Volume 8, Issue 5
October 2019

ITEE Journal
Information Technology & Electrical Engineering

ISSN: - 2306-708X

Segmentation of Railings and Bars from RGB-Depth Images

[1]M. Varalakshmamma and [2]T. Venkateswarlu

[1,2] Department of Electronics and Communication Engineering,
S.V University College of Engineering, Tirupati, A.P., India,

E-mail: [1]varalaxmi.39@gmail.com, [2]warlu57@gmail.com

ABSTRACT

This paper presents a new approach for the removal of bars and railings from RGB-Depth images. In recent days, obtaining the RGB-Depth images has become inexpensive and easier by Microsoft Kinect. The main aim of this paper is to segment bars or railings from the RGB-Depth images. By using only depth information it is problematic to meet the desired segmentation. In this paper, threshold-based segmentation is used for an initial segmentation of railings or bars. Depth information and morphological techniques are used to improve the initial segmentation result. The railings or bars are removed from the image. The gaps in the image after removal of undesired structures are filled by a modified wavelet-based exemplar inpainting technique. To demonstrate the proposed approach verity of images are considered. Results are compared with the state-of-art image de-fencing techniques

Keywords: *Image segmentation, Image de-fencing, Image Inpainting, Depth map, Morphology.*

1. INTRODUCTION

When armature photographers take a picture, it may contain certain undesirable structures or objects. For example, when a photo of a beautiful outdoor scene or a historic building is captured, there may be some fence or fence-like structures located for protection. Such undesirable structures may affect the aesthetic appeal of the picture. So many editing tools and inpainting techniques are proposed, for the restoration of the damaged portion and holes left after removal of the undesired structures from the image.

In recent years, depth cameras became cheap, so researchers can use depth cameras in many applications. In computer vision, depth images can provide more useful information. However, there are some problems in depth images: such as holes, incomplete edge and temporal random fluctuations. Recently, Microsoft released a Kinect depth sensor inserted camera [1] that can provide both RGB color information and depth information. But, the RGB information provided by Kinect has lower resolution than conventional cameras.

The Kinect depth sensor depends on AC power and it has a limited measurable depth range. It shows that future cameras will collect both RGB and depth information in high resolution would become the default and basic function. This aspect will reshape the current computer vision research by incorporating additional depth information. For example, recently Kinect has been successfully used to address the activity recognition and challenging person-tracking problems.

Segmentation of the undesired objects from the image is so difficult by considering the depth information alone. The complications that make this segmentation difficult are as follows: the limited resolution of the depth sensor usually accurate only in the range from 1.2 to 3.5 m. It cannot capture the depth information of the thin and very small objects accurately. The depth sensor is usually very sensitive to highly reflective surfaces, e.g., mirrors or polished surfaces. The depth information provided by Kinect is inaccurate or even lost (setting as the maximum value).

In this paper, for accurate detection and segmentation of the fence-like structures from the RGB-Depth (RGBD) images, both depth information and color information is used. After segmentation, the fence mask is refined by using morphological techniques. After the removal of the fence from the image, the holes are filled by using a modified wavelet-based exemplar inpainting technique.

2. RELATED WORK

The main challenge in image de-fencing is the segmentation of the undesired structures from the image. After removal of the undesired structures from the image the gaps are filled by using a modified wavelet-based exemplar inpainting technique. In this section, a brief overview of these two steps is explained.

Liu, et al. introduced the word "Image de-fencing" [2]. The fence in the foreground is searched in the form of deformed lattices separated according to the regularity of appearance. Fence area in the image is filled nearby area using the Criminisi inpainting algorithm [3]. The fence detection [4] uses online learning and classification and a multi-view inpainting technique is proposed for the restoration. In [5, 6], the fence-like structures are removed from multi-focus images. These techniques require multiple images with a flashlight on and off and foreground image focused and out of focused. In [7], the fence is separated by frequency domain processing before applying wavelet transformation and

ITEE Journal
Information Technology & Electrical Engineering
ISSN: - 2306-708X

segmented through a support vector machine classifier. The occluded fence is filled with the Criminisi inpainting algorithm. The detection technique proposed [8] will detect the fences having both regular and irregular patterns. For the restoration of fence occlusions hybrid inpainting algorithm is proposed. In [9], the fence like structures is removed from RGBD images. In this approach, candidate regions are identified in fence-like structures using the depth information. These help to find strokes in the foreground and background in the RGBD images and which are used by graph cut algorithm to segment the undesired structures. After removing the unwanted fence-like structure's image is restored by using the exemplar-based inpainting technique. In [10], a multimodal approach is used to remove the fence from a video sequence.

Image inpainting is used in image editing, restoration and completion applications. After removing the undesired structures form the image the appearance of the image is enhanced by inpainting techniques these are categorized as three types [11, 12]. The first method uses statistical methods and correspondence maps [13, 14]. These methods produce good results when the background around the hole is having regular texture. Second one is based on diffusion; which uses partial differential equations [15, 16, 17, 18]. In these methods, the filled area is assumed as a non-textured region. The information near the boundary is propagated smoothly into the hole using the diffusion process. These methods may introduce blurring artifacts when we want to fill the larger textured region. In the third method, the removed foreground is filled with the matching patch in the background area. The matching patch is searched in the entire image region. In [3, 19], the filling order is determined by considering gradient and confidence values. Many exemplar-based algorithms are used for texture synthesis [20, 21, 22, 23]. Multi-scale analysis such as wavelets [24] and frame lets [25] are also used for image inpainting.

3. PROPOSED METHOD

The proposed image de-fencing technique works in two steps, starting with fence detection and segmentation by using the color information and depth information. The morphological methods are used for the refinement of the segmented fence [8]. In the second step, the fence region is inpainted by using the wavelet-based exemplar inpainting technique [26].

3.1 Segmentation of the Fence

The initial mask of the railings or bars is extracted by converting the RGB image into HSV color space [27]. The histogram is found for each channel of the HSV image to identify the intensity variations present in the color image. Based on these variations threshold values are found and the image is segmented using the Eqn. (1).

$$g(x,y) = \begin{cases} 1, & \text{if } T_{i1} \leq f(x,y) \leq T_{i2} \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

Where T_{i1} and T_{i2} are the minimum and maximum threshold values, Ti is used to specifying each layer of the color space {i=1, 2, 3}. The selection of the threshold values depends on user interpretation. A wrong selection of the threshold values may lead to undesired segmentation. The mask obtained by segmenting Fig.1a is shown in Fig.1c.

The depth information of image is used to extract the mask accurately. The depth image obtained by the Kinect has holes. The holes are filled by using the counter-based segmentation. The mask obtained by segmenting Fig.1b is shown in Fig.1d. The ground truth image of the Fig.1a is shown n Fig.1e. The ground truth image is obtained by carefully marking the boundary of the image manually. The fence masks obtained by RGB image and depth image are added. This is considered a finial mask image.

The final mask has contained some misclassifications that occurred due to the color similarity between the foreground and background. These misclassifications are categorized as false positives and false negatives. These misclassifications are reduced using morphological operations [8]. False positives are reduced by using connected component analysis and false negatives are included by dilation. The final fence mask is refined and the mask obtained after refinement is shown in Fig.1f. To remove the unwanted bars from the image, the input RGB image in Fig.1a is masked with the refined mask. One can observe the masked image obtained by the proposed approach in Fig.1g and the masked image obtained by [9] in Fig.1h

.

3.2. Image Inpainting

Restoration of the occluded region is very challenging in digital images. These occlusions may extend along the entire region of the image. Inpainting is one of the widely used techniques for the restoration of the missing portions in the images. In this paper, the modified wavelet-based exemplar inpainting technique is used for the restoration of the occlusion.

The masked image and mask are resized into 'n' decomposition levels using Eqn. (2) [31].

$$W_{\phi}(j_0,m,n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)\phi_{j_0,m,n}(x,y)$$
$$W_{\psi}^{i}(j,m,n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)\psi_{j,m,n}^{i}(x,y), i = \{H,V,D\} \quad (2)$$

The restoration of the occluded region in the image is done first at the higher levels by using a modified exemplar-based inpainting technique [27]. The modified exemplar-based inpainting technique uses local window constraints. In this approach, the best patching match is searched in the local window instead of the whole image. This reduces the computational cost and also reduces the selection of the wrong exemplars.

ITEE Journal
Information Technology & Electrical Engineering

ISSN: - 2306-708X

Leftover pixels at the higher levels are restored in the lower levels. Lower levels are obtained by inverse wavelet transformation shown in Eqn. (3) [31].

$$f(x,y) = \frac{1}{\sqrt{MN}}\sum_{m}\sum_{n} W_{\phi}(j_{0},m,n)\phi_{j_{0},m,n}(x,y) + \frac{1}{\sqrt{MN}}\sum_{i=H,V,D}\sum_{j=j0}\sum_{m}\sum_{m} W_{\psi}^{i}(j,m,n)\psi_{j,m,n}^{i}(x,y) \quad (3)$$

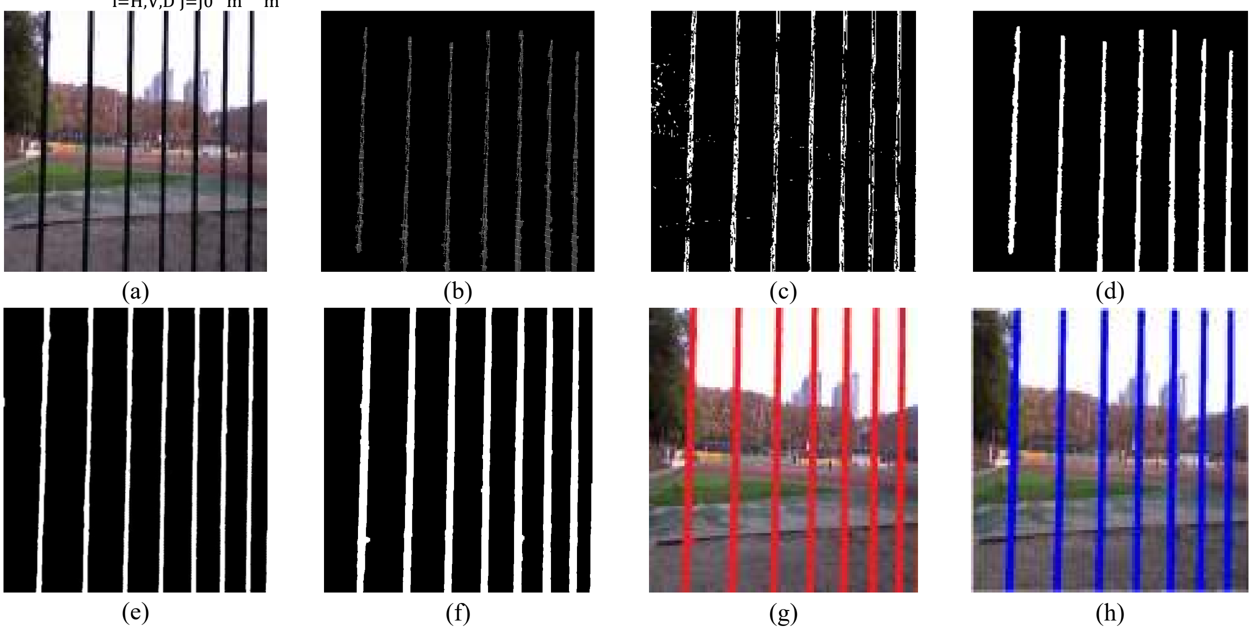

Fig.1Extraction of the fence mask: (a) Original RGB image, (b) Depth map image, (c) Fence mask obtained from the RGB image, (d) Refined depth map fence mask, (e) Ground truth image, (f) Fence mask obtained after refinement (g) Masked fence image obtained by proposed approach, (h) Masked fence image in [9].

4. RESULTS AND DISCUSSIONS

The proposed algorithm is tested on a set of 120 Kinect images [28]. The values used in the experiments are found empirically. The size of the structuring element used for the dilation is ranging from 2×2 and 5×5. The number of decomposition levels n is considered as 2. The searching window used for the selection of the matching patch is 14×14. The patch size used for the restoration of the fence occluded area is 8×8.

Due to the limitations of Kinect Sensor [9], it is unable to provide the depth information of edges in the image. Let us consider examples of Fig.1a such as the disconnected vertical bars that are not available at the right and left edge portions of the image. It can be observed as in Fig.1g and Fig.1h that the proposed algorithm and [9] can get the complete fence mask.

Let us consider other examples of Fig.2a and Fig.2c such as the disconnected vertical bars that are available at the right edge portion of the image. The depth information of the 9th vertical bar in Fig.2a and the 13th vertical bar in the Fig.2c are not available. One can observe this in Fig.2b and Fig.2d. In this case, fence detection technique [9] unable to detect the 9th vertical bar of Fig. 2a and the 13th vertical bar of the Fig.2c. One can observe this in Fig.2e and Fig.2g. However, the proposed algorithm can detect. This can be observed in Fig.2f and Fig.2h, in the fence mask images.

In the Fig.2i, the portion of the last vertical bar is not inpainted. However, by comparing Fig.2a and Fig.2j the last vertical bar is inpainted effectively using the proposed algorithm. In the Fig.2k the last vertical bar is not inpainted. However, comparing Fig.2c and Fig.2l the last vertical bar is inpainted effectively using the proposed algorithm.

Let us consider some other examples of Fig.3 (a-d), the railings are connected either horizontal or vertical and extended up to the edges of the image. One can observe these in Fig.3 (e-h) a major portion of the depth information of the images are available. In this case, the proposed algorithm and [9] can detect the complete fence mask. Fig.3 (i-l) Ground truth images of Fig.3 (a-d) respectively. For the examples considered in Fig.3 (a-d) the fence masks obtained by the proposed approach are shown in Fig.3 (m-o).

It is shown that the proposed algorithm works for the three different cases as discussed above. The first case, when

Volume 8, Issue 5
October 2019

ITEE Journal
Information Technology & Electrical Engineering

ISSN: - 2306-708X

the disconnected bars are not at the right and left edge side of the image and the depth information of all bars are available. In the second case, when the disconnected bars are at the right and edge of the image. The depth information of the edge side bars is not available. However, In this case, the algorithm [9] is not able to get the complete mask. In the third case, when the connected railings in the image are considered and depth information of the fence like structures are available.

4.1 Quantitative Evaluation

The performance of the proposed fence detection technique is evaluated by using some statistical parameters. For this, the ground truth fence mask of all the images is obtained manually by carefully marking the boundaries of the fence.

4.4.1 Quantitative evaluation parameters

The performance of the proposed fence-like structure detection technique is evaluated using different statistical parameters. Sensitivity (recall or true positive rate) is the average of properly classified positives to actual positives. Specificity or true negative rate is the average of the properly classified negatives to actual negatives. These are computed using Eqn. (4) [29]

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}\ ; \quad \text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{4}$$

A higher value of Sensitivity and Specificity indicates better classification. The degree of closeness and repeatability can be calculated using *Precision* and *Accuracy* respectively. These are computed using the following Eqn. (5) [29]

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}\ ; \quad \text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}} \tag{5}$$

However, accuracy and precision measure deceptive in the cases when the positive and negative classes are distributed. In these cases, F-measure is computed. It is trade-off between precision and recall. This is the weighted harmonic mean. This is computed using the following Eqn. (6) [29].

$$\text{F} - \text{measure} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \tag{6}$$

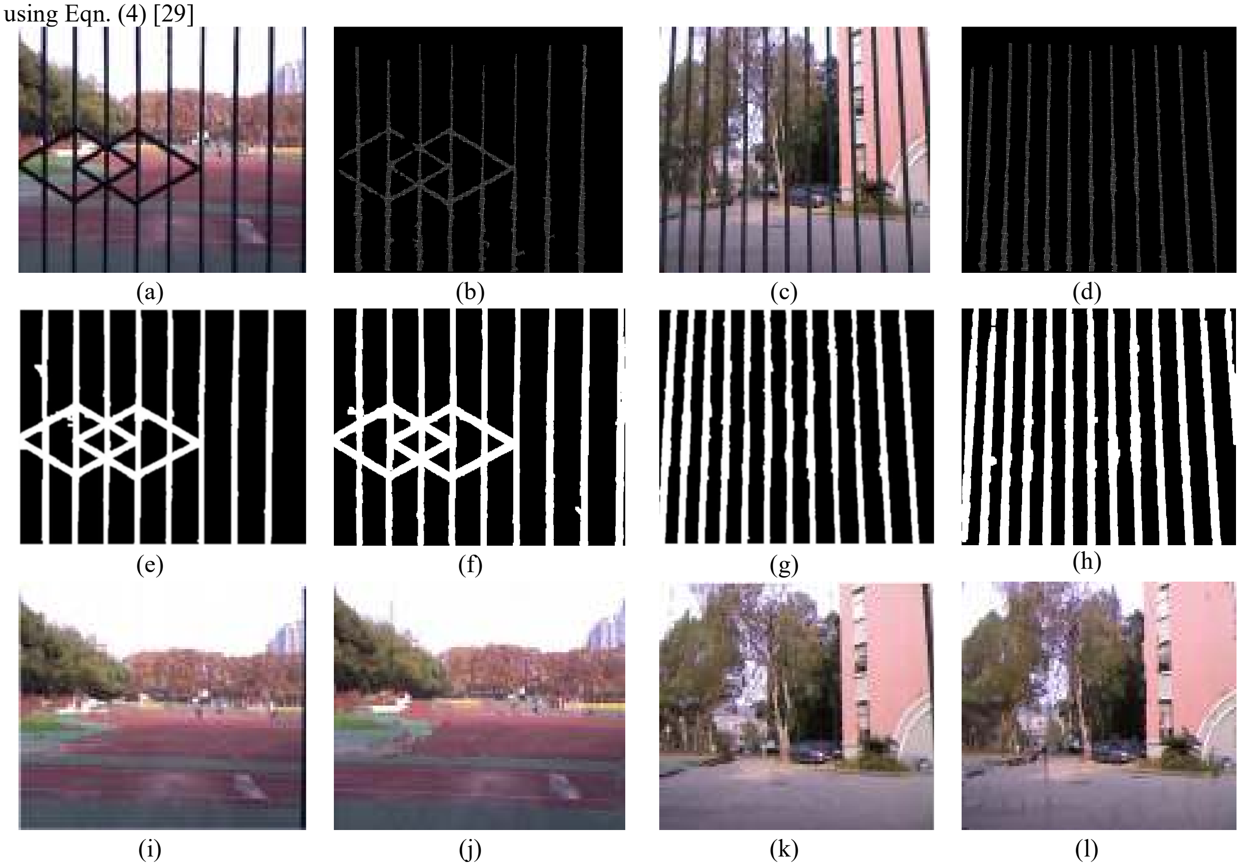

Fig.2.Comparison of the results: (a, c) Input RGB images, (b, d) Depth map images, (e, g) Fence mask in [9], (f, h) Fence mask obtained with proposed approach, (i, k) Inpainted images in [9], (j, l) Inpainted images obtained with the proposed approach.

Volume 8, Issue 5
October 2019

ITEE Journal
Information Technology & Electrical Engineering

ISSN: - 2306-708X

4.4.2 Quantitative evaluation results

The statistical measures presented in the previous section are computed for all the images described in the paper. The results are presented in *Table 1*. The analysis of statistics reveals that the proposed method achieved more than 97 % average accuracy and 99% average precision. The analysis of the accuracy and precision proves that the proposed method gives a significantly lower deviation of measurement from the ground truth data. The performance analysis of a fence like structure segmentation technique is reflected by sensitivity and specificity parameters. It is evident from the results that the proposed segmentation technique is reliable with an average sensitivity of more than 90% and specificity of more than 99% of all the figures presented in the paper. The average F_1 score of all the figures is more than 95% for the proposed method.

Table 1: Quantitative analysis measured in terms of Accuracy, Precision, Specificity, F- measure and Sensitivity.

Exp	Accuracy	Precision	Specificity	F-measure	Sensitivity
Fig.1a	98.97	99.96	99.76	96.06	92.46
Fig.2a	92.51	99.95	99.81	93.09	87.11
Fig.2c	96.63	99.74	99.25	90.56	82.93
Fig.3a	96.58	99.84	99.43	94.7	90.06
Fig.3b	99.34	99.88	99.17	96.48	93.3
Fig.3c	97.94	99.99	99.99	96.68	93.57
Fig.3d	99.81	99.93	99.62	99.04	93.57
Average	**97.39**	**99.89**	**99.57**	**95.23**	**90.42**

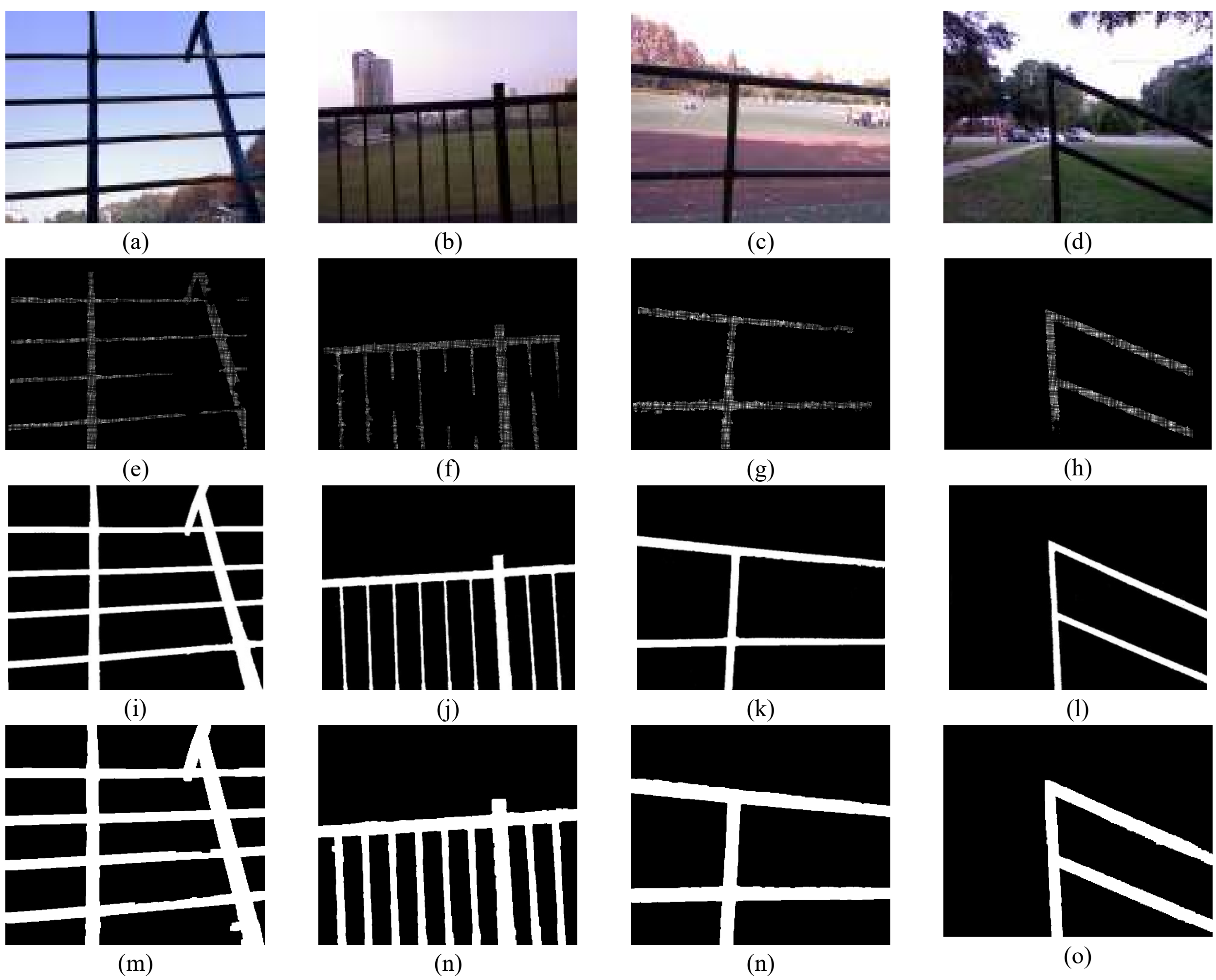

Fig.3. (a-d) Input RGB images, (e-h)Depth map images, (i-l) Ground truth images,(m-o)Fence mask obtained with the proposed approach.

Volume 8, Issue 5
October 2019

ITEE Journal
Information Technology & Electrical Engineering

ISSN: - 2306-708X

5. CONCLUSIONS

This paper presents a new approach for semi-automatic segmentation and restoration of railings or bars from RGBD images provided by Kinect. In this approach, railings or bars are segmented using the threshold based technique from the RGB image. The holes in the depth image are filled using counter-based segmentation. Both the segmented results are combined to form the mask of railings or bars. The misclassifications are arisen due to the color similarity of the fence-like structures and background. These misclassifications are mitigated or eliminated using morphological connected components and dilation operation. The mask obtained after the elimination of misclassifications is used as the final mask. The railings or bars in the RGB image are masked by the final mask. Finally, the masked railings or bars in the image are restored by using a modified wavelet-based exemplar inpainting technique.

As a sample application, the proposed method is tested on a set of Kinect images for the removal of railings or bars in the RGBD images. The proposed approach achieved good results in terms of segmentation of railings or bars from the image.

REFERENCES

[1] Shotton, J., Fitzgibbon, A., Cook, M., Sharp, T., Finocchio, M., Moore, R., Kipman, A., Blake, A.: Real-time human pose recognition in parts from single depth images. In: CVPR (2011)

[2] Liu, Y., Belkina, T., Hays, J., Lublinerman, R.: Image de-fencing. In: Proceedings of the IEEE CVPR, pp. 1–8 (2008)

[3] Criminisi, A., Perez, P., Toyama, K.: Region filling and object removal by exemplar-based image inpainting. IEEE Trans. Image Process. 13(9), pp. 1200–1212 (2004)

[4] Park, M., Brocklehurst, K., Collins, R.T., Liu, Y.: Image de-fencing revisited. In: Proceedings of the Asian Conference on Computer Vision (ACCV), pp. 422–434 (2011)

[5] Yamashita, A., Matsui, A., Kaneko, T.: Fence removal from multi-focus images. In: International Conference on Pattern Recognition (ICPR'10) (2010)

[6] Yamashita, A., Tsurumi, F., Kaneko, T., Asama, H.: Automatic removal of foreground occlude from multi-focus images. In: IEEE International Conference on Robotics and Automation (ICRA'12) (2012)

[7] Hettiarachchi, R., Peters, J., Bruce, N.: Fence-like quasi-periodic texture detection in images. Theory and Applications of Mathematics & Computer Science 4 (2), pp. 123-139 (2014)

[8] Farid, M.S., Mahmood, A., Grangetto, M.: Image de-fencing framework with hybrid inpainting algorithm. Signal, Image Video Process. 10(7), 1193–1201 (2016)

[9] Qin Zou, Yu Cao, Qingquan Li, Qingzhou Mao, Song Wang: Automatic inpainting by removing fence-like structures in RGBD images. Machine Vision and Applications 25:1841–1858 (2014)

[10] Jonna, S., Voleti, V.S., Sahay, R.R., Kankanhalli, M.S.: A multimodal approach for image de-fencing and depth inpainting. In: Proc. Int. Conf. Advances in Pattern Recognition. (2015) 1-6.

[11] Bugeau, A., Bertalmío, M., Caselles, V., Sapiro, G.: A comprehensive framework for image inpainting. IEEE TIP19(10), 2634–2645 (2010)

[12] Kwatra, V., Schodl, A., Essa, I., Turk, G., Bobick, A.: Graphcut textures: image and video synthesis using graph cuts. In: SIGGRAPH, pp. 277–286 (2003)

[13] Ashikhmin, M.: Synthesizing natural textures. In: ACM Symp. Interactive 3D Graphics (2001)

[14] Demanet, L., Song, B., Chan, T.: Image inpainting by corresponding dence maps: a deterministic approach. In: UCLACAMR, Tech.Rep. (2003)

[15] Bertalmío, M., Sapiro, G., Caselles, V., Ballester, C.: Image inpainting. In: SIGGRAPH (2000)

[16] Cai, J., Chan, R., Shen, Z.: A framelet-based image inpainting algorithm. Appl. Comput. Harmon. Anal. 24(2), 131–149 (2008)

[17] Chan, T.F., Shen, J.: Nontexture inpainting by curvature-driven diffusions. J. Vis. Commun. Image Rep. 12(4), 436–449 (2001)

[18] Tschumperlé, D.: Fast anisotropic smoothing of multi-valued images using curvature-preserving PDE's. IJCV 68(1), 65–82 (2006)

[19] Criminisi, A., Pérez, P., Toyama, K.: Object removal by exemplar-based inpainting. In: CVPR (2003)

[20] Efros, A., Freeman, W.: Image quilting for texture synthesis and transfer. In: SIGGRAPH, pp. 341–346 (2001)

[21] Efros, A., Leung, T.: Texture synthesis by non-parametric sampling. In: ICCV, pp. 1033–1038 (1999)

ITEE Journal
Information Technology & Electrical Engineering

ISSN: - 2306-708X

[22] Liang, L., Liu, C., Xu, Y., Guo, B., Shum, H.: Real-time texture synthesis by patch-based sampling. In: SIGGRAPH (2001)

[23] Wei, L., Levoy, M.: Fast texture synthesis using tree-structured vector quantization. In: SIGGRAPH (2000)

[24] Dong, B., Ji, H., Li, J., Shen, Z., Xu, Y.: Wavelet frame based blind image inpainting. Appl. Comput. Harmon. Anal. 32, 268–279 (2012)

[25] Cai, J., Chan, R., Shen, Z.: A framelet-based image inpainting algorithm. Appl. Comput. Harmon. Anal. 24(2), 131–149 (2008)

[26] Padmavathi, S., Priyalakshmi, B., Soman, K.P., "Hierarchical Digital Image Inpainting Using Wavelets", Signal & Image Processing: An International Journal (SIPIJ) 3(4), pp. 85–93(2012)

[27] M.Vralakshmamma, T.Venkateswarlu, "Image De-fencing using Wavelet Based Inpainting Technique" Int. J. Imag. Robot. (IJI). 18(4), pp. 49-63 (2018).

[28] Kinect test data set https://sites.google.com/site/qinzoucn/document

[29] Farid, M.S., Lucenteforte, M., Grangetto, M., "DOST: a distributed object segmentation tool" Multimed.Tools Appl. (2017) 1-24.

Mrs**. M.Vaalakshmamma** received the B.Tech and M.Tech Degrees in E.C.E from Jawaharlal Nehru Technological University Anantapur, Ananthapuramu, India in 2011 and 2014 respectively. Currently pursuing a Ph.D. at S.V. University College of Engineering, Tirupati. She had few publications in international journals. Her research areas include Image inpainting and image de-fencing.

Dr. T.Venkateswarlu received the B.Tech and M.Tech Degrees in E.C.E from S.V. University College of Engineering Tirupati, India in 1979 and 1981 respectively. He received a Ph.D. degree in Electrical Engineering from the Indian Institute of Technology, Madras in 1990. Currently working in the department of E.C.E, SVUCE, Tirupati. During 1986-89 he was a QIP Research Scholar at the Department of Electrical Engineering, Indian Institute of Technology, Madras. He had many publications in IEEE transactions and various international reputed journals. His teaching and research interest are in the areas of digital systems, communications, and image processing and multidimensional digital filters.