A Novel Approach for Depth Image Based Rendering using Depth map Smoothening

Asha Jayachandran

PG Scholar, Electronics and Communication Engineering,
Sree Chitra Thirunal College of Engineering, Trivandrum, Kerala

Preetha V.H

Asst. professor, Electronics and Communication Engineering,
Sree Chitra Thirunal College of Engineering, Trivandrum, Kerala

Abstract- Depth Image Based Rendering is a 2-D to 3-D conversion technology that is widely used in applications like 3D TV and auto-stereoscopic displays. Holes will be present in the virtual views generated if the depth map is not complete. This work proposes a new method to reduce the holes produced in the virtual views generated in Depth Image Based Rendering. Since holes occur in the rendered view when depth map changes significantly smoothening methods were proposed. Gaussian filtering affected the edges and caused distortions in the rendered view. Gaussian filtering also costs a lot in terms of computation time. This work proposes a median filter to pre-process the depth map. Median filtering preserves the edge information of the depth map and results in a high quality rendered view compared to Gaussian filtering methods. The result shows that the proposed method results in high PSNR and SSIM values compared to existing methods. The number of holes in the rendered view also reduces along with the computation time.

Keywords-Depth Image Based Rendering, 3D TV, holes, symmetric smoothening, asymmetric smoothening, median filtering

I.   INTRODUCTION

Humans have stereoscopic vision, i.e., we are able to perceive depth from what we see. This is because human eyes are placed about 7cm apart, each eye captures a slightly different views of the same scene. The brain performs some complex geometry on the left and right views to give a three dimensional viewing experience. All current 3D viewing technologies utilise this ability of the brain, by generating right virtual view and left virtual view and projecting them simultaneously.

3D-TV, which offers a realistic viewing experience, is one of the most happening viewing technologies. Currently there exist three methods to retrieve the depth map: stereo metric camera, image structure analysis and depth camera [1]. Stereo metric camera captures the same scene from two different angles. But it requires large transmission bandwidth and the depth effects are not easy to customise. Image structure analysis utilises the characteristics of the image like texture, relative positions etc. to generate the corresponding depth map. This technology has two main problems: (1) switching view angles may generate holes in the 3D Images (2) relative movement between scenes and objects may increase the computational cost and slow down the speed [2]. Depth camera makes use of depth sensors to estimate the depth. When a 3D view is generated using more than two colour images, it demands a lot in terms of bandwidth for transmission. American Three-Dimensional Television System Technologies (ATTEST) of European Information Society Technology proposed a 2D-3D conversion method that uses a colour image and its associated depth image called Depth Image Based Rendering (DIBR). Depth image is a grey scale image which conveys information related to the distance of the surfaces of the object from a view point or camera.

Occlusion and disocclusion poses a major challenge to DIBR. Occlusion occurs due to regions that are visible in the original view and become invisible in the virtual view. Holes are due to accretion (Disocclusion) of portions/regions of objects or background that would have been visible only from the new view point but not from the original location that was used in capturing the original image [3]. No information corresponding to the disoccluded regions will be available in the original image and hence would appear as empty or holes in the virtual image. Holes can be reduced by either pre-processing the depth map or post-processing the virtual views synthesised. There are several methods for addressing the hole filling problem in synthesised 3D-views. such as linear interpolation, linear

extrapolation, mirrors and image inpainting [1]. This approach introduces obvious artefacts in object boundaries with large sized holes and high texture areas known as rubber-sheet artefact.

Since holes occur when the depth changes sharply, smoothening methods were proposed earlier. Tam et al. [4] proposed a symmetric smoothening method for pre-processing the depth map. But this method affected the vertical edges in the depth map and produced geometric distortions [1]. Zhang et al. [3] proposed an asymmetric smoothening filter which preserves the vertical edges but it affected the intensity of non-vertical lines and increased the computational complexity [1]. To address these issues an adaptive edge oriented smoothening method was proposed by Lee and Effendi [5]. This method applies symmetric filtering to areas having vertical edges and asymmetric filter to regions having non-vertical edges. Although this method preserves the vertical edges the original information of the depth map is lost because of the various size and direction of holes that are created by different sizes and direction of rapid changes in the depth map, which results in a deterioration of 3-D view [1]. Chih-hsien in [1] proposed an adaptive compensation method which preserves both vertical edges and original information of the depth map. The positions of the holes are determined using a fast parallax look up table. The image is then divided into 16*16 blocks and classified into three modes according to the number of holes. Mode 1 blocks which covers region having large number of holes are inpainted in accordance with analysis of texture direction. Adaptive smoothening method is applied to Mode 2 blocks having medium number of holes. Mode 3 blocks, which contains no holes, are left as such. Adaptive compensation method produces low quality rendered images if large holes are present.

II.   PROPOSED METHOD

The flow chart of the proposed method is as follows, where a colour image and its associated depth map are input to the system [1]. The colour image and pre-processed depth image are given as input to image warping. Warping performs a mapping from the colour image to the right and left virtual image using the depth information from the depth map.
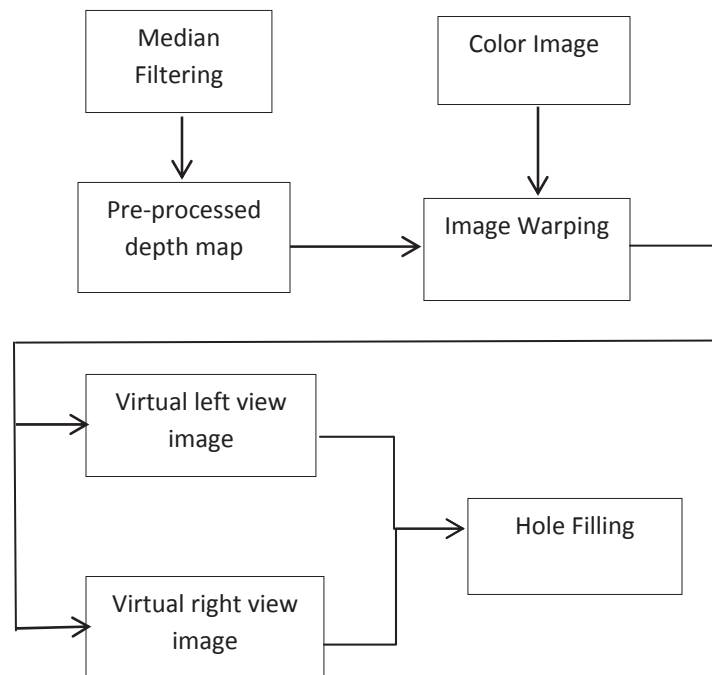

Figure 1. Flowchart of the proposed DIBR system

1. THE PRE-PROCESSING OF DEPTH MAP

Pre-processing of depth map aims at reducing the number of holes generated in the virtual left and right views. Since holes occur when depth map changes abruptly, smoothening methods were proposed initially. Since symmetric filtering destroys vertical lines and introduces geometric distortions, asymmetric filtering was proposed. Asymmetric filtering affected the non-vertical lines of the depth map. Adaptive edge oriented filtering applied asymmetric filtering to regions with vertical edges and symmetric filtering to other regions. Adaptive compensation method divided the depth map into 3 modes according to the number of holes. Mode 1 regions with maximum number of holes were inpainted in accordance with the texture direction. Adaptive filtering was applied to medium sized hole regions and Mode 3 regions were left as such. Adaptive compensation method gave poor results when the number of holes in the depth map was very high.

This work proposes a median filter to pre-process the depth map. A median filter is applied to the entire depth map to eliminate sharp change in intensities. Since median filter replaces each pixel with the median of the neighbouring values, it helps in eliminating regions with abrupt depth changes thus resulting in a high PSNR values. The edge preserving property of median filter ensures that the edges are not distorted thus resulting in a high SSIM value compared to existing methods. The blurring effect caused by Gaussian filtering is absent in this type of pre-processing techniques. This method produces virtual views having reduced number of holes compared to other method. The time taken for computation is also reduced significantly.

2. IMAGE WARPING

Image warping is a fundamental concept in DIBR, using the original depth map to compute the parallax of each pixel and then the left and right shift distance are calculated [1]. Lastly, using a shift sensor, two virtual left and right view images are generated, as [6]

$$X_l = X_c + \frac{t_x}{2}\frac{f}{Z}$$
$$X_r = X_c - \frac{t_x}{2}\frac{f}{Z}$$

(1)

where X_l and X_r represents the virtual left and right view images pixel coordinates respectively, and X_c is the original color image coordinate, where t_x denotes the baseline distance, and f represents focal length, Z is the depth distance from object to camera.

3. HOLE FILLING

Even though pre-processing of depth map reduces the number, holes will be present in the virtual left and right views synthesized via image warping. The holes can be filled by can be reduced by linear interpolation, linear extrapolation or neighbour mirroring.

III. RESULTS AND DISCUSSION

The input images are obtained from the Middlebury Stereo Vision database [7]. The input image used is as in Fig 2, where Fig 2(a) and Fig 2(b) represents the colour and depth image of Cones respectively and where Fig 2(c) and Fig 2(d) represents the colour and depth image of Woods. The input images used are of resolution 696*556. The simulation was carried out using MATLAB R2014b in MacBook Air with 1.4 GHz Intel Core i5 processor.
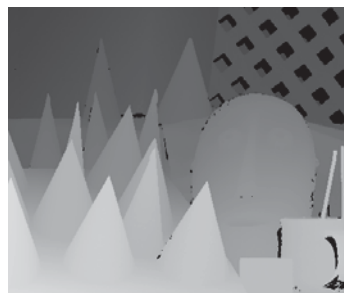
(a)   (b)

(c)   (d)

Fig 2.Test input images (a) Cones color image (b) Cones depth image (c) Woods color image (d) Woods depth image

The depth image is pre-processed applying symmetric Gaussian filter, asymmetric Gaussian filter, adaptive compensation method and the proposed method. The results of pre-processing operation are as in Fig 3 and Fig 4.
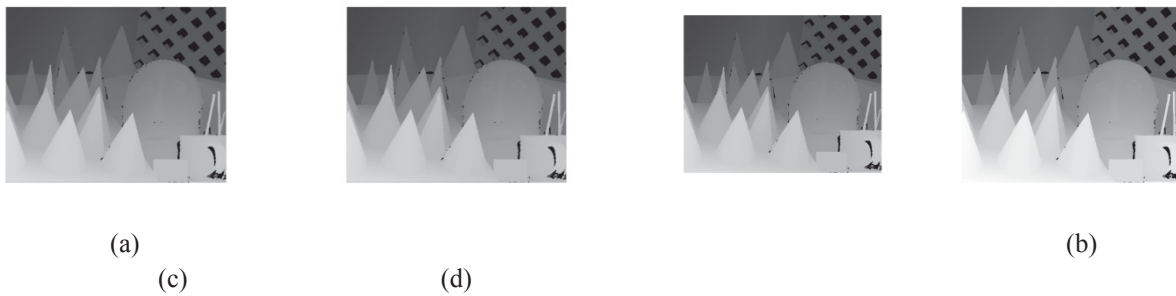
(a)   (b)
(c)   (d)

Fig 3. Preprocessed depth map of cones by (a) symmetric smoothening (b) asymmetric smoothening (c) adaptive compensation method (d) proposed method

(e)   (f)   (g)   (h)
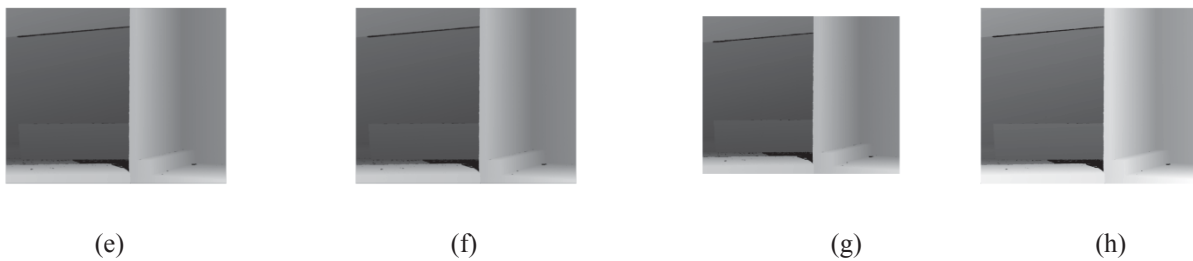
Fig 3.Preprocessed depth map of woods by (a) symmetric smoothening (b) asymmetric smoothening (c) adaptive compensation method (d) proposed method

The pre-processed depth map and color image are given as input to image warping to generate the right and left virtual views. The left virtual views generated after pre-processing the depth map using symmetric Gaussian filter,

asymmetric Gaussian filter, adaptive compensation method and the proposed method for cones and woods are as in Fig 4 and Fig 5.

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

Fig 4.Virtual left view image of cones using depth image pre-processed by (a) symmetric smoothening (b) asymmetric smoothening (c) adaptive compensation method (d) proposed method

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

Fig 5.Virtual left view image of woods using depth image pre-processed by (a) symmetric smoothening (b) asymmetric smoothening (c) adaptive compensation method (d) proposed method

Fig 4(d) and 5(d) indicates that both the vertical and horizontal edges of the virtual views are preserved and the original information in the depth map is also retained. Table I shows the comparison of the performance of symmetric filtering, asymmetric filtering, adaptive compensation method and the proposed method.

TABLE I: Comparison of PSNR, SSIM, no: of holes and computation time of various methods

Test Image	Methods	PSNR	SSIM	No:of holes	Computaion time(s)
Cones	Symmetric Smoothening	16.6818dB	0.7323	5257	**0.092521**
	Asymmetric Smoothening	18.6895dB	0.7527	5301	0.104961
	Adaptive Compensation Method	**32.1279dB**	**0.9918**	6603	0.208563
	Proposed Method	29.4466dB	0.7829	**5062**	0.140163
Woods	Symmetric Smoothening	20.1008dB	0.8450	2243	**0.090608**
	Asymmetric Smoothening	20.1022dB	0.8735	2248	0.108658
	Adaptive Compensation Method	**37.5981dB**	**0.9946**	2600	0.214334
	Proposed Method	36.3071dB	0.8876	**2222**	0.128221

From Table I it is evident than the proposed method results in high PSNR and SSIM value compared to that of Gaussian filtering method. The number of holes generated in the virtual views are considerabley small compared to all the methods. The computation time is also considerabley less compared to adaptive compensation method.

From the comparison , it can be noted that the proposed method reduces the number of holes and still maintains the vertical and horizontal edges of the depth map. This method results in a high qualty rendered view and can be used to obtain high quality stereoscopic view.

IV.   CONCLUSION

In order to minimize texture artifacts appearing in the newly exposed areas of the virtual image, smoothing of the depth maps is done. Symmetric smoothing can create geometric distortions, leading to vertical straight boundaries becoming curved. To reduce this distortion, asymmetric smoothing of depth map is proposed. Asymmetric smoothing and adaptive smoothening destroyed the original information of the depth map and caused blur. Adaptive compensation method could not produce satisfactory results for depth map having large number of holes. The proposed median filtering of depth map results in high quality rendered view and less number of holes. This pre-processing technique results in a comparatively high PSNR and SSIM values and low computation time.

REFERENCES

[1]   Chih-Hsien Hsia, "Singularity Improved Depth Image-Based Rendering Using an Adaptive Compensation Method on an Autostereoscopic 3-D Display for a Kinect Sensor," *IEEE SENSORS JOURNAL*, vol. 15, No.2, Feb. 2015.
[2]   Ming-Fu Hung, Shaou-Gang Miaou, and Chih-Yuan Chiang, "Dual Edge-Confined Inpainting of 3D Depth Map Using Color Images Edges and Depth Images Edges," *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2013 Asia-Pacific*, Nov. 2015, pp. 1-9.
[3]   L. Zhang and W. J. Tam, "Stereoscopic image generation based on depth images for 3D TV *IEEE Trans. Broadcast, .*, vol. 51, no. 2, pp. 191-199, Jun. 2005.
[4]   W. J. Tam, G. Alain, L. Zhang, T. Martin, and R. Renaud," Smoothing depth maps for improved stereoscopic image quality," *Proc. SPIE*, vol. 5599, pp. 162-172, Oct. 2004.
[5]   P.-J. Lee and Effendi, "Nongeometric distortion smoothing approach for depth map preprocessing," *IEEE Trans. Multimedia*, vol. 13, no. 2, pp. 246-254, Apr. 2011.
[6]   C. Fehn, K. Hopf, and Q. Quanta, "Key technologies for an advanced 3D TV system" *Proc. SPIE*, vol. 5599, pp. 66-80, Oct. 2004.
[7]   Middlebury Stereo Vision Database. [Online]. Available: http://vision.middlebury.edu/stereo/data/