



**HAL**  
open science

# A Novel RGBD Image Superpixel Segmentation Intergrated Depth Map Quality

Weiwei Wei, Wenxia Chen, Hong Tao

► **To cite this version:**

Weiwei Wei, Wenxia Chen, Hong Tao. A Novel RGBD Image Superpixel Segmentation Intergrated Depth Map Quality. 12th International Conference on Intelligent Information Processing (IIP), May 2022, Qingdao, China. pp.444-456, 10.1007/978-3-031-03948-5\_36 . hal-04178748

**HAL Id: hal-04178748**

**<https://inria.hal.science/hal-04178748v1>**

Submitted on 8 Aug 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

# A Novel RGBD Image Superpixel Segmentation Intergrated Depth Map Quality

Weiyi Wei, Wenxia Chen<sup>1</sup>, Hong Tao

College of Computer Science and Engineering, Northwest Normal University,  
Lanzhou, Gansu, China  
2020211950@nwnu.edu.cn

**Abstract.** Superpixel segmentation is to gather adjacent pixels into sub-regions with certain semantics, which can greatly reduce the computational complexity of subsequent image processing. The existing SLIC-D algorithm of RGBD image is a superpixel segmentation method that integrates depth information based on SLIC algorithm. However, for some images with complex background, it cannot stick to the weak edge of the image, and even has the defect of wrong segmentation. Therefore, this paper proposes a novel RGBD image superpixel segmentation intergrated depth map quality. Firstly, RGB map is enhanced by bilateral filtering to suppress and smooth noise, joint bilateral filtering is used to fill the depth map hole. Secondly, Holistically-Nested Edge Detection method based on convolutional neural network is used to obtain the RGB contour map, the edge consistency map is obtained by Hadamard product operation with depth gradient map. Finally, superpixel segmentation is performed by clustering based on LAB color space, spatial location and high-quality depth information similarity. Through the verification on NYU dataset, the advantages of this algorithm in edge fit degree are fully demonstrated.

**Keywords:** Superpixel segmentation, RGBD, depth map quality, Clustering.

## 1 Introduction

As an image preprocessing step, superpixel segmentation gathers adjacent pixels with similar attributes into a class of sub-regions, which reduces the computational complexity of subsequent image processing. Superpixel segmentation has been widely used in image segmentation [1], target recognition and tracking [2], and saliency detection [3-6]. Among the algorithms for generating superpixels, the SLIC (Simple Linear Iterative Clustering) algorithm proposed by Achanta et al. [7] often achieves the best results, the algorithm generates regular shape and controllable number of superpixel blocks based on the 5-dimensional features of CIELAB color space and pixel position. LSC [8] (Linear Spectral Clustering) developed from SLIC, replaces clustering pixels with high-dimensional features. Guo et al. [9] based on the SLIC algorithm, proposed an improved algorithm combining adaptive multi-threshold LBP texture features, which solved the inaccurate segmentation problem of the SLIC algorithm considering only color and spatial information. In view of the time-consuming

problem of SLIC algorithm, Lei et al. [10] proposed a fast SLIC superpixel segmentation method. Liao et al. [11] aiming at the fact that the above superpixel segmentation algorithm cannot automatically determine the appropriate number of superpixels, proposed a superpixel segmentation method using local information. Xiao et al. [12] proposed an adaptive content-based superpixel segmentation method, which can iteratively evaluate the importance of color, contour, texture and spatial location features based on the discriminant method of clustering, and adaptively adjust the weight of each feature.

With the development of imaging technology, many researchers began to turn to the RGBD image superpixel segmentation method for research. In the depth map, the depth sensor may not detect some pixels of depth information, and these pixels constitute depth map holes. Zhang et al. [13] proposed a joint filtering method to fill the hole and up-sampled depth information to generate RGBD image capture system. The depth map quality was often neglected in previous work, Zhao et al. [14] mentioned the importance of depth map quality and introduced contrast loss of RGB stream into depth stream to improve the quality of depth map. Wang et al. [15] proposed three quality assessment methods of depth map, the selective fusion of these three depth qualities improves the accuracy of saliency detection in RGBD image. In the case of very low light, can't rely on color information for superpixel segmentation, Agoes [16] proposes a superpixel generation algorithm rely on depth information, the depth and gradient direction as an alternative LAB color space features, to solve the problem of semantic inconsistency in depth map segmentation due to the same depth value near the boundary. Li et al. [17] proposed RGBD image superpixel segmentation, namely, SLIC-D algorithm, the depth information is not only used to solve the image with similar foreground and background color, but also to achieve the co-segmentation of multiple images. Therefore, this paper proposes a novel RGBD image superpixel segmentation method intergrated depth map quality. Firstly, the quality of depth map is evaluated according to the edge consistency of RGB contour map and depth gradient map. In this paper, HED (Holistically-Nested Edge Detection method) [18] based on convolutional neural network is used to extract RGB contour map. Secondly, the results of depth map quality evaluation are weighted and fused with depth map to obtain high-quality depth map. Finally, superpixel segmentation is achieved by clustering based on LAB color space, spatial location and high-quality depth information similarity. The main contributions are summarized as follows.

- This paper proposes a high-quality depth map generation method. The edge consistency map is obtained from RGB contour map and depth gradient map by Hadamard product operation, high-quality depth map is obtained by weighted fusion of edge consistency map and depth map.
- This paper proposes a novel superpixel segmentation method for RGBD images. Superpixels are generated by clustering with the similarity of color, similarity of space location and similarity of high-quality depth information.

## 2 RGBD Image Superpixel Segmentation Intergrated Depth Map Quality

In this paper, firstly, the RGB map is enhanced by bilateral filtering, and the hole of depth map is filled by joint bilateral filtering enhancement. Secondly, Depth map quality assessment method was carried out based on the edge consistency map between RGB profile map and depth gradient map, however, in the space weighting operation, to reduce the computational complexity, considering neighborhood information at the same time, this paper uses an adaptive nonlocal random walk algorithm for segmenting superpixel proposed by Wang et al. [19]. Thirdly, edge consistency map are weighted and fused with the depth map to obtain a high-quality depth map. Lastly, the color, space location and high-quality depth information are integrated for clustering to obtain the RGBD image superpixel segmentation method. The algorithm flow is shown in Fig. 1.

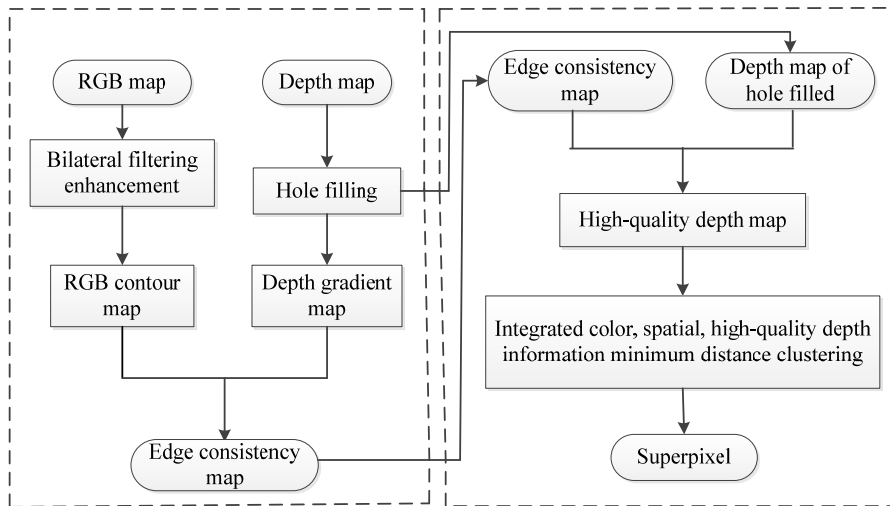


Fig. 1. The algorithm flow of the proposed superpixel segmentation.

### 2.1 Color Map Bilateral Filtering Enhancement

Bilateral filtering enhancement can not only eliminate image noise, but also preserve image edge information, which is convenient for subsequent image superpixel segmentation. Its principle is to calculate the gaussian function of the space distance and gray distance between two pixels, determine the weight size through the range of the Gaussian kernel function, and then multiply the two weight coefficients for filtering. The expression is as follows.

$$g(x, y) = \frac{\sum_{(k,m) \in S(x,y)} I(k, m)w(x, y, k, m)}{w(x, y, k, m)} \quad (1)$$

$$w(x, y, k, m) = w_s(x, y, k, m) * w_r(x, y, k, m) \quad (2)$$

$$w_s(x, y, k, m) = \exp\left(-\frac{(k-x)^2 + (m-y)^2}{2\delta_s^2}\right) \quad (3)$$

$$w_r(x, y, k, m) = \exp\left(-\frac{\|I(k, m) - I(x, y)\|^2}{2\delta_r^2}\right) \quad (4)$$

where  $g(x,y)$  is the image processed by bilateral filtering,  $S(x,y)$  is the collection of  $(2N+1)(2N+1)$  neighborhood pixels centered on pixel point  $(x,y)$ ,  $I(k,m)$  denotes the pixel value at  $(k,m)$ .  $w(x,y,k,m)$  denotes the value calculated by two gaussian functions,  $w_s(x,y,k,m)$  is based on the similarity of spatial distance, which decreases with the increase of the distance between pixels.  $w_r(x,y,k,m)$  is the gray value similarity calculated by square of the absolute value difference between two pixels, which decreases with the increase of the gray value difference. The larger the  $\delta_s$  is, the more weighted pixels will lead to the more blurred the image. The limit of  $\delta_r$  will keep the edge information of the image, so as to eliminate the image noise and preserve the edge information of the image, making the image prepare for the subsequent super-pixel segmentation.

## 2.2 Depth Map Hole Filling

To fill depth map hole using joint bilateral filtering algorithm, joint bilateral filter is developed by the bilateral filtering, the difference between them is that the joint bilateral filtering algorithm uses color map as guide map to calculate the domain weight, and the filtering results of the bilateral filtering algorithm are optimized to have the unstable effect of flipping near the edge. The expression is as follows.

$$U_y = \frac{\sum_{x \in \Omega} w_c(x, y) \cdot w_s(x, y) \cdot O_x}{\sum_{x \in \Omega} w_c(x, y) \cdot w_s(x, y)} \quad (5)$$

$$w_c(x, y) = \exp\left(-\frac{\|D_y - N_x\|^2}{2\delta_c^2}\right) \quad (6)$$

$$w_s(x, y) = \exp\left(-\frac{\|y - x\|^2}{2\delta_s^2}\right) \quad (7)$$

where  $U_y$  is the pixel depth value after joint filtering,  $O_x$  is the depth value of the processed depth map, and  $\Omega$  denotes the filtering neighborhood range,  $W_c(x,y)$  is the color similarity weight,  $W_s(x,y)$  is the spatial position similarity weight,  $D_y$  is the pixel depth value of the position  $y$  of the point to be filled in the color map, and  $N_x$  is the pixel value of the position  $x$  in the neighborhood.  $\delta_c$  is the color similarity weight parameter, and  $\delta_s$  is the spatial location similarity weight parameter. Color similarity means that the more similar the pixels to be filled are in the color image, the more weight the pixels to be filled will have. Spatial location similarity means that the closer the pixels to be filled are, the more weight they will have.

### 2.3 High-quality Depth Map Generation Algorithm

In the D (Depth) map of RGBD image, the area around the edge of the image is a high-quality depth area. Therefore, in this paper, RGB contour map and depth gradient map are used to obtain edge consistency map, and it is weighted and fused with depth map to obtain high-quality depth map. The algorithm steps are as follows.

**Algorithm 1.** High-quality depth map generation algorithm.

Input: RGB map, Depth map.

Output: High-quality depth map.

Step 1. RGB contour map and depth gradient map are generated to obtain the rough edge consistency map, and then gaussian smoothing operation is carried out to roughly locate high-quality depth areas.

In this paper, the HED method based on convolutional neural network is used to obtain RGB contour map. Compared with the traditional edge detection method, this method can not only show the object boundary, but also suppress the irrelevant edges inside the object.  $EC$  (Edge Consistency map) of  $HED$  (RGB contour graph) and  $DG$  (Depth gradient map) can be expressed as equation (8).

$$EC = HED \odot DG \quad (8)$$

where  $\odot$  denotes the Hadamard product operation of the elements corresponding to image matrix  $HED$  and image matrix  $DG$ , and the resulting of edge consistency map is shown in Fig. 2.

Since high-quality depth regions tend to be located near those pixels with large edge consistency values, a defined threshold is used to determine a subgroup of "anchor pixels", the APC in Fig. 2. and these anchor pixels are used to roughly locate high-quality D regions ( $EC'$ ) through equation (9).

$$EC' = G(f(EC - T_c)) \quad (9)$$

where,  $T_c$  is the determined threshold,  $G$  denotes Gaussian smoothing operation,  $f$  is a function that assigns its negative value to zero.

Step 2. In this paper, the adaptive non-local random walk algorithm is used for superpixel segmentation to replace image pixels and generate a complete edge consistency map.

The depth map quality feature map should be complete, so a novel spatial weighting operation is used for regions not near the edge, it mainly includes similarity measure and space weighted range of features,  $\exp$  is the similarity measure of features, it is implemented with the common thread [20]. The euclidean distance is used to determine the spatial weighted range (11) in equation (10), that is, APA in Fig. 2. is determined adaptively by equation (11), where  $T_c$  is set to 20,  $T_a$  is set to 30. In this paper, in order to consider the global neighborhood relationship between image pixels and reduce time complexity, an adaptive nonlocal random walk algorithm is used to generate superpixel blocks.

$$EC(P_i) \leftarrow \frac{\sum P_j \phi EC^-(P_j) \exp(w \|V(P_i), V(P_j)\|_2)}{\sum P_j \in \exp(w \|V(P_i), V(P_j)\|_2)} \quad (10)$$

$$\phi: \|c(p_i, p_j)\|_2 \leq \min(\{ \|p_i, f(DG - T_a)\|_2 \}) \quad (11)$$

where  $P_i$  denotes the  $i$ -th superpixel,  $T_a$  is determined threshold,  $w$  is the weight parameter,  $\|\cdot\|$  is  $l_2$  norm,  $V(\cdot)$  and  $P(\cdot)$  returns the average and center of the given RGB image respectively. The final edge consistency map ( $EC$ ) is at the lower right corner of Fig. 2.

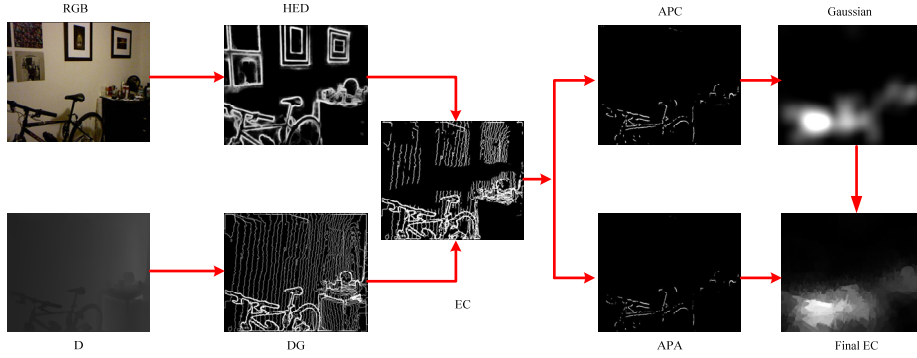


Fig. 2. Edge consistency map

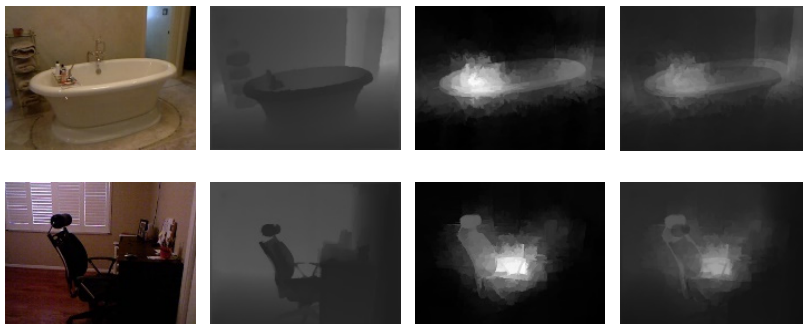
Step 3. The complete edge consistency map is weighted and fused with depth map to obtain high-quality depth map, as the equation 12.

$$h(x) = (1 - \partial) f_0(x) + \partial f_1(x) \quad (12)$$

where  $\partial$  sets to 0.6, namely, the weight of edge consistency map is set to 0.6, and the weight of depth map is set to 0.4 will get high-quality depth map to extract the high-quality depth information, combined with the similarity of high-quality depth infor-



information, color similarity, similarity of space location clustering, get superpixel segmentation results will be more accurate,  $f_0(x)$  and  $f_1(x)$  respectively denote the depth map and the corresponding edge consistency map,  $h(x)$  denotes the weighted fusion of the depth map and edge consistency map, high-quality depth map generation process is shown in Fig. 3.



**Fig. 3.** The process of generating high-quality depth map. From left to right are RGB map, Depth map, Edge consistency map, High-quality depth map.

## 2.4 Superpixel Segmentation Intergrated Depth Map Quality

The method in this paper is to transform the 5-dimensional feature composed of CIELAB color space and pixel position from the traditional SLIC algorithm to the 6-dimensional feature composed of CIELAB color space, the depth information after quality evaluation and pixel position to cluster adjacent pixels. The algorithm steps are as follows.

**Algorithm 2.** Superpixel segmentation intergrated depth map quality.

Input: RGB map, high-quality depth map.

Output: Superpixel block.

Step 1. The interval  $S$  is used to initialize the cluster center  $C_k$ , in order to prevent the clustering center from being at the edge of the image, the clustering center was moved to the position of the lowest gradient of the pixel in the 3 neighborhood. Set each pixel label  $l(i)$  to -1, and set the distance  $d_i$  to  $\infty$ .

Step 2. For each clustering center  $C_k$ , calculate the distance  $D_{i,k}$  from the clustering center  $C_k$  to each pixel  $i$  in  $2S$  region. If  $D_{i,k} < d_i$ , assign the values of  $D_{i,k}$  to  $d_i$ , and set the label  $l(i)$  is  $k$ , the pixels are divided into the central region with the highest similarity in turn. Where interval  $S$  and the similarity distance formula can be calculated as.

$$S = \sqrt{Num / K} \quad (13)$$

$$d_{lab} = \sqrt{(l_i - l_k)^2 + (a_i - a_k)^2 + (b_i - b_k)^2} \quad (14)$$

$$d_d = \sqrt{(d_i - d_k)^2} \quad (15)$$

$$d_{xy} = \sqrt{(x_i - x_k)^2 + (y_i - y_k)^2} \quad (16)$$

where  $Num$  denotes the total number of pixels in an image, and  $K$  denotes the number of presegmented pixels.  $d_{lab}$  denotes the distance measurement from  $i$  to  $k$  in the Lab color space,  $d_d$  denotes the distance calculation of high-quality depth value from  $i$  to  $k$ ,  $d_{xy}$  denotes the calculation of the spatial distance from  $i$  to  $k$  pixel coordinate values. The final similarity distance measure  $D_{ik}$  is shown in equation (17).

$$D_{i,k} = \sqrt{\left(\frac{d_{lab}}{m}\right)^2 + \left(\frac{d_d}{n}\right)^2 + \left(\frac{d_{xy}}{S}\right)^2} \quad (17)$$

where  $m$ ,  $n$  and  $S$  are weight coefficients, which are used to balance the proportion of color information, high-quality depth information and spatial location information in similarity measurement.

### 3 Experimental Results and Analysis

#### 3.1 Datasets and Evaluation Metrics

This paper selects 654 images contained in NYU dataset, which are composed of video sequences of various indoor scenes recorded by RGB and Depth cameras of Microsoft Kinect.

Both  $UE$  (Under-segmentation Error) [7] and  $ASA$  (Achievable segmentation Accuracy) [21] are used for experimental verification of the proposed algorithm.  $UE$  and  $ASA$  are indexes to evaluate the edge fitting ability of superpixel segmentation results, where  $UE$  denotes the percentage of pixels that overflow from the truth boundary, and  $ASA$  denotes the upper limit of segmentation accuracy, which is the highest accuracy that can be obtained when using superpixels as basic units for image segmentation, its definition is shown in equations (18)–(19).

$$UE = \frac{\sum_m \left( \sum_{C_k \cap G_m > B} |C_k| \right) - N}{N} \quad (18)$$

$$ASA(S) = \frac{\sum_k \max_m (C_k \cap G_m)}{\sum_m |G_m|} \quad (19)$$

where  $C_1, C_2, C_3 \dots C_n$  denotes  $n$  superpixel blocks generated by the superpixel segmentation algorithm,  $N$  is the number of pixels in the image,  $G$  represents the truth segmentation region,  $B$  is 5%,  $m$  denotes the number of truth segmentation regions, and  $k$  denotes the number of superpixel segmentation regions.

### 3.2 Influence of Parameters

Given that the number of superpixels is 300 and  $n$  is 500, the influence of  $m$  on the result of superpixel segmentation is shown in Table 1.

**Table 1.** The influence of  $m$  on the result of superpixel segmentation.

$m$	$ASA$	$UE$
10	0.893	0.216
20	<b>0.898</b>	<b>0.212</b>
30	0.892	0.222
40	0.897	0.243

It can be observed from Table 1 that when  $m$  is 20,  $ASA$  takes the maximum value and  $UE$  takes the minimum value. When the value of  $m$  is greater than or less than 20, the segmentation accuracy will decrease and the under-segmentation error will increase compared with that when  $m$  is 20. So  $m$  is going to take 20.

When the given number of superpixels is 300 and  $m$  is 20, the influence of  $n$  on the result of superpixel segmentation is shown in Table 2.

**Table 2.** The influence of  $n$  on the result of superpixel segmentation.

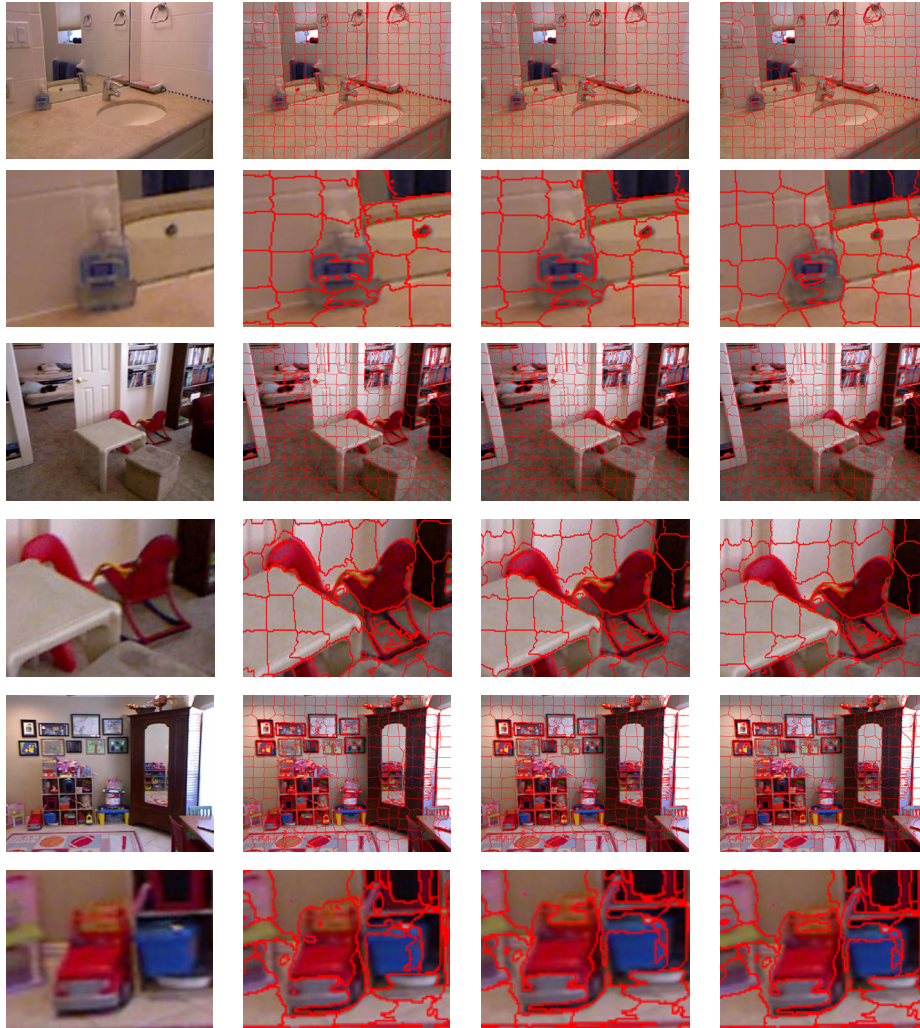
$N$	$ASA$	$UE$
50	0.848	0.297
200	0.854	0.286
500	<b>0.869</b>	<b>0.254</b>
600	0.865	0.354

As can be seen from Table 2, when  $n$  is less than 500, the accuracy of superpixel segmentation increases with the increase of  $n$  value, and the under-segmentation error decreases with the increase of  $n$  value. When the value is an integer above 500, the segmentation accuracy will decrease and the under-segmentation error will increase compared with that when  $n$  is 500. So  $n$  is going to take 500 in this paper.

### 3.3 Comparison Experiment

On the one hand, the method in this paper has the advantages of fast running speed and clear contour of SLIC algorithm. On the other hand, for some images with complex background, by incorporating high-quality depth information, the superpixel segmentation region of the image stick to the image boundary better. The comparison algorithms involved in verification in this paper are: the SLIC algorithm [7] and the SLIC-D algorithm [17].

**Visually Compare.** In Fig. 4. shows the comparison of the superpixel segmentation results visual effect map in NYU dataset when the number of superpixels is set to 300. The original map and the superpixel segmentation result map are given, and local magnification of them is also given.

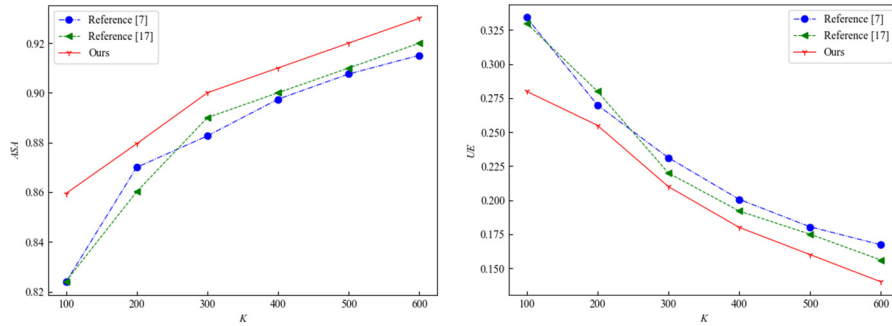


**Fig. 4.** Visual effect map. From left to right are Origin, Reference [7], Reference [17], Ours.

From Fig.4. the first group of picture, as can be seen from the superpixel segmentation result of hand sanitizer on the wash table, the algorithms in Reference [7] and Reference [17] do not distinguish the outer contour edges of hand sanitizer, while Our algorithm can not only segment the rough contour of the outer contour of the hand sanitizer, but also it can separate the weak edges of the target object inside the

image. The second group of picture, it can be seen from the superpixel segmentation result of the left chair, Reference [7] algorithm and Reference [17] algorithm do not stick to the edge of chair, and only Our algorithm can stick to the edge of chair closely. The third group of picture can be observed that Reference [7] algorithm can only divide the outer contour of the stool, but the weak edge part of the internal target in the image cannot be recognized, Reference [17] algorithm mistakenly regarded the contour of the internal target of the stool as the outer contour of the stool. The algorithm in this paper not only segmented the rough edges of the outer contour of the stool, but also accurately segmented the weak edge part of the internal target of the stool. In conclusion, RGBD image superpixel segmentation method integrated depth map quality has more advantages in edge fitting degree. We analyzed theoretically, RGB profile map and depth gradient map by Hadamard product operation to get edge consistency map, high-quality depth map is obtained by weighted fusion between edge consistency map and depth map, because clustering is based on high-quality depth information, color information and spatial location information, therefore, compared with the SLIC algorithm and the SLIC-D algorithm, the algorithm in this paper can significantly improve the image edge fitting degree.

**Objective Evaluation.**  $UE$  and  $ASA$  are indicators to evaluate the edge fitting ability of superpixel segmentation results. In this paper, under the condition that the number of pixels is given, the mean value of the superpixel segmentation results on  $ASA$  and  $UE$  indexes of all images is selected.



**Fig. 5.** Contrast data diagram. From left to right are  $ASA$ ,  $UE$ .

In general, when the number of superpixels is given, Our algorithm has higher segmentation accuracy and lower under-segmentation error than Reference [17] algorithm, Reference [17] algorithm has higher segmentation accuracy and lower under-segmentation error than Reference [7] algorithm. However, when the number of superpixels is set to 200, the average segmentation accuracy of Reference [7] algorithm is 0.870, and the under-segmentation error is 0.269, the average segmentation accuracy of Reference [17] algorithm is 0.861, and the under-segmentation error is 0.280. However, the segmentation accuracy of the proposed algorithm is 0.879 and the un-

der-segmentation error is 0.255. That is to say, when the number of superpixels is set to 200, Reference [17] algorithm has lower segmentation accuracy and higher under-segmentation error than Reference [7] algorithm, while the proposed algorithm has higher segmentation accuracy and lower under-segmentation error than the previous two algorithms. It shows that the proposed algorithm can not only stick to the weak edge of the image with a given number of superpixels, but also improve the image edge fitting degree. For example, the proposed algorithm can well stick to the weak boundary region of hand sanitizer and stool, so the proposed algorithm can maintain good contour details and have a good edge fit degree.

## 4 Conclusion

In this paper, firstly, the RGB map is enhanced by bilateral filtering to suppress and smooth the noise, joint bilateral filtering enhancement is used to fill the depth map hole. Secondly, contour map was obtained from the enhanced RGB map, and depth gradient map was obtained from the hole filled depth map. Then the edge consistency map and depth map are weighted and fused to get a high-quality depth map. Finally, clustering is based on the similarity of color, spatial location and high-quality depth information, a novel superpixel segmentation method for RGBD image was obtained. The algorithm is verified by experiments on NYU dataset and compared with the existing SLIC algorithm and SLIC-D algorithm. The results show that for images with complex background, the superpixel blocks generated by the algorithm in this paper can not only recognize the weak edges of the image, but also closely stick to the edges of the image.

**Acknowledgement.** This work was supported by the Science and Technology Plan-natural Science Foundation project of Gansu (grant numbers 20JR5RA518), and the Cultivation plan of major Scientific Research Projects of Northwest Normal University (grant numbers NWNLU-LKZD2021-06).

## References

1. Arbelaez, P., Maire, M., Fowlkes, C., et al.: Contour detection and hierarchical image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **33**(5), 898-916 (2010). doi: 10.1109/TPAMI.2010.161
2. Yang, F., Lu, H., Yang, M.: Robust superpixel tracking. *IEEE Trans. Image Process.* **23**(4), 1639-1651 (2014). doi: 10.1109/TIP.2014.2300823
3. Li, C., Cong, R., Hou, J., et al.: Nested network with two-stream pyramid for salient object detection in optical remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **57**(11), 9156-9166 (2019). doi: 10.1109/TGRS.2019.2925070
4. Cong, R., Lei, J., Fu, H., et al.: An iterative co-saliency framework for RGBD images. *IEEE Trans. Cybern.* **49**(1), 233-246 (2017). doi: 10.1109/TCYB.2017.2771488

5. Cong, R., Lei, J., Fu, H., et al.: Video saliency detection via sparsity-based reconstruction and propagation. *IEEE Trans. Image Process.* **28**(10), 4819-4831 (2019). doi: 10.1109/TIP.2019.2910377
6. Cong, R., Lei, J., Fu, H., et al.: Going from RGB to RGBD saliency: A depth-guided transformation model. *IEEE Trans. Cybern.* **50**(8), 3627-3639 (2019). doi: 10.1109/TCYB.2019.2932005
7. Achanta, R., Shaji, A., Smith, K., et al.: SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**(11), 2274-2282 (2012). doi: 10.1109/TPAMI.2012.120
8. Li, Z., Chen, J.: Superpixel segmentation using linear spectral clustering. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1356-1363. IEEE Press, Boston (2015). doi: 10.1109/CVPR.2015.7298741
9. Guo, Y.J., Yang, M., Hou, Y.C.: Application of an Improved SLIC Algorithm in Color Image Segmentation. *Journal of Chongqing University of Technology (Natural Science)* **34**(2), 158-164(2020). doi: 10.3969/j.issn.1674-8425(z).2020.02.022
10. Lei, T., Lian, Q., Jia, X.H., et al.: Fast Simple Linear Iterative Clustering for Image Superpixel Algorithm. *Comput. Sci.* **47**(2), 143-149 (2020).
11. Liao, M., Li, Y., Zhao, Y.Q., et al.: A new method for image superpixel segmentation. *Journal of Electronics & Information Technology* **42**(2), 83-89 (2020).
12. Xiao, X., Zhou, Y., Gong, Y.J.: Content-adaptive superpixel segmentation. *IEEE Trans. Image Process.* **27**(6), 2883-2896 (2018). doi: 10.1109/TIP.2018.2810541
13. Zhang, Y., Ding, L., Sharma, G.: Local-linear-fitting-based matting for joint hole filling and depth upsampling of RGB-D images. *J. Electron. Imaging.* **28**(3), 033019 (2019). doi: 10.1117/1.JEI.28.3.033019
14. Zhao, J.X., Cao, Y., Fan, D.P., et al. Contrast prior and fluid pyramid integration for RGBD salient object detection. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3927-3936. IEEE Press, Long Beach (2019). doi: 10.1109/CVPR.2019.00405
15. Wang, X., Li, S., Chen, C., et al.: Depth map quality-aware selective saliency fusion for RGB-D image salient object detection. *Neurocomputing* **432**, 44-56 (2021). doi: 10.1016/j.neucom.2020.12.071
16. Agoes, A.S., Hu, Z., Matsunaga, N.: DSLIC: A Superpixel Based Segmentation Algorithm for Depth Image. In: Chen, C.S., Lu, J., Ma, K.K. (eds.) *ACCV 2016. LNCS*, vol. 10117, pp. 77-87. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-54427-4\\_6](https://doi.org/10.1007/978-3-319-54427-4_6)
17. Li, X.Y., Wan, L.L., Li, H.N., et al.: RGBD Image Co-segmentation via Saliency Detection and Graph Cut. *J. Simul.* **30**(7), 2558 (2018). doi: 10.16182/j.issn1004731x.joss.201807016
18. Xie, S., Tu, Z.: Holistically-nested edge detection. In: *Proceedings of the IEEE international conference on computer vision*, pp. 1395-1403. IEEE Press, Santiago (2015). doi: 10.1109/ICCV.2015.164
19. Wang, H., Shen, J., Yin, J., et al.: Adaptive nonlocal random walks for image superpixel segmentation. *IEEE Trans. Circ. Syst. Vid.* **30**(3), 822-834 (2019). doi: 10.1109/TCSVT.2019.2896438
20. Chen, C., Li, S., Wang, Y., et al.: Video saliency detection via spatial-temporal fusion and low-rank coherency diffusion. *IEEE Trans. Image Process.* **26**(7), 3156-3170 (2017). doi: 10.1109/TIP.2017.2670143
21. Stutz, D., Hermans, A., Leibe, B.: Superpixels: An evaluation of the state-of-the-art. *Comput Vis. Image Underst.* **166**, 1-27 (2018). doi: 10.1016/j.cviu.2017.03.007