2D to 3D Conversion Using Depth Estimation

Hemali Dholariya
Student of
Integrated M.Sc.
(IT)
At UTU University
Bardoli, Gujarat.

Jayshree Borad
Student of
Integrated M.Sc.
(IT)
At UTU University
Bardoli, Gujarat.

Pooja Shah
Student of
Integrated M.Sc.
(IT)
At UTU University
Bardoli, Gujarat.

Archana Khakhariya
Student of
Integrated M.Sc.
(IT)
At UTU University
Bardoli, Gujarat.

Juhi Patel
Teaching Assistant of
Department of Computer Science and
Technology
At UTU University
Bardoli, Gujarat.

Abstract - **"Image" is used to indicate the image data that is sampled, Quantized, and readily available in a form suitable for further processing by digital Computers in image processing. For high quality stereoscopic images, the conversion of 2D images to 3D achieves the growing need.Robotics branch is the main area of depth-map application. In this review paper, we relate how 2D to 3D conversion using Depth Estimation works and where it is convenient in actual world. We compare different algorithms likeMarkov Random Field(MRF), Modulation Transfer Function(MTF), Image fusion, Local Depth Hypothesis,Predicted Semantic Labels,3DTV Using Depth Map Generation and Virtual View Synthesis.We find out some issues of 2D to 3D conversion.**

Keywords: MRF, MTF, Image Fusion, Squeeze function,SVM

1. INTRODUCTION

User-defined strokes correlate to a rough depth estimate values in the scene are explained for the image of interest is said to be 2D to 3D conversion.

Figure 1: Difference between 2D and 3D image

3D Reconstructions are carry out by two ways such as,

1) Single image 3D Reconstruction
2) Multiple image 3D Reconstruction

Single image 3D Reconstructionis the design of a set of images.

Multiple image 3D Reconstruction is the design of three-dimensional models from a set of images. It is the alter procedure of acquiring 2D images from 3D scenes.

it is not possible to find which point on this line correlate with the image point from a single image. Then the position of a 3D point can be found as the interchange of the two prediction spark. This procedure is called as triangulation.

In Computer vision and Computer graphics, 3D reconstruction is the procedure of capturing the shape and appearance of real objects.

The stereoscopic images give information on details of every object in the picture in three proportions and help to notice the image in a best way. Stereoscopic images are also referred as 3D image.

Now a days, iphone and htc smart phone provide the built in facilities for the generate 3D image and its resolution. That's why reduction in the size of the 3D scanner and the high resolution camera which is generate the 2D to 3D image.

Paper is organized as below. In section 2 we have done literature of different papersdescribing working of Depth estimation and algorithms used in it. Section 3 shows the methods andalgorithm for the conversion of 2D to 3D image and finally in section 4 we give conclusion of the Literature review.

Disadvantage of 2D:

A single two dimensional (2D) image does not contain depth information. An infiniteNumber of points in the three dimensional (3D) spaces are projected to the same pointin the image plane. But a single 2D image has some monocular depth cues, by which wecan make a hypothesis of depth variation in the image to generate a depth map.

Application of 3D conversion

3D models and 3D viewing is catching great pace in the field of computer vision due to its applicability in diverse fields of heath, aerospace, textile etc.

3D modelers having been used in a wide variety of industries. The medical industry operates them to build detailed models.

The movie industry operates them to manipulate characters and objects for animation and real-life moving pictures.

The video game production uses them to make resources for video games.

The science section manages them to build highly detailed models of chemical fusion. The architecture area uses them for making models of advanced buildings and scenery.

The engineering community make use of them to design new tools, motor vehicles and constructions.

2. RELATED WORK
2.1 Literature Review

Two Eyes = Three Dimensions (3D)!

For processing, both eyescapture its own vision and the two different images are forward to the brain. When two images reach at the same time in the back of the brain, they are unified into one picture. By coordinating with the similarities and joining the small differences the mind combines the two images. The small differences between the two pictures combine to a big change in the absolute picture. The final picture is more than the sum of its piece. It is now called three-dimensional *stereo* image.
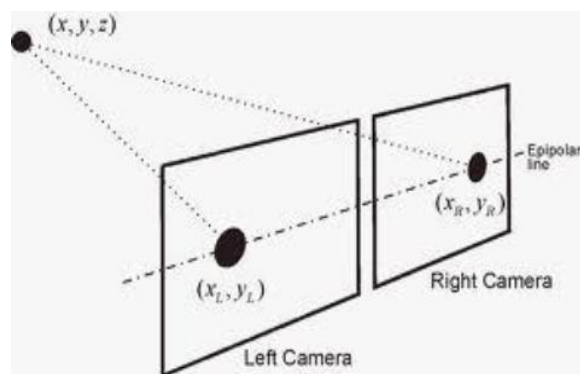
The word "stereo" means firm or solid. In stereo vision you can see an object as well assemble in three spatial dimensions width, height and depth. It is the added realization of the depth dimension that build stereo vision so elegant and notable.

In Novel algorithm For transforming 2D image to stereoscopic image with Depth control using Image Fusion method a faster 2D to 3D conversion algorithm is evolved. This algorithm defines two images one as left eye vision and second one as right eye vision with source to the user defined depth. These left view image and right view image are combined and fixed at mean value. The combined image and left view image are reserved in the 3D image format MPO and can be viewed in 3D capable device.

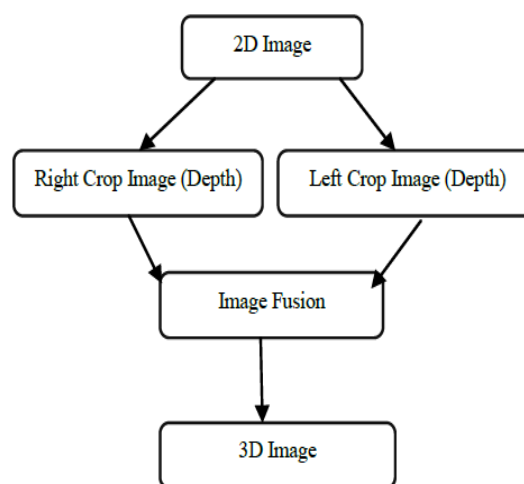Steps used for 2D to 3D conversion are as shown in below figure:

Figure 2:Proposed 2D to 3D conversion algorithm

As input 2D image and depth are taken from the user. The 2D sample image taken for the experiment is shown in the below figure:
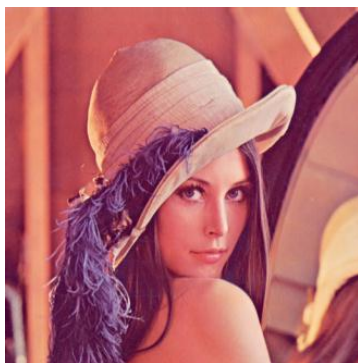
Figure 3:Input image sample 2D image

Figure 4:Left crop image

Figure 5:Right crop image

The 2D picture is taken using single view lenses. The human vision system is a natural made perfect system of 3D with two eyes is apart in a fixed distance. The 3D pictures are taken using two lenses kept apart at a fixed distance. The distance between the lenses are calculated using,

Stereo=1/30 x distance of object

In 3-D Depth Reconstruction from a Single Image,

From a single still image we consider the job of 3-d depth estimation. We begin by collecting a training set of monocular images (of not structured indoor and outdoor locations that include forests, sidewalks, trees, buildings, etc.) and their corresponding floor-truth depth maps.Our model make use ofordered, Multiscale Markov Random Field (MRF) that integrate multiscale local- and global-

image features, and models the depths and the relation dissimilar depths at differentpoints in the figure.

We can see that, even on not structured scenes, our conclusion is frequently able to retrieve fairly truthfuldepth maps. We further propose a representation that incorporates both stereo cues and monocular cues, to obtain significantly more precise depth estimates than is probable using either monocularor stereo cues by yourself.
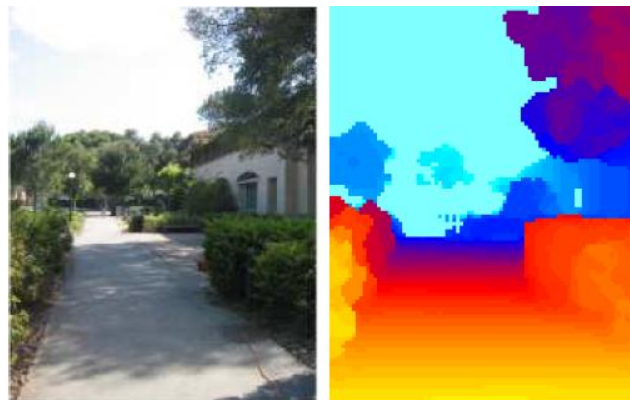

Figure 6:(a) A single still image, and (b) the corresponding (ground-truth) depthmap. Colors in the depthmap
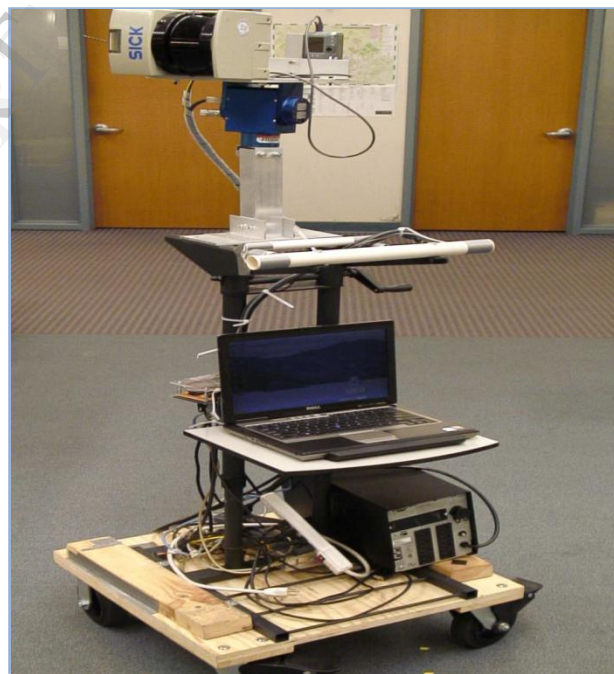

Figure 7: The 3D-Scanner used for collecting images and the corresponding image map

Conversion 2D Image to 3D based on Squeeze function and Gradient Map,

The three-dimensional (3D) displays required the depth information which is unavailable in the conventional 2D content. This work describes a novel method that automatically converts 2D images into 3D ones. The proposed algorithm is composed of the estimation of depth levels by using modulation transfer function (MTF) squeeze model and determination of gradient map related to each depth level. The grouping is based on the pixels

having similar colors and spatial locality. Based on a depth gradient map, a depth level is assigned. Next, the depth map is assigned by cooperating with a cross bilateral filter to diminish the blocky artifacts.

The conversion process of existing 2D images to 3D is commercially viable and is fulfilling the growth of high quality stereoscopic images. The dominant technique for such content conversion is to develop a depth map for each frame of 2D material. When observing the world, the human brain usually integrates the heuristic depth cues for the generation of the depth perception. The major depth perceptions to be noted are binocular depth cues from both eyes and monocular depth cues from a single eye.

In order to overcome these two challenges, this paper presents an algorithm that uses a simple depth theory to allocate the depth of everyset instead of retrieving the depth value directly from the depth cue based on area of interest. MTF Squeeze model is used in order to model the property of sampling artifacts on goalgratitude and identification presentation. Secondly, identifying the region of interest in an image is necessary for performing useful post-processing on the image for research and treatment.

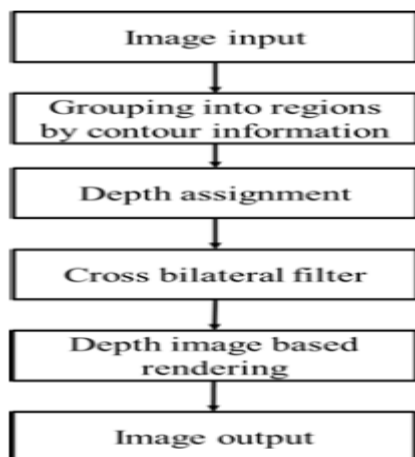The appropriate layers are defined as that depth will be displayed on the both sides with six parallelograms.

Figure 8:2D to 3D conversion process

In 2D to 3D Conversion in 3DTV using depth map generation,the perceptual depth information from monocular images was estimated by the best use of comparative-height. Depth assignment operation was followed to generate the initial depth map.The advantage of relative-height was that it could be utilized in the majority of the scene and did not need large computation.

They first separated the image into foreground objects and the background, and then refined the foreground object by using gradient vector flows. Depth values were assigned to the fore-ground objects according to the motion analysis. In the background, depth values were assigned by using the linear perspective. Relative-height was one type of well-known depth recovery cues, especially in landscape scene.
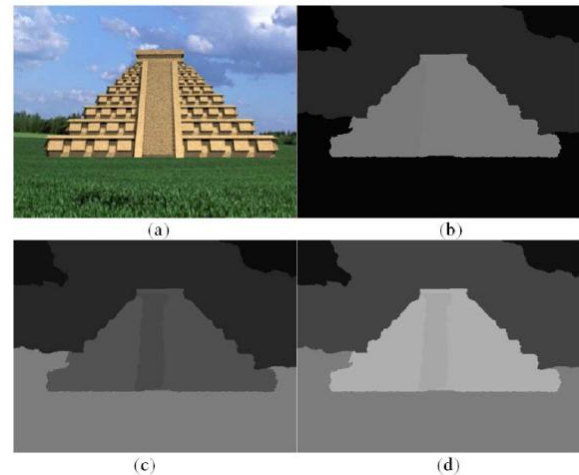

Figure 9:Depth map generation results in the 'Temple' sequence. (a) Original image. (b) Depth map from motion parallax. (c) Depth map from relative-height. (d) Final fusion depth map.

In Single Image Depth evaluationon Predicted Semantic tickets,We have described that the semantic and geometric context on the images which is defines that the different-different classes that based on that the colour used to that images.
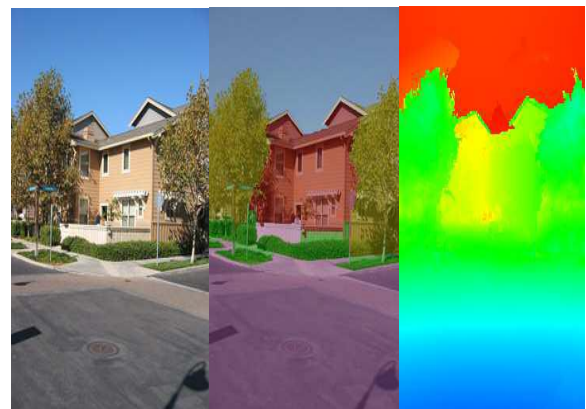
Above figure defines that the some qualitative Depth reconstruction from out left to right the image, semantic overlay, ground truth.

In Depth Map Generation from a Single Image Using Local Depth Hypothesis,

In depth map, depth is expressed in greyscale.The vanishing point represents the farthest point. Scene grouping use a graph-based segmentation algorithm in grouping similar regions in order to improve salient segmentation and assign the same depth value.
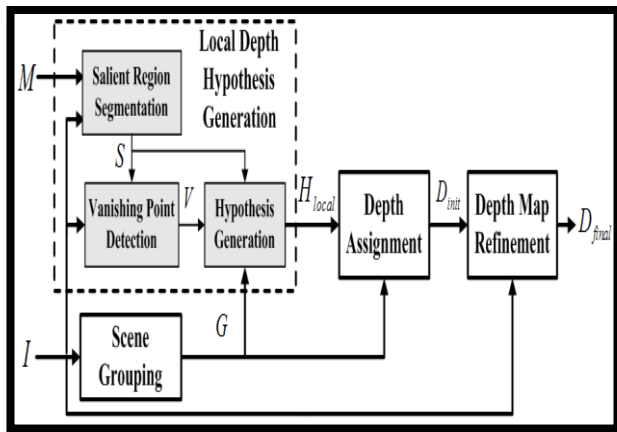
Figure 10:Depth generation from single image using local depth hypothesis

All this paper has some methods for converting 2D image to 3D imagevia some methods, that all methods are described in the next session.

3. Methods and Algorithms

3.1 Methods

During Our literature review we study the basic method for 2D to 3D conversion using Depth Estimation; all that methods are describe bellow:

3.1.1 Local Depth Hypothesis

This proposed method groups an input image into similar regions toreserve details and segments the image into salient regions with user intersection. The method describing the below steps:

1. Scene grouping
2. Depth hypothesis generation
3. Depth assignment and refinement

3.2 Algorithms

During Our literature review we study the basic algorithm for 2D to 3D conversion using Depth Estimation;those entire algorithms aredescribing bellow:

3.2.1 Image fusion

This Image fusion algorithm uses that the left view camera and right view camera. There are many steps to implement this algorithm:

1. Capturing 2D image
2. Capture 2D image through left view camera
3. Capture 3D image through right view camera
4. Image fusion
5. Generate the 3D image

3.2.2 Support Vector Machine (SVM)

The SVM algorithm to identify that the specific sections that belong in the region of interest. It's defines many steps to performing the find depth:

1. To cluster the pixels by intensity, assign them a new value based on the cluster they belong to, and store the resulting image.
2. Sectioning the image involves finding edges.
3. Region identification.
   a. Region attributes.
   b. Region labeling.
   c. Training and Classification.
   d. Replacing a ROI with a machine-learning identified ROI is to turn the region of connected pixels back into an ROI boundary created as a Bezier spline curve.

3.2.3 MTF Squeeze Model

This model based on that the pixels having similar colors and spatial locality. This model describes the 2 steps:

1. MTF squeeze model is used when we want to model the effects of sampling artifacts on goal recognitionpresentation.
2. Identifying the region of interest in an image is necessary for performing useful post-processing on the image for research and treatment.

The depth generation algorithms are roughly classified into three categories which utilize different kinds of depth cues: The binocular, monocular, and pictorial depth cues.

4.CONCLUSION

From all this literature review we have get some problem like Depth estimation and stereo vision.So, we conclude that we want to solve the find accuracy depth of image for converting 2D to 3D image.

4. REFERENCES

1. http://en.wikipedia.org/wiki/2D_to_3D_conversion. Date : 18/12/2014
2. Na-Eun Yang, Ji Won Lee, Rae-Hong Park. " Depth Map Generation from a Single Image Using Local Depth Hypothesis.", 2012 IEEE.
3. Ashutosh Saxena, Sung H. Chung, and Andrew Y. Ng. "Learning Depth from Single Monocular Images. ", Computer Science Department Stanford University Stanford,CA94305,asaxena@stanford.edu, fcodedeft,angg@cs.stanford.edu.
4. Beyang Liu, Stephen Gould, Daphne Koller ."Single Image Depth Estimation From Predicted Semantic Labels.", in Proc. IEEE Int.
5. Cheolkon Jung1, Xiaohua Zhu1, Lei Wang1, Tian Sun1. "2D to 3D Conversion in 3DTV Using Depth Map Generation and Virtual View Synthesis",rd International Conference on Multimedia Technology(ICMT 2013).
6. Hong YR, Tseng YC, Chang TS (2010) Stereoscopic images generation with directional Gaussian filter. In: Proc. IEEE ISCAS, pp.2650-2653.
7. Zhang L, Vazquez C, Knorr S (2011) 3D-TV content creation: automatic 2D-to-3D video conversion. IEEE Transactions on Broadcasting 99:1-12.
8. Park YK, Jung K, Oh Y, Lee S, Kim JK, Lee G, Lee H, Yun K, Hur N, Kim J, (2009) Depth-image-based rendering for 3DTV service

over T-DMB. Signal processing: Image communi-cation 24:122-136.

9. Fehn C (2004) Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV. In: Proc. SPIE.

10. Jung YJ, Baik A, Kim J, Park D (2009) A novel 2D-to-3D conversion technique based on relative height depth cue. In: Proc. SPIE Electronics Imaging, Stereoscopic Displays and Applications.

11. Yu F, Liu J, Ren Y, Sun J, Gao Y, Liu W (2011) Depth generation method for 2D to 3D con-version. In: Proc. 3DTV-Con.

12. Lai YK, Lai YF, Chen YC (2012) An effective hybrid depth-perception algorithm for 2D-to-3D conversion in 3D display systems. In: Proc. IEEE ICCE, pp.612-613.

13. C.-C. Cheng, C.-T. Li, and L.-G. Chen, "A novel 2D-to-3D conversionsystem using edge information," *IEEE Trans. Consumer Electronics*, vol 56, no. 3, pp. 1739–1745, Aug. 2010.

14. K. Han and K. Hong, "Geometric and texture cue based depth-mapestimation for 2D to 3D image conversion," in *Proc. 2011 IEEE Int.Conf. Consumer Electronics*, pp. 651–652, Las Vegas, NV, Jan. 2011.

15. V. Cantoni, L. Lombardi, M. Porta, and N. Sicard, "Vanishing pointdetection: Representation analysis and new approaches," in *Proc. 11ᵗʰInt. Conf. Image Anal. and Process.*, pp. 90–94, Palermo, Italy, Sept.2001.

16. Bajcsy, R.K.; Lieberman, L.I. (1976) "Texture Gradient as a Depth Cue", CGIP,5(1): 52-67.

17. Maar, D.; Poggio, T. (1976) "Cooperative Computation of Stereo Disparity",Science, Volume 194, Pages: 282-287.

18. Cutting, J.; Vishton, P. (1995) "Perceiving Layout and Knowing Distances: theIntegration, Relative Potency, and Contextual Use of Different Information about Depth",

19. In Epstein, W; Rogers, S (editors.) Perception of Space and Motion, Pages 69-117.Academic Press, San Diego.URL: http://pmvish.people.wm.edu/cutting&vishton1995.pdf

20. Saxena, A; Chung, S.H.; Ng, A. Y. (2005) "Learning Depth from Single MonocularImages", Proceedings, 19th Annual Conference on Neural Information ProcessingSystems (NIPS 2005).

21. Torralba, A.; Oliva, A. (2002), "Depth Estimation from Image Structure", IEEETransactions on Pattern Analysis and Machine Intelligence, Volume 24 , Issue 9, Pages:1226 – 1238.

22. S. H. Lee, D. W. Park, J. P. Jeong and K. I. Moon, "Conversion 2D Image to 3D Based on Squeeze function and Gradient Map", On Application to Partial Differential Equations of Warranty Reclaims, Proceedings of the 1th International Symposium, ISAAC 2013 in conjunction with ICACT 2013 Seoul, Korea, Revised Selected Papers, **(2013)** November.

23. G. D. Ronald, V. Richard and O. Barbara, "Sampled imaging sensor design using the MTF squeeze model to characterize spurious response", Part of the SPIE Conference on Infrared lmaing Systems: Design, Analysis, Modeling, and Testing X. Orlando, Florida, SPIE, vol. 3701, (1999), pp. 61-73.