International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249 – 8958, Volume-5 Issue-2, December 2015

Color Segmentation Based Resolution Enhancement of Depth Image and Stereoscopic Image Synthesis

Sreejith K.M., K. Suresh, K. Rajeev

Abstract — In this paper a color segmentation based resolution enhancement of depth image is proposed. The resolution enhancement technique is combined with depth image based rendering (DIBR) method to generate stereo images. The major problem in transmission of stereo images is bandwidth. By using DIBR the transmission efficiency can be increased. Also if we are using low resolution depth map for transmission, the bandwidth for transmission can be further reduced. A color segmentation based interpolation procedure is used to enhance the resolution of the depth map. This depth resolution enhancement method sharpens depth image using the color information from the high resolution color image. From this resolution enhanced depth image the stereoscopic images are synthesized using DIBR method. For more textured region, we are using an edge-guided image interpolation algorithm to enhance the performance of the color segmentation based interpolation algorithm.

Index Terms—DIBR, depth image, enhanced depth map, color segmentation

I. INTRODUCTION

To improve the quality of TV perception, scheme providing more realistic and immersive viewing experience has been proposed in recent years. The proposed methods targets for higher depth perception in addition to enhancement in visual quality of the content. 3D television (3DTV) is a television that conveys depth perception to the viewer by employing techniques such as stereoscopic display, multi-view display, 2D-plus-depth, or any other form of 3D display. The stereoscope is working with the principle that when two pictures are viewed stereoscopically, they are combined by the brain to produce 3D depth perception. Previously many 3D-TV systems, with two-view-video representation format, has been proposed [1] [2] . The main drawbacks of these systems are low efficiency of data transmission and poor flexibility of depth reproduction. In 2002, the European IST project Advanced Three-Dimensional Television System Technologies (ATTEST) proposed a novel 3D-TV system which uses the video-plus-depth representation format to effectively overcome these drawbacks [3]. In this new format of 3D-TV content, virtual left-view and right-view images can be synthesized by means of Depth Image Based Rendering (DIBR) technologies according to the view's stereoscopic preference.

Revised Version Manuscript Received on November 04, 2015.

Sreejith K.M., HCL Technologies, Bangalore, India.

Dr. K. Suresh, Department of Electronics and Communication Engineering, Government Engineering College, Barton Hill, Thiruvananthapuram, India.

K. Rajeev, Quest Global Engineering Services Private Limited, Thiruvananthapuram, India.

DIBR is an emerging technique to create realistic images from virtual viewpoint using an image taken from nearby viewpoint and corresponding per pixel depth information. This technique has recently received much attention as a promising technology for three-dimensional television (3D TV) systems. In this method, a monoscopic image and an associated depth map are transmitted instead of the two stereoscopic image pair [4]. At the receiver, the stereoscopic images are created using this monoscopic image and the depth image. Depth map is a two dimensional function consisting of distance between objects corresponding to a pixel to the camera. Depth information can be obtained by depth estimation algorithms using binocular disparity, de - focus, motion etc., or range sensors in depth cameras [5]. The range sensors works with the principle of time of flight of infrared beam. For a high resolution stereo images, high resolution depth image is needed. For increasing the depth resolution, more sensors are needed and this will result in cost escalation. Also low resolution depth image further reduce the bandwidth required for transmission. If an appropriate resolution enhancement algorithm is employed to improve the depth resolution, high resolution depth image can be created using less number of sensors and transmission efficiency can be increased. We can use commonly used resolution enhancement of methods such as bi-linear interpolation and bi-cubic convolution interpolation, which are simple and fast [6]. But they will not preserve sharp changes in depth and blur the edges and suffer from defects, including block effects, and ringing artifacts around edges.

In DIBR scheme high resolution color image is available along with depth image. Hence we can use the color information to enhance the resolution of depth image. We consider the homogeneous regions in the color image as smooth surface in the depth image and there will not be any sharp changes in those regions. Color edge dependent interpolation method uses color information of the RGB image to enhance the resolution of the corresponding depth image [7]. Depth edge gradients are detected with the Sobel operator and regions does not have any edges will directly interpolated. But this method will create artifacts at the edges for images having large edge width.

We proposes a depth image resolution enhancement method based on segmentation of the color image associated with the depth image. Each segment will be mapped and unknown depth values will be found using interpolation by the known depth values in that particular segment. For more textured images, the segments will be small and the information for interpolation will be less. This will create defective in-painting of depth pixels. In order to avoid this

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

we are combining this method with a super resolution method based on linear minimum mean square error estimation (LMMSE) [8].

This paper organized as follows. Section 2 discusses about DIBR method and steps required for 3D propagation. Next section reviews about a couple of resolution enhancement methods and explains about the proposed method based on color segmentation. Experimental validation of the proposed method is included in section 4 and conclusions are given in section 5.

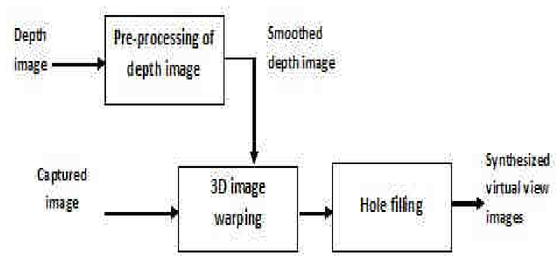


Figure 1: Block Diagram of DIBR System

II. DEPTH IMAGE BASED IMAGE RENDERING SYSTEM

DIBR based depth propagation algorithm consists of 3 steps [9]. The flow chart describing the rendering steps is given in figure 1.

A. *Preprocessing of Depth map*

In the pre-processing step, the depth image is smoothened using low-pass filter to remove sharp edges and so as to reduce the number of large size holes present in the stereoscopic images. After 3D warping step, due to the difference in the view point, some regions which are occluded in the original image will be visible in the left or right image. These regions are called dis-occlusion or holes, which does not have any texture information. In depth image, only horizontal sharp changes will induce holes due to the horizontal positioning of stereo cameras and traditional smoothing of the whole depth image will result in vertical distortion in the generated image after 3D image warping. Liang Zhang [9] proposed an asymmetric 2-D Gaussian filter to reduce vertical distortion. In this method the smoothing strength of depth maps in the horizontal direction would be weaker than that in the vertical direction. Wan-Yu Chen [10] proposes that edge filtering method based on depth image can preserve edge information and improve the overall image quality.

B. 3D warping

For generating virtual stereoscopic images we consider the commonly used parallel camera configuration consisting of one center image associated with one depth map. In figure

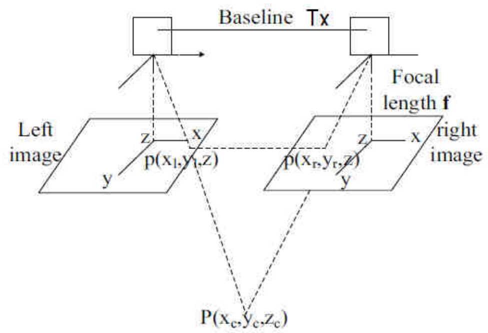


Figure 2: Stereo Imaging Model

2, f is the focal length of three cameras and T_x is the baseline distance between two virtual cameras. From this camera configuration one point P with the depth Z, the disparity can be calculated by, $T_x f / Z$ and that point is projected onto the image plane of left, center and right cameras at pixels (x_l, y),(x_c, y) and (x_r, y) respectively. If the information about the x_c is known, x_l and x_r can be calculated from the disparity [9]. So this can be represented by

$$x_l = x_c + (T_x f)/(2 Z) \quad (1)$$

$$x_{lr} = x_c - (T_x f)/(2 Z) \quad (2)$$

C. Hole filling

Even though we smoothen the depth image, after 3D warping, occluded regions or holes will be created. We fill the newly exposed areas by averaging textures from neighborhood pixels, and this process is called hole-filling. These holes can also be removed by any appropriate interpolation methods as suggested in [11].

III. RESOLUTION ENHANCEMENT USING EDGE GUIDED IMAGE INTERPOLATION ALGORITHM

An edge-guided image interpolation algorithm proposed by Lei Zhang and Xiaolin Wu [8] is based on linear minimum mean square-error estimation (LMMSE) technique. A missing sample is interpolated in not one but two mutually orthogonal directions. The two interpolation results are treated as two estimates of the sample and adaptively fused using the statistics of a local window. Specifically, we partition the neighborhood of each missing sample into two oriented subsets in orthogonal directions. The hope is that the missing sample has higher correlation with its neighbor in the edge direction. The pixel is finally interpolated by combining the two directional estimates in the principle of LMMSE [8]. The edge direction is the most important information for the interpolation process. To extract and use this information, we partition the neighboring pixels of each missing sample into two directional interpolation is made, and then the two interpolated values are fused to arrive at an LMMSE estimate of the missing sample. We recover the high resolution image I_h in two steps. First, those missing samples $I_h(2n, 2m)$ at the center locations surrounded by four low resolution samples are interpolated. Second, the other missing samples $I_h(2n-1, 2m)$ and $I_h(2n, 2m-1)$ are interpolated with the help of the already recovered samples $I_h(2n, 2m)$.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

A. Interpolation of samples $I_h(2n, 2m)$

We can interpolate the missing high resolution sample $I_h(2n, 2m)$ along two orthogonal direction: 45° diagonal and 135° diagonal denote by I_{45} (2n,2m) I_{135} (2n,2m) by some linear methods. Consider the direction interpolation output as the noisy measurements of the missing high resolution sample

$$I'_{45}(2n,2m) = I'_h(2n,2m) + v_{45}(2n,2m)$$
$$I'_{135}(2n,2m) = I'_h(2n,2m) + v_{135}(2n,2m) \qquad (3)$$

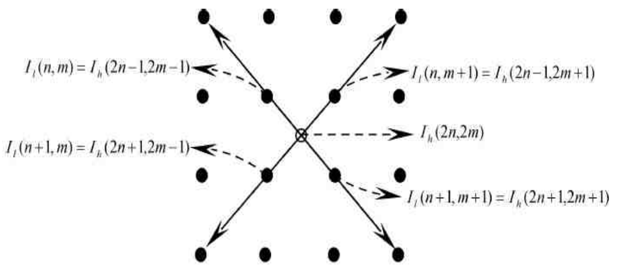


Figure 3: Interpolation of the high resolution samples Ih(2n,2m)

where the random noise variables υ_{45} and υ_{135} represent the interpolation errors in the corresponding direction. We write into matrix form

$$\mathbf{Y} = \mathbf{1}.\ \mathbf{I}_h + \mathbf{V} \qquad (4)$$

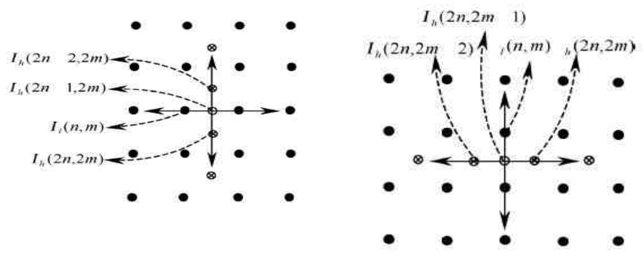
where $\mathbf{Y} = [\ I'_{45}\ \ I'_{135}]^T$, $\mathbf{1} = [\ 1\ 1]^T$, and $\mathbf{V} = [v_{45}, v_{135}]$

The interpolation problem is to estimate the unknown sample I_h from the noisy observation Y. This estimation can be optimized in linear minimum mean square error sense (LMMSE). It can be derived that [12]

$$I'_h = \mu_h + \sigma^2_h 1^T(1.\sigma^2_h.1^T + R_v)^{-1}(Y-1.\mu_h)(5)$$

where σ^2_h variance of I_h and R_v is the variance of **V.**

B. Interpolation of samples $I_h(2n-1, 2m)$ and $I_h(2n, 2m-1)$



a b

Figure 4: Interpolation missing high resolution samples (a) Ih(2n-1, 2m) and (b) Ih(2n, 2m-1)

After the missing high resolution samples $I_h(2n, 2m)$ are estimated, the other missing samples $I_h(2n-1,2m)$ and $I_h(2n, 2m-1)$ can be estimated similarly, but now with the aid of the just estimated high resolution samples. The missing sample $I_h(2n-1, 2m)$ or $I_h(2n, 2m-1)$ can be estimated in one direction by the original pixels of the low resolution image and in the other direction by the already interpolated high resolution samples. The two directional approximations of the missing samples are considered as the noisy measurement of $I_h(2n-1, 2m)$ or $I_h(2n, 2m-1)$, and then the LMMSE of the missing sample can be computed in a similar way as described in the previous section.

IV. COLOR SEGMENTATION BASED RESOLUTION ENHANCEMENT

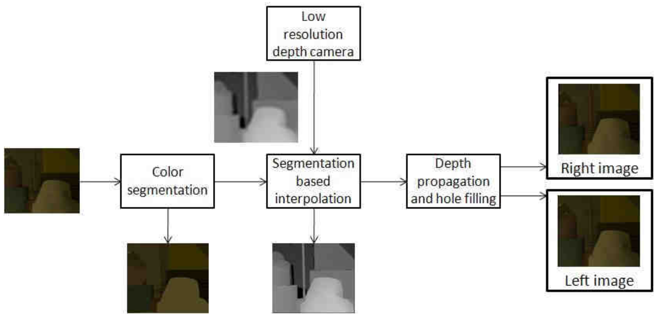


Figure 5: Block diagram of color segmentation based resolution enhancement

In order to create a high resolution stereoscopic image we need to know the depth value of each pixel in the RGB image. We can increase the resolution by interpolation using bilinear or bi-cubic interpolation. The main drawback of simple interpolation is that it will smoothen the depth map and introduce blurring effect in the edges that comes out as significant and visually annoying artifact in the rendered image. Also, it does not use the color information available in the high resolution RGB image. Another possibility to increase the resolution is through interpolation after depth propagation to the color view. But the virtual views created after depth propagation contain occluded regions and this will degrade the virtual image due to the higher number of unknown pixels in the image. Even though the hole filling techniques preserve sharpness of edges, blurring of edges in the low resolution depth case causes too many visual artifacts in the final stereoscopic images. Our proposed method for creating high resolution depth image from low resolution depth image and high resolution color image is based on color segmentation. Block schematic of the color segmentation based depth image resolution enhancement algorithm is shown in Figure 5. The first step is to decompose the reference image into regions of homogeneous color. The algorithm assumes that depth values vary smoothly in those regions and that depth discontinuities occur only on region boundaries. Over-segmentation is preferred, since it helps to meet this assumption in practice. Hence mean-shift color segmentation technique suggested by Comaniciu and Meer [13] is used. The mean-shift analysis approach is essentially defined as a gradient ascent search for maxima in a density function defined over a high dimensional feature space. The feature space includes a combination of the spatial coordinates and all its associated attributes that are considered during the analysis. The main advantage of the mean-shift approach is based on the fact that edge information is incorporated as well. After obtaining a region that did not have any sharp depth variation, the depth values of the unknown points in that region is estimated using interpolation technique. For that we mapped the segmented areas to the depth image. Then the unknown depth values are found using the known depth values of the particular segment by any interpolation

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

method. Segmentation is done in such a way that each segment contains at least one known depth value. Finally using this high resolution depth image the stereoscopic images are created using the DIBR procedure discussed in section II.

V. RESOLUTION ENHANCEMENT METHODS FOR TEXTURED IMAGES

For more textured image after segmentation there may be regions having smaller size. Direct interpolation in those region create artifacts because of the absence of sufficient data for interpolation. Here the number of known depth values will be very less and in order to fill those we need to interpolate depth pixel outside of each regions. So this will smoothen the depth map.

For this, we are proposing a method by combining above two methods namely, color segmentation and LMMSE based edge guided image interpolation algorithm. Color segmentation based resolution enhancement method is a simplest method having least number of operations and best performance in the edges of the image because we are directly taking region edges from its corresponding color image. The computation complexity of edge guided image interpolation algorithm is high compared to usual linear interpolation. So those regions having plane surface edge guided interpolation is not efficient. In order to reduce the cost of the algorithm we are using this algorithm only in the highly textured area and plane surface will be directly interpolated. Here the unknown pixels in each segments will be directly interpolated, while regions having small segments which is considered as textured region will interpolate using edge guided interpolation.

VI. RESULTS AND DISCUSSIONS

In order to evaluate the performance of proposed algorithms we used some low resolution depth images which is created by down sampling high resolution depth images and the high resolution images were reconstructed using proposed method. Thus we can compare the original images with the simulated results and measure the PSNR of those reconstructed image. We compared the proposed method with

Table 1: PSNR (dB) Values of reconstructed high resolution images by different algorithms

Images	Chairs	Bowling 1	Bowling 2	Laundry	Lamp shade
Bilinear Interpolation	19.93	18.42	19.68	18.73	19.98
LMMSE	21.73	19.7	21.72	19.2	21.06
Color Segmentation	21.29	19.68	21.83	20.02	21.75
Proposed Method	22.06	20.07	21.86	20.35	21.84

bilinear interpolation, color segmentation based interpolation and LMMSE based edge guided interpolation.

We used five sets of images to evaluate the performance of the proposed algorithm. Datasets “Lampshade", “Laundry", “Bowling1", and “Bowling2" are created by Brad Hiebert-Treuer, et.al. [14] and were published in conjunction with two papers [15],[16] and the depth map is generated from the disparity map of stereo images. Data set “chairs" is taken from [17], which is created by depth camera. The resolution of depth image are 600 x 500 and that of color images are 1200x1000. From these two images a depth image and two stereo images having resolution 1200 x1000 was created. Also we were used an image “Chess" to test 3D propagation and to create stereoscopic image. The color image “chair" was divided into homogeneous regions by using mean shift segmentation by giving parameter values h_s = 10, h_r = 6 and M =30. Figure 6 shows the results of intermediate steps in which figure 6(b) shows the color segmented image. Using this image the low resolution image was interpolated using color segmentation based bilinear interpolation and the region having more textures were interpolated using edge guided interpolation algorithm. The resulting enhanced depth images are shown in figure 6(d). In order to shown the increase in sharpness a region shown as enlarged. Figure 6(e) and 6(f) are the images of the enlarged regions of low resolution and high resolution images. Those figure shows the sharpness of the edges increased. We tested this algorithm in 4 another different databases and the results are given if figure 7

In order to evaluate the performance of the proposed algorithm we calculated the PSNR of reconstructed algorithm by using the original high resolution which is used to down-sample. We took 5 images to evaluate PSNR and we compared the results with bilinear interpolation, color segmentation based interpolation and LMMSE based edge guided image interpolation. The results are shown in figure 8. Also the PSNR values are shown in table 1 and the table shows that the proposed method have the highest PSNR value. So we can conclude that the performance of the proposed algorithm is better than other methods and this algorithm is the best choice among them for resolution enhancement of a depth image.

From the high resolution depth image the stereo images were synthesized using DIBR method by taking the baseline distance of 20 pixels, which is shown in figure 9. Figure 9(a) is the given RGB image corresponding to the center image and using high resolution depth image the virtual left and right images were constructed using DIBR based depth propagation. In order to avoid vertical distortion we used asymmetric Gaussian smoothing. From the result we can see that shift in the pixels of rendered image corresponding to near objects is more than that of far objects.

VII. CONCLUSION

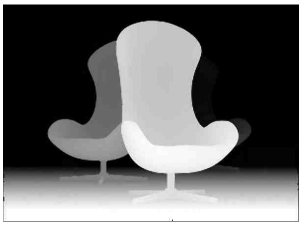
We have demonstrated a method to synthesize high resolution stereoscopic image from a low resolution depth map and corresponding high resolution color image. For this high resolution depth image is constructed from low resolution depth image by extracting the color information from high resolution color image. Due to the combination of

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

(a) High Resolution image “chairs”

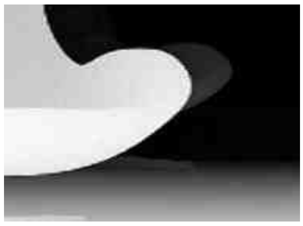
(b) Segmented image

(c) Low resolution depth image of “chairs”



(d) High Resolution depth image by proposed algorithm

(e) Enlarged portion of low resolution image



(f) Enlarged portion of high resolution image

Figure 6: Resolution enhancement of depth image "chairs"

(a) Original Image

(b) Corresponding depth

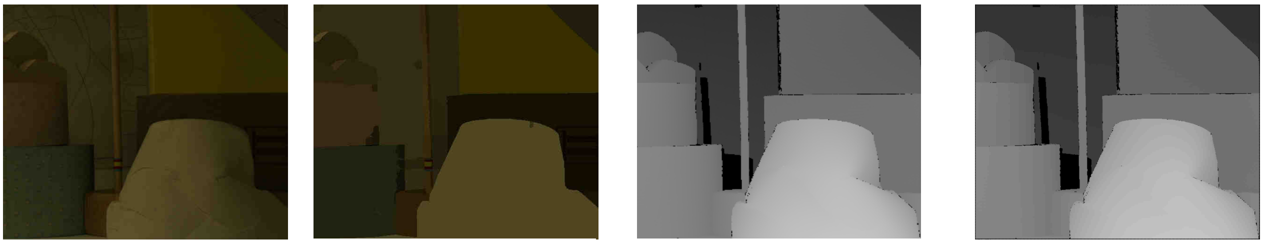
(c) Virtual left

(d) Virtual right image

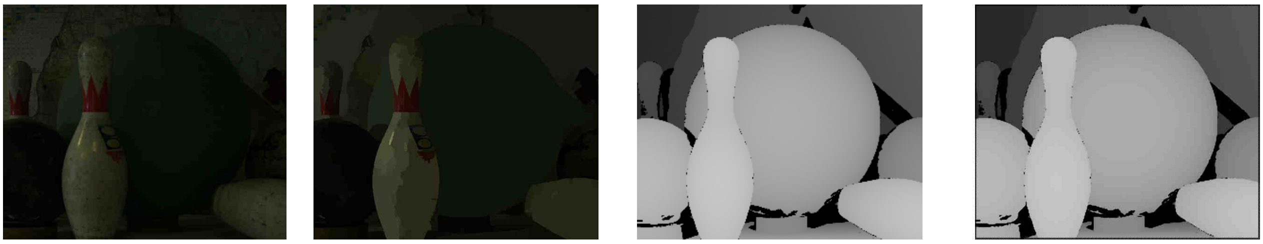
Figure 7: Virtual left and right images generated

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

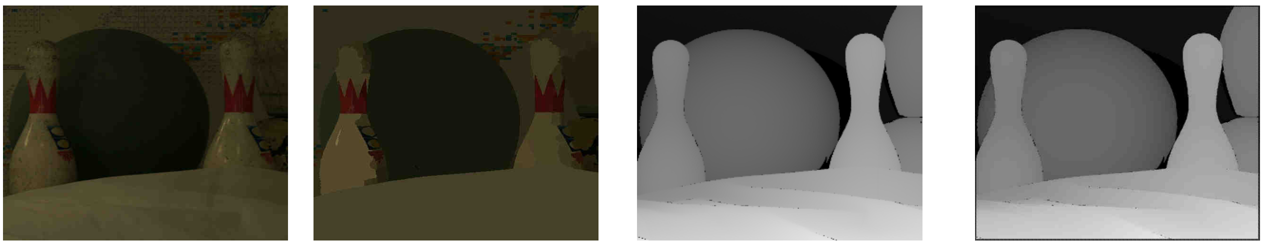
(b) Color image of "Laundry" **(a) Segmented image** **(d) Low resolution depth image** **(c) Resolution enhanced depth image**



(e) Color image of "Lampshade" **(f) Segmented image** **(g) Low resolution depth image** **(h) Resolution enhanced depth image**



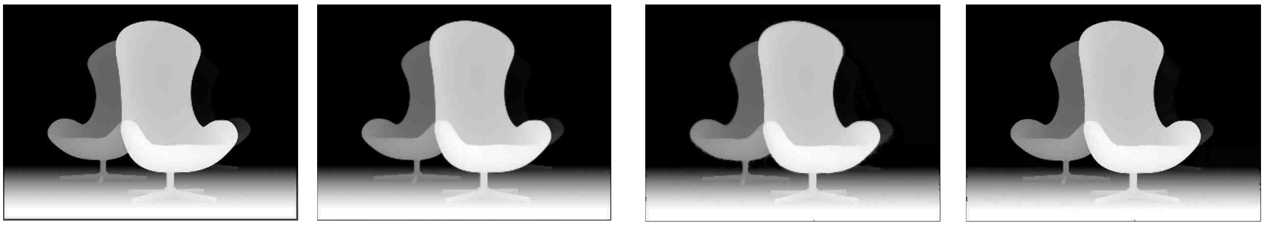
(i) Color image of "Bowling1" **(j) Segmented image** **(k) Low resolution depth image** **(l) Resolution enhanced depth image**



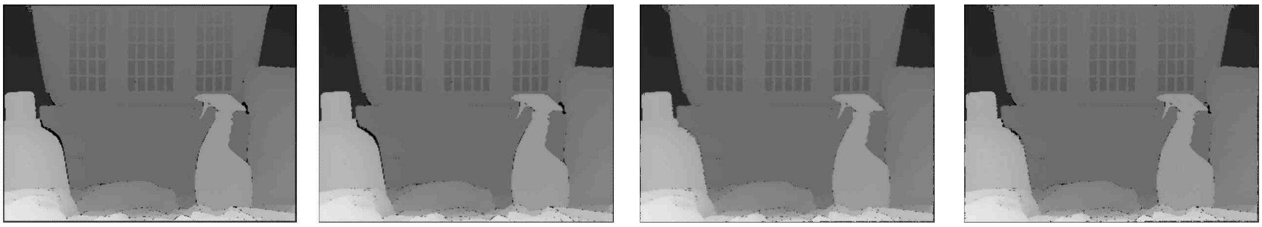
(m) Color image of "Bowling2" **(n) Segmented image** **(o) Low resolution depth image** **(o) Resolution enhanced depth image**

Figure 8: Proposed implementation in different images

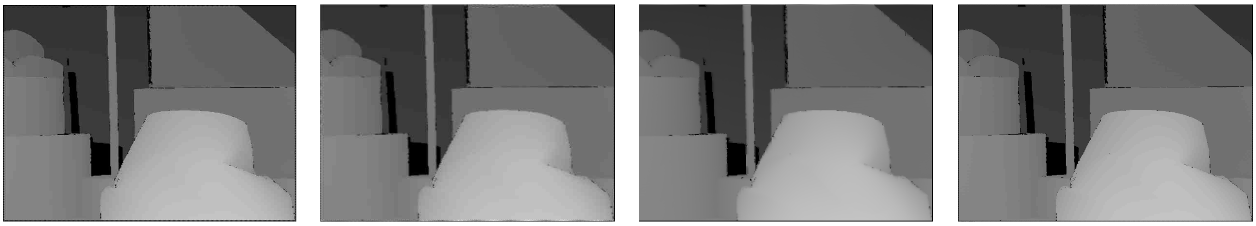
Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.



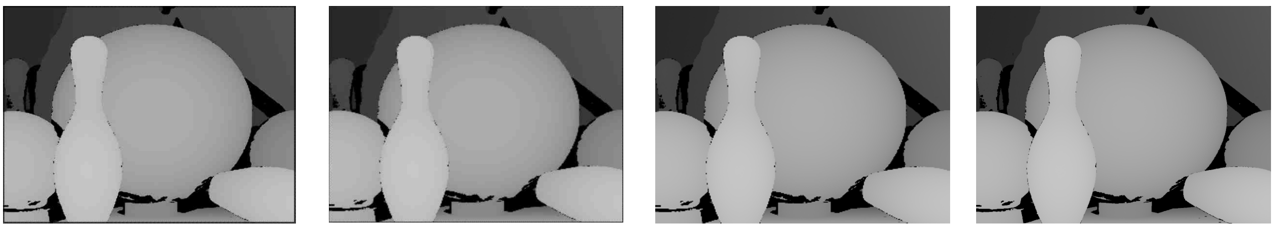
(a) Depth image "Chairs" by bilinear interpolation | (b) LMMSE algorithm | (c) Color segmentation | (d) Proposed Method



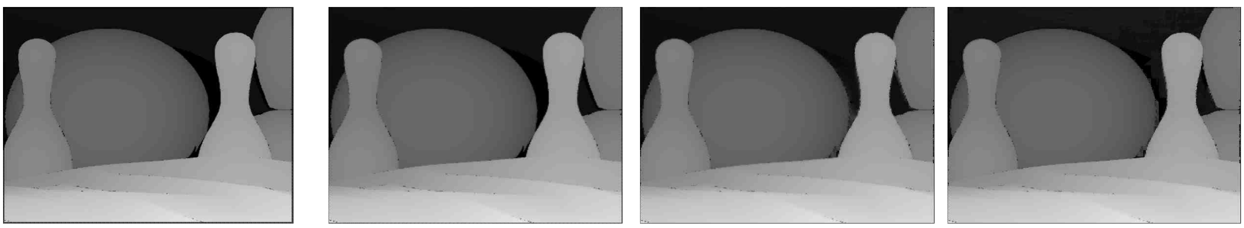
(e) Depth image "Laundry" by bilinear interpolation | (f) LMMSE algorithm | (g) Color segmentation | (h) Proposed method



(i) Depth image "lampshade" by bilinear interpolation | (j) LMMSE algorithm | (k) Color segmentation | (l) Proposed method



(m) Depth image "bowling1" by bilinear interpolation | (n) LMMSE algorithm | (o) Color segmentation | (p) Proposed method



(q) Depth image "bowling" by bilinear interpolation | (r) LMMSE algorithm | (s) Color segmentation | (t) Proposed method

Figure 9: Comparison of different interpolation methods on different images

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication Pvt. Ltd.

edge guided interpolation with color segmentation based interpolation, this algorithm performs well for depth images having texture in its corresponding color image. Only low resolution depth image need to be transmitted along with the high resolution color image and hence the bandwidth required is less compared to the case where stereo images are transmitted. Also, the cost for generating low resolution depth image is less due to less number of sensors needed. The computational complexity is less compared to [8] because we are applying edge-guided image interpolation algorithm on highly textured regions only. The sharpness of the enhanced image is better than other methods and the calculation of PSNR values shows that the proposed resolution enhanced method is the best option for resolution enhancement of depth images.

REFERENCES

1. K. Gomi, Y. Nishino, K. Tai, M. Yasuda, and N. Tetsutani, “Stereoscopic video transmission and presentation system for ISDN,” IEEE Transactions on Consumer Electronics, vol. 36, no. 3, pp. 759–766, 1990.
2. J.Flack, P.V.Harman, and S.Fox,“Low-bandwidth stereoscopic image encoding and transmission,” in Proceedings of SPIE Conference on Stereoscopic Displays and Virtual Reality Systems X, 2003, vol.5006,pp. 206–214.
3. C. Fehn,“A 3D-TV system based on video plus depth information,” Conference Record of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers, 2003. IEEE, 2003, vol. 2, pp. 1529–1533. in
4. C. Fehn, “Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV,” in Proceedings of SPIE, 2004, vol. 5291, p. 93.
5. Q. Wei, “Converting 2D to 3D: A survey,” in International Conference on Information and Communication Theory, 2005, vol. 7, p. 14.
6. H. Hou and H. Andrews, “Cubic splines for image interpolation and digital filtering,” IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 26, no. 6, pp. 508–517, 1978.
7. Q.H. Nguyen, M.N. Do, and S.J. Patel, “Depth image-based rendering with low resolution depth,” in 16th IEEE International Conference on Image Processing (ICIP), 2009, 2009, pp. 553–556.
8. L. Zhang and X. Wu, “An edge-guided image interpolation algorithm via directional filtering and data fusion,” IEEE Transactions on Image Processing, vol. 15, no. 8, pp. 2226–2238, 2006.
9. L. Zhang and W.J.Tam, “Stereoscopic image generation based on depth images for 3D TV,” IEEE Transactions on Broadcasting, vol. 51, no. 2, pp. 191–199, 2005.
10. W.Y. Chen, Y.L. Chang, S.F. Lin, L.F. Ding, and L.G. Chen, “Efficient depth image based rendering with edge dependent depth filter and interpolation,” in IEEE International Conference on Multimedia and Expo, 2005. ICME 2005., 2005, pp. 1314–1317.
11. M. Schmeing and X. Jiang, “Time-consistency of disocclusion filling algorithms in depth image based rendering,” in 3DTV Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON), 2011. IEEE, 2011, pp. 1–4.
12. F. Malgouyres and F. Guichard, “Edge direction preserving image zooming: a mathematical and numerical analysis,” SIAM Journal on Numerical Analysis, pp. 1– 37, 2002.
13. D. Comaniciu and P. Meer, “Mean shift: A robust approach toward feature space analysis,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pp. 603–619, 2002.
14. “Middlebury stereo datasets,” http://vision.middlebury.edu/stereo/data/
15. D. Scharstein and C. Pal, “Learning conditional random fields for stereo,” in IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR’07., 2007, pp. 1–8.
16. H. Hirschmuller and D. Scharstein, “Evaluation of cost functions for stereo matching,” in IEEE Conference on Computer Vision and Pattern Recognition, 2007. CVPR’07., 2007, pp. 1–8.
17. “Depth image of ’chairs’, http://www.dofpro.com/cgigallery.htm

Sreejith K. M. is a Lead Engineer at HCL Technologies Bangalore, INDIA. His area of interest is Image Processing . He developes algorithms for various image and video processing applications. He has experience in Open CV, CUDA, GLSL and OpenGL. He also worked in real time video processing and Ethernet camera drivers. He received M. Tech.(Signal Processing) Degree from Kerala University in 2012 and B. Tech (Electronics & Communication) from Calicut University in 2009

Dr. K. Suresh is Associate Professor at Government Engineering College Barton Hill, Thiruvananathapuram, INDIA. His areas of interest are Signal Processing Theory, Audio Signal Processing and Image Processing. He received PhD. (2010) and M. E. (Signal Processing, 2000) from IISc Bangalore, and B. Tech. (Electronics&Communication,1994) from Kerala University.

K. Rajeev is a Technology Solutions Leader at QuEST Global, Thiruvananthapuram, INDIA. He Specializes in Algorithm development in Signal Processing, Computer Vision and Medical Imaging, Pattern Recognition, M achine Learning, Support Vector Machines and Algorithm optimization for GPU and Embedded Systems . He received M. E. (Signal Processing, 2005) from IISc Bangalore and B. Tech (Electronics & Communication, 1990) from Kerala University.