Journal of Field Robotics*, vol. 26, no. 2, pp. 145–175, 2009.
- [8] I. Kostavelis, L. Nalpantidis, and A. Gasteratos, "Collision risk assessment for autonomous robots by offline traversability learning," *Robotics and Autonomous Systems*, vol. 60, no. 11, pp. 1367–1376, 2012.
- [9] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, Nov 2012.
- [10] M. Van den Bergh, X. Boix, G. Roig, and L. Van Gool, "Seeds: Superpixels extracted via energy-driven sampling," *International Journal of Computer Vision*, vol. 111, pp. 298–314, 2015.
- [11] P. Neubert and P. Protzel, "Compact watershed and preemptive slic: On improving trade-offs of superpixel segmentation algorithms," in *2014 22nd International Conference on Pattern Recognition (ICPR)*, Aug 2014, pp. 996–1001.
- [12] M. J. Procopio, "Hand-labeled DARPA LAGR datasets," Available at <http://www.mikeprocopio.com/labeledlagrdata.html>, 2007.
- [13] X. Zhu, H. Lu, X. Yang, Y. Li, and H. Zhang, "Stereo vision based traversable region detection for mobile robots using u-v-disparity," in *32nd Chinese Control Conference*, July 2013, pp. 5785–5790.
- [14] B. Musleh, A. de la Escalera, and J. Armingol, "U-v disparity analysis in urban environments," in *Computer Aided Systems Theory EUROCAST 2011*, ser. LNCS, 2012, vol. 6928, pp. 426–432.
- [15] A. Geiger, M. Roser, and R. Urtasun, "Efficient large-scale stereo matching," in *Computer Vision ACCV 2010*, ser. LNCS, 2011, vol. 6492, pp. 25–38.
- [16] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, Jul 2002.
- [17] M. Heikkilä, M. Pietikäinen, and C. Schmid, "Description of interest regions with local binary patterns," *Pattern Recognition*, vol. 42, no. 3, pp. 425–436, 2009.
- [18] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Transactions on Image Processing*, vol. 19, no. 6, pp. 1635–1650, June 2010.
- [19] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikäinen, X. Chen, and W. Gao, "Wld: A robust local image descriptor," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1705–1720, Sept 2010.
- [20] H.-D. Cheng and Y. Sun, "A hierarchical approach to color image segmentation using homogeneity," *IEEE Transactions on Image Processing*, vol. 9, no. 12, pp. 2071–2082, Dec 2000.
- [21] U. von Luxburg, "A tutorial on spectral clustering," *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, 2007.