Journal of AI and Data Mining
Vol 7, No 4, 2019, 563-574
DOI: 10.22044/JADM.2019.7278.1864

Depth Improvement for FTV Systems Based on the Gradual Omission of Outliers

H. Hosseinpour[1], A. Mousavinia[2]* and M. A. Pourmina[3]

1. Department of Electrical Engineering, Science and Research branch, Islamic Azad University, Tehran, Iran.
2. Department of Computer Engineering, K.N Toosi University, Tehran, Iran.
3. Department of Electrical Engineering, Science and Research branch, Islamic Azad University, Tehran, Iran.

Received 17 July 2018; Revised 29 December 2018; Accepted 26 March 2019
**Corresponding author: moosavie@eetd.kntu.ac.ir(M. Mousavinia).*

Abstract
The virtual view synthesis is an essential part of computer vision and 3D applications. A high-quality depth map is the main problem with the virtual view synthesis because as compared to the color image, the resolution of the corresponding depth image is low. In this paper, an efficient and confided method based on the gradual omission of outliers is proposed to compute the reliable depth values. In the proposed method the depth values that are far from the mean are gradually omitted. By comparison with other state of the art methods, the simulation results show that, on average, PSNR is 2.5dB (8.1%) higher, SSIM is 0.028 (3%) more, UNIQUE is 0.021 (2.4%) more, the running time is 8.6s (6.1%) less, and the wrong pixels are 1.97 (24.8%) less.

Keywords: *Virtual View Synthesis, Epipolar Line, Reliable Depth, Gradual Omission of Outliers, Hole Filling.*

1. Introduction

Recently, free viewpoint television (FTV) has been more considered because this technology allows the user to freely choose the desired view of the scene and receive more intense feelings while watching TV [1]. In FTV systems in order to achieve the different views, many cameras are required around the scene; therefore it leads to having a large memory and high bandwidth for data transmission. Also with a large number of cameras, it cannot cover any arbitrary view of the scene. A key technical approach to the solution of this problem is the virtual view synthesis. This important technique in FTV systems decreases the number of cameras and covers any desired view of the scene [2]. The depth-image-based rendering (DIBR) algorithm is a known and popular method for rendering virtual views [1,3]. In the DIBR algorithm, a depth map is obtained by different stereo-matching methods [4]. Using the depth map and image warping algorithm, the virtual view image is synthesized. In most cases, the depth map is improved by different methods [5,6]. In [7], the depth map has been reported to improve by a median filter, and then the textures are warped to virtual view using depth information. In [8], edge detection has been implemented using morphological operations to refine the disparity map and then holes are filled using the refined disparity map. In [9], after depth pre-processing and depth-based 3d warping, illumination and color differences between the two images are reduced by histogram matching. After that, the intermediate view image is obtained based on the view blending, and it is completed by the hole-filling process using depth-based in-painting. In [10], a real-time IBR system is presented for GPU using a stream-centric local stereo-matching algorithm. In [11], after applying the morphological operations on the depth images, holes are filled with image in-painting, and the virtual view is synthesized using the improved image blending. A probability-based rendering method is presented in [12] by formulating the rendering process as an image fusion so that the intermediate view is synthesized by adaptively blending textures of all possible matching pixels. In [13], the depth values have been improved by the probabilistic weighting of each depth pixel. In [14], the depth blending problem has been resolved using the depth super-resolution method. In [15],

the depth map fusion strategy has been used to refine low-resolution depth maps. In [16], the view synthesis is based on a minimum spatial distance and correspondence field.

The main purpose in the DIBR algorithm is to gain an accurate and high- quality virtual view image. In this way, stereo-matching is an important part. Since this step is the first step and affects the next steps so that a small error of disparity results in wrong depth value, the wrong position is found in the virtual view image, and the image is distorted. Also a low-resolution depth map may be caused by noise and occlusion [15]. In this paper, an efficient and confident method based on the gradual omission of outliers has been presented to obtain reliable depth values among n-views. The rest of the paper is organized as what follows. In Section 2, the proposed method is completely described. Section 3 shows the simulation results for the proposed method. Section 4 is devoted to the concluding notices.

2. Proposed method description

In this section, the virtual view synthesis is described based on the optimization of depth values. Assume that n-views of the scene are achievable, as shown in figure 1, and the purpose is to synthesize of virtual view with high quality. The proposed method consists of several steps, as shown in figure 2.

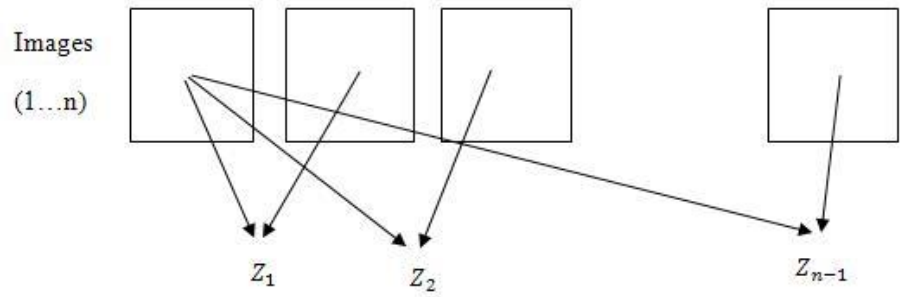


Figure 1. Different views arrangements and corresponding depth values.

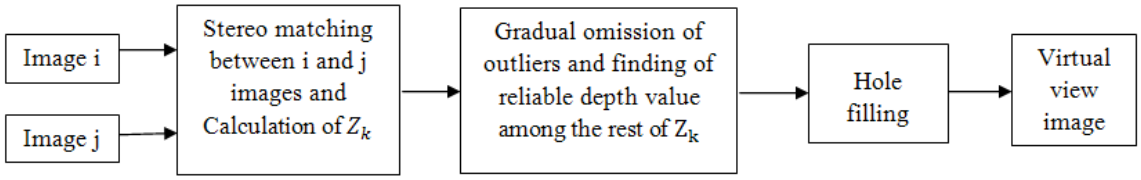


Figure 2. Block diagram of virtual view synthesis by the proposed method.

The first step is stereo-matching based on the Sum of Absolute Difference (SAD) algorithm [17] and epipolar line constraint. Then the depth values are found. These depth values are related to the same point in space, and they should be closed to each other. Therefore, the reliable depth value is chosen based on the gradual omission of outliers in the next step. There are two notices for choosing a reliable depth value: 1) the depth values should be located in a specific range 2) if there is just one depth value, it is chosen as a reliable one. In the last step, holes are filled based on the horizontal and vertical interpolation algorithm, and the virtual view image is synthesized. In what follows, the details of the proposed method are explicitly explained.

2.1. Reliable depth value computation

A reliable depth value is an important quantity for gaining a high-quality virtual view. The depth values are calculated using the accurate disparity of the corresponding pixels between the i and $i+1$ images, as in Equation (1).

$$Z_i = \frac{Bf}{D_i} \qquad i=1,2,3,.....,n-1 \tag{1}$$

where the parameters D_i , f , B , and Z_i represent the disparity value, focal length, baseline, and depth value, respectively. Then the mean of the depth values is calculated and the depth values that are close to it are chosen for the next steps, as follows:

$$|Z_i - \mu| < T \qquad i=1,2,3,.....,n-1 \tag{2}$$

In Equation (2), μ and T are the mean of the depth values and threshold, respectively. Suppose that k depth values satisfy Equation (1). Therefore, the reliable depth value is computed by Equation (3).

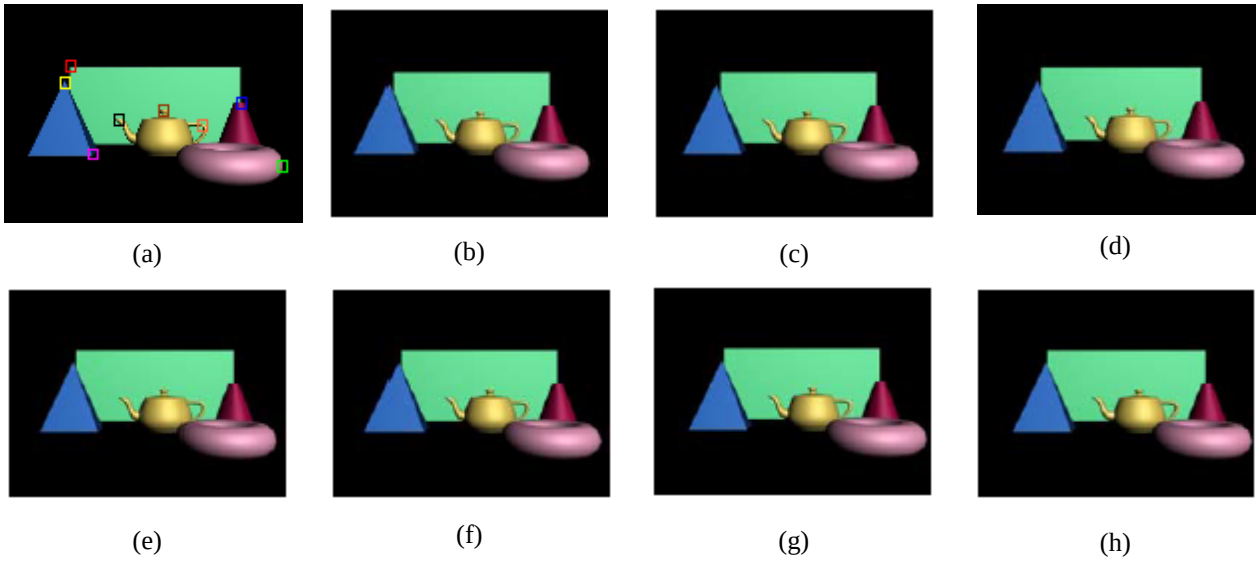
$$Z_{rel} = \frac{1}{W_t} \sum_{j=1}^{k} W_j Z_j \quad (3)$$

$$W_j = \left| Z_j - \mu \right|^{-1} \quad (4)$$

In this equation, W_j is the weighted coefficient of the depth values that are calculated by Equation (4). Based on this equation the depth values that are closer to the mean of the depth values have larger weighted coefficients.

W_t is the sum of the weighted coefficients, computed as follows:

$$W_t = \sum_{j=1}^{k} \left| Z_j - \mu \right|^{-1} \quad (5)$$



(a) (b) (c) (d)

(e) (f) (g) (h)

Figure 3. Different views of "Collection".

2.2. Influence of threshold on reliable depth value

In the proposed algorithm, the threshold is an important quantity for determination of the reliable depth value so that by changing the threshold value, the number of depth values that satisfy Equation (1) is changed and the reliable depth value is varied. In this section, 8-views of "Collection" obtained by the 3DS Max software are used, as shown in figure 3. 3DS Max simulates a camera with a focal length of 35mm and image resolution of 640×480 pixels. The SAD algorithm was applied to stereo-images (Figure 3-(a)) and the other views to find the corresponding pixels considering the horizontal disparity constraint and then depth map was computed for each one of the stereo-images. The reliable depth values were calculated for eight pixels by applying the different threshold values, as shown in table 1. As illustrated in this table, the proposed method can calculate the depth value properly and precisely. Based on this table, the threshold value is important, and it should be selected accurately. If it becomes large, the algorithm accepts the invalid depth values that are far from the mean and if it becomes small, the algorithm omits the valid depth values that are near the mean. To solve this problem, the gradual omission of the outliers method can be used.

Table 1. Reliable depth value using the different thresholds for "Collection".

	Reliable depth (cm)			
Points	**T=10**	**T=5**	**T=3**	**Real depth (cm)**
Red	189.39	189.39	190.31	189.85
Blue	153.37	153.58	154.55	153.9
Yellow	144.68	145.3	145.73	145
Green	120.64	120.74	121.52	121
Pink	136.5	136.7	136.17	137.2
Black	139.46	139.68	140.66	140
Brown	138.14	138.57	139.42	138.8
Orange	138.55	138.72	139.83	139.2
The mean error	0.5275	0.3837	0.6875	

2.3. Gradual omission of outliers method

In the gradual omission of the outliers method (GOOM) the invalid depth values that are far from the mean of depth values are gradually omitted. To implement this method, firstly, an initial threshold value is selected (T_1) and the depth values that are very far from the mean are omitted (Equation (6)).

$$\left|Z_i - \mu_1\right| < T_1 \quad i=1,2,3,.....,n-1 \tag{6}$$

Assume that the k depth values satisfy Equation (6); therefore, it can be written as:

$$\left|Z_j - \mu_2\right| < T_2 \quad j=1,2,3,.....,k \tag{7}$$

In Equation (7), μ_2 and T_2 are the mean of the k-depth values and the second threshold value, respectively. In this step, the depth values that are far from the mean are omitted. This process can be continued insofar as all the invalid depth values are omitted. Table 2 shows the influence of the gradual omission of the outliers method on the reliable depth value. By comparing table 1 with table 2, it can be realized that the reliable depth value has been improved by GOOM because the invalid depth values are gradually omitted and the mean of the depth values is computed precisely. Due to the digital image property, this amount of error is unavoidable. In table 3, the proposed gradual omission of outliers is compared with the other state-of-the-art methods. The reliable depth values were calculated for $T_3 = 5.5$, $T_2 = 5$, and $T_3 = 3$. It could be observed that the reliable depth values obtained by the proposed method were much better than those computed by the other methods.

Table 2. Reliable depth value using the gradual omission of outliers method for "Collection".

Points	Reliable depth (cm)		Real depth (cm)
	T_1 =10 , T_2 =5	T_1 =10 , T_2 =5 , T_3 =3	
Red	189.56	189.56	189.85
Blue	153.64	153.75	153.9
Yellow	145.3	145.21	145
Green	120.86	120.86	121
Pink	136.7	136.94	137.2
Black	139.76	139.76	140
Brown	138.57	138.57	138.8
Orange	138.82	138.94	139.2
The mean error	0.2925	0.2225	

Table 3. Comparison of the proposed gradual omission of outliers method with the other methods for "Collection".

Points	Methods					Real depth
	Used in [18]	Used in [5]	Used in [19]	Used in [15]	GOOM for T_1 =10 , T_2 =5 , T_3 =3	
Red	189.18	189.31	189.36	189.42	189.56	189.85
Blue	153.23	153.36	153.41	153.52	153.75	153.9
Yellow	144.24	144.32	144.48	144.58	145.21	145
Green	120.34	120.48	121.35	121.43	120.86	121
Pink	136.42	136.53	136.61	136.76	136.94	137.2
Black	139.13	139.28	139.41	139.53	139.76	140
Brown	138.14	138.2	138.32	138.46	138.57	138.8
Orange	138.3	138.42	138.53	138.7	138.94	139.2
The mean error	0.7462	0.6312	0.5225	0.4262	0.2225	

2.4. Hole filling

In the virtual view image, the occluded areas may be seen but they cannot be seen on both reference images. Therefore, some holes are made in the virtual view image. Large baseline and wrong depth values can yield holes in the synthesized images [19,8]. Also the digital sampling property of the image can make holes.

In this work, holes are filled using both the horizontal and vertical interpolation algorithms. In the horizontal interpolation algorithm each pixel of hole H is filled by the left and right boundary pixels, as defined below:

$$I[n]=(1-W)I_l[n]+WI_r[n] \quad \forall n \in H \tag{8}$$

In Equation (8), $I_l[n]$ and $I_r[n]$ are the intensities of the left and right boundary pixels, respectively, and W is calculated as follows:

$$W=\frac{n-m_l}{m_r-m_l} \tag{9}$$

In Equation (9), m_l and m_r are the position of the left and right boundary pixels, respectively, and W is the normalized distance between the hole left border and the pixel to be filled [20].

3. Simulation results

In this section, the proposed method is evaluated for the Middlebury Dataset images. Also some real images are captured using a canon camera with a focal length of 4.3mm and image resolution of $\mathbf{640 \times 480}$ pixels. Different algorithms were implemented by a computer with AMD Athlon 64 X2 Dual Core 5000+ of CPU and 2GB of RAM. Figure 4 shows 6-views of "Art", and the proposed method uses different stereo-images consisting of figure 4b and other views to find the corresponding pixels. In figure 5, the Receiver Operating Characteristic (ROC) curve is plotted for T_1 and T_2 selection. In the ROC curve, an error less than 1% is acceptable. Based on the ROC curve, the best threshold occurs for $T_1 = 5.5$ and $T_2 = 5$. Testing different thresholds, the best value for T_3 is 1. Finally, a reliable depth map has been shown in figure 6. In table 4, the quality of reliable depth map for "Art" has been evaluated considering different thresholds. Wrong pixels are pixels whose disparity error is larger than 6 pixels.

(a) (b) (c) (d) (e) (f)

Figure 4. Different views of "Art".

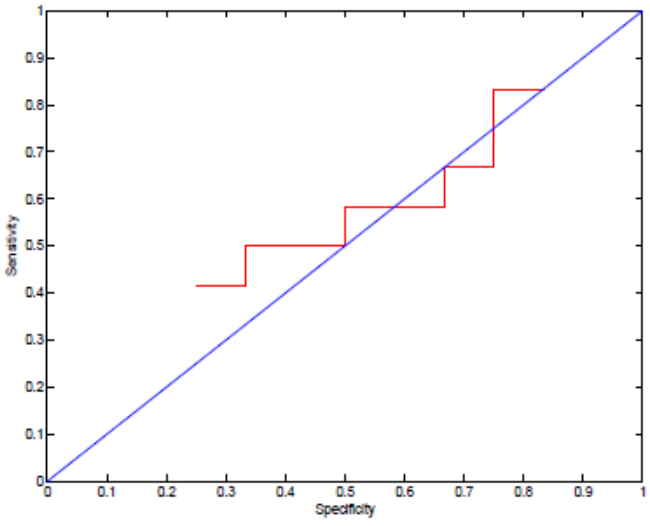


Figure 5. ROC curve for "Art" considering different thresholds of T_1 and T_2 .

(a) (b) (c) (d) (e)

Figure. 6. Reliable depth map by the proposed method considering different thresholds (a) T=10 (b) T=5 (c) T_1 =5.5 , T_2 =2 (d) T_1 =5.5 , T_2 =2 , T_3 =1 (e) ground truth.

Table 4. Reliable depth map quality for "Art" considering different thresholds.

	Different thresholds			
Image quality	**T=10**	**T=5**	**T_1 =5.5 , T_2 =2**	**T_1 =5.5 , T_2 =2 , T_3 =1**
PSNR (dB)	30.42	31.23	32.94	33.38
Wrong pixels	7.34%	6.54%	5.04%	4.52%

Figure 7 illustrates the virtual view synthesis by the proposed method for "Art" considering different thresholds. Table 5 shows the objective measure of the virtual view by the proposed method for "Art" considering different thresholds. In this table, the values for PSNR, SSIM, and UNIQUE [21] have been reported for the common area of the virtual view between the proposed method and the real one. SSIM, and UNIQUE are metrics that determine the similarity between the desired image and the ground truth. When two images are the same the value for them is 1. Moreover, wrong pixels are those whose intensity value error is more than 5% of the maximum pixel intensity value. The number of these pixels in every 1000 pixels has been mentioned in table 5.

(a) (b) (c) (d) (e)

Figure 7. Virtual view synthesis by the proposed method for "Art" considering different thresholds (a) T=10 (b) T=5 (c) T_1 =5.5 , T_2 =2 (d) T_1 =5.5 , T_2 =2 , T_3 =1 (e) ground truth.

Table 5. Objective measures of virtual view by the proposed method for "Art".

	Thresholds			
Quantity	**T=10**	**T=5**	**T_1 =5.5 , T_2 =2**	**T_1 =5.5 , T_2 =2 , T_3 =1**
PSNR (dB)	31.18	31.76	32.87	33.52
SSIM	0.96	0.96	0.97	0.98
UNIQUE	0.95	0.95	0.96	0.98
Running time (s)	112.67	110.38	126.47	131.84
Wrong pixels	7.83	7.12	6.24	5.13

Figure 8 shows the virtual view image using the other state-of-the-art methods for "Art". It is noticeable that for virtual view synthesis, different methods have been used to obtain the depth map, and the other parts are the same as the proposed method. In table 6, the virtual view image synthesized by different methods have been evaluated. Based on this table, quality of virtual view by the proposed method is higher than the other state-of-the-art methods. Also the proposed method is faster than the other methods.

(a) (b) (c) (d)

Figure 8. Virtual view synthesis by different methods for "Art" (a) used in [18] (b) used in [5] (c) used in [14] (d) used in [15].

Table 6. Objective measures of virtual view by different methods for "Art".

Quantity	Methods			
	Used in [18]	**Used in [5]**	**Used in [14]**	**Used in [15]**
PSNR (dB)	28.96	29.31	30.12	31.42
SSIM	0.93	0.93	0.94	0.95
UNIQUE	0.91	0.91	0.92	0.93
Running time (s)	145.32	146.43	147.61	148.21
Wrong pixels	9.56	8.53	7.76	6.51

Figure 9 shows different views of "Books" and the best threshold for T_1 and T_2 based on the ROC curve, as shown in figure 10 are 8.6 and 7, respectively. Also the best value for T_3 is 5.4. Figure 11 illustrates the virtual view images by different methods for "Books". Tables 7, 8, and 9 compare the proposed method with the other state-of-the-art methods for "Books", "Dolls", and "Laundry", respectively. In all cases, the best threshold values have been obtained using the ROC curves.

(a) (b) (c) (d) (e) (f)

Figure 9. Different views of "Books".

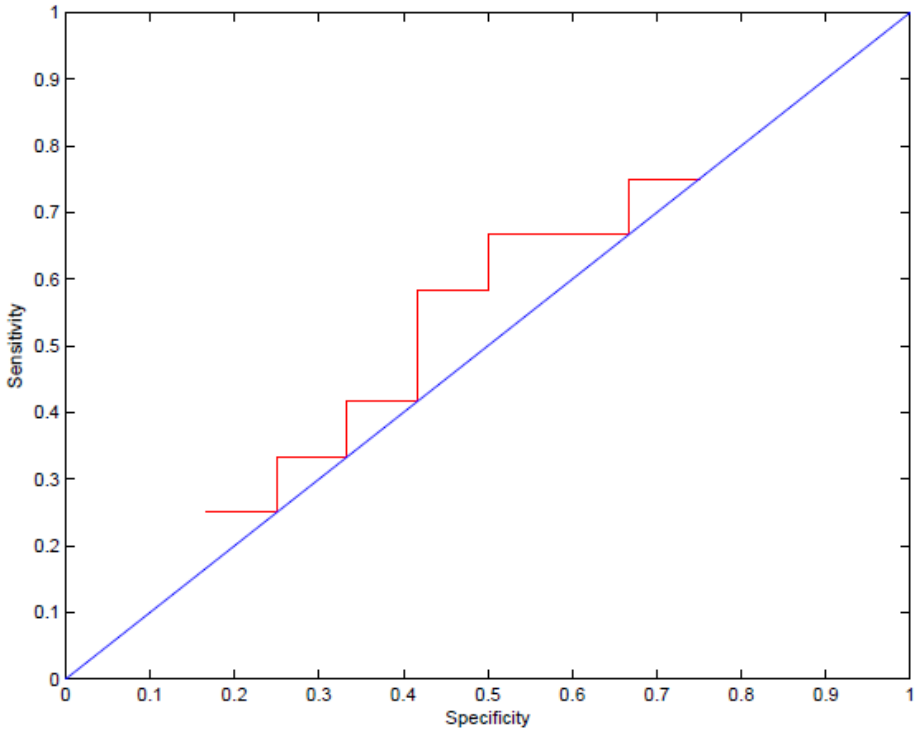


Figure. 10. ROC curve for "Books" considering different thresholds of T_1 and T_2 .

(a) (b) (c) (d) (e) (f)

Figure 11. Virtual view synthesis by different methods for "Books" (a) used in [18] (b) used in [5] (c) used in [15] (d) the proposed method for T_1 =8.6 , T_2 =7 (e) the proposed method for T_1 =8.6 , T_2 =7 , T_3 =5.4 (f) ground truth.

Table 7. Comparison of the proposed method with the other methods for "Books".

Quantity	Methods				
	Used in[18]	Used in [5]	Used in[15]	GOOM for T_1 =8.6 , T_2 =7	GOOM for T_1 =8.6 , T_2 =7 , T_3 =5.4
PSNR (dB)	29.54	30.43	31.26	32.83	32.92
SSIM	0.94	0.94	0.95	0.98	0.98
UNIQUE	0.92	0.92	0.93	0.97	0.98
Running time (s)	134.56	138.43	141.12	123.3	128.43
Wrong pixels	8.74	8.14	7.32	5.85	5.64

Table 8. Comparison of the proposed method with the other methods for "Dolls".

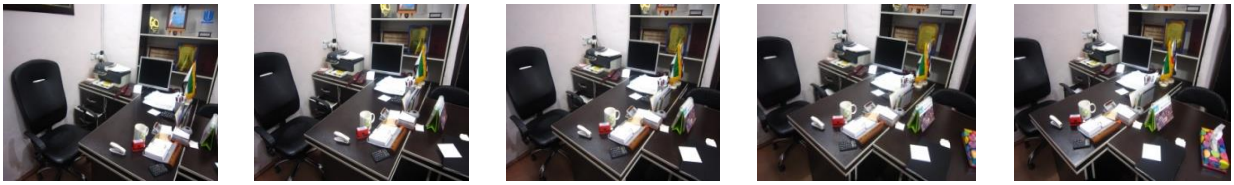
Quantity	Methods				
	Used in [18]	**Used in [5]**	**Used in [15]**	**GOOM for T_1 =8 , T_2 =6.5**	**GOOM for T_1 =8 , T_2 =6.5 , T_3 =5.4**
PSNR (dB)	28.35	29.08	30.23	31.88	32.34
SSIM	0.93	0.93	0.94	0.97	0.97
UNIQUE	0.92	0.92	0.93	0.96	0.97
Running time (s)	136.62	141.34	144.76	125.74	129.86
Wrong pixels	9.78	9.21	8.54	7.18	6.24

Table 9. Comparison of the proposed method with the other methods for "Laundry".

Quantity	Methods				
	Used in [18]	**Used in [5]**	**Used in [15]**	**GOOM for T_1 =6.8 , T_2 =5.4**	**GOOM for T_1 =6.8 , T_2 =5.4 , T_3 =3.5**
PSNR (dB)	30.12	30.56	31.45	33.14	33.38
SSIM	0.95	0.95	0.96	0.98	0.99
UNIQUE	0.93	0.93	0.94	0.97	0.98
Running time (s)	131.46	137.82	142.08	124.6	127.43
Wrong pixels	8.24	7.83	7.31	6.25	5.94

In the second part of the simulation results, the proposed method uses the real images captured by a canon camera. Figure 12 shows different views of "Office 1" and the best values for T_1 and T_2 based on the ROC curve, as shown in figure 13, are 5.8 and 2.8, respectively. Also the best value for T_3, by testing different thresholds, is 1.5. Figure 14a and figure 14b represent the virtual view images synthesized by the method used in [16] and the proposed method, respectively. In figure 14b, GOOM has been applied to the spatial positions obtained by the method used in [16], and then the reliable depth value has been calculated. Finally, the virtual view image has been synthesized using the reliable depth values. The real virtual view image has been shown in figure 14c. Figure 15 shows different views of "Office 2", and similar to "Office 1", the best values for T_1, T_2, and T_3 have been obtained. Figure 16 illustrates the virtual view images by different methods and the real one. Table 10 compares the proposed method with the method used in [16]. According to this table, it is realized that the proposed method has synthesized the virtual view image with higher quality and faster than the method used in [16].

As compared to other methods, the simulation results show that the proposed method is faster and more effective. As a result, on average, PSNR is 2.5dB (8.1%) higher, SSIM is 0.028 (3%) more, UNIQUE is 0.021 (2.4%) more, the running time is 8.6s (6.1%) less and the wrong pixels are 1.97 (24.8%) less than the other methods.



(a) **(b)** **(c)** **(d)** **(e)**

Figure. 12. Different views of "Office 1".

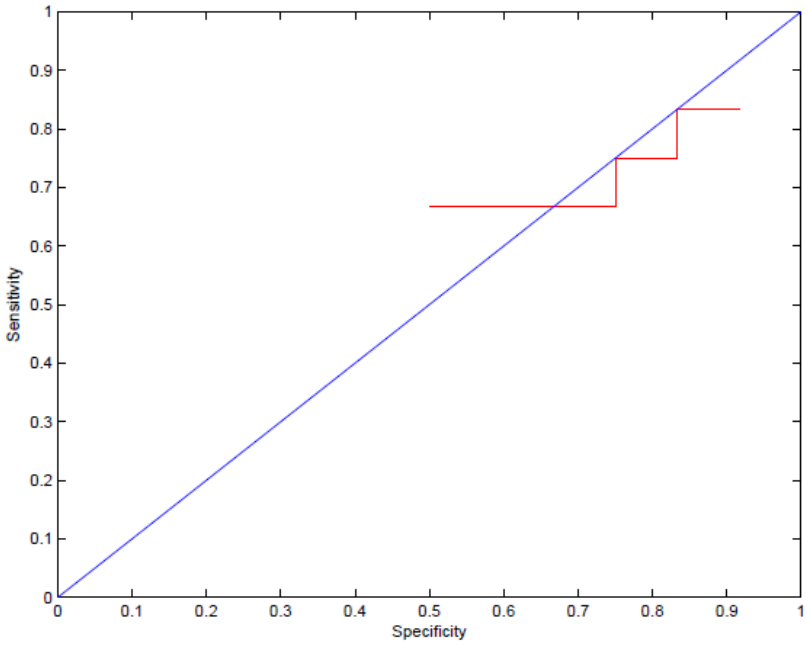
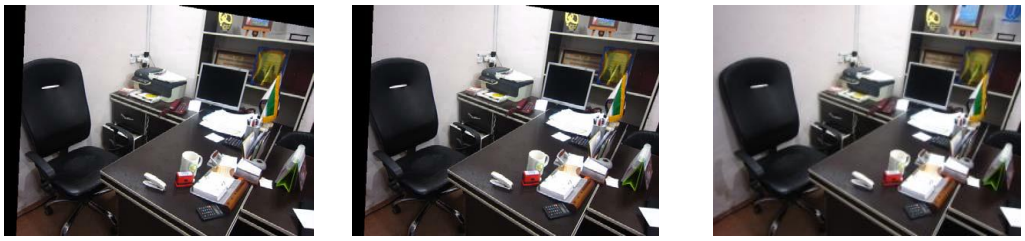
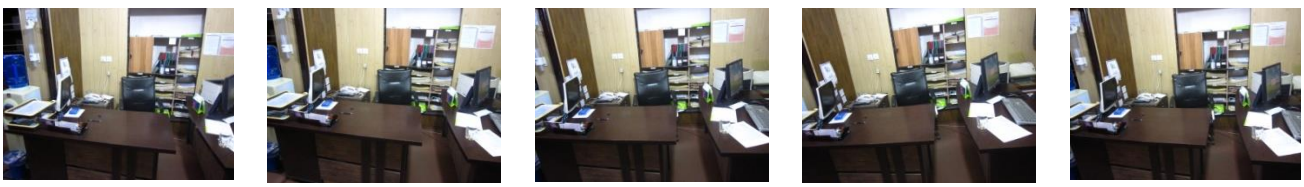


Figure 13. ROC curve for "Office 1" considering different thresholds of T_1 and T_2 .



(a) **(b)** **(c)**

Figure 14. Virtual view synthesis by different methods for "office 1" (a) the virtual view used in [16] considering $n=5$, (b) the virtual view by the proposed method considering $n=5$, $T_1=5.8$, $T_2=2.8$, $T_3=1.5$, and (c) the real image.



(a) **(b)** **(c)** **(d)** **(e)**

Figure 15. Different views of "Office 2".

(a) **(b)** **(c)**

Figure 16. Virtual view synthesis by different methods for "office 2" (a) the virtual view by the method used in [16] considering n=5 , (b) the virtual view by the proposed method considering n=5 , T_1 =3.5 , T_2 =2.1 , T_3 =1.2 , and (c) the real image.

Table 10. Comparison of the proposed method with the method used in [16] for "Office 1" and "Office 2".

Quantity	Office 1				Office 2			
	Used in [16]		GOOM for T_1 =5.8 , T_2 =2.8 , T_3 =1.5		Used in [16]		GOOM for T_1 =3.5 , T_2 =2.1 , T_3 =1.2	
	n=5	n=7	n=5	n=7	n=5	n=7	n=5	n=7
PSNR (dB)	32.93	33.41	33.34	34.47	33.46	34.45	34.38	35.17
SSIM	0.97	0.97	0.97	0.98	0.98	0.99	0.99	0.99
UNIQUE	0.66	0.67	0.69	0.72	0.97	0.98	0.98	0.99
Running time (s)	153.26	161.34	150.67	158.7	153.42	163.76	151.4	161.23
Wrong pixels	7.24	6.84	6.53	5.68	6.71	6.35	5.64	5.04

4. Conclusion

In this paper, an efficient and confident depth improvement algorithm based on the gradual omission of outliers has been presented for FTV systems. In this algorithm, at first, the corresponding depths are computed using different views, and then the reliable depth value is calculated based on the gradual omission of outlier. After that the virtual view image is synthesised. Finally, holes are filled using both the horizontal and vertical interpolation algorithms. The simulation results show that the quality of virtual view image synthesized by the proposed method is higher, and also the proposed method is faster than the other state-of-the-art methods.

5. Reference

[1] Paradiso, V., Lucenteforte, M. & Grangetto, M. (2012). A novel interpolation method for 3D view synthesis. IEEE 3DTV-Conference: The True Vision - Capture Transmission and Display of 3D Video, pp. 1-4.

[2] Zhu, C. & Li S. (2016). Depth image based view synthesis: new insights and perspectives on hole generation and filling, IEEE Transactions on Broadcasting, vol. 62, no. 1, pp. 82-93.

[3] Park, J., et al. (2012). Universal View Synthesis Unit for Glassless 3DTV, IEEE Transactions on Consumer Electronics, vol. 58, no. 2, pp. 706-711.

[4] Manap, N. & J. Soraghan. (2012). Disparity refinement based on depth Image layers separation for stereo matching algorithms, Journal of Telecommunication, Electronic and Computer Engineering, vol. 4, no. 1, pp. 51-64.

[5] Choi, J., Min, D. & Sohn, K. (2014). Reliability-based multiview depth enhancement considering interview coherence, IEEE Transactions on Circuits and Systems for Video Technology, vol. 24, no. 4, pp. 603-616.

[6] Min, D., Lu, J., & Do, M. N. (2012). Depth video enhancement based on weighted mode filtering, IEEE Transactions on Image Processing, vol. 21, no. 3, pp. 1176-1190.

[7] Lee, C., & Ho, Y. (2011). View Extrapolation Method using Depth Map for 3D Video Systems, in APSIPA ASC.

[8] Jain, A. K., et al. (2011). Efficient stereo-to-multiview synthesis, IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 889-892.

[9] Oh, K., et al. (2010). Virtual view synthesis method and self-evaluation metrics for free viewpoint television and 3D video, Imaging system and technology, vol. 20, no. 4, pp. 378-390.

[10] Lu, J., et al. (2009). Stream-centric stereo matching and view synthesis: a high-speed approach on GPUs. IEEE Transactions on Circuits and Systems for Video Technology, vol. 19, no. 1, pp. 1598-1611.

[11] Gao, Y., et al. (2013). Virtual View Synthesis Based on DIBR and Image Inpainting, PSIVT 6th Pacific-Rim Symposium, pp. 172-183.

[12] Ham, B., et al. (2014). Probability-based rendering for view synthesis. IEEE Transaction on image processing, vol. 23, no. 2, pp. 870-884.

[13] Rana, P., et al. (2015). Probabilistic Multiview Depth Image Enhancement Using Variational Inference. IEEE Journal of Selected Topics in Signal Processing, vol. 9, no. 3, pp. 435-448.

[14] Ham, B., Min, D., & Sohn, K. (2015). Depth superresolution by transduction. IEEE Transactions on Image Processing, vol. 24, no. 5, pp. 24(5), 1524-1535.

[15] Lei, J., et al. (2017). Depth Map Super-Resolution Considering View Synthesis Quality, IEEE Transactions on Image Processing, vol. 26, no. 4, pp. 1732-1745.

[16] Hosseinpour, H., & Mousavinia, A. (2019). View synthesis for FTV systems based on a minimum spatial distance and correspondence field, Multidimensional Systems and Signal Processing, vol. 30, no. 1, pp. 275-294.

[17] Manap, N., & Soraghan, J. (2011). Novel view synthesis based on depth map layers representation, 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, pp. 1-4.

[18] Min, D., Lu, J. & Do, M. N. (2012). Depth video enhancement based on weighted mode filtering, IEEE Transactions on Image Processing, vol. 21, no. 3, pp. 1176-1190.

[19] Bosc, E., et al. (2011). Towards a new quality metric for 3-D synthesized view assessment, IEEE Journal of Selected Topics in Signal Processing, vol. 5, no. 7, pp. 1332-1343.

[20] Po, L., et al. (2011). A new multidirectional extrapolation hole-filling method for Depth-Image-Based Rendering, IEEE 18th International Conference on Image Processing, pp. 2589-2592.

[21] Temel, D., Prabhushankar, M. & Alregib, Gh. (2016). UNIQUE: Unsupervised Image Quality Estimation. IEEE Signal Processing Letters, vol. 23, no. 10, pp. 1414-1418.

نشریه هوش مصنوعی و داده کاوی

بهبود عمق برای سیستم‌های FTV بر اساس حذف تدریجی داده‌های غیرمرتبط

حسام الدین حسین پور[1]، امیر موسوی نیا[2,*] و محمد علی پورمینا[1]

[1] دانشکده مهندسی برق، دانشگاه آزاد اسلامی، تهران، ایران.

[2] دانشکده مهندسی کامپیوتر، دانشگاه خواجه نصیر، تهران، ایران.

ارسال ۲۰۱۸/۰۷/۱۷؛ بازنگری ۲۰۱۸/۱۲/۲۹؛ پذیرش ۲۰۱۹/۰۳/۲۶

چکیده:

سنتز نمای مجازی یک بخش مهم در بینایی کامپیوتر و کاربردهای سه بعدی است. یک نقشه عمق با کیفیت بالا مشکل اساسی در سنتز نمای مجازی است، زیرا در مقایسه با تصویر رنگی رزولوشن تصویر عمق متناظر پایین است. در این مقاله یک روش مطمئن و موثر بر اساس حذف تدریجی داده‌های غیرمرتبط ارائه می‌شود تا مقادیر عمق مطمئن را محاسبه نماید. در روش ارائه شده مقادیر عمق که از مقدار میانگین دور هستند به طور تدریجی حذف می‌شوند. در مقایسه با روش‌های دیگر، نتایج شبیه‌سازی نشان می‌دهد به طور میانگین PSNR به مقدار 2.5dB (8.1%) بالاتر، SSIM به مقدار 0.028 (3%) بیشتر، UNIQUE به مقدار 0.021 (2.4%) بیشتر، زمان اجرا 8.6s (6.1%) کمتر و پیکسل‌های غلط 1.97 (24.8%) کمتر می‌باشد.

کلمات کلیدی: سنتز نمای مجازی، خط اپی پولار، عمق مطمئن، حذف تدریجی داده‌های غیرمرتبط، پر کردن حفره‌ها.