Recovering 3D Shape of Weak Textured Surfaces

K. Kanthamma and Dr. G. N. Kodandaramaiah

Abstract: 3D shape recovery of the object from its 2D images based on image focus has been an important field of research shape from focus is one of the passive method to recover the shape of the object. Mostly, existing approaches work well with dense textured objects. However, they cannot compute depth of week textured scenes with great precision. In this paper, we propose a new shape from focus algorithm which improves the recover shape of weak textured objects. The proposed method is experimented and its performance is tested using different image sequences of synthetic and real objects with varying texture. The proposed approach provides better results as compared to previous approaches.

Keywords: shape from focus, dense texture, weak texture.

I. INTRODUCTION

Focusing the image system automatically or obtaining the sparse depth information from the observed scene has been an active area of research. Shape from focus is one of the optical passive methods for 3D shape recovery of an object from its 2D images. In SFF a sequence of images is used, taken by a single camera at different focus levels, to compute the depth of the object in the scene. In literature, a variety of SFF technique including traditional method, neural networks and dynamic programming have been reported. In SFF , focus measure is applied to compute the focus quality for each pixel in image sequence and then an approximation technique is applied to recover the shape of the object. Laplacian, modified laplacian, sum of the modified laplacian tenenbaum gray level variance and M_2 are some of the famous focus measures.

Mostly focus measures like SML and tenenbaum are gradient based and provide poor response in the less textured areas. Other statistical based focus measures, like GLV, depend on how sharply the pixel values differ in spatial domain. Their performance also deteriorates in estimating correct focus values when the objects have less variation in texture. In nature, some parts of objects generally have weak textures. Therefore mostly SFF techniques work well with dense textured objects. However, they are poor in computing the depth of weak textured precisely, as they are based on assumptions about the presence of prominent texture in the scene.

K.Kanthamma is working a Associate professor dept. of ECE, SRIT,Anantapur,Andrapradesh, and Dr.G.N.Kodandaramaiah is working as Professor dept of ECE,kuppam engineering college,kuppam,Andrapradesh, Kantha.srinivas@yahoo.in kondandramaiah@yahoo.com

In this paper we propose a new SFF method to approximate the 3D shape of the objects with weak texture surfaces. An initial depth map is computed by applying a focus measure on the image sequence and maximizing the value of the focus curve along the optical axis. Due to illumination, noise and/or other factors, but mainly due to presence of weak texture, maximizing the focus curve results in false depth detection of some points. Their estimated depth is discarded by threshold selection method. The missing parts are then computed by using line interpolation. The proposed method is experimented and evaluated using different image sequences of synthetic and real objects with varying texture.

II. SHAPE FROM FOCUS

In Shape from Focus (SFF) methods, a sequence of images is used, taken by a single camera at different focus levels, to compute the depth of the object in the scene. Then, the entire image sequence is searched to find the best focused image frame for a particular point in the image space; and, camera parameter settings for that image frame are used to compute the distance of corresponding object point. In figure 1, the point '*P*' is best focused at image distance '*v*' from the lens.

$$\frac{1}{f}=\frac{1}{u}+\frac{1}{v} \qquad \text{.....1}$$

To compute the distance '*u*' of the corresponding object-point, Thin Lens Formula is used, where '*f*' is the focal length of the lens. The equation for the blur Radius '*R*' caused by the focusing of the object surface is given by the(2).

$$R=s\frac{D}{2}\left(\frac{1}{f}-\frac{1}{u}-\frac{1}{v}\right) \qquad \text{.....2}$$

Where '*D*' is the diameter for the lens aperture as shown in the figure 1.

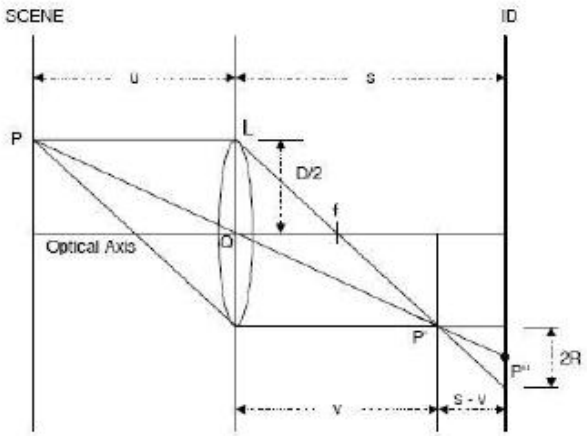


Figure 1. Camera System for SFF.

III .MAGE ACQUISITION SYSTEM

For experiments we have used three different image sequences (two simulated cone image sequences with different textures, and LCD-TFT image sequence). The Simulated cone images are generated by computer simulations. We have used algorithm to generate image sequence. The thick ring pattern on the first image sequence of simulated cone provides rich texture for shape recovery. Whereas, the second image sequence of the simulated cone have less variation in texture. The LCD images are microscopic images of LCD-TFT filter, obtained by microscopic control system (MCS). Figure 1 shows sample images in the image stacks for simulated cone, LCD-TFT, and the original depth map of simulated cone. All the images are obtained by varying the object plane in small steps, and are stored in a sequence on every step.

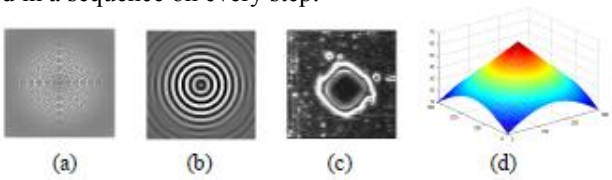


Figure 2. (a & b) 60th frame of each image sequecne of simulated cone, (c) 10th frame of LCD-TFT image sequecne, (d) original depth map of simulated cone

III. EFFECT OF TEXTURE ON SHAPE RECOVERY

An inherent weakness of focus-based method is that they require the imaged scene to have significant texture. In many real-world applications, surfaces can be smoothly shaded and lacking in detectable texture. In such cases, existing SFF techniques generate inaccurate and sparse depth maps. For accurate calculation of depth map, we consider the effects of Texture on shape recovery. The interaction between light and surface is complex and introduces many unwanted artifacts into the image. For example, shading, shadows, and inter reflections, all make it more difficult to achieve basic visual tasks such as image segmentation, object recognition, and tracking. In SFF the images are taken with one light source. Thus, the regions in the image which are well illuminated and have dense texture give more accurate depth map; whereas, the weak textured regions give less accurate depth map. Using the Thin-Lens-Model, when the point '*P*' on the object is best focused in the focused plane, its corresponding pixel intensity value in the image is its true value (i.e. in a gray scale image, for white points on the body the pixel intensity value in the image will be near to maximum and for the black points the pixel intensity value will be near to minimum), whereas, when the point is defocused it will have the defused value (i.e. between white and black). The change in pixel intensity follows a '*Gaussian Model*'. This phenomenon is observed when the neighborhood (of the object point) has different values than that of the object point. Figure 3 shows the typical pixel behavior for two different pixels of simulated cone in the image sequence.

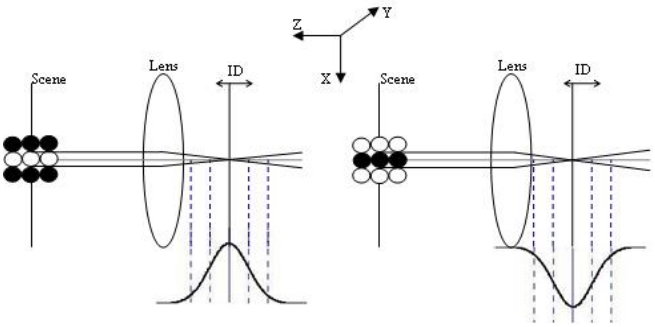


Figure 3 intensity variation of a pixel with different neighborhood in the image stack.

When the neighborhood of the object point has similar values, then its corresponding pixel intensity value in the image will not change with the change in focus. The values of the corresponding pixel will be same when focused or defocused. This phenomenon is due to the lack of texture in the observed scene. Figure 4 demonstrates the change of pixel values with similar neighborhood. The depth of these types of object points is the difficult to detect. In literature this problem is solved by increasing the window size for summation of the focus measure. and window size on the 3D shape recovery.

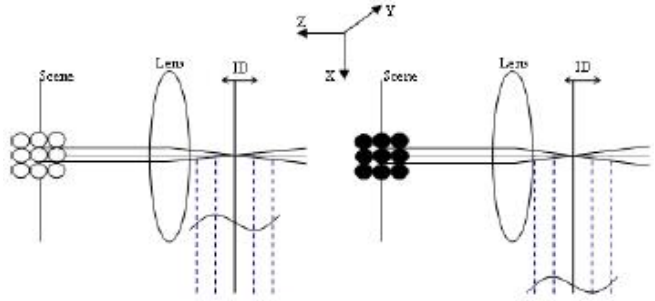
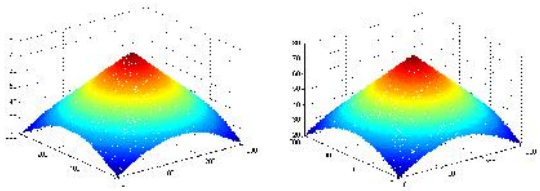
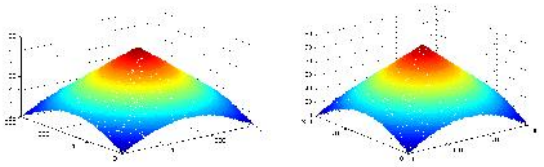


Figure 4. Intensity variation of a pixel with similar neighborhood in the image stack.

Figure 5 and 6 show the reconstruction of simulated cone by different FMs using image stack 1 and 2. It is clear from these figures that the images of first stack have high texture and are easy to reconstruct where as the others have weak textures, and thus the reconstruction becomes difficult. Two global statistical metrics namely, root mean squared error (*RMSE*) and correlation (*Cor.*) are applied to evaluate the performance of the focus measure with high and low textures.



SML TEN



GLV M_2

figure 5. Reconstruction of simulated cone using image stack 1

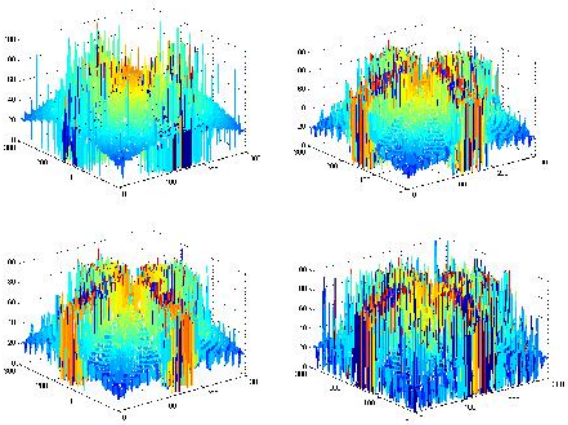


Figure 6. reconstruction of simulated cone using image stack

V. PROPOSED ALGORITHM

We observed the presence of spikes in the initial depth estimation. These spikes may be due to the illumination, noise of similar effects; but are mainly due to the areas on the surface where the object has weak texture. These spikes result in false depth detection of some points. This fact established in previous section.

We proposed a solution for this problem by discarding or eliminating these wrong depth values and later estimating them from accurate depth values of their neighbors. For this purpose segmentation based on threshold selection method is proposed to separate the noise (spikes) from the initial depth map ($D_{initial}$) calculated by using a focus measure.

The initial depth map $D_{initial}$ is divided into the small windows patches P_a , each of size w x w (thus the total number of small windows patches is N=X X Y/w, and, a=1,…..N)), and the histogram h(k) of each window is calculated by below equation

$$h(k) = \sum_{i=1}^{\omega}\sum_{j=1}^{\omega} c_k(i,j),$$

$$c_k(i,j) = \begin{cases} 1 & if \quad P_a(i,j) = z \\ 0 & otherwise \end{cases} \qquad \text{……..3}$$

Where h(k) is the number of points within depth range ($z \in [(k-1)\beta, k\beta]$), β is the total number of bins given by

$$\beta = \frac{P_a^{max} - P_a^{min}}{n}$$

(where n=10 in our case), k=1,….. β , and ω^2 is the total number of object points in the window patch. If the histogram of the window patch includes some peaks, then we can separate them by threshold. The histogram h(k) indicates the occurrence frequency of the object points with similar depth values. Let '$P_k(z)$' be the probability of object points belonging to depth range 'z' and is defined as '$P_k(z)=h(k)/\omega^2$' and C_k be the K^{th} cluster of the depth range z in ascending order. Then, the probability of object points belonging to the cluster C_k is given by below equation

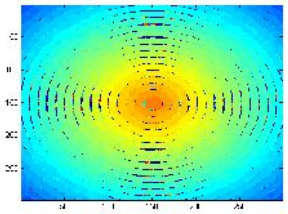
$$P(C_k) = \sum_{z=(k-1)\beta}^{k\beta} p_k(z), \qquad \sum_{k=1}^{\beta} P(C_k) = 1 \qquad \text{……..4}$$

The adjacent clusters with higher probabilities are kept and the rest discarded. The missing parts of the resultant depth map are then approximated by using line interpolation, and the complete 3D shape is recovered.

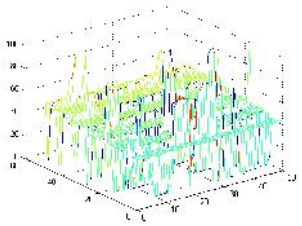
VI. RESULTS AND DISCUSSIONS

The initial depth map calculated by using a focus measure is divided into small windows each of size ω X ω,After calculating the histogram using clustering algorithm.

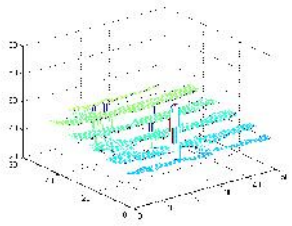
Figure7 a shows simulated cone reconstructed using SML. In figure 7b and 7c, an example of window patch is shown before and after discarding the wrong estimated depths. Figure 7e shows the histogram of the window patch for simulated cone, the cluster threshold is calculated. The depth values between these threshold values are kept and the rest are discarded . figure 9 shows the depth maps of simulated cone after applying the proposed algorithm. The empty patches show the areas of discarded depth values. Figure 7 shows the reconstruction of the simulated cone after interpolating the missing data.



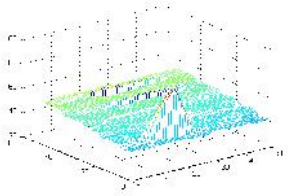
(a)



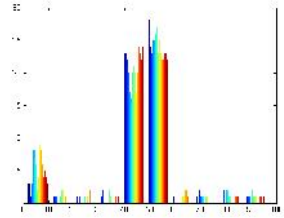
(b)



(c)

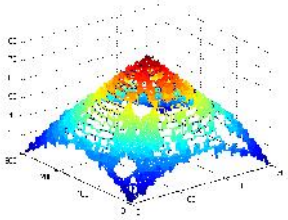


(d)

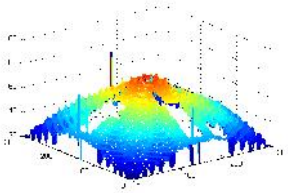


(e)

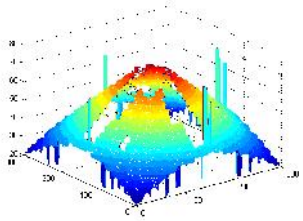
Figure 7. Reconstruction of simulated cone using SML. (a) Top view, (b) Window patch example, (c) with discarded depth, (d) after Interpolating missing data, (e) histogram of the window patch in'b'.



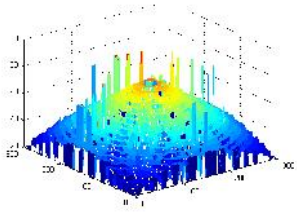
SML



TEN

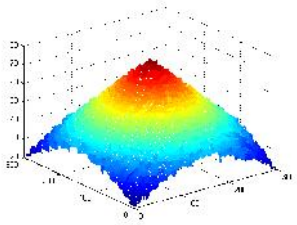


GLV

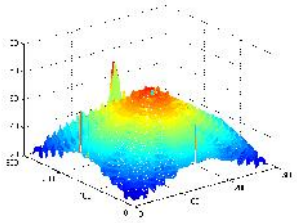


M_2

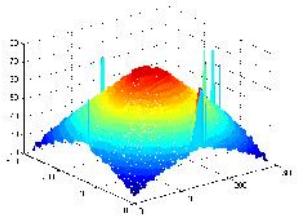
Figure 8.Reconstruction of simulated cone by proposed method before interpolating missing data



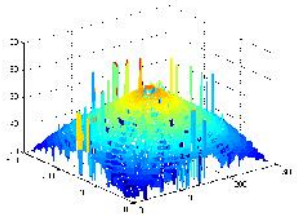
SML



TEN



GLV



M_2

Figure 9. Reconstruction of simulated cone by proposed method after interpolating missing data.

VII. CONCLUSION

In this paper we have proposed a new SFF method to approximate the 3D shape of the object with less texture surfaces. An initial depth map is computed by using well known focus measure and maximizing the focus curve along the optical axis. Experimentally, it is observed that the depth map has spike-like-noise due to the points where the focus measure has failed to determine accurate depth. The proposed solution to this problem is to discard the depth of these points through segmentation. The missing parts are computed by using line interpolation. The proposed method works well for the less textured surfaces

REFERENCES

[1] Ahmad M. B., Choi T. S., "A heuristic approach for finding best focused shape", IEEE Trans. CSVT, Vol. 15, No 4, pp 566-574, April 2005.

[2] Asif M., Choi T. S., "Shape from focus using Multilayer Feed-forward Neural Networks", IEEE TIP, Vol. 10,No. 11 November 2001.

[3] Subbarao M., Choi T. S., "Accurate Recovery of Three Dimensional Shape from Image Focus", PAMI, vol 17,no 3, 1995.

[4] Subbarao M., Lu M. C., "Image-sensing Model and Computer Simulation for CCD Camera systems", Machine Vision and Applications, vol. 7 pp. 277-289,

[5] Ahmad M.B.,choi T.S., "A heuristic approach for finding best focused shape" IEEE trans. CSVT, vol. 15, No 4, pp 566-574,april 2005.

[6] Aamir malik and Tae-sun choi, "Consideration of illumination effects and optimization of window size for accurate calculation of depth map for 3D shape recovery" pattern recognition, vol. 40,no.1,pp. 154-170,2007.

[7] Arifin A.Z., Asano A., "Image segmentation by histogram thresholding using hierarchical cluster analsys", pattern recognition letters 27, pp.1515-1521, 2006.

[8] A. S. Malik and T. S. Choi, "Application of Passive Techniques for Three Dimensional Cameras," *IEEE Transactions on Consumer Electronics,* vol. 53, pp. 258-264, 2007.

[9] T. Aydin and Y. S. Akgul, "An occlusion insensitive adaptive focus measurement method," *Opt. Express,* vol. 18, pp. 14212-14224, 2010.

[10] A. Thelen*, et al.*, "Improvements in Shape-From-Focus for Holographic Reconstructions With Regard to Focus Operators, Neighborhood-Size, and Height Value Interpolation," *IEEE Transactions on Image Processing,* vol. 18, pp. 151-157, 2009.

[11] A. S. Malik and T. S. Choi, "Consideration of illumination effects and optimization of window size for accurate calculation of depth map for 3D shape recovery," *Pattern Recognition,* vol. 40, pp. 154-170, 2007.

[12] S. K. Nayar and Y. Nakagawa, "Shape from focus," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol.16, pp. 824-831, 1994.

[13] S.O. Shim and T. S. Choi, "Depth from focus based on combinatorial optimization," *Opt. Lett.,* vol. 35, pp. 1956-1958, 2010.

[14] M. T. Mahmood and T. S. Choi, "Focus measure based on the energy of high-frequency components in the S transform," *Opt. Lett.,* vol. 35, pp. 1272-1274, 2010.

[15] M. B. Ahmad and T. S. Choi, "Application of Three Dimensional Shape from Image Focus in LCD/TFT Displays Manufacturing," *IEEE Transactions on Consumer Electronics,* vol. 53, pp. 1-4, 2007.

[16] J. Weickert, *Anisotropic Diffusion in Image Processing*: B. G. Teubner Stuttgart, 1998.

[17] P. Saint-Marc*, et al.*, "Adaptive smoothing: a general tool for early vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 13, pp. 514-529, 1991