ISSN: 2278 – 909X
International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)
Volume 5, Issue 2, February 2016

A review paper on Swift & Novel Algorithm for 2D to 3D Conversion of HD Image/Video Using Energy Reduction Approach

Anamika Patre[1], Ravi Tiwari[2]

1 M.Tech Scholar, Dept. of Electronics & Communication Engg.
Dr. C.V. Raman University Bilaspur
Chattisgarh-India

2 Asst. Prof, Dept. of Electronics & Communication Engg.
Dr. C.V. Raman University Bilaspur
Chattisgarh-India

***Abstract*— Video, audio and multimedia offer powerful means of communication. Moving pictures are excellent for showing how things change or how something is done, for establishing a context for information (such as a landscape or a working environment) to make it easier for an audience to relate to what you are saying. Three-dimension (3D) technology increases the visual quality as compared to (2D) technology. In present era every multimedia device needs 3D technology. The depth information needed for the generation of 3D form 2D content. Therefore, conversion of 2D video/image into 3D has become an important issue in emerging 3Dapplication. 2D to 3D conversion is basically based on accurate algorithm. This work presents Swift & Novel algorithm for 2D to 3D Conversion of HD Image/video. This approach will automatically converts 2D content into 3D content. In this work a novel approximate algorithm for 2D to 3D conversion. That algorithm include depth map generation unit and depth image based rendering (DIBR) process.**

Keywords: SPAA, DIBR, Gaussian, 2D, 3D, Approximation, HD.

I. INTRODUCTION

1.1 Overview

WITH the recent advancement in 3-D displays, 3-D content visualization has become very popular in our daily life [1]. Recently, 3-D contents are being used in 3-D photography [2], 3-D broadcasting system [3], 3-D movie [4], etc. In contrast to the traditional 2-D content visualization, 3-D content provides an impressive visual experience and heightens a sense of realism with depth perception of the observed scene. 3D video is getting immense public attention recently because of vivid stereo visual experience over conventional 2D video. There are several methods to produce 3D content, such as active depth sensing, stereo camera recording, and 3Dgraphics rendering using of this approach 3D content is generated.

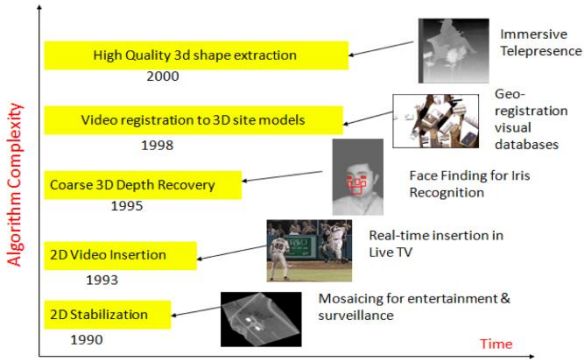


Figure 1.1: Vision algorithm performance over time

As we can see in fig. 1.1vision algorithm performance flow according to increase in time. Quality of video increases now in present era Video has be- come more realistic with heavy resolution, at present 720P, 1080P and 1280P are known as HD video, This HD video are now convert into from of 3D HD video so for market demand 2D to 3D conversion is needed. Their are many area which require 2D to 3D conversion like medical, entertain- ment and lm industry. At present era film industry are mainly using 2D to 3D conversion technique.

1.2 Glass Work

Why 3D glass is needed? 3D image is combination of two image i.e left & right image and by help of 3D glass we can feel 3D effect because these both image are mask with red and blue color and 3D glass project depth on out mind.

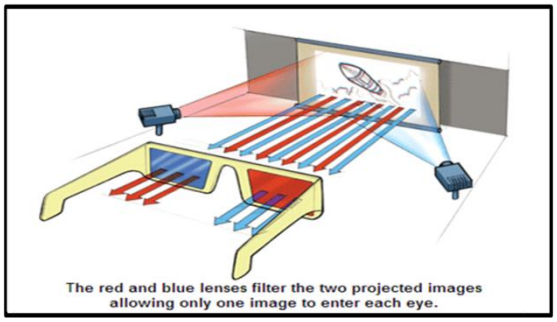


Figure 1.2: How 3D Glass Work

2D to 3D conversion is mostly based on accurate algorithm and accurate algorithm is faced with time complexity problem. So for reduction of this problem many researchers are working in this area. For reduction of time complexity issue we was apply approximation. it is a key for solution of time complexity related problem at algorithm . As we know approximation means we can apply some modification on design and based on that design we will get some output and error at particular output is neglect by human eye. Human eye accuracy level is up to 90% to 95% so why we design those logic which give 100% result.

1.3 Motivation

1. Need of 2D to 3D conversion-

- As we see that 3D technology is uses in many image/video application and at present image/movie are captured in the form of 2D.
- If we are using 3D technology for capturing of image/video so it is an expensive process.
- In every 3D application, inputs are taken in the form of 2D image/video and by using some conversion technique output is converted into 3D image/video.
- Example: This phenomena are mainly using in 3DTV concept because our input image/video are captured in the form of 2D and by satellite we are receiving 2D data, then these 2D data are converted into 3D. With the help of 3D glass we can watch that 3D data.

II. LITERATURE REVIEW

The author in [1]- (Image Quality Assessment: From Error Visibility to Structural Similarity, 2004);

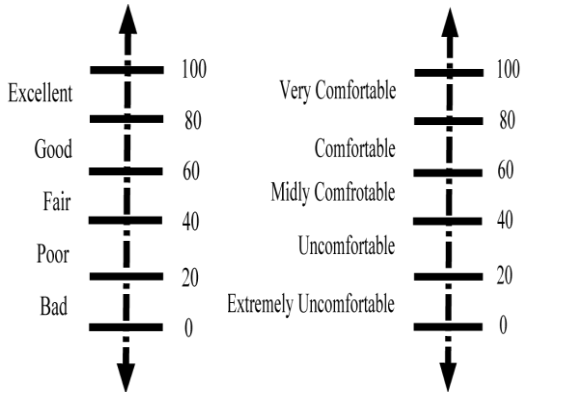
In this paper objective methods for assessing perceptual image quality traditionally attempted to quantify the visibility of errors (differences) between a distorted image and a reference image using a variety of known properties of the human visual system. Under the assumption that human visual perception is highly adapted for extracting structural information from a scene, we introduce an alternative complementary framework for quality assessment based on the degradation of structural information. As a specific example of this concept, we develop a Structural Similarity Index and demonstrate its promise through a set of intuitive.

- Objective Analysis

1. Peak signal to noise ratio
2. Structural similarity index (SSIM) [4]
3. RIESZ transformed based feature similarity index [5]
4. Feature similarity index (FSIM) [6]
5. Gradient magnitude standard deviation (GMSD) [7]
6. Absolute similarity
7. Correlation

- Subjective Analysis

Subjective analysis [32] is a method which is used to check the quality and Visuality of the generated output. In 2D to 3D conversion subjective analysis is also a very import part. Using of this analysis we can check visual quality and depth of generated 3D content. Here rating scale for different parameter are shown in below.



(a) Depth rating scale (b) Visual rating scale

Figure 2.1: Rating Scale for evaluation of generated 3D content

The author in [2]- Efficient Depth Image Based Rendering with Edge Dependent Depth Filter and Interpolation [2005]; An efficient depth image based rendering with edge dependent depth filter and interpolation is proposed. The proposed method can solve the hole-filling problem in DIBR
system efficiently with high quality. The PSNR of the proposed method is better than the previous work by 6 dB. And the subjective view shows the quality is better. In addition to that, the number of instruction cycles is 3.7 percent compared with the previous work

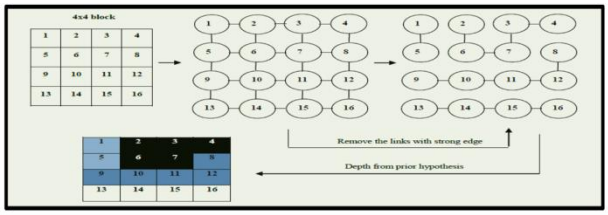
The author in [3]- Edge Information (Transaction on Consumer Electronics, 2010);

A. Approach-

1.Block Generation-

Create group of 4X4 block than Estimate mean of current and neighbour.

Calculate Difference and compare from threshold value.



2.Depth Generation-

$$Depth(R) = 128 + 255\left\{\sum_{pixel(x,y)} Wrl\frac{x-\frac{width}{2}}{width} + Wud\frac{y-\frac{height}{2}}{height}\right\} \Big/ pixel_{num(R)}$$

$$Where\ |Wrl| + |Wud| = 1$$

Generated Depth map image is apply to Bilateral Filter, this filter smoothen the edge of output depth-image.

The author in [4]- Optimized DIBR Process (Trust Com, 2011);

A. Approach-

In Previous Basic DIBR approach for Hardware point of view there is major draw back is designing of Divider logic for reduction of this problem new approach using shifting logic and replace (1=depth) with (256-Depth)=Depth.

B. Problem-

examples, as well as comparison to both subjective ratings and state-of-the-art objective methods on a database of images compressed with JPEG and JPEG2000.

In this approach there is a problem of large time complexity issue, Because in this process there is need of two separate unit which generate left and right image.

Modified left and right image generation formula:

$$D = 256 - depth$$

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

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

The authors in [5]- (2D to 3D Conversion system 2010);

3D world is projected onto each retina of two eyes with slightly different 2D views. Therefore, objects appear in different positions relative to each other and this is called "binocular disparity". Mechanism of HVS is that the brain uses to combine the information from both eyes to make a sense of depth is known as stereopsis. The basis of the 2D-to-3D conversion starts from finding robust monocular depth cues of a given image. In the following, we briey review the previous related works according to their main depth cues.

Currently, 3D content can be generated in one of three ways. One approach Triangular stereo vision is to use two or more camera systems. Several sports events such as the world cup and the super bowl have been globally broad-casted in this manner due to the high quality of the captured stereoscopic images. Although this approach allows viewers to get realistic depth perception by directly acquiring the observed scene, there are some inherent di-culties in producing 3-D contents.

1. Cut-paste technique

The easiest way to create 3D image by using of the original 2D image as the left eye view and generate the right eye image by horizontally shifting local regions of the original image, using a cut-and-paste technique. For the isolation of the local regions image segmentation techniques can be used, but problem with this approach is objects are well separated, but for images with small objects, large areas with low textures and little gradations of depth. The depth is assigned to each object but depth is not to each pixel[15].

2. Depth map generation

When observing the world, the human brain integrates various heuristic depth cues to generate the depth perception. The major depth perceptions are binocular depth cues from two eyes and monocular depth cues from a single eye. binocular visual system. In this system both eyes are used together. human eyes to converge and accommodate the object at the right distance monocular visual system In this effect each eyes are used separately. Including focus/defocus, motion parallax, relative height/size, and texture gradient, provide various depth perceptions based on human experience. Therefore, humans can also perceive depth from the single-view image/video. Depth map generation method is combination of two approach which are single frame and multi frame. In this approach generated Depth map is real because of generated 5 depth map is for each pixel.

2.1. Single frame

Depth from color/intensity People feel red (warm) color is nearer, and blue (cold) color is farther in visual perception. Besides, objects with higher luminance feel like nearer than those with lower luminance. Therefore, color can be used as a depth cue to enhance the depth perception on both edge and color domains. Based on the concept, a novel combination of Y, Cr, and Cb color channels to generate the ne-grained depth map,but problem with this technique is generated depth map is look like virtual depth [16].

2.2 Depth from edge information

This approach basically reduce computation complexity of depth mapping. In this approach author divide one image into group of 4x4 block. Estimate mean of that group and similar calculate mean of neighbour group than taking difference between both group and pass that result from one threshold value that threshold value decide region of groups, after getting region value apply depth map formula using that formula depth image is generated but this approach have many problems like time complexity, memory consumption [21].

3.Depth image based rendering

The purpose of DIBR is to generate 3D left and right virtual views from the intermediate view which comes from the original 2D view image. DIBR consists of two parts: 3D warping and in painting (hole filling). 3D Warping 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 [25],[35].

$$Xl = Xc - \frac{tc * f}{2 * Depth} + \frac{-tc * f}{2 * Zc}, \quad Xr = Xc + \frac{tc * f}{2 * Depth} - \frac{-tc * f}{2 * Zc}$$

Here Xl 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. f is camera focal length and tc is distance between left and right eye retina. Zc is convergence distance. The formula shows that 3D warping maps pixel of intermediate view to left and right view in horizontal direction [25], [35].

4. Hole filling

Linear Interpolation, Edge dependent depth filter or Average filter interpolation [25] method is a common method for Hole-Filling in DIBR. However, using average filter only would result in artifacts at highly-textured areas. Besides, hole-size in DIBR is so huge that it is needed to using average filter with large window size. At the same time, average filter with large window size can not preserve edge information for the reason that edge information is blurred. If the current pixel is a hole, so by use of this filter technique add color into the hole.

The authors in [6]- (A Novel 2D-to-3D Conversion System Using Edge Information, 2010);

developed a three-dimensional (3D) displays enhance visual quality more than two-dimensional (2D) displays do, the depth information required for 3D displays is unavailable in the conventional 2D content. Therefore, converting 2D videos into 3D ones has become an important issue in emerging 3D applications. This work presents a novel algorithm that automatically converts 2D videos into 3D ones. The proposed algorithm utilizes the edge information to segment the image into object groups. A depth map is then assigned based on a hypothesized depth gradient model. Next, the depth map is block-based assigned by cooperating with a cross bilateral filter to generate visually comfortable depth maps efficiently and also diminish the block artifacts. A multiview video can be readily generated by using a depth image based rendering method.

The authors in [7] Vivid-DIBR Based 2D–3D Image Conversion System for 3D Display(2014);

The multi-view perspectives are captured around a scene by camera array at the same time, and multi-view images cannot only keep two dimensions information but also has the depth architecture of a scene. Most of 3D displays support the multi-view contents, while there is a bottleneck of the 3D TV system to broadcast the larger storage and bandwidth for end user. For saving bandwidth, a new TV broadcasting channel transfers third dimensional information, i.e., depth map has been utilized. The 3D display renders intermediate image and a depth map to form different views, as well as synthesizes these images into one image; this technology is the renowned Depth Image Based Rendering (DIBR). In other words, any 2D content can be transformed into multi-view images according to their corresponding depth map. Our system processes 2D images to capture a depth map of the intermediate view image during 3D contents capturing. In this paper, we utilize the vanishing point detection and color image segmentation technique to find the objects and deepest point (i.e., vanishing point) in the image, and then we assign depth value by comparing the vanishing point of object with image. After depth map generation, we propose the Vivid-DIBR system to imitate how human eyes see things. On the other hand, this system solves the holes (warping error points) problem by redistributing corresponding depth map. The design aims to convert 2D images into multi-view images which will be suitable for any interlacing 3D display by adjusting the focal plane location.

The authors in [8] (A Real-Time 1080p 2D-to-3D Video Conversion System, 2011);

In this paper, a 2D-to-3D video conversion system capable of real time conversion of 1920×1080p 2D video to 3D video is presented. System fuses global and local depth generation modules to generate depth image, and use depth image based rendering(DIBR) algorithm to render 3D video. The system is implemented both on software and hardware. Software is based on multi-core system with CUDA platform. To optimize performance, several techniques are proposed, including unified streaming dataflow, multi-thread schedule synchronization, and CUDA acceleration. Real-time 1920×1080p 2D-to-3D video conversion system running at 30fps is achieved.

The author in [9]- (3D-TV Content Creation: Automatic 2D-to-3D Video Conversion, 2011);

Three-dimensional television (3D-TV) is the next major revolution in television. A successful rollout of 3D-TV will require a backward-compatible transmission/distribution system, inexpensive 3D displays, and an adequate supply of high-quality 3D program material. With respect to the last factor, the conversion of 2D images/videos to 3D will play an important role. This paper provides an overview of automatic 2D-to-3D video conversion with a specific look at a number of approaches for both the extraction of depth information from monoscopic images and the generation of stereoscopic images. Some chall enging issues for the success of automatic 2D-to-3D video conversion are pointed out as possible research topics for the future.

III. RESEARCH GAP

Research gap of 2D to 3D conversion process on Algorithm level:

- Time Complexity
- Lack in quality depth of object
- Lack in quality of visual perception
- Virtual Depth

IV. APPLICATIONS AND FUTURE SCOPE

As we know in present era every one need realistic view , and that view is only possible by using of 3D logic. We also know there is lots of approaches are there which will generate 3D image/video. So there is followings application where we are using 2D to 3D Conversion:

1. Medical Image Processing
2. Digital Image Processing
3. Video image Processing
4. Mobile Application
5. 3D TV Technology
6. 3D Printer Technology

V. CONCLUSION

As we already saw there is need of 3D technology due to high vision quality requirement. 2D to 3D conversion is a combination of depth map, DIBR and hole filling. For depth

map generation edge detection, Gaussian filter etc process are require. Depth image based rendering process is most important process to generate 3D image because at initial stage using Depth map generation we generate one more dimension. Now using DIBR we can generate left and right image which is require for generation of 3D technology. According to previous existing approach there is lots of issues which is still a main problem. Those problems are like time complexity and memory complexity. Still there is lots of scope for future object for generation of 3D content using 2D image/video.

ACKNOWLEDGMENT

The authors would like to thank the Department of ECE at DR. C. V. RAMAN UNIVERSITY for facilitating the development of the project, making available resources for testing the system and also for final deployment.

REFERENCES

[1] N. Holliman, N. Dodgson, G. Favalora, and L. Pockett, "Three dimensional displays: A review and applications analysis", IEEE Trans. Broadcasting, vol. 57, no. 2, pp. 362,371, Jun. 2011.

[2] Stereo Photography [Online]. Available: http://www.3dphoto.net

[3] C. Fehn, R. Barre, and S. Pastoor, "Interactive 3D-TV: Concepts and key technologies", Proc. IEEE, vol. 94, no. 3, pp. 524â€“ 538, Mar. 2006.

[4]] IMAX [Online]. Available: http://en.wikipedia.org/wiki/IMAX

[5] M. Kim, J. Nam, and W. Baek, "The adaptation of 3-D stereoscopic video in MPEG-21 DIA", Signal Proc. Image Commun., vol. 18, no. 8, pp. 685,697, 2003.

[6] Y. Taguchi, T. Koike, and K. Takahashi, TransCAIP: "A live 3DTV system using a camera array and an integral photography display with interactive control of viewing parameters", IEEE Trans. Vis. Comput. Graph., vol. 15, no. 5, pp. 841,852, Sep.â€“ Oct. 2009.

[7] A. Woods, T. Docherty, and R. Koch, "Image distortions in stereoscopic video systems", in Proc. SPIE (Stereoscopic Displays Appl.), vol. 1915. 1993, pp. 36,48.

[8] H. Yamanoue, M. Okui, and F. Okano, "Geometrical analysis of puppettheater and cardboard effects in stereoscopic HDTV images", IEEE Trans. Circuits Syst. Video Technol., vol. 16, no. 6, pp. 744,752, Jun.2006.

[9] C. Fehn, "A 3D-TV system based on video-plus-depth information", in Proc. Asilomar Conf. Signals Syst. Comput., vol. 2. 2003, pp.1529,1533.

[10] L. Zhang and W. Tam, "Stereoscopic image generation based on depth images for 3D-TV", IEEE Trans. Broadcasting, vol. 51, no. 2, pp. 191,199, Jun. 2005.

[11] D. Nagahara and S. Takahashi "Mobile robot control based on information of the scanning laser range sensor", in Proc. IEEE Int. Workshop Advanced Motion Control, Mar. 2010, pp. 258,261.

[12] J. Zhu, L. Wang, R. Yang, and J. Davis, "Fusion of time-of-ight depth and stereo for high accuracy depth maps", in Proc. IEEE Int. Conf. Comput. Vision Pattern Recognit., Jun. 2008, pp. 1,8.

[13] J. Zhu, L. Wang, J. Gao, and R. Yang, "Spatial-temporal fusion for high accuracy depth maps using dynamic MRFs", IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 5, pp. 899,909, May 2010.

[14] L. Zhang, C. Vazquez, and S. Knorr, "3D-TV content creation: Automatic 2D-to- 3D video conversion", IEEE Trans. Broadcasting, vol. 57, no. 2, pp. 372,383, Jun. 2011. 75

[15] Stelmach, L.; Wa James Tam; Meegan, D.; Vincent, A., "Stereo image quality: effects of mixed spatio-temporal resolution," Circuits and Systems for Video Technology, IEEE Transactions on , vol.10, no.2, pp.188,193, Mar 2000

[16] Sung-Fang Tsai; Chao-Chung Cheng; Chung-Te Li; Liang-Gee Chen, "A real-time 1080p 2D-to-3D video conversion system," Consumer Electronics (ICCE), 2011 IEEE International Conference on , vol., no., pp.803,804, 9-12 Jan. 2011

[17] Jungwoo Park ; Changick Kim; "Extracting focused object from low depth-of-eld image sequences". Proc. SPIE 6077, Visual Communications and Image Processing 2006, 60771O (January 18, 2006); doi:10.1117/12.642319.

[18] Yi-Min Tsai; Chang, Yu-Lin; Liang-Gee Chen, "Block-based Vanishing Line and Vanishing Point Detection for 3D Scene Reconstruction," Intelligent Signal Processing and Communications, 2006. ISPACS '06. International Symposium on , vol., no., pp.586,589, Dec. 2006

[19] Sebastiano Battiato ; Salvatore Curti ; Marco La Cascia ; Marcello Tortora ; Emiliano Scordato; "Depth map generation by image classication". Proc. SPIE 5302, Three-Dimensional Image Capture and Applications VI, 95 (April 16, 2004); doi:10.1117/12.526634.

[20] Yong Ju Jung ; Aron Baik ; Jiwon Kim ; Dusik Park; "A novel 2D-to-3D conversion technique based on relative height-depth cue". Proc. SPIE 7237, Stereoscopic Displays and Applications XX, 72371U (February 18, 2009); doi:10.1117/12.806058.

Anamika Patre has received her Bachelor of Engineering degree in Electronics & Telecommunication from CSVTU ,in the year 2012. At present she is pursuing M.Tech. with the specialization of Digital Electronics CVRU Bilaspur. Her area of interest Embedded system & by using microcontroller design embedded systems.

Ravi Tiwari has received his Bachelor of Engineering degree in Electronics & Telecommunication GGU Bilaspur and M.Tech in Digital communication from RGPV. At present he is Asst. Prof. in ECE department CVRU Bilaspur. His area of interest wireless mobile communication, digital communication, Wireless chip desining.