The International Journal of Engineering and Science (IJES)
|| Volume || 9 || Issue || 01 || Series II || Pages || PP 15-20| 2020 ||
ISSN (e): 2319-1813 ISSN (p): 23-19-1805

Converting 2D Pictures In Jpeg Into 3D Objects In Space

Baki Koyuncu[1]Rafal Hameed [2]
Electrical & Electronic Engineering Dept.
Istanbul Gelisim university, Istanbul, Turkey

--***ABSTRACT***--
This work proposes a method to reconstruct the 3D object model from 2D images. First, A depth map is calculated which is an image that contains information about the distance between the surface of the objects from a given viewpoint and then the imageisreconstructedin 3D vision . The 3D reconstruction can be deployed that must be a part of numerous highlights, for example, surface, edge, shading, and size of the item, and so forth. These elements rely upon the scientists choosing to extract the significant highlights of the item in order to make the 3D object reconstruction

Keywords-*3D Reconstruction; 2D image; depth calculation; stereoscope; Parallax Computation.*

Date of Submission: 15-01-2020 Date of Acceptance: 31-01-2020

I. INTRODUCTION

With the development of innovative techniques, image processing has been stretched out from 2D to 3D measurements. The applicationsof the 2D picturesare utilized generally in our life such as inobjectidentificationtodistinguish objects, individual face identification, vehicle permit ID and item examination in different frameworks.

Because of the quick advances in PC development and innovative image processing techniques, we can reconstruct the pictures from 2D to 3D. The contrast between the 2D and 3D pictures is the depth information.Hencethe separation of the depth information is the issue of the 3D reconstruction technology. The 3D reconstruction can be divided into contact and non-contact types. The contact typeshould contact the object surface by a scriber to extract shape data from the object surface.

The CMM (Coordinate Measuring Machine) is one of the contact type 3D scanners. Although the 3D scanner needs to contact the object surface, it will sometimes cause damages on the surface. Another important issue is the long-time taken by the 3D scanner to travel on the objectsurface to generate the 3D model.

The non-contact type 3D scannercan be separated into an active and passive constructive metsys. The active one identified as the dynamic kind uses a laser to capture the depth data [1]. The passive kind uses a camera to catch the 2D pictures. It usesthese 2D pictures to ascertain the depth and surface contour data tocreate the 3D model of the objects.

Contemporary 3D imaging strategies give significant benefits over customary 2D imaging techniques. Furthermore, most 3D innovations are competing to offer another type of 3D facility that has regular shading and full parallax without the need to wear particular glasses [2].

3D vision is getting progressively significant, in major films just as in-home entertainment. Depth map-generating plays a significant role in the 2D-to-3D transformation and becomes a precursor stage to Depth Image Based Rendering (DIBR).

A depth map is a picture that contains data about the distance between the outside of the objectdna a given perspective. A depth map is made from a source picture and is normally in a grayscaleformat.

II. THE PROPOSED METHOD

i. The depth calculations

In a characteristic picture, objects in the center plane are sharper than those out of center because of the depth of the objects. Those calculations that show the sharpness or obscurity of object points as their htped information is classified as “depth From Focusing” or “depth From Defocusing” algorithms. Either “depth From Focusing” or "depth From Defocusing” calculation needs center measure which discovers the sharpness of an object pointto decide its depth. There are a lot of approaches to measure the sharpness, either in spatial space or frequency domain. Every one of them is simply extraordinary aspects of some kind of high pass channels from various points of view. Many methods are proposed to calculate the depth map depending on focusing.

1. Xiao-Ling Deng's work [3]introduceda new methodology to measure depth map constructing from monocular indoor environments. Here, the idea of the planar model of 2D-to-3D transformation strategy is acquainted and the approach of removingplanes and progressively measure 2D picture depth is proposed. An indoor picture is,first of all, partitioned into a closer view (foreground) and foundation(background). The foundation is then translated into one or two horizontal and a few vertical planes. Finally, the depth guide of the picture is obtainedaccording tothe plane grouping machine and geometric highlights. Every pixel is appointed relative depth esteem as per the situation in it's credited plane after the plane grouping machine [3]. The plane extraction algorithm isn't still, advanced and there is a limitation to manage broad pictures and sequencesprogressively. See Figure 1.

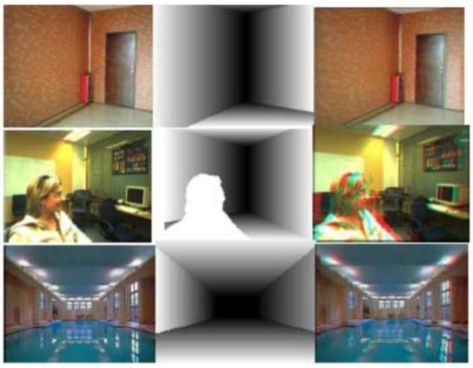

Figure 1. Examples of depth map generation proposed by [3].

2.In S. Bharathi's work [4], the Hypothetical depthgradient model for depth map estimation is proposed.Division of picture into object groups of dependent edge data is deployed with the help of edge data. Pixels are gathered based on comparable shading orcolors or exceptional areas [4]. The same depthidentity is denoted to amass pixels and it is filled up by depth gradient. Blockedold pixels in the subsequent depth map are reduced by applying the Cross Bilateral filter (channel). The last Filtered picture is prepared by DIBR (Depth Image Based Rendering) model.

3.Yea-Shuan Huang, Fang-Hsuan Cheng [5] proposed a technique that generates the depth map of an image, supported by scene classification. Color elements are separated in terms of chromatic and achromatic data. Depth is measured by sharpness, contrast, chrominance,vand estimation is finished in line with vanishinglines.

RGB to HIS (Hue, saturation, intensity) color house conversion is deployed because it is more appropriate for the image process. The scene is classified into a personal image, out of doors image and close-up image. when generating the depth map of the scene, the stereoscopic pictures can be generated by the left and right shifting of the original image [5]. The projected technique [5] still has some constraints. For example, during the taking of a private image, the garmentsmust be pure and the clothes also must have less texture. ecneh, the amount of object points within the scene is restricted to at least one within the projected technique.

4. Yi-Che Chen [6] proposed a pragmaticmethodology to produce a depth map from a progression of pictures taken at various focal points of a similar scene. This method requires Gaussian channels with some expansion/subtraction utilitarian blocks. A progression of pictures of various centers captured from a similar scene is handled by focal measurements. Afterward, each prepared information is compared based on the supposition that pixels closer to the center picture ought to have greater center qualities (keener) in the frontal area than those out of sight [6]. Along these lines, the last depth map produced is a sort of a record map that shows which center plane the corresponding pixel lies on.

ii. Depth Map-based Rendering

1. Pseudo Depth-map:

The atmospheric light must be estimated asA for the imageI_{haze} . A can be estimated withmost of the algorithms from the highest intensities pixels quickly but not very accurately.

The depth map can be calculatedbased on the degradation model of the image [11] and the proposed dark channel prior [12, 13].

Initial depth map,$\widehat{m}(x,y)$,is estimated.This process can be written as:

$$\widehat{m}(x,y) = 1 - \omega_2 \, min_{c\in\{R,G,B\}} \left(min_{(X',Y')\in\Omega(X,Y)} \left(\frac{I^c_{haze\,(X',Y')}}{A^c} \right) \right)$$

(1)

Where I^c_{haze} is a color channel of I_{haze}, $\Omega(x,y)$ local patch centered at (x,y), and $(X',Y')$is the pixel location that belongs to $\Omega(X,Y)$.ω_2 is a constant parameter for adjusting the amount of haze for distant objects.
In order to handle these deficiencies of the redundant details and the obvious block effects in the initial depth map, The bilateral filter and guided filter [12] are used to refine the initial depth map.
The following steps describe the estimation process of the final depth-map.

Step 1. For the initial depth map, the linear coefficients a_k and $b_k$are computed first for the guided filter:

$$a_k = \frac{\frac{I}{|\omega|}\sum_{(x,y)\in\omega_k} I_{haze}(x,y)\widetilde{m}(x,y) - u_k\overline{m}_k}{\sigma_k^2+\varepsilon}, \qquad (2a)$$

$$b_k = \overline{m}_k - a_k u_k \quad (2b)$$

where I_{haze} is the guidance image and $\widehat{m}$ is the input image of the guided filter.The filter is a general linear translation-variant filtering process, which involves a guidance image and an input image [14].
In Eq. (2), $\varepsilon$is a regularization parameter preserving a_k from being too large. u_k and σ_k^2 are the mean and variance of I_{haze} in a window ω_k that centered at the pixel k. $|\omega|$ is the number of pixels in ω_k, and $\overline{m}_k = (I/|\omega|)\sum_{i\in\omega_k}\widetilde{m}_i$ is the mean of$\widetilde{m}$ in ω_k.

Step 2. Once the linear coefficients (a_k, b_k) are obtained, we can compute the filter output by:

$$m'(x,y) = \overline{a_k}\widetilde{m}(x,y) + \overline{b_k}, \quad (3)$$

Where $\bar{a}_k = (I/|\omega|)\sum_{i\in\omega_k} a_i$and
$\bar{b}_k = (I/|\omega|)\sum_{i\in\omega_k} b_i$.$\widetilde{m}$ is the initial depth map, and the filter output m' is the refined depth map.

Step 3. A bilateral filter is used here to remove the redundant details for the refined depth map m' since the bilateral filter can smooth images whilepreserving edges. Thus, the redundant details of the refined depth map m' estimated by the algorithm, presented above, can be effectively removed. This process can be written as:

$$\widehat{m}(u) = \frac{\sum_{p\in N(u)} W_c(\|p-u\|)W_s(|m'(u)-m'(p)|)m'(p)}{\sum_{p\in N(u)} W_c(\|p-u\|)W_s(|m'(u)-m'(p)|)}, \qquad (4)$$

Where$m'(u)$ is the refined depth map corresponding to the pixel $u = (x,y)$, $N(u)$ is the neighbors of u. The spatial domain similarity function Wc(x) is a Gaussian filter with the standard deviation ofσ_c . It can be defined as $W_c(x) = e^{-x^2/2\sigma_c^2}$.The intensity similarity function Ws(x) is a Gaussian filter with the standard deviation ofσ_s.It can be defined as$W_s(x) = e^{-x^2/2\sigma_s^2}$. Thus, the final depth map $\widehat{m}(x,y)$ can be obtained. Figure 3(c) shows the final depth map for the original image.

2. Left and right view:

Once the depth map is obtained, the left-view and the right-view images can be synthesized by the following steps. Firstly,the parallax value,Parallax(x, y), is computed from each pixel (x, y) in the estimated depth map. The computation of the parallax value can be written as

$$parallax(x,y) = \omega_3 \times \left(1 - \frac{\widehat{m}(x,y)}{zpp}\right), \qquad (5)$$

where $\widehat{m}(x,y)$ is the final depth map for the single image, ω_3 is the parallaxmaximumvalue. As can be seen in Figure 2(a), one can get the value of ω_3 by a similar triangle.
The zero parallax plane (ZPP) is set as the region with the depth value of Th, which is computed by
Th=max(m'(x,y))-10 to prevent separation and loss of artifacts[7].

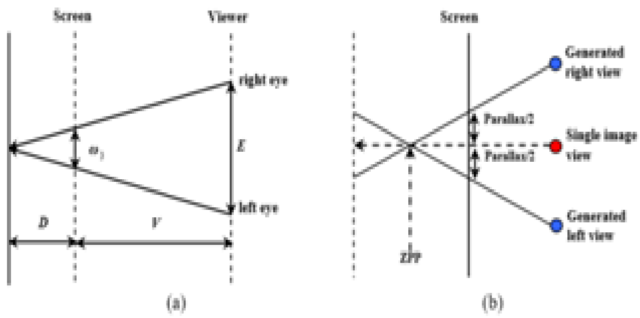

Figure 2. Generating Stereoscopic (a) Parallax Computation (b) Generating the Right Left Views

At that point,the information picture is considered as the middle perspective on the stereoscopic pair, as appeared in Figure 2(b). In order to create the left or right-see picture, every pixel of the information picture is moved by the measure$parallax(x,y)\ /\ 2$ to the left or right bearing.

Rendering is carried out on the estimated depth map in Figure 3(c) which is obtained from the proposed system to assess ecnamrofrep eht. In most cases, the pseudo depth map is produced dependent on the visual consideration. However, it ought to be seen that the assessed depth territory corresponding to certain individuals in Figure 3(c) appears to be erroneous by saliency identification.Theapproach may give a 3D observation, as shown in Figure 3(f), and the overall procedure of converting 2D image into a stereoscopic image can be seen in Figure 4.

Figure 3. (a) Input Image (b) Simulated Haze Image (c) Estimated Depth Map, (d) and (e) Left-view and Right-view Images, Respectively (f) Stereoscopic Conversion Result

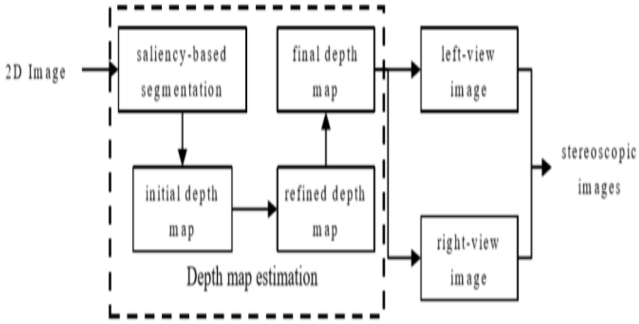

Figure 4. The Overall Procedure for 2D-To-stereoscopic images.

iii. 3D reconstruction

The 3D recreation calculation in a stereo picture pair for acknowledging common impediments and collaborations between the genuine and virtual world in a picture synthesis is proposed. A two-stage calculation, consisting of uniqueness estimation and regularization is used to find a smooth and exact divergence vector.The hierarchal inequality estimation technique expands the productivity and unwavering quality of the estimation procedure.Edge-preservingdisparityfieldregularization produces disparity fields while safeguarding discontinuities that result from object boundaries. The synthesis system is shown in Figure 5.

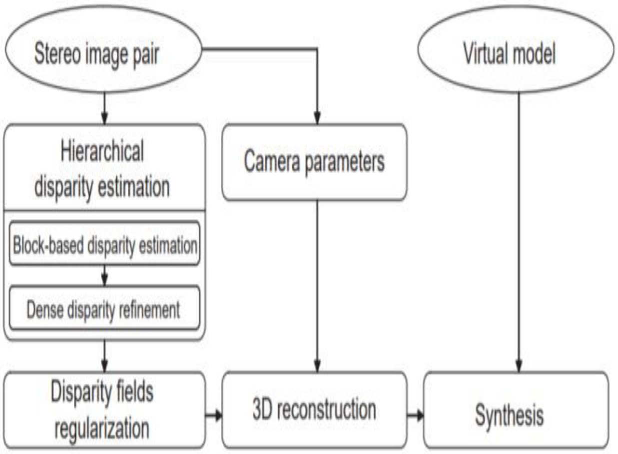

Figure 5. Block diagram of the synthesis system.

Atwo-stage algorithm consisting of hierarchical dense disparity estimation and vector field regularization is proposed to evaluate the 3D reconstruction.

1. EstimatingHierarchical dense disparity:

Dense disparity vectors are at first hierarchicallyestimated utilizing a district isolating procedure [8]. The region-isolating method depends on the requesting constraint [9]. The procedure performs point matching in the request for the plausibility of right matching and partitions the area into two sub-regions at the genuine coordinating point.

2. Regularization ofEdge-preserving disparity:

The uniqueness vectors assessed by the [10] strategy give commonly dependable information. Be that as it may, spatial connection of the estimated vector fields isn't considered, and, as a result, bogus vectors can be acquired by a wrong estimation or an engendering error.

***3. 3D reconstruction and synthesis*:**

A 3D model of test picture setsare reproducedfrom the remade 3D directions and unique texture data with 3D studioMAX. Figure 6 shows the locations of the remade models from several perspectives. Stereo image sets don't provide complete surface data for the model. Hence the surface of the frontal(foreground) areais diffusedto the background so as to increase the 3D impact. As a result, 3D scenes of the test sets are normally produced. In the event that this system is applied to multiview pictures, it is conceivable to remake a nearly complete 3D model by making up for missing structure and surface information. To confirm the exactness of the reproduced depth data from the assessed difference fields, calculations are carried out for the scene with clearly known depth data in a virtual environment.

Figure 6. 3D reconstruction results (a)lamp and head(b) saw tooth, (c) man .

III. CONCLUSIONS

In this paper,a method for 2D image conversion to 3D objects has been overviewed. There are a lot of procedures.Every procedure has its own points of interest and inconveniences. These procedures evaluate the 3D object in specific methodology and there is a lot of different methodologies proposed in different places depending on different outcomes.

The generation of the depth map shows the separation between the object surfaces from a given perspective. A depth map is made from a source picture which is commonly in gray scale.Transfering the source picture into gray scale picture is the main step of 2D-to-3D conversion. The next step is the transferring of depth map into stereoscope (stereo pairs)pictureconsisting of right and left pictures.

This is carried out by Rendering the depth mapandthan estimating the construction of 3D object from the data of stereoscope image with 3D studioMAX. Different correlations of various techniques have been deployed to improve the nature of the 3D picture reconstruction.Some of the 2D-to-3D conversion applications areMedical, Robotics, film industry, gaming, virtual environment,animation, human PC cooperation, reverse engineering and so forth.

REFERENCES

[1]. J. Wang, B. Li, W. Chen and L. Rong, "3DReconstruction Embedded System Based on Laser Scanner for Mobile Robot". 3rd IEEE Conference on ICIEA, pp. 697-701, June. 2008.
[2]. A.E, A.A,A.Maysam, Swash. M.R, A.F.O,andF.J, "Scene depth extraction from Horoscopic imaging technology," IEEE,2008,pp.1–4.
[3]. Xiao-Ling Deng, Xiao-Hua Jiang1, Qing-Guo Liu, Wei-Xing Wang, Department of Electronics, South China Agricultural University, Guangzhou, "Automatic Depth Map Estimation of Monocular Indoor Environments", 2008 International Conference on Multimedia and Information Technology.
[4]. S. Bharathi, PG Scholar, Kumaraguru Collegeof Technology, Coimbatore 641049, Tamil Nadu, India. A. Vasuki Professor, "2D-To-3D Conversion of Iimages using Edge information", (ICON3C 2012) Proceedings published in International Journal of Computer Applications (IJCA).
[5]. Yea-Shuan Huang, Fang-Hsuan Cheng, and Yun-Hui Liang, Department of Computer Science &Information Engineering of Chung Hua University, Taiwan, "Creating Depth Map from 2D Scene Classification", The 3rd International Conference on Innovative Computing Information and Control (ICICIC'08), 2008 IEEE.
[6]. Yi-Che Chen, Yi-Chin Wu, Chih-Hung Liu,Wei-Chih Sun and YungChang Chen, Department of Electrical Engineering., National Tsing Hua University, Hsinchu, Taiwan, "Depth Map Generation Based on Depth From focus", 2010 International Conference on Electronic Devices, Systems and Applications, ICEDSA2010.
[7]. Q. Hao, R. Cai, Z. Li, L. Zhang, Y. Pang, F. Wu and et al. "Efficient 2D to 3D Correspondence Filtering for Scalable 3D object Recogntion," 2013 IEEE Conference on Computer Vision and Pattern Recognition, Portland, Oregon USA, 23-28 June 2013.
[8]. G.J.Iddan, G. Yahav," 3D imaging in the studio and elsewhere", Proc. SPIE 4298 (2001) 48–55.
[9]. A.L.Yuille, T. Poggio, "A generalized ordering constraint for stereo correspondence", A.I. Memo 777, AI Lab, MIT, 1984
[10]. R. Dutra da Silva, R. Minghim and H. Pedrini,"3D Edge Detection ased on Boolean Functions and Local Operators," International Journal of Image and Graphics, Singapore, vol.15, n. 1, January 2015, pp. 1550003-1-1550003-21.
[11]. N.Hautiere, J.-P. Tarel, J. Lavenant and D. Aubert, Machine Vision and Applications, vol. 1, no. 17, (2006).
[12]. K. He, J. Sun and X. Tang, "Single image hazeremoval using dark channel prior", Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, (2009) June 20-25, Miami, United states.
[13]. K. M. He, J. Sun and X. O. Tang, IEEE Trans.Pattern Anal. Mach. Intell., vol. 12, no. 33, (2011).
[14]. K. M. He, J. Sun and X. O. Tang, "Guided Image Filtering", Proceedings of 11th European Conference on Computer Vision, (2010) September 5-11, Berlin, Germany.

Baki Koyuncu, et.al. "Converting 2D Pictures In Jpeg Into 3D Objects In Space" *The International Journal of Engineering and Science (IJES)*, 9(01) (2020): 15-20.