ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X
International Journal of Advanced Research in Biology, Engineering, Science and Technology (IJARBEST)
Vol. 2, Issue 5, May 2016

AN EFFICIENT APPROACH FOR 2D TO 3D IMAGE CONVERSION USING FUZZY C-MEANS SEGMENTATION

M.SHEIK MANSOOR,
M.Phil Scholar,
Sadakathullah Appa College,
Tirunelveli.
sheik_haya@yahoo.co.in

Dr.M.MOHAMED SATHIK,
Principal,
Sadakathullah Appa College,
Tirunelveli.
mmdsadiq@gmail.com.

Abstract— **The main objective of this paper is to design an efficient 2D-to-3D conversion method based on the use of edge information. The edge of an image has a high probability as it can be the edge of the depth map. Once the pixels are grouped together, a relative depth value can be assigned to each region. Initially, the input RGB image is converted into HSV color space. And then it is converted into gray image. And then the block-based image is considered to segment it into multiple groups. To segment the image this paper uses the Fuzzy C-Means Segmentation Approach. Then the depth of each segment is assigned with the help of an initial depth hypothesis. Next, the blocky artifacts have to be removed using cross bilateral filtering. Finally, multi-view images are obtained by the method of DIBR. As a result, the input 2D image is converted into visually comfortable 3D image without the presence of artifacts enhancing the quality of the image in the display. To analyse the performance of the proposed method several performance metrics are used. This paper uses PSNR,SSIM, MSE and RMSE to analyses the performance. From the experimental results it is shown that the proposed method performs better than the other existing methods.**

Keywords— ***2D-to-3D conversion, depth boundaries, depthmap,***

nonlocal neighbors, nonlocal edge information.

Introduction

3DTV is widely anticipated as the next revolution of television technology. '3D' (three dimensional) here means 'stereoscopic', which provides viewers with feeling of immersion. The promotion of 3DTV depends not only on technological advances in 3D displays, but also on availability of large amount of 3D content. However, it's both costly and time-consuming to make 3D content of high quality directly with stereoscopic cameras, so the shortage of 3D content becomes one of severe bottlenecks for 3D industry. Converting 2D images or videos to 3D is one way to alleviate the problem in the early stage of 3DTV development, because it not only can create 3D content with lower cost and less time, but also makes full use of large amount of existing 2D materials.

Generally, the existing 2D to 3D conversion approaches can be classified as two categories: human assisted conversion and automatic conversion. The human-assisted approach is to convert 2D images or videos to 3D with some corrections made "manually" by an operator [1]. Even though this approach has relatively better performance, it's still impractical in many scenarios. To convert the vast collection of available 2D material into 3D in an economic manner, an automatic approach is desired [2]. The automatic approach utilizes the depth cues in a single monocular image to generate another or more virtual views without any human assistance.

There is several research works are progressing on 2D to 3D conversion of images which shall be used in the motion pictures [4] and [5]. 3D imaging system has been incorporated in the televisions, cameras etc. In the health system the 3D body scanners help surgeons to determine the accurate status of various diseases. The 3D hardware is expensive compared with 2D hardware system. Therefore, it is necessary to develop a fast and accurate algorithm for converting 2D images to 3D images. In this research article, a new simple algorithm is proposed for converting 2D image to 3D image using image fusion. Xiaoyang Mao, lbsiyasu L. Kunii, Hierarchical was proposed G-octree as an extension of G-quadtree to 3D grey-scale images. They did the program in C on VAX 11/750. Application to the color coding of macro-auto-radiography images of rat brains demonstrated the advantages of the approach [6]. Chin-Tung Lin, Chiun-Li Chin, Kan-Wei Fan, and Chun-Yeon Lin was presented a 2D to 3D effect image conversion architecture integrated image segmentation system and depth estimation. They tested many 640*480 RGB format color images. They generated left view and right view image and displayed the 3D stereo image [7]. H. Murata, X Mori, S. Yamashita, A. Maenaka, S. Okada, K. Oyamada, and S. Kishimoto, proposed a system for converting all kinds of 2D images into 3D images. The method is used adaptively by computing the depth of each separated area of the 2D images with their contrast, sharpness, and chrominance [8]. Wa James Tam and Liang Zhang provided an overview of the fundamental principle of 2D to 3D conversion techniques, short note on approaches for depth extraction using a single image, and depth image based rendering [9]. Chao-Chung Cheng, Chung-Te Li, and Liang-Gee Chen presented an automatic system for converting 2D videos to 3D videos. They grouped the regions into blocks using the edge information and applied bilateral filter to generate depth map [10]. Zhebin Zhang, Yizhou Wang, Tingting Jiang, and Wen Gao described an approach which estimated a 2.5D depth map by leveraging motion cues and photometric cues in video frames [11]. Ching-Lung Su, Kang-Ning Pang, Tse-Min Chen, Guo-Syuan Wu, Chia-Ling Chiang, Hang-Rnei Wen, Lung-Sheng Huang, Ya-Hsin Hsueh, and Shau-Yin Tseng, presented an algorithm for conversion of 2D to 3D in real time. The 2D video accompanied with a depth image was stored to create 3D video [12]. Yeong-Kang Lai, Yu-Fan Lai, and Ying-Chang Chen proposed a hybrid algorithm for 2D to 3D conversion. They used motion information, linear perspective, and texture characteristic for depth estimation. They used bilateral filter for depth map smoothing and noise removal [13].

Two approaches to 2D to 3D conversion can be loosely defined: quality semiautomatic conversion for cinema and high quality 3DTV, and low-quality automatic conversion for cheap 3DTV, VOD and similar applications. [14] In semiautomatic conversion a skilled operator assigns depth to various parts of an image or video. Based on this sparse depth assignment, a computer algorithm estimates dense depth over the entire image or video sequence. In the case of automatic methods, no operator intervention is needed and a computer algorithm automatically estimates the depth for a single image or video. Automatic methods estimates shape from shading, structure from motion or depth from defocus. Electronics manufacturers use stronger assumptions to develop real-time 2D-to-3D converters. Such methods may work well in specific scenarios. But generally it is very difficult to construct heuristic assumptions that cover all possible background and foreground combinations. An important step in any 3D system is the 3D content generation. Several specialcameras have been designed to generate 3D model directly. For example, a stereoscopic dual-camera makes use of a co-planar configuration of two separate, monoscopiccameras, each capturing one eye's view, and depth information is computed usingbinocular disparity. A depth-range camera is another example. It is a conventional videocamera enhanced with an add-on laser element, which captures a normal two-dimensionalRGB image and a corresponding depth map. A depth map is a 2D function that gives thedepth (with respect to the viewpoint) of an object point as a function of the imagecoordinates. Usually, it is represented as a gray level image with the intensity of eachpixel registering its depth. The laser element emits a light wall towards the real worldscene, which hits the objects in the scene and reflected back. This is subsequentlyregistered and used for the construction of a depth map.All the techniques described above are used to directly generate 3D content, which certainly contribute to the prevalence of 3D-TV. However, the tremendous amount of current and past media data is in 2D format and should be possible to be viewed with a stereoscopic effect. This is where the 2D to 3D conversion method comes to rescue. This method recovers the depth information by analyzing and processing the 2D image structures.

In this paper an efficient 2D-to-3D conversion method is developed based on the use of edge information. The edge of an image has a high probability as it can be the edge of the depth map. Once the pixels are grouped together, a relative depth value can be assigned to each region. Initially, the input RGB image is converted into HSV color space. And then it is converted into gray image. And then the block-based image is considered to segment it into multiple groups. To segment the image this paper uses the Fuzzy C-Means Segmentation Approach. Then the depth of each segment is assigned with the help of an initial depth hypothesis. Next, the blocky artifacts have to be removed using cross bilateral filtering. Finally, multi-view images are obtained by the method of DIBR. As a result, the input 2D image is converted into visually comfortable 3D image without the presence of artifacts enhancing the quality of the image in the display.

The remainder of the paper is organized as follows: In Section II, the overview of proposed method is presented. In Section III, the proposed method is specifically depicted, including its design idea and practical implementation approach. In Section IV, the performance of the proposed method is evaluated. Finally, conclusions are made in Section V.

2D to 3D image conversion using edge information

The overall block diagram of this approach is shwon in Fig.1. This work uses an efficient 2D-to-3D conversion method based on the use of edge information. Importantly, the edge of an image has a high probability as it can be the edge of the depth map. Once the pixels are grouped together, a relative depth value can be assigned to each region. Figure 1 schematically depicts the proposed conversion system. Initially, the block-based image is considered to segment it into multiple groups. Then the depth of each segment is assigned with the help of an initial depth hypothesis. Next, the blocky artifacts have to be removed using cross bilateral filtering. Finally, multi-view images are obtained by the method of DIBR. As a result, the input 2D image is converted into visually comfortable 3D image without the presence of artifacts enhancing the quality of the image in the display. The further details of these modules are discussed below:

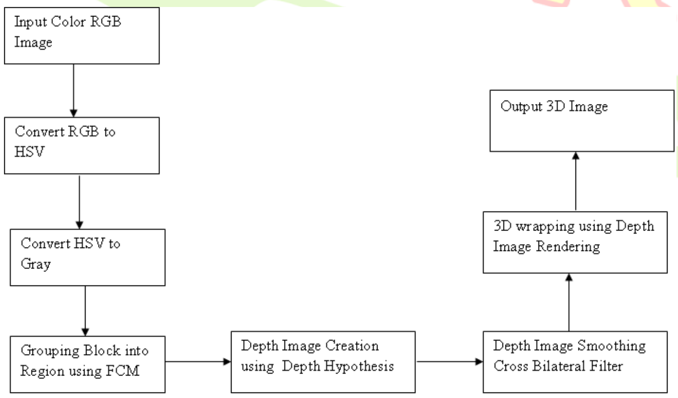


Fig. 1. Overall Block Diagram of Method1

Proposed approach

A. RGB to HSV Conversion

In image processing there are various color models: RGB model, CMY model, HSV model and YIQ model. RGB color model is used for color monitors and color video cameras, CMY model for printers, YIQ model is used in color TV broadcasting. Here RGB color model is converted into HSV color model. RGB (R-Red, G-Green and B-Blue) is a color model appears in primary colors and based on a Cartesian coordinate system. HSV is one of the frequently used color model. In HSV, H stands for hue which specifies the position of pure color on color wheel. Thus hue is related with dominant wavelength in a mixture of light wave. S is saturation, gives measure of the degree to which pure color is diluted with white. V is value called as lightness of color. Sometimes it is represented as I (intensity) or B (brightness). HSV model is having advantages over RGB model. HSV is strong model than RGB because it offers a more intuitive representation of the relationship between colors; it selects more specific color. RGB is costly in terms of computation time. The way in which human beings perceive color; hue and saturation components from HSV relates same way.

B. HSV to Gray Conversion

Convert HSV color image to grayscale image by averaging the H, S and V color space pixels.

C. Block-Based Region Grouping

Computational complexity is reduced mainly by block-based algorithm. This implies each pixel in the same block has the same depth value. A 4-by-4 block is used as an example. Each node is a 4-by-4 pixel block, and is four-connected. The value of each link is calculated by considering the absolute difference of the mean of neighboring blocks. A smaller value obtained implies a higher similarity between the two blocks. Following calculation of the absolute difference of the mean of the neighboring blocks, the blocks are then segmented into multiple groups by applying the fuzzy c-means (FCM) segmentation.

ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X
International Journal of Advanced Research in Biology, Engineering, Science and Technology (IJARBEST)
Vol. 2, Issue 5, May 2016

D. Depth from Prior Hypothesis

The extraction of depth is the crucial one in the conversion process. The greatest difference between 2D and 3D image is the depth information. The object can jump out of the screen and look like a real life due to the depth information. If we extract these depth signals and integrate them together, we will build a strong foundation to make 3D images of better and higher quality. The depth generation algorithms are roughly classified into three categories which utilize different kinds of depth cues: the binocular, monocular and pictorial depth cues. Each signal represents different depth information. In this conversion process, following the generation of the block groups, the corresponding depth for each block is then assigned by the depth gradient hypothesis. The process includes the generation of gradient planes, depth gradient assignment, consistency verification of the detected region, and finally the depth map generation. When each scene change is detected, the linear perspective of the scene can be analyzed with the help of line detection algorithm using Hough transform.

E. Bilateral Filtering

The bilateral filter is non-iterative and also achieves satisfying results with only a single pass. This makes the filter‟s parameters relatively intuitive as their effects are not cumulated over several iterations. The bilateral filter has proven to be much useful although it is slow. It is nonlinear and also its evaluation is computationally expensive because the traditional accelerations like performing convolution after an FFT, are not applicable. Nonetheless, solutions have been proposed later in order to speed up the evaluation of the bilateral filter. Unfortunately, these methods seem to rely on approximations that are not grounded on firm theoretical foundations. Among the variants of the bilateral filter, this conversion method has selected the cross bilateral filtering. In some applications like computational photography, it is often useful to decouple the data to be smoothed to define the edges to be preserved. The chosen cross bilateral filter is a variant of the classical bilateral filter. This filter is used to smoothen the image to locate the edges to preserve. The depth map generated by block-based region grouping contains blocky artifacts.

F. Depth Image Based Rendering

The filtered depth map has a comfortable visual quality because the cross bilateral filter generates a smooth depth map inside the smooth region with similar pixel values and preserves sharp depth discontinuity on the object boundary. Following filtering by the cross bilateral filter, the depth map is then used for the generation of the left/right or multi-view images using depth image-based rendering (DIBR) for 3D visualization. DIBR for advanced 3D TV System can be illustrated by the following block diagram. This system includes three parts, pre-processing of depth map, 3D image Warping and Hole- Filling. Smoothing filter is first stage applied to smooth the depth image. Then the 3D image warping generates the left and right view according to the smoothed depth map and also intermediate view. If there are still holes in the image, hole-filling is then applied to fill color into these holes.

G. 3D Image Warping

3D image warping maps the intermediate view pixel by pixel to left or right view according to the pixel depth value. In the other words, 3D image warping transforms the location of pixels according to depth value.

1. Expiremental Images

Experiments were conducted on a group of color images to verify the effectiveness of the proposed scheme. For the experimental purpose several standard, 512 × 512 cover images are taken. Some of these images , i.e., Lena, Barbara, Babbon, Peppers, Sailboat, and Tiffany, are shown in Figure 2.

(a) (b)

Fig. 2. Expiremental Images (a) Building (b) Boat

B. Performance Analysis

To evaluate the performance of the steganography techniques several performance metrics are available. This paper uses the PSNR,SSIM, MSE and RMSE to analyses the performance.

1. Peak Signal-to-Noise-Ratio

The peak signal-to-noise ratio (PSNR) is used to evaluate the quality between the 3D image and the original 2D image. The PSNR formula is defined as follows:

ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X
International Journal of Advanced Research in Biology, Engineering, Science and Technology (IJARBEST)
Vol. 2, Issue 5, May 2016

$$\text{PSNR} = 10 \times \log 10 \frac{255 \times 255}{\frac{1}{H \times W} \sum_{x=0}^{H-1} \sum_{y=0}^{W-1} [f(x,y) - g(x,y)]^2} \text{dB}$$

where H and W are the height and width of the image, respectively; and f(x,y) and g(x,y) are the grey levels located at coordinate (x,y) of the original image and attacked image, respectively.

2. Structural Similarity Index

The structural similarity index is a method for measuring the similarity between the 3D image and the original 2D image.

$$SSIM\ (y, \hat{y}) = \frac{\left(2_{\mu_y \mu_{\hat{y}}} + c_1\right)(2\sigma_{y\hat{y}} + c_2)}{(\mu_{\hat{y}}^2 + \mu_{\hat{y}}^2 + c_1)(\sigma_{\hat{y}}^2 + \sigma_{\hat{y}}^2 + c_2)}$$

where, $\hat{Y}$ is the 3D image, the Y is the original 2D image, μ is the mean and the is the variance.

3. Mean Square Error

The mean square error (MSE) is used to evaluate the difference between a 3D image and the original 2D image. The MSE can be calculated by,

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2$$

where, $\hat{Y}$ is the 3D image and the Y is the original 2D image.

4. Root Mean Square Error

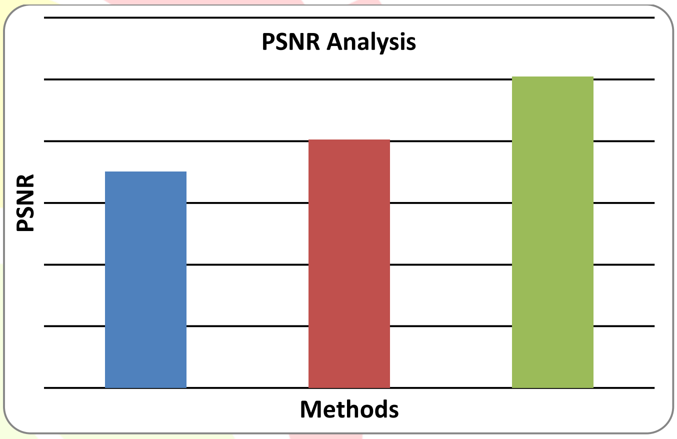
The Root Mean Square Error (RMSE) is a frequently used measure of the difference between 3D image values and the original 2D image values.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\hat{Y}_i - Y_i)^2}{n}}$$

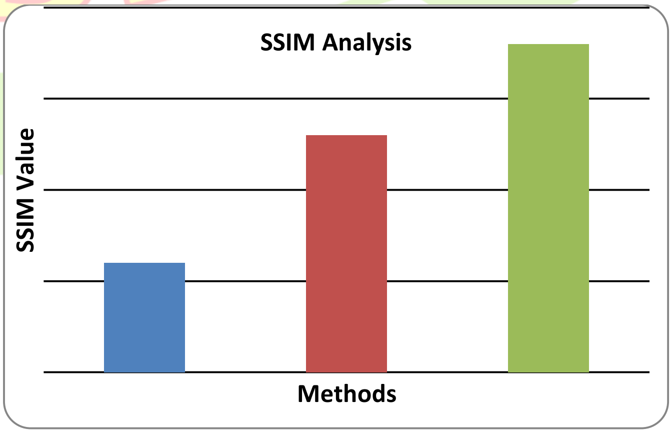
where $\hat{Y}$ is 3D image and Y is original 2D image.

To analysis the performance of the three methods by using the performance metrics which are mentioned above. This is shown in the below tables and graphs

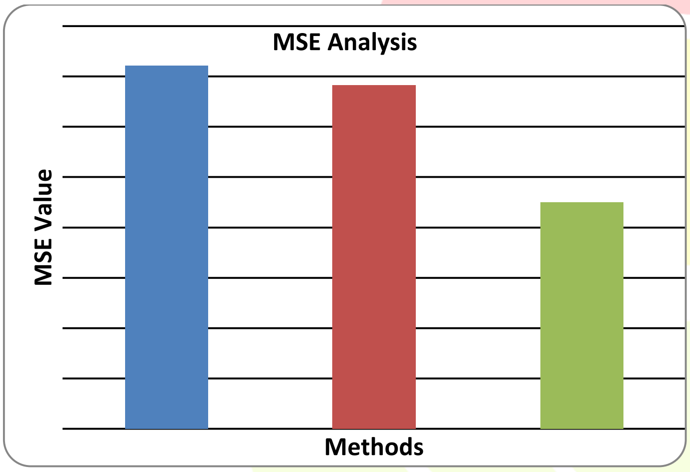
Methods	PSNR
Depth with RGB	17.5
Depth with Fusion	20.1
Proposed Method	25.2



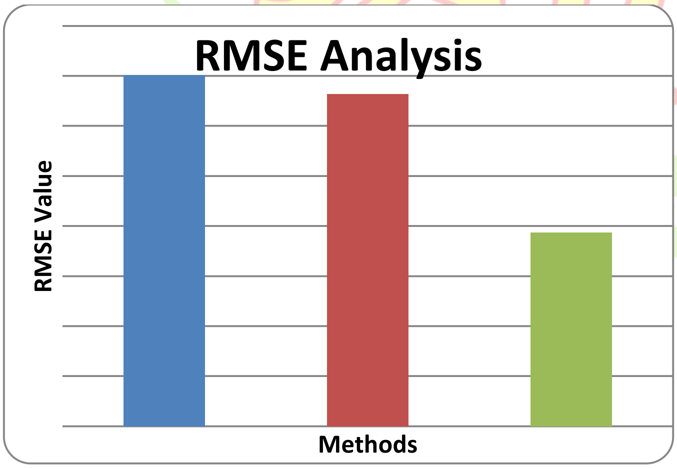
Methods	SSIM
Depth with RGB	0.61
Depth with Fusion	0.68
Proposed Method	0.73



Methods	MSE
Depth with RGB	7.21
Depth with Fusion	6.82
Proposed Method	4.50



Methods	RMSE
Depth with RGB	3.50651252
Depth with Fusion	3.32036033
Proposed Method	1.9362638



Conclusion:

This paper developed an efficient 2D-to-3D conversion method based on the use of edge information. The edge of an image has a high probability as it can be the edge of the depth map. Once the pixels are grouped together, a relative depth value can be assigned to each region. Initially, the input RGB image is converted into HSV color space. And then it is converted into gray image. And then the block-based image is considered to segment it into multiple groups. To segment the image this paper uses the Fuzzy C-Means Segmentation Approach. Then the depth of each segment is assigned with the help of an initial depth hypothesis. Next, the blocky artifacts have to be removed using cross bilateral filtering. Finally, multi-view images are obtained by the method of DIBR. As a result, the input 2D image is converted into visually comfortable 3D image without the presence of artifacts enhancing the quality of the image in the display. To analyse the performance of these method several performance metrics are used. This paper uses PSNR,SSIM, MSE and RMSE to analyses the performance. From the experimental results it is shown that the proposed method performs better than the other two methods..

References

[1] Wikipedia the Free Encyclopedia: Stereoscopy [Online]. Available: http://en.wikipedia.org/wiki/Stereoscopy
[2] Getting Started in Stereo Photography. [Online]. Available: http://home.comcast.net/~dssweb/intro_to_stereo.htm
[3] Wikipedia the Free Encyclopedia: 3D film. [Online]. Available: http://en.wikipedia.org/wiki/3D_film
[4] J. Caviedes and J. Villegas, "Real time 2D to 3D conversion: Technical and visual quality requirements," *IEEE International Conference on Consumer Electronics*, pp. 897-898, 2011.
[5] T. N. Lu and S. Z. Tan, "An algorithm based on bilinear interpolation of converting 2D video to 3D video," *IEEE*, pp. 4813-4815, 2011.
[6] X. Y. Mao and L. K. lbsiyasu, "Hierarchical representations of 2D/3D Gray-Scale Images and their 2D/3D two way conversion," *IEEE*, pp. 37-44, 1987.
[7] T. L. Chin, C. L. Chin, K. W. Fan, and C.Y. Lin, "A novel architecture for converting single 2D image into 3D effect image," *IEEE*, pp. 52-55.
[8] H. Murata, X Mori, S. Yamashita, A. Maenaka, S. Okada, K. Oyamada, and S. Kishimoto, "A real-time 2-D to 3-D image conversion technique using computed

ISSN (ONLINE): 2395-695X
ISSN (PRINT): 2395-695X
International Journal of Advanced Research in Biology, Engineering, Science and Technology (IJARBEST)
Vol. 2, Issue 5, May 2016

image depth," *SID Symposium Digest of Technical Papers*, vol. 29, no. 1, pp. 919-923, 1998.

[9] W. J. Tam and L. Zhang, "3D-TV content generation: 2D-TO-3D Conversion," *ICME*, pp.1869-1872, 2006.

[10] C. C. Cheng, C. T. Li, and L. G. Chen, "A 2D-to-3D conversion system using edge information," in *Proc. Digest of Technical Papers International Conference on Consumer Electronics*, 2010, pp. 377-378.

[11] Z. B. Zhang, Y. Z. Wang, T. T. Jiang, and G. Wen, "Visual pertinent 2D-TO-3D video conversion by multi-cue fusion," in *Proc. 18th IEEE International Conference on Image Processing*, 2011, pp. 909-912.

[12] C. L. Su, K. N. Pang, T. M. Chen, G. S. Wu, *et al.*, "A real-time Full-HD 2D-to-3D conversion system using multicore technology," in *Proc. fifth FTRA International Conference on Multimedia and Ubiquitous Engineering, IEEE*, 2011, pp. 273-276.

[13] Y. K. Lai, Y. F. Lai, and Y. C. Chen, "An effective hybrid depth-generation algorithm for 2D-to-3D conversion in 3D displays," *Journal of Display Technology*, vol. 9 no. 3, pp. 154-161, March 2013.

[14] Jean Maria Dominic," Recent Trends in 2D to 3D Conversion: A Survey", MG University, IJRASET Vol.2 IV, April 2014.

[15] GUAN, S., KLETTE, R. Belief Propagation on edge image for stereo analysis of image sequences. In Proceedings Robot Vision. LNCS 4931, 2006,

[16] BENCO, M., HUDEC, R. The advances image segmentation techniques for broadly useful retrieval in large image database. In NSSS IX. TatranskeZruby (Slovakia), 2006,

[177] HE, R., ZHU, Y. A hybrid image segmentation approach based on Mean Shift and fuzzy C – Means. In Asia – Pacific Conference on Information Processing., 2009.

[18] WANG, G., JU, H. A disparity map extraction algorithm for lunar rover BH2. In IEEE International Conference on Intelligent Computing and Intelligent Systems ICIS 2009. Shanghai, 2009, vol. 4.

[19] McKINNON, B., BALTES, J. Practical region – based matching for stereo vision. In IWCIA. 2005, vol. 3322 of Lecture Notes in Computer Science, Springer,

[20] KUHL, A. Comparison of stereo matching algorithms for mobile robots. Centre for Intelligent Information Processing System. 2005, University of Western Australia,