IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-7, July 2018 E-ISSN: 2347-2693

Recent Trends in Image 2D to 3D: Binocular Depth Cues

Disha Mohini Pathak[1*], Swati Chauhan[2], Dayanand[3]

[1,2,3]*Department of Computer Science,KIET group of Institutions, Ghaziabad, INDIA,*

Available online at: www.ijcseonline.org

Accepted: 03/Jul/2018, Published: 31/Jul/2018

Abstract- As 3D images and videos are grabbing more attention of people these days, it became very important concern for researchers. Recent area of interest growing in converting 2D images into 3D images. Many researchers have worked upon different methods to bridge this gap. This paper addresses various methodologies and recent trends in Image from 2D to 3D conversion. We have bestow on Binocular Depth Cues Techniques like binocular Disparity Motion, Focus, Defocus, silhouette and other concepts based on several aspects and considerations. This paper also discusses strength and limitations of these algorithms to give a broader review on which various techniques to be used in different cases.

Keywords: Binocular Depth Cues, 2D to 3D Images, Strength, Limitations, Conversion Algorithms.

I. INTRODUCTION

Depending on the number of input images, we can label the existing conversion algorithms into two groups: algorithms based on two or more images and algorithms based on a single still image. In the first case, the two or more input images could be taken either by multiple fixed cameras located at different viewing angles or by a single camera with moving objects in the scenes. We generate depth cues used by the first group the multi-ocular depth cues. The second group of depth cues operates on a single still image, and they are referred to as the monocular depth cues [3].
The Table 1 summarizes the depth cues used in 2D to 3D conversion algorithms and their representative works. [1]
Until now we have been concerned with monocular techniques. The camera we considered had one lens and the properties of arrangements. Two eyes provide two views. However, two eyes are placed in front of head, have fields of view that coincide considerably (field of view is 208º with 130º overlap). Thus eyes provide slightly different views of almost same scene. These slightly different images of the same scene enable to estimate the depth of 3 dimensional scene. The world of 3D incorporates the third dimension of depth, which can be perceived by the human vision in the form of binocular disparity. The brain is then able to reconstruct the depth information from these different views. A 3D display takes benefit of these circumstances, creating two slightly different images of every scene and then presenting them to the individual eyes. With an appropriate disparity and calibration of parameters, a correct 3D perception can be realized.

TABLE 1: Depth cues and their representative algorithms. [1]

	Depth Cues	***Representative Works***
One single image (monocular)	Binocular	Correlation-based, feature-based
	Disparity	correspondence; triangulation [23][24]
	Motion	Optical flow [23]; Factorization[30]
	Defocus	Kalman filter [31] Local image decomposition using the Hermite polynomial basis [25]; Inverse filtering [32]; S-Transform [33]
	Focus	A set of images of different focus
		level and sharpness estimation [26]
	Silhouette	Voxel-based and deformable mesh
		model [27]
One single image (monocular)	Defocus	Second Gaussian derivative [28]
	Linear perspective	Vanishing line detection and gradient plane assignment [29]
	Atmosphere Scattering	Light scattering model [13]
	Shading	Energy minimization [14]
	Patterned texture (Incorporates relative size)	Frontal texel [15]
	Symmetric patterns	Combination of photometric and geometric constraints [16]
	Occlusion	

	- Curvature - Single Transform	Smoothing curvature and isophote [17] Shortest path [18]
	Statistical patterns	Color-based heuristics [29], Statistical estimators [19]

Many 3D instruments have entered into our live, such as 3D display, stereoscopic capture, games and so on. The discernment of 3D images is due to the parallax between the viewer's two eyes. Therefore, traditional stereo vision generation requires at least two images with slightly different projections.[2] Thus changing 2D images into 3D images requires different techniques to get the perfection.The conversion methods are also categorized into automatic method and semi-automatic method. In automatic method human intervention is not required, but in semi-automatic method human operator is involved. Computational time and design cost are the metrics that should be considered while designing these algorithms.

II. Conversion Techniques

Binocular Disparity

With two images of the alike scene captured from slightly different viewpoints, the binocular disparity can be used to recover the depth of an object. This is the main mechanism for depth perception. Those of us blessed with pair of eye need not to move in order to gain the benefits of parallax for estimating depth. Eyes are separated by about 6 and half cm and with significantly overlapping field of view, see slightly different views of any object they look at. The difference between the views of the two eyes (binocular disparity) thus provides a way to determining the distance of the object in sight. This difference in response leads to the viewer's perception and these features lie at different depths.[40]

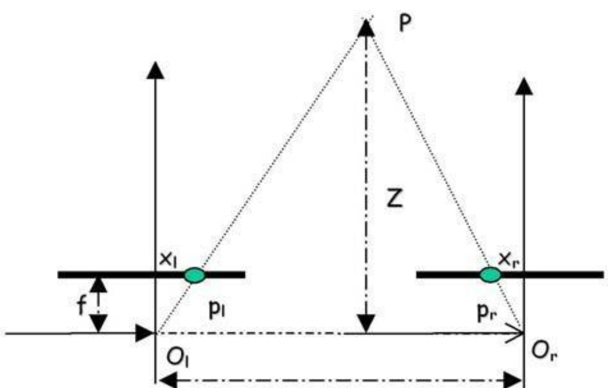

FIGURE 1: Binocular Disparity.

Binocular Disparity is useful for depth estimation only if your eyes see different images. Distant scene also presents essentially the same view to your two eyes, for such cases binocular disparity is of little use. But for close objects it is extremely effective way of gauging depth. The amount of Depth perceived depends on binocular disparity between your two views. The wider the separation of the points of views, the greater the apparent depth is.

Assume p_l and p_r are the projections of the 3D point P on the left image and right image; O_l and O_r are the origin of camera coordinate systems of the left and right cameras. Based on the relationship between similar triangles (P , p_l , p_r) and (P,O_l ,O_r) shown in Figure 2, the depth value Z of the point P can be obtained:

$$Z = f\,\frac{T}{d} \qquad (2.1)$$

where $d = x_r - x_l$, which measures the difference in retinal position between corresponding image points. The disparity value of a point is often interpreted as the inversed distances to the observed objects. Therefore, finding the disparity map is essential for the construction of the depth map.[3]

Advantages:

Increase Depth Perception

With both eyes of a creature with binocular vision are in front and near each other, it results to a highly improved perception of depth. This is because a creature is able to distinguish how near or far is the subject from it. This is not the same in the case of a creature with monocular vision.

Flexibility

Another advantage of having two eyes that can focus on an image directly is being able to use only one eye in case the other one is damaged or blinded. Although this might not be easy at first and needs getting used to, having only one functioning eye is possible. This is because the eyes are located in front and it will still be possible for a person or an animal with only one eye to see what is in front.

Visibility Beyond an Obstacle

This is explained by Leonardo Da Vinci when he said that if there is a vertical column closer to a person than the image or subject of focus, there column might block a portion of the image from one eye but that part of the image can be visible to the other eye.

Disadvantages:

Decreased Visual Field

A creature with binocular vision might have the ability of depth perception but cannot enjoy a wider visual field unlike preys with monocular vision that can see from behind. Since the eyes are both located in front, a creature with binocular vision can also see in front and from the peripheral view but not beyond these two.

Limited Focus
The placement of the eyes which are proximal to each other allows for a creature to see an image and focus on a single image. This means that the creature will not be able to see another object from the sides or behind since the brain sends signals to the eyes to focus on only one image.

$$v_x \square \frac{T_z x - T_x f}{Z} - \omega_y f \square \omega_z y \square \frac{\omega_x xy}{f} - \frac{\omega_y x^2}{f}$$

$$v_y \square \frac{T_z x - T_y f}{Z} \square\square\omega_x f - \omega_z y - \frac{\omega_y xy}{f} \square \frac{\omega\, xy^2}{f}$$

Prone to Disorder
Binocular vision can suffer from different visual disorders and anomalies. These include visual confusion, suppression and diplopia. Also, these anomalies can result to blurred vision, headache and eye pain, among others.

Motion
By moving your head you can change your view sufficiently so that you can see the object. To estimate depth you can rely much more heavily on the fact that your view is different as you are at different position. This phenomenon known as parallax. So you view of different objects changes as you move, according to the distance from you (Figure 2)

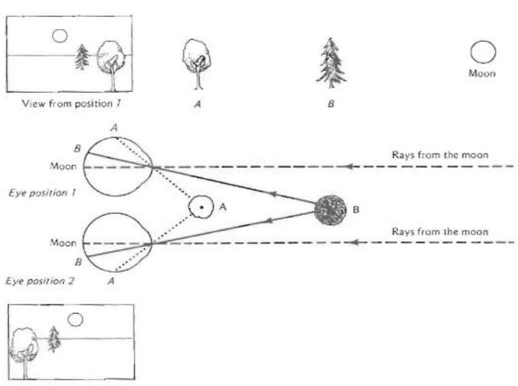

FIGURE 2: Perception of image w.r.t motion

The relative motion between the viewing camera and the observed scene provides an important cue to depth perception: near objects move faster across the retina than far objects do. The extraction of 3D structures and the camera motion from image sequences is termed as structure from motion. The motion may be seen as a form of "disparity over time", represented by the concept of motion field. The motion field is the 2D velocity vectors of the image points, induced by the relative motion between the viewing camera and the observed scene. The basic assumptions for structure-from-motion are that the objects do not deform and their movements are linear. Suppose that there is only one rigid relative motion, denoted by V, between the camera and scenes. Let $P = [X, Y, Z]^T$ be a 3D point in the conventional camera reference frame. The relative motion V between P and the camera can be described as

$$V = -T - \omega \times P \qquad (2.2)$$

where T and ω are the translational velocity vector and the angular velocity of the camera respectively. The connection between the depth of 3D points and its 2D motion field is incorporated in the basic equations of the motion field, which combines equation (2.2) and the knowledge of perspective projection:

(2.3)

(2.4)

Where v_x and v_y are the components of motion field in x and y direction respectively; Z is the depth of the corresponding 3D point; and the subscripts x, y and z indicate the component of the x-axis, y-axis and z-axis directions.[6]

Advantages: when the average disparity between frames is large, the depth reconstruction can be done in a way as that of binocular disparity (stereo). The motion field becomes equal to the stereo disparity map only if the spatial and temporal variances between frames are sufficiently small. [6]

Disadvantages: It is worth to note that the sufficiently small average spatial disparity of corresponding points in consecutive frames is beneficial to the stability and robustness for the 3D reconstruction from the time integration of long sequences of frames. [6]

Defocus Using More Than Two Images
Depth-from-Defocus (DFD) technique is an elegant passive auto focusing method. It only needs two or three images recorded with different camera settings to recover the depth of certain object by computing the degree of blurring.

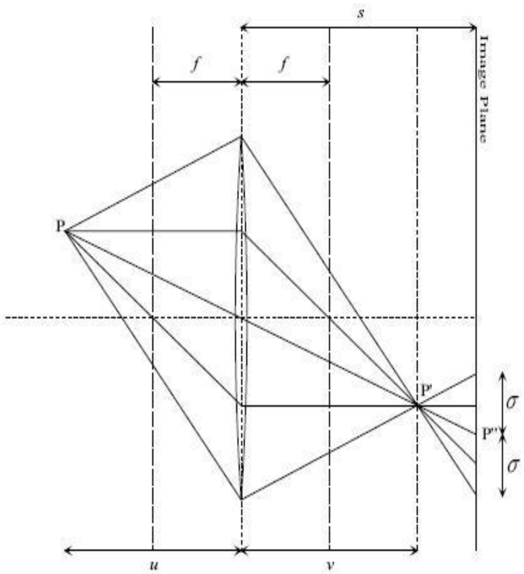

FIGURE 3: Thin Lens Model [6]

There are generally two categories of DFD algorithms: statistical and deterministic. Statistical approaches like Maximum likelihood [7] and Markov Random field methods [8] require more image computing. Deterministic algorithms can be classified as frequency domain approaches [9][10] and spatial domain approaches.[11][12] The frequency domain approaches are generally computation expensive and yield lower depth-map density. While spatial domain approaches only use a small image region, thus require less computation and generate piecewise depth-map. Due to the inherent advantage of being local in nature, spatial domain approach is more suitable for real-time auto focusing applications.[39] [6]
Depth-from-defocus methods generate a depth map from the degree of blurring present in the images. In a thin lens system, objects that are in-focus are clearly pictured whilst objects at other distances are defocused, i.e. blurred. Figure 3 shows a thin lens model of an out-of-focus real world point P projected onto the image plane. Its corresponding projection is a circular blur patch with constant brightness, centered at P'' with a blur radius of σ. [6]
The blur is caused by the convolution of the ideal projected image and the camera point spread function (PSF) $g(x, y, \sigma(x, y))$ where (x, y) are the coordinates of the image point P''. It is usually assumed that $\sigma(x, y) = \sigma$, where σ is a constant for a given window, to simplify the system and Gaussian function is used to simulate the PSF:

$$g_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \qquad (2.5)$$

Pentl and [14] has derived a relationship between the distance u (Figure 3) and the blur σ in equation (2.7):

$$u = \frac{fs}{s-f-kf\sigma} \quad \text{if } u>v$$

$$u = \frac{fs}{s-f+kf\sigma} \quad \text{if } u<v \qquad (2.7)$$

where u is the depth, v is the distance between the lens and the position of the perfect focus, s is the distance between the lens and the image plane, f is the focal length of the lens, and k is a constant determined by the lens system. Of these, s, f and k are camera parameters, which can be determined by camera calibration. Please note that the second case $u < v$ is possible to happen, for example, when $f < u < 2f$, based on the fundamental equation of thin lenses (2.6), we can obtain $v > 2f$, which yields thus $u < v$.[6]
The depth-from-focus algorithms focus thus on the blur radius estimation techniques.We do not normally experience changes in the precision of depth estimates behind and in front of where we are looking. This state is often achieved by using other depth cues to fill in the gaps left by disparity.

Advantages:
We do not normally experience changes in the precision of depth estimates behind and in front of where we are looking. This state is often achieved by using other depth cues to fill in the gaps left by disparity.[39]
Blur is nearly always informative. It makes the depth estimation significantly more precise throughout visual space[39]
[16] Shows that disparity and depth-of-field blur have the same underlying geometry and therefore that blur is roughly a fixed proportion of disparity.[39]

Disadvantages:
Defocus blur does not in any obvious way indicate the sign of a change in distance: i.e., whether an out-of-focus object is nearer or farther than an in-focus object. However, the visual system does clearly solve the sign-ambiguity problem. For depth estimation, the system solves the problem by using other depth cues that do not provide metric depth information[16][39]
The relationship between distance and blur depends on pupil size. There is no evidence that people can measure

their own pupil diameter, so the relationship between measured blur and specified distance is uncertain. But steady-state pupil size does not vary much under typical daylight conditions. Specifically, intrasubject pupil diameters vary over a range of 2.8 mm for luminance levels between 0.40 and 1,600 cd/m^2 [15], yielding an uncertainty in the estimate of z_0 of only 66% over a luminance range of 200,000%.[16]

Focus

The Depth-from-focus approach is closely related to the family of algorithms using depth from defocus. The main difference is that the depth-from-focus requires a series of images of the scene with different focus levels by varying and registering the distance between the camera and the scene, while depth-from-defocus only needs 2 or more images with fixed object and camera positions and use different camera focal settings. [6]

Depth of focus is a lens optics concept that measures the tolerance of placement of the image plane (the film plane in a camera) in relation to the lens. In a camera, depth of focus indicates the tolerance of the film's displacement within the camera, and is therefore sometimes referred to as "lens-to-film tolerance."The phrase depth of focus is sometimes erroneously used, to refer to the depth of field (DOF), which is the area in front of the lens in acceptable focus, whereas the true meaning of depth of focus refers to the zone behind the lens wherein the film plane or sensor is placed to produce an in focus image.

Depth of focus can have two slightly different meanings. The first is the distance over which the image plane can be displaced while a single object plane remains in acceptably sharp focus, the second is the image-side conjugate of depth of field. With the first meaning, the depth of focus is symmetrical about the image plane; with the second, the depth of focus is greater on the far side of the image plane, though in most cases the distances are approximately equal. Where depth of field often can be measured in macroscopic units such as meters and feet, depth of focus is typically measured in microscopic units such as fractions of a millimetre or thousandths of an inch.

The same factors that determine depth of field also determine depth of focus, but these factors can have different effects than they have in depth of field. Both depth of field and depth of focus increase with smaller apertures. For distant subjects (beyond macro range), depth of focus is relatively insensitive to focal length and subject distance, for a fixed f-number. In the macro region, depth of focus increases with longer focal length or closer subject distance, while depth of field decreases. [17]

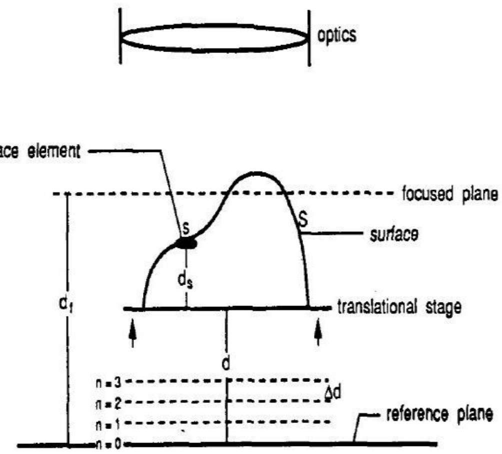

FIGURE 4: Depth from Focus [6]

Figure 4 illustrates the principle of the depth-from-focus approach [18]. An object with an arbitrary surface is placed at the translational stage, which moves towards the camera (optics) starting from the reference plane. The focused plane is defined by the optics. It is located at the position where all points on it are focused on the camera sensor plane. Let ‘s’ be a surface point on the object. When moving the stage towards the focused plane, the images of ‘s’ become more and more focused and will obtain its maximum sharpness when ‘s’ reaches the focused plane. After this, moving ‘s’ furthermore makes its image defocused again. During this process, the displacements of the translational stage are registered. If we assume that the displacement is $d_{focused}$ when ‘s’ is maximally focused and the distance between the ‘focused plane’ and the reference plane is d_f, then the depth value of ‘s’ relative to the stage will be determined as $d_s \square d_f - d_{focused}$. Applying this same procedure for all surface elements and interpolating the focus measures, a dense depth map can be constructed. [6]

Advantages:

In [19] the authors discuss a method in which the blur is evaluated from the intensity
change along corresponding pixels in the multi-focus images instead of using window based blur estimation operators.

Recently Hasinoff and Kutulakos [20] have shown that for very high resolution images the
depth from focus can be seen to be reduced to color comparison with regions of an
Aperture-focus image representation for each pixel.

Disadvantages:

The fundamental weakness of the DFF method is, however, the time required for image
acquisition.

In practice about ten or even more images are required to estimate the depth of a scene
for a reasonable level of accuracy.

Silhouette

A silhouette of an object in an image refers to the contour separating the object from the background. Shape-from-silhouette methods require multiple views of the scene taken by cameras from different viewpoints. Such a process together with correct texturing generates a full 3D model of the objects in the scene, allowing viewers to observe a live scene from an arbitrary viewpoint.

Shape-from- silhouette requires accurate camera calibration. For each image, the silhouette of the target objects is segmented using background subtraction. The retrieved silhouettes are back projected to a common 3D space (see Figure 5) with projection centers equal to the camera locations. Back-projecting a silhouette produces a cone-like volume. The intersection of all the cones forms the visual hull of the target 3D object, which is often processed in the voxel representation. This 3D reconstruction procedure is referred to as shape-from-silhouette.

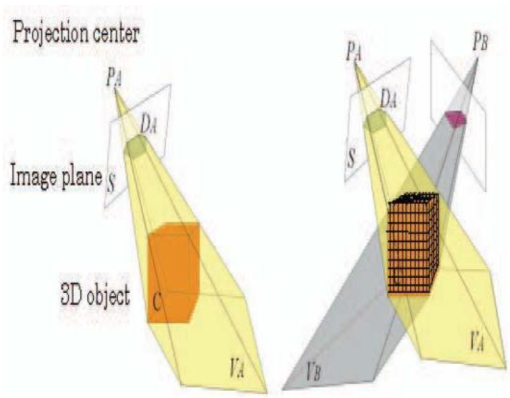

FIGURE 5: Silhouette volume intersection [6]

A shape-from-silhouette algorithm is often followed by a texturing algorithm. The visual hull is a geometry that encloses the captured object, but it does not capture the concave portion of the object that is not visible on the silhouette. Moreover, the number of views is often limited to make the processing time reasonable. This leads to a coarse geometry of the visual hull. Texturing assigns colors to the voxels on the surface of the visual hull and is therefore an indispensable step in creating realistic renderings.[6]

Advantages:

First of all, silhouettes are readily and easily obtainable, especially in indoor environment where the cameras are static and there are few moving shadows. [22] The implementation of most SFS methods is also relatively straightforward, especially when compared to other shape estimation methods such as multi-baseline stereo] or space carving[22]. Moreover, the inherently conservative property of the shape estimated using SFS is particularly useful in applications such as obstacle avoidance in robot manipulation and visibility analysis in navigation[22]

Disadvantages:

Shape-from-silhouette methods require multiple views of the scene taken by cameras from different viewpoints. SFS suffers from the limitation that the shape estimated by SFS (the VH) can be a very coarse approximation when there are only a few silhouette images, especially for complex objects. [22]

III. COMPARISION

On the basis of some parameters these conversion techniques has been compared.

S.N.	Conversion Techniques	Advantage	Disadvantage
1	Binocular Disparity	Increase Depth Perception, Flexibility Visibility Beyond an Obstacle	Decreased Visual Field Limited Focus Prone to Disorder
2	Motion	motion field becomes equal to the stereo disparity map	time integration of long sequences of frames
3	Defocus Using More Than Two Images	Blur is nearly always informative. It makes the depth estimation significantly more precise throughout visual space	The system solves the problem by using other depth cues that do not provide metric depth information
4	Focus	For very high resolution images the depth from focus can be seen to be reduced to	Ten or even more images are required to estimate the depth of a scene

		color comparison	for a reasonable level of accuracy
5	Silhouette	silhouettes are readily and easily obtainable, useful in applications such as obstacle avoidance in robot manipulation	Shape-from-silhouette methods require multiple views

IV. CONCLUSION

Binocular Depth Cues Techniques like binocular Disparity Motion, Focus, Defocus, silhouette are important techniques for 2D to 3D conversions. The binocular disparity can be used to recover the depth of an object. To estimate depth you can rely much more heavily on the fact that your view is different as you are at different position is done by Motion. Depth-from-Defocus (DFD) technique is an elegant passive auto focusing method. The main difference is that the depth-from-focus requires a series of images of the scene with different focus levels by varying and registering the distance between the camera and the scene. A silhouette of an object in an image refers to the contour separating the object from the background

V. REFERENCES

[1] M. Galabov, Research Conference In Technical Disciplines, *2D to 3D conversion algorithms*, November, 2014

[2] T-Ying Kuo, Y-Chung Lo, and C Lin, ICASSP IEEE 2012 *2D-TO-3D CONVERSION FOR SINGLE-VIEW IMAGE BASED ON CAMERA PROJECTION MODEL AND DARK CHANNEL MODEL*,March, 2012

[3] M. Galabov , IJESIT Volume4, Issue 1,*A Real Time 2D to 3D Image Conversion Techniques*, January, 2015

[4] D.Mohini Pathak, Tanya Mathur, IJEDR, Volume 5, Issue 2, *Recent trends in image 2D to 3D: Monocular depth cues*, May 2017

[5] Scott Squires." 2D to 3D Conversions", *http://effectscorner.blogspot.in/2011/08/2d-to-3d-conversions.html#.WUNpxohEmUk*, August, 2011

[6] Q. Wei, "Research Assignment for Master Program Media and Knowledge Engineering of Delft University of Technology" , Converting 2D to 3D: A Survey, December,2005

[7] Xue Tu, Youn-sik Kang and Murali Subbarao, "Article Proceedings of SPIE - The International Society for Optical Engineering", Depth and Focused Image Recovery from Defocused Images for Cameras Operating in Macro Mode,Sep 2007

[8] A. N. Rajagopalan and S. Chaudhuri," IEEE trans image process", A recursive algorithm for maximum likelihood-based identification of blur from multiple observations,1998

[9] A. N. Rajagopalan and S. Chaudhuri," ICCV, pp. 1047–1052" , Optimal recovery of depth from defocused images using an MRF model, 1998.

[10] Y. Xiong and S. Shafer, "DARPA93, pp. 967" ,Depth from focusing and defocusing,1993.

[11] M. Watanabe and S. Nayar, "Tech. Rep. CUCS-035-95, Dept. of Computer Science, Columbia University ", Rational filters for passive depth from defocus, Sept. 1995.

[12] M. Subbarao and G. Surya, "IJCV **13**, pp. 271–294" , Depth from defocus: A spatial domain approach,December 1994.

[13] D. Ziou and F. Deschenes, "Computer Vision and Image Understanding : CVIU **81**(2), pp. 143–165", Depth from defocus estimation in spatial domain, 2001.

[14] Pentland, A. P. "IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 9, No.4, Page(s) 523-531", Depth of Scene from Depth of Field", 1987

[15] Spring, K.H., and Stiles, W.S. "Br. J. Ophthalmol. 32, 340–346 ",Variation of pupil size with change in the angle at which the light stimulus strikes the retina, 1948

[16] R. T. Held,[1,]* A. Cooper,[2] and S. Banks," Current Biology 22, 426–431, March 6, 2012 ª2012 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2012.01.033", Blur and Disparity Are Complementary Cues to Depth, March, 2012.

[17] WIKIPEDIA, "Depth of Focus" https://en.wikipedia.org/wiki/Depth_of_focus#cite_ref-Larmore1965p167_1-0

[18] Nayar, S.K.; Nakagawa, Y, "Pattern Analysis and Machine Intelligence, IEEE Transactions on Volume 16, Issue 8, Page(s): 824 – 831", Shape from Focus, 1994

[19] T. M Naoki Asada, H. Fujiwara, "*International Journal of Computer Vision*, 28(2):153–163",
Edge and depth from focus, 1998.

[20] S.W. Hasinoff and K.N. Kutulakos." In *Proc. Ninth European Conference on Computer Vision*, pages 620–634", Confocal stereo. May 2006.

[21] Matsuyama, T. "Informatics Research for Development of Knowledge Society Infrastructure, ICKS 2004, International Conference, Page(s) 7-14",Exploitation of 3D video technologies ",2004

[22] K. Cheung, S. Baker and T. Kanade, "International Journal of Computer Vision, May 2005 , **Volume** 62, **Issue** 3, pp 221–247", Shape-From-Silhouette Across Time, May 2005.

[23] Michel B., "*La Stéréoscopie Numérique*,Eyrolles, Chapter 5", La conversion 2D–3D",2011.

[24] Scharstein, D.; Szeliski, R, "International Journal of Computer Vision 47(1/2/3), 2002,7-42",
A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms, 2002

[25]DA SILVA V.,"Depth image based stereoscopic view rendering forMATLAB",http://www.mathworks.com/matlabcentral/fileexchange/27538-depthimage-
based-stereoscopic-view-rendering, 2010.

[26] Guan-Ming Su, Yu-Chi Lai, Andres Kwasinski and Haohong Wang," First Edition, Published by John Wiley &Sons, Ltd", 3D Visual Communications, 2013.

[27] Matsuyama, T. "Informatics Research for Development of Knowledge Society Infrastructure, ICKS 2004, International Conference, 2004, Page(s) 7-14",Exploitation of 3D video technologies, 2004

[28] Wong, K.T.; Ernst, F., "Single Image Depth-from-Defocus", Master thesis, Delft university of Technology & Philips Natlab Research,Eindhoven, The Netherlands, 2004.

[29] Battiato, S. ; Curti, S. ; La Cascia, M.; Tortora, M.; Scordato, E. "SPIE Proc. Vol 5302, EI2004
conference „Threedimensional image capture and applications VI", Depth map generation by image classification", 2004.

[30] Han, M; Kanade, T. “IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 25, Issue 7, 2003, Page(s): 884 – 894”, Multiple Motion Scene Reconstruction with Uncalibrated Cameras, 2003

[31] Franke, U.; Rabe, C., “Intelligent Vehicles Symposium, Proceedings. IEEE, 2005, Page(s): 181 – 186”, Kalman filter based depth from motion with fast Convergence, 2005.

[32] Kao, M.A., Shen, T.C. “*IDW'09*, p. 203, 2009”, A novel real time 2D to 3D conversion technique using depth based rendering, 2009

[33] Bleyer M., Gelautz M., “Proceedings of the 6th International Symposium on Image and Signal Processing and Analysis (ISPA),Salzburg, pp. 383–387, 16–18”, Temporally consistent disparity maps from uncalibrated stereo videos”, September 2009.

[34] XU F., LAM K.M., DAI Q., “*Image and Vision Computing Journal*, vol. 29, no. 2–3, pp. 190–205”, Video-object segmentation and 3Dtrajectory estimation for monocular video sequences, 2011.

[35] Kang, G; Gan, C.; Ren, W., “, International Conference on Machine Learning and Cybernetics, Volume 8, 2005, Page(s): 5165 – 5169”,Shape from Shading Based on Finite-Element,2005

[36] Loh, A.M.; Hartley, R., “Proceedings, the British Machine Vision Conference”,Shape from Non-Homogeneous, Non-Stationary, Anisotropic, Perspective texture, 2005.

[37] Shimshoni, I.; Moses, Y.; Lindenbaumlpr, M.,“ Proceedings, International Conference on Image Analysis and Processing, 1999, Page(s): 76 – 81“, Shape reconstruction of 3D bilaterally symmetric surfaces”, 1999

[38] Redert, A., “Creating a Depth Map”, Royal Philips Electronics, the Netherlands, Patent ID: WO2005091221 A1, 2005.

[39] V.P. Namboodiri, “Novel diffusion based techniques for depth estimation and image restoration from defocused images.” Doctor of Philosophy thesis, IIT, Bombay, India, 2008

[40] N. Qian,” Neuron, Vol. 18, 359–368, March, 1997, Copyright ã1997 by Cell Press”, Binocular Disparity Review and the Perception of Depth, 1997.

Authors Profile

Disha Mohini pathak has completed bachelors in Technology in Computer Science Engineering from VIET, Gautam Buddh Nagar Masters from SRM University, Chennai in 2013 and currently she is pursuing Doctorate from JIIT, Noida Uttar Pradesh. She has an experience of 4 years in academics and currently working with KIET group of Institutions, Ghaziabad.. She has published numerous research papers in various conferences and journals. Her area of interest is in image processing.

Ms.Swati Chauhan is working as an Assistant Professor in Computer science and engineering department since 2014 at KIET Group of Institutions, Ghaziabad, Uttar Pradesh. She completed her post-graduation from Uttar Pradesh Technical University and B.Tech from HNB Garhwal University, Garhwal. She did her research internship at IIT Delhi in Computer Vision. Her research interest includes Digital Image Processing, Computer Vision and Machine Learning.

Dayanand has completed bachelors in Technology in Computer Science Engineering from SHIATS, Allahabad, Masters from Birla Institute of Technology, Ranchi in 2013 and currently he is pursuing Doctorate from Sam Higginbottam University of agricultural technology and Sciences, State University, Uttar Pradesh. He has worked as manager IT in Govt. of Delhi and has done a number of govt. projects . He has an experience of 4 years in academics and currently working with KIET group of Institutions, Ghaziabad. He has authored books namely Foundation of Computer Science, Discrete Mathematics and Information Security. He has been awarded Dr. Rajendra Prasad Teachers award 2016. He has published more than 40 research papers in various conferences and journals.

