IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 1, January 2015.
www.ijiset.com
ISSN 2348 – 7968

# Fast and approximate processing unit for depth image based Rendering process

**Neetesh Nema[1], Brajesh Patel[2]**
[1]M.Tech Scholar, Department of CSE, SRIT, Jabalpur (M.P.) RGPV, India
[2]H.O.D, Department of CSE, SRIT, Jabalpur (M.P.), India

***Abstract -*** Three dimension (3D) technology increases the visual quality as compared to two-dimensional (2D) technology. In Currently time every multimedia device needs 3D technology. So for generation of 3D content there is need of Depth image based rendering (DIBR) process which will generate left and right image through depth image and original image. Basically DIBR is following the concept of actual 3D recording camera setup. Through original camera setup there is virtual camera formula is generated which will create left and right image. As we already know for any image processing application time complexity is main issue. So in this work I will propose a fast and approximate DIBR algorithm which will reduce the time complexity issue. For image quality measurement there is some scientific parameter will use which will check the quality of generated left and right image through proposed DIBR algorithm. Those parameters are like PSNR, SSIM, RFSIM and FSIM. Algorithm will implement on Matlab.

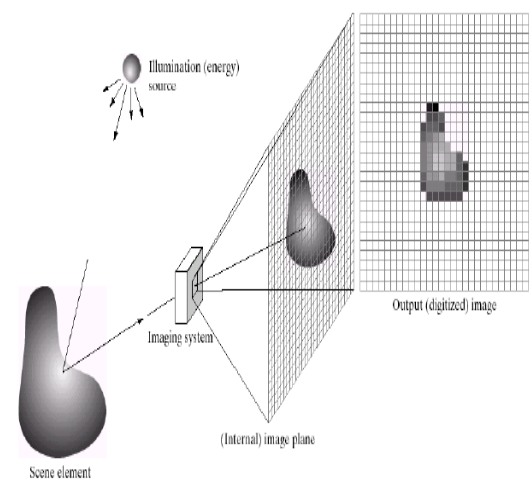
***Keywords -*** Depth image based rendering, Adaptive smoothing, Terrestrial DMB, 3D service, Image Processing and Analysis

## 1. INTRODUCTION

The 3D video signal processing has received considerable attention in visual processing. Given advances in 3D display technology, humans aspire to experience more realistic and unique 3D effects. Depth-Image-Based-Rendering (DIBR) is a key technology in advanced three dimensional television system (ATTEST 3D TV System) [2][3]. Depth image based rendering contains three steps [4].Pre-processing of depth map is applied for reducing the sharp horizontal transitions on depth map. Second, 3D image warping renders left and right images based on the pre-processed depth map and intermediate color image. On the basis of my Survey Paper the traditional 3D game has some information of the third dimension, but it is still two-dimensional when it is displayed on screen, while the stereoscopic game has stereoscopic view objects in the images stand out of the screen[1]. The depth preprocessing method is applied before transmission. Once both the original image and the accompanying depth image are received, the auto stereoscopic image is produced by applying 3D Wrapping, hole filling, and merging simultaneously[2][3].

### 1.1 Introduction to Digital Image processing:

A digital image is a form of two dimensional matrixes with a finite set of values in a particular range, the set of values called as pixel elements. Pixels values represent the colors of the values as gray and heights, objects etc.



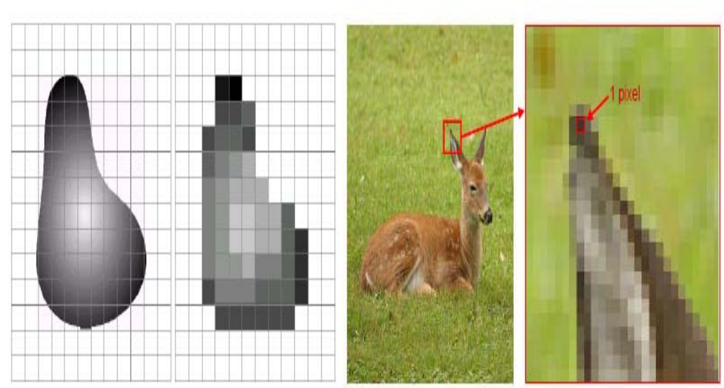




**Figure1: Digital illustration of picture**

Digital image processing concentrates on two important aspects:

1. Improvement of pictorial data for human translation.

ISSN 2348 – 7968

2. Processing and reducing the image data for storage, transmission and representation for autonomous machine discernment some contention about where image processing ends and fields such as image analysis and computer vision start .The image processing sequence to computer vision can be divided as low, mid and high level process. The utilization of digital image processing techniques/procedures has shattered and they are now used for numerous kinds of assignments in different sort of regions.

- Image enhancement.
- Artistic effects
- Medical visualization
- Industrial inspection
- Law enforcement
- Human computer interfaces
- 3D image for 3D videos

### 1.2 Introduction to 2D to 3D Conversion:

The historical backdrop of stereoscopy, stereoscopic imaging or three-dimensional (3d) imaging can be followed back to 1833 when Sir Charles Wheatstone created a mirror device that provides to the view the depth which is illusion, in his description of the "Phenomena of Binocular Vision". The process consists in bending two slightly different views of the same painting or drawing into one stereoscopic image, bringing about a convincing 3d view of the original picture. In 1839, after realization of still photography it was only years before paintings and drawings replaced by photographs in his stereoscopic device. In 1844, Sir David Brewster further developed the stereoscope by utilizing prismatic lenses to enlarge and fuse the stereo images Three-dimensional television (3DTV) has a long history, and over the years a consensus has been reached that a successful introduction of 3DTV broadcast services can only reach success if the perceived image quality and the viewing comfort are at least comparable to conventional two-dimensional television (2DTV).

### 1.3 3D video application scenarios:

The denomination 3D broadly refers to any visual or additive system that attempts to maintain or represent the illusion of 3D sphere of human viewing or hearing, including model representation with a collection of points in 3D space connected by various geometric entrees. In the case of video application scenarios, we will refer in this work to the pseudo- 3D denomination that refers to the ability of the viewer to perceive an illusion of depth, or the illusion to navigate in a 3D space [5].

**3D displays -** Any 3D display system (anaglyph glasses, shutter glasses, auto stereoscopic, holographic, and volumetric, *etc*) which provides different perspective views to the left and right eye, will create a an interest and efficient sensation of depth. Ideally, such 3D displays provide stereo sis (*i.e.* binocular perception of depth), kenosis (*i.e.* depth perception from motion parallax), and accommodation (*i.e.* depth perception through focusing). 3D displays that provide all of these depth points are called multi view *auto stereoscopic* or *auto multi scope* displays. They allow wild viewing (*i.e.* without glasses) of high-resolution stereoscopic images from arbitrary positions. Modern auto multi scope displays use either holographic, volumetric or parallax technology.

**Three-dimensional television (3DTV) -** After the black-and-white models evolved into high-quality color television, 3DTV is believed to be the next major revolution in the history of television by providing to the viewer a feeling of introducing in the movie. In the sequel of the Advanced Three-dimensional Television System Technologies (ATTEST) project , the 3DTV Network of Excellence [4] and the 3D4YOU project [3], 3DTV is entrusted to be designed as a open, flexible and modular system (see Fig), which can be used in a broadcast environment. Essential requirements are the backward compatibility with existing 2D broadcast system and flexibility to support a wide range of different 2D and 3D displays.

## 2. INTRODUCTION TO DIBR

The depth-based 3D system is one of the strong entrants of the second generation 3Dtv, followed by the stereoscopic 3D-TV. The data formats involve one or several group of coupled texture images and depth maps known as image-plus-depth (2D+Z), multi-view video plus depth (MVD), and layered depth video (LDV). With information of depth, novel views at arbitrary viewpoints can be synthesized with a depth-image-based rendering (DIBR) technique. In such a way, the depth-based 3D-TV system can provide stereoscopic pairs with an adjustable baseline or multiple views for auto stereoscopic displays.
The principal function for conversion of 2D to 3D consists of two parts:
(1) Generation of Depth map information.
(2) 3D rendering based on the depth information.

IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 1, January 2015.
www.ijiset.com
ISSN 2348 – 7968

Multi view images can obtain once the original and its depth are obtained by projecting back the real image points to the respective 3D places and then to the respective desired plane of image. The depth image based rendering (DIBR) is considered as one of the important technology in developing and generating stereo scope images used in 3D images and TV's. DIBR algorithm is used to produce stereoscope images which are obtained after completion of real 2D and the respective depth map are generated. In the given system, the DIBR algorithm is optimized to achieve high density and lesser power in terms of hardware. Some unknown places of the imaginable view are loaded out after 3D warping, so in this process the hole-filling is used to fill these places. Using linear interpolation which uses neighbor pixels is the common method for hole-filling [6].

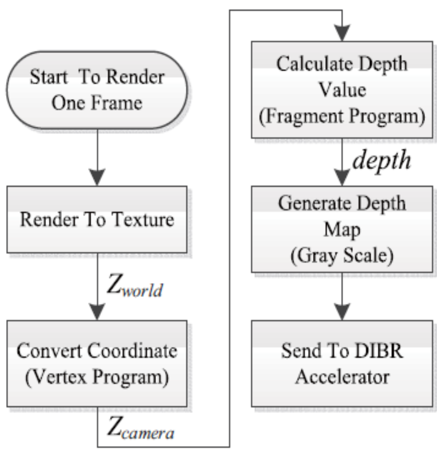



**Figure2: Flowchart of the depth map creation**

## 3. LITERATURE REVIEW

### 3.1 Depth-Image Based Rendering (DIBR):

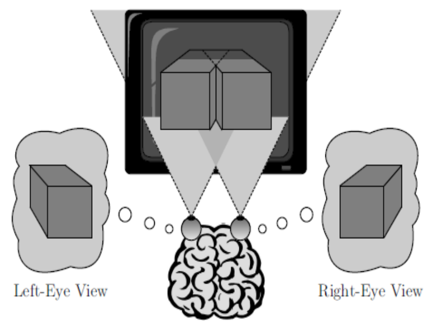
Depth-image-based rendering (DIBR) is the procedure of synthesizing "virtual" views of a frame from still or moving images and respective per-pixel depth information. Theoretically, this view generation can be understood as the following two-step process:

1. The real image points are re projected into the 3D virtual world, utilizing their depth data.
2. These 3D space points are projected into the image plane of a "virtual" camera, which is located at the required viewing position.

The concatenation of re projection (2D-to-3D) and next sequence projection (3D-to-2D) is usually called 3D image warping in the Computer Graphics (CG).

### 3.2 Stereoscopic Image Creation:

On a stereoscopic- or auto stereoscopic 3D-TV display, two slightly different perspective views of a 3D scene are reproduced simultaneously on a joint image plane. The horizontal differences between these left- and right-eye views, is called as *screen parallax* values, which are interpreted by the human brain and then the two images are merged into a single, 3D percept.



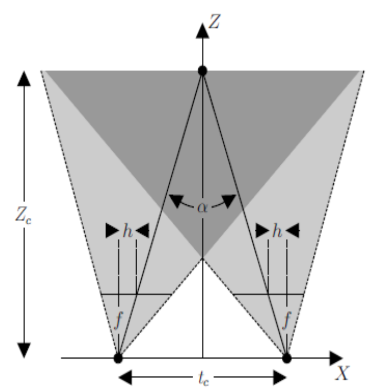




**Figure3: Shift-sensor stereo camera setup**

### 3.3 DIBR Algorithm:

DIBR is the process of generating and producing virtual views from the mono scope and the associated and respective depth. For example consider a virtual three D space point M, and then the two side by side calculations of the equations in two views are result to:

$$\widetilde{m} = AP_n\widetilde{M}, \widetilde{m}^* = A^*P_nD\widetilde{M}.$$

From the above formula m and m* represents two two D image points in the left and right view. The rotational matrix D and translational matrix t that transforms the three D from points the regular angle unit into the camera angle units. The matrices A* and A specify the essential values of the cameras. Then the identity matrix Pn authorizes the normal perspective calculated matrix. The below fig shows a virtual camera setup for altering of virtual angle viewing, parameters *f* and *tc* denotes the focal length and the baseline distance between 2 virtual angle of camera cr and cl . ZC is known as the depth value of the ZPS . The setup of camera placed side by side is not diverged setup used to generate vertical inconsistency and it is very easier to implement with the DIBR, the position of the 3D space point M is depended on its depends on its depth value , the 3D image warping equation can be obtained [11]:

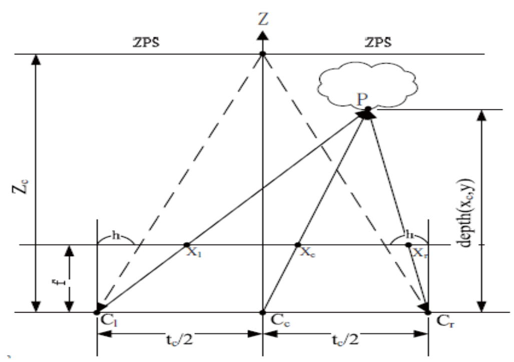



**Figure4: Virtual Stereoscopic Camera setup**

$$\widetilde{m}^* = \widetilde{m} + \frac{A^*t}{depth} + \begin{bmatrix} h \\ 0 \\ 0 \end{bmatrix}, with \quad t = \begin{bmatrix} t_x \\ 0 \\ 0 \end{bmatrix}.$$

The calculation of pixel position (x,y) of wrapped image point are

$$x^* = x + \frac{\alpha_u t_x}{depth} + h, with \quad y^* = y$$

Where $\alpha u$ is a parameter which represent the distance of left-right-eye and screen. The following relationship is the pixel position ($xc$, $y$), ($xl$, $y$) and ($xr$, $y$) of the reference view and two virtual views for corresponding to the point *P* with *depth:*

$$x_l = x_c + \frac{t_c f}{2depth(x_c, y)} + h,$$
$$x_r = x_c - \frac{t_c f}{2depth(x_c, y)} - h$$

The offset h between reference view and target view can be computed by:

$$h = -\frac{t_c f}{2Z_c}$$

### 3.4 Algorithm Optimization:

The next step after the hole-filling is the complete left-and-right view images which are to be transformed into horizontally, by replacing h pixels for addition and subtraction. The division operation is one which consumes time and lower area by hardware implementation, the DIBR algorithm is optimized to remove the division operation. 1/*depth* ranges between 0 and 1. The value (256-depth)/256 can be replaced by the 1/depth , because the value of it represents the relative distance of the pixel but not the accurate real distance. The people are cared for sensitive objects which have the smaller depth values .

$$x_l = x_c + k\frac{t_c f}{2}\frac{256 - depth(x_c, y)}{256},$$
$$x_r = x_c - k\frac{t_c f}{2}\frac{256 - depth(x_c, y)}{256}$$

During the implementation for each depth value which lies between 0 and 255 the average drifted betwenn the practical and the conceptual value is about having a difference of 3.7%. The division operation can be replacable and implemented by shift operation in hardware. The nearest plane is represented by white and the farthest color is represented by balck which is oppsite to the real depth obtained from W. *D os used* to represent the value that obtained from the depth map and get:

$$D = 256 - depth$$

L_ of f set and R_ of f set represent the value of off set in the left-and-right view images, respectively. From this while implementing the DIBR in hardware the above equation is used to compute the left and right offset.

$$\begin{cases} l\_offset = D \cdot pos/256 \\ r\_offset = pos - l\_offset. \end{cases}$$

From above equation the *pos* = *k tcf* 2 is a parameter value which is related to the eye screen ditance and the frame or scene size.pos to 1/32 of the width of the screen. Simple shifting operation can be implemented by shifting operation by giving real view.

$$\begin{cases} lpic(x,y) = pic(x - l\_offset, y) \quad left-view \\ rpic(x,y) = pic(x - r\_offset, y) \quad right-view. \end{cases}$$

From the above equation lpic and rpic are the left and right view images are represented. But picture represents real 2D image, and (*x, y*) the pixel position in the images.[11]

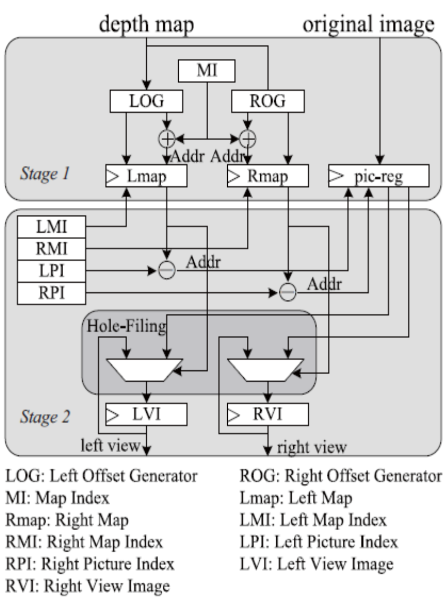



**Figure5: DIBR accelerator architecture Pre-Processing of Depth Image**

### 3.5 3D Image Wrapping:

3D image wrapping maps intermediate view pixel by pixel to left or right view according to pixel depth value. In the other words, 3D image warping transforms pixel location according to depth value. The 3D image warping formula is as following:

$$x_l = x_c + \left(\frac{t_x}{2}\frac{f}{Z}\right),$$

$$x_r = x_c - \left(\frac{t_x}{2}\frac{f}{Z}\right)$$

The x*l* is the horizontal coordinate of the left view, and *xr* is the horizontal coordinate of the right view. Besides, *xc* is the horizontal coordinate of the intermediate view. *Z* is depth value of current pixel, *f* is camera focal length and t*t x* is eye distance. The formula shows that 3D warping maps pixel of intermediate view to left and right view in horizontal direction.

### 3.6 Hole Filling:

Average filter interpolation method is a common method for Hole-Filling in DIBR, which only would result in artifacts at highly textured areas. Apart from this the hole-size in DIBR is so big that it is needed to using average filter with large window size. At the same time, average filter with large window size cannot store edge information for the reason that edge information is blurred. The formula used to do Hole-Filling is as below [7][8]:

$$\frac{\sum_{v=-w}^{v=w}\left\{\sum_{\mu=-w}^{w} s(x-\mu, y-v) \times non\_hole(x-\mu, y-v)\right\}}{\sum_{v=-w}^{v=w}\left\{\sum_{\mu=-w}^{w} non\_hole(x-\mu, y-v)\right\}}$$

$$non\_hole(x,y) = \begin{cases} 0 & \text{If } (x, y) \text{ is a hole} \\ 1 & \text{Otherwise} \end{cases}$$

### 3.7 Approximation:

The approximation can increase performance and reduce the time complexity issues especially in image processing applications. For applications related to human senses, approximate arithmetic can be used to generate sufficient results rather than absolutely accurate results. Approximate design exploits trade-off accuracy in computation versus performance. However, required accuracy varies according to applications. In present era for multimedia application hardware complexity and memory conception are main issue. Human eye sense accuracy of 90-95% so why we need accurate algorithm for reduction of complexity issue we can apply approximation unit.

## 4. IMPLEMENTATION DETAIL OF DIBR:

From the block diagram the complete process of DIBR is shown. Initially a color image is taken to calculate the depth .The color image is processed through RGB to YCbCr to convert into grey scale image such that the computation time is reduced by one third of a color image. In next step the offset of image is calculated by sending a depth image which is preprocessed by mentioned method previous chapters. From there the right and left image are calculated by proposed formula The next step is preprocessing of the depth of the image is smoothing of it by smoothing filter.

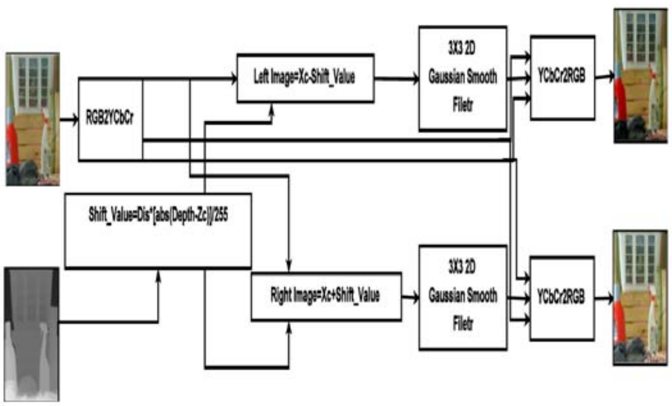



**Figure6: Block diagram of complete process to calculate DIBR of image**

The smoothing is step processed because the depth image with horizontal sharp transition to reduce the number of big holes. Eventually if we blur the image instead of big holes small holes are wrapped view. The depth map of non whole area is smoothed. After both images are extracted they are turned back into color by using the original formula which inverse transform of original method.

## 5. RESULT AND ANALYSIS

The comparative analysis is perform for different existing and proposed RGB to YCbCr and Depth image based rendering process. Analysis is performing by using of MATLAB. Here generated image from different approach is passing from image quality measurement parameter which are:

- PSNR
- SSIM
- RFSIM
- FSIM
- Absolute Similarity (%)

### 5.1 Image Quality Parameter Results: Average Calculation

| PARAMETER | 3DTV | AES | FAPU |
|---|---|---|---|
| PSNR | 16.31 | 0.854 | 21.18 |
| SSIM | 0.836 | 0.854 | 0.923 |
| RFSIM | 0.1325 | 0.3801 | 0.4277 |
| FSIM | 0.9059 | 0.9302 | 0.958 |
| SIMMILARITY % | 71.005 | 75.86 | 82.83 |

**Table1: Image Quality Parameter Results: Average Calculation**

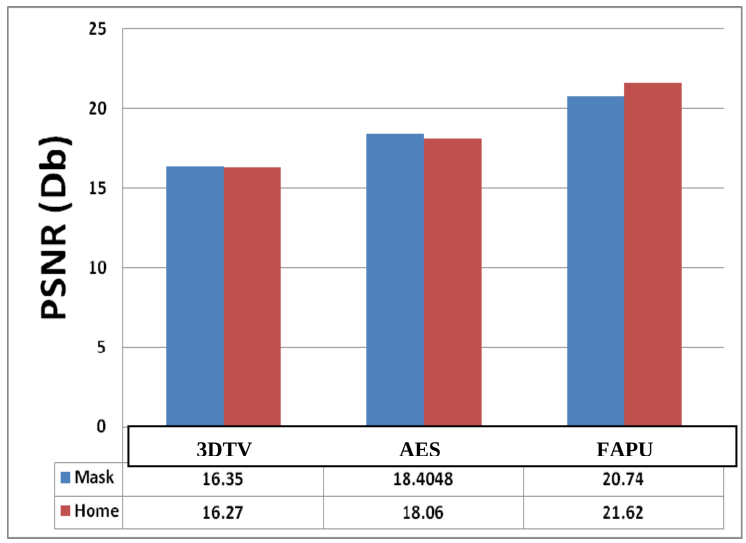



**Figure7: Comparative PSNR Analysis**

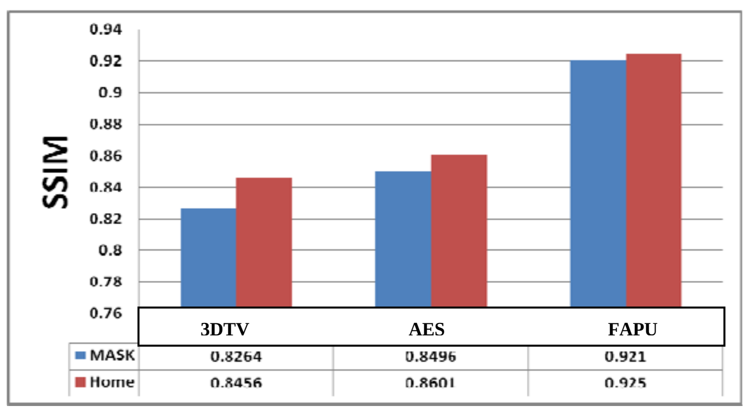



**Figure8: Comparative SSIM Analysis**

ISSN 2348 – 7968

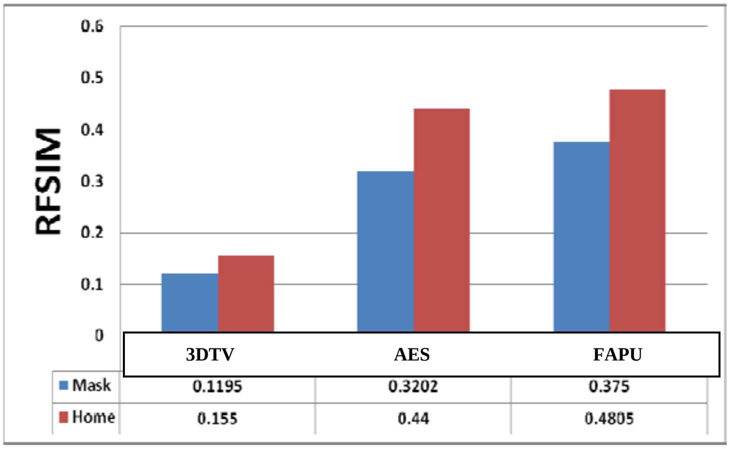



**Figure9: Comparative RFSIM Analysis**

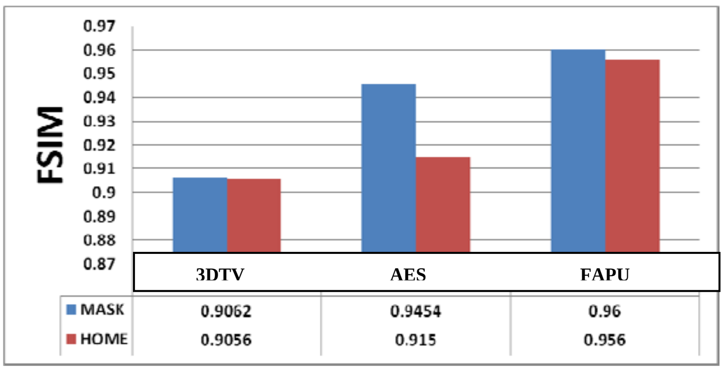



**Figure10: Comparative FSIM Analysis**

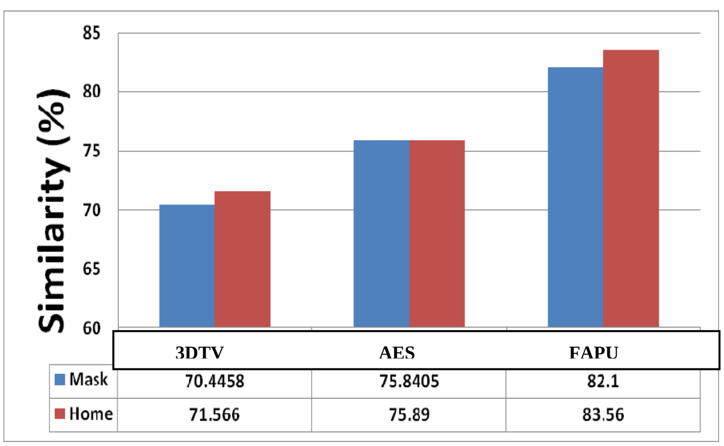



**Figure11: Comparative Absolute Similarity Analysis**

As we can see in this result proposed DIBR approach generated image quality is better than previous approaches like DIBR AES[3] and DIBR 3DTV[19].

## 6. CONCLUSION

Depth Image Based Rendering (DIBR) process is basically used in 2D to 3D conversion. In this approach we are proposed a new algorithm for RGB to YCbCr conversion and DIBR. In proposed approach time complexity is reduce and generated output image quality is measured by using of some image quality parameter like PSNR, SSIM, FSIM, RFSIM, Similarity (%). Simulation results shows that time complexity of DIBR is reduced by 38% compare to existing process. Image quality of generated image is better than existing approach.

## 7. REFERENCE

[1] Neetesh Nema, Bragesh Patel, "A Survey: Fast & Approximate Algorithm of Depth Image Based Rendering Process for Application of 2D to 3D Conversion" IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 1 Issue 10, pp 744-747, December 2014.

[2] Jui-Sheng Lee; Guo-An Jian; Cheng-An Chien; Peng-Sheng Chen; Jiun-In Guo, "3D depth map generation for embedded stereo applications," Visual Communications and Image Processing (VCIP), 2012 IEEE , vol.no. 1, pp. 27-30 Nov. 2012.

[3] Hao Liu; Wei Guo; Chao Lu; Jizeng Wei, "An Efficient Stereoscopic Game Conversion System for Embedded Platform," Trust, Security and Privacy in Computing and Communications (TrustCom), 2011 IEEE 10th International Conference on, vol. no.5, pp. 16-18 Nov. 2011.

[4] Lin Zhang; Zhang, D.; Xuanqin Mou; Zhang, D., "FSIM: A Feature Similarity Index for Image Quality Assessment," Image Processing, IEEE Transactions on, vol. no. 20, pp. 231-235, Aug. 2011.

[5] H. Emrah Tasli, Kemal Ugur, "Interactive 2d 3d Image Conversion Method For Mobile Devices", 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON), IEEE, Vol 5, no. 4, pp. 1-4, 2011.

[6] Lin Zhang; Zhang, D.; Xuanqin Mou, "RFSIM: A feature based image quality assessment metric using Riesz transforms," Image Processing (ICIP), 2010 17th IEEE International Conference on, vol. no 7, pp. 26-29 Sept. 2010.

[7] Chao-Chung Cheng; Chung-Te Li; Liang-Gee Chen, "A novel 2Dd-to-3D conversion system using edge information," Consumer Electronics, IEEE Transactions on , vol. no. 56, pp. 39-43, Aug. 2010.

[8] Nguyen, Q.H.; Do, M.N.; Patel, S.J., "Depth image-based rendering with low resolution depth," Image Processing (ICIP), 2009 16th IEEE International Conference on, vol. no. 8, pp. 7-10 Nov, 2009.

[9] Hang Shao; Xun Cao; Guihua Er, "Objective quality assessment of depth image based rendering in 3DTV system," 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, vol. no 7, pp. 4-6 May 2009.

[10] Xue, W.; Zhang, L.; Mou, X.; Bovik, A., "Gradient Magnitude Similarity Deviation: A Highly Efficient Perceptual Image Quality Index," Image Processing, IEEE Transactions on, vol no 3, pp-99-105, 2008.

[11] Jianmin Jiang and Guoqiang Xiao, Object-based pseudo-threedimensional conversion of two-dimensional videos, Opt. Eng. 48, 077007,doi:10.1117/1.3159873, 2008.

[12] Kwanghee Jung; Young Kyung Park; Joong Kyu Kim; Hyun Lee; Kugjin Yun; Hur, N.; Jinwoong Kim, "Depth Image Based Rendering for 3D Data Service Over T-DMB," 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, vol. no 2, pp. 28-30 May 2008.

[13] K. Lee, M. Kim, N. Dutt, and N. Venkatasubramanian, “Error-exploitingvideo encoder to extend energy/QoS tradeoffs for mobile embeddedsystems,” vol. 271, pp. 23–34, 2008.

[14] Wan-Yu Chen; Chang, Yu-Lin; Hsu-Kuang Chiu; Shao-Yi Chien; Liang-Gee Chen, "Real-Time Depth Image based Rendering Hardware Accelerator for Advanced Three Dimensional Television System," Multimedia and Expo, IEEE International Conference on, vol no 6, pp. 9-12 July 2006.

[15] Wan-Yu Chen; Chang, Yu-Lin; Shyh-Feng Lin; Li-Fu Ding; Liang-Gee Chen, "Efficient Depth Image Based Rendering with Edge Dependent Depth Filter and Interpolation," Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on, vol. no. 5, pp.145-150, July 2005.

[16] C. Fehn, K. Hopf, Q. Quante, *Key Technologies for anAdvanced 3D-TV System*, in Proc. of SPIE Three-Dimensional TV, Video and Display, vol no. 5, pp. 66-80, Oct. 2004,

[17] Zhou Wang; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P., "Image quality assessment: from error visibility to structural similarity," Image Processing, IEEE Transactions on, vol. no. 13, pp.606-612, April 2004.

[18] C. Fehn, Depth-Image-Based Rendering (Dibr), compression and transmission for a new approach on 3D-TV, in: Proceedings of the SPIE Stereoscopic Displays and Virtual Reality Systems XI, San Jose, CA, USA, vol. no. 3, pp. 93-104, January 2004.

[19] Fehn, Christoph. "A 3D-TV approach using depth-image-based rendering (DIBR)." Proc. of VIIP. Vol. 3, pp. 7 -12, 2003.

[20] Rec. ITU-R BT.601-5, Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide-screen 16:9 Aspect Ratios, (1982-1986-1990-1992-1994-1995).

[21]http://www02.lps.ufrj.br/~eduardo/MMP/3D_results.html.

[22] http://www.ni.com/white-paper/13306/en.

[23]http://homepages.inf.ed.ac.uk/rbf/HIPR2/gsmooth.html.