Shape from focus through Laplacian using 3D window

Youngeun An[1], Gwangwon Kang[1], Il-Jung Kim[2], Hyun-Sook Chung[1], and Jongan Park[1]
[1]*Chosun University, South Korea*
[2]*Korea Electronics Technology Institute, South Korea*
japark@chosun.ac.kr

Abstract

One of the fundamental objectives of computer vision is to reconstruct a three-dimensional (3D) structure of objects from two-dimensional (2D) images. The basic idea of image focus is that objects at different distances from a lens are focused at different distances. Shape from Focus (SFF) is the problem of reconstructing the depth of the scene changing actively the optics of the camera until the point of interest is in focus. The point in focus gives information about its depth through the thin lens Gaussian law. An effective focus measure operator should be a high-pass filter. Usually, the variation of frequency components are not enough that focus measure could be computed pixel-wise, therefore, sum of pixels in small 2D windows are used for detecting the high frequency components. In this paper, we propose to use 3D windows instead of 2D windows for detecting the high frequency components in the images. The proposed algorithm using 3D window gives better depth map than the previous algorithms using 2D windows.

1. Introduction

One of the fundamental objectives of computer vision is to reconstruct a three-dimensional (3D) structure of objects from two-dimensional (2D) images. The 3D depth maps can be used for 3D features extraction, range segmentation, and object distances from camera in image sequences. The 3D depth map is used for checking the solder volume of printed circuit boards, the 3D shape of the microbiological species, etc.

In literature, there are two types of techniques based on the image focus analysis: Shape from focus (SFF) [1]-[9], and Shape from defocus (SFD) [10]-[13]. SFF is a search method which searches the camera parameters that correspond to focusing the object. SFF uses a sequence of images taken by changing the focus setting of the imaging optics in small steps. For each pixel, the focus setting that maximizes image contrast is determined. This, in turn, can be used to compute the depth of the corresponding scene point.

Most previous research on Shape From Focus (SFF) concentrated on the developments and evaluations of different focus measures [2]-[9]. From the analysis of defocused image [2], it is shown that the defocusing is a LFP, and hence, focus measure should respond to high frequency variations of image intensity and produce maximum values when the image is perfectly focused. Therefore, most of the focus measure in the literature somehow maximizes the high frequency variations in the images. The common focus measure in the literature are; maximize high frequency energy in the power spectrum using FFT [2], variance of image gray levels [5], image gradient [4], energy of Laplacian [2], Modified Laplacian [7], Sum-Modulus-Difference [5], etc. There are other focus measures based on moments [14], [15], wavelet [16], etc.

2. Shape from focus

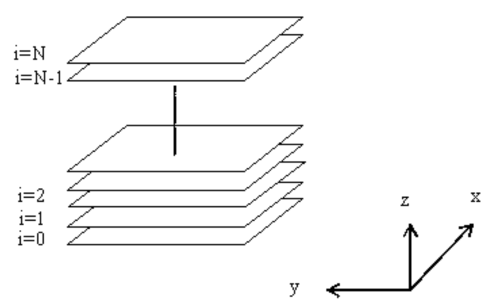


Figure 1. Sequence of images recorded with small interval of lens

In SFF, a sequence of images is obtained by continuously varying the distance between the lens and the image detector as shown in Fig. 1. At each pixel, the image frame which gives the maximum sharpness measure is determined as shown in Fig. 2. The whole image sequence can be viewed as image volume $V_{i,x,y}$

978-0-7695-3431-2/08 $25.00 © 2008 IEEE
DOI 10.1109/FGCN.2008.139

as shown in figure 3, where x, y, and i denotes the number of rows, columns and image frames respectively. For each image in the sequence, focus measure is computed at each pixel and focus measure volume $O_{i,x,y}$ is obtained as shown in figure 4. Sum of Focus Measure (SFM) in 2D window (about 5x5) around the pixel is used as focus measure. From SFM, the image frame among the image sequence that gives a maximum sharpness measure is deter-mined. The gray level (which is proportional to image irradiance) of the pixel in the image frame thus determined is taken to be the gray level value of the focused image for that pixel. The camera parameter values for this image frame are used to compute the distance of the object point corresponding to that pixel.

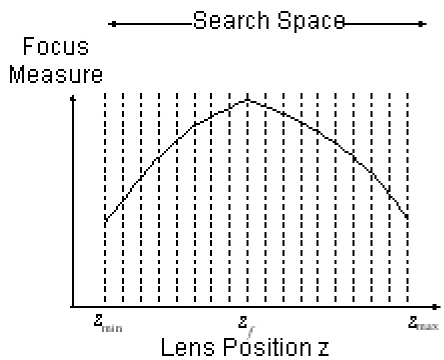


Figure 2. Position of sharpest image frame

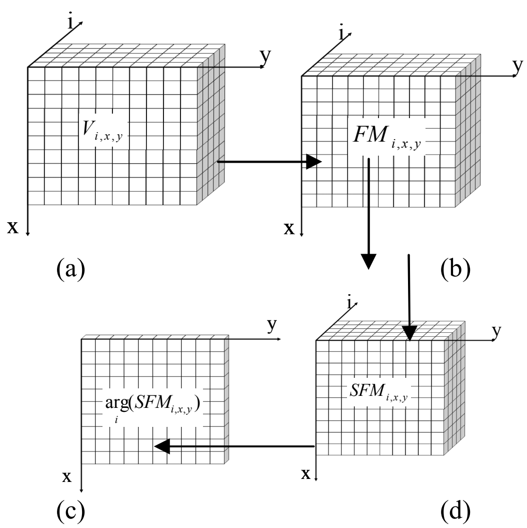


Figure 3. Basic diagram of SFF (a) Input Image Sequence (b) Focus Measure Image Volume (c) Sum of Focus Measure Volume (d) Final Depth Map

2.1. Focus measure operator

From the analysis of defocused image [2], it is shown that the defocusing is a low-pass filtering, and hence, focus measure should respond to high frequency variations of image intensity and produce maximum values when the image is perfectly focused. Therefore, most of the focus measure in the literature somehow maximizes the high frequency variations in the images. The most common focus measure operators are described below.

2.1.1. Gradient magnitude maximization. Since the quality of focus affects edge characteristics, it is natural to use an edge detector for computing the quality of focus. The thresholded gradient magnitude scheme described below has been investigated by Tenenbaum [4], and Schlag et al. [3], who call it the Tenengrad. The gradient of one point does not give any meaning. But the sum of the gradient in a given area called window can be used as a focus measure. So the focus measure using the gradient of an image in a given window can be computed as:

$$\sum_x \sum_y S(x,y)^2 \qquad for \quad S(x,y) \geq T \qquad \ldots(1)$$

where T is a threshold.

2.1.2. High pass filtering. One technique for passing the high spatial frequencies is to determine its second derivative, such as Laplacian, given as

$$\nabla^2 I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \qquad \ldots(2)$$

The Laplacian masks of 4-neigbourhoods and 8-neighborhoods are given in Fig. 4.

0	-1	0
-1	4	-1
0	-1	0

4-neigbourhoods

-1	-1	-1
-1	8	-1
-1	-1	-1

8-neigbourhoods

Figure 4. Laplacian masks

Laplacian is computed for each pixel of the given image window and the criterion function can be stated as:

$$\sum_x \sum_y \nabla^2 I(x,y) \qquad for \quad \nabla^2 I(x,y) \geq T \quad \ldots(3)$$

Nayar [7] noted that in the case of the Laplacian the second derivatives in the x and y directions can have opposite signs and tend to can-cel each other. He, therefore, proposed the Modified Laplacian (ML) as:

$$\nabla_M^2 I = \left|\frac{\partial^2 I}{\partial x^2}\right| + \left|\frac{\partial^2 I}{\partial y^2}\right| \qquad \ldots(4)$$

The discrete approximation to the Laplacian is usually a 3 x 3 operator. In order to accommodate for possible variations in the size of texture elements, Nayar computed the partial derivatives by using a variable spacing (step) between the pixels used to compute the derivatives. He proposed the discrete approximation of the ML as:

$$\nabla^2_{ML} I(x,y) = \left|2I(x,y) - I(x-step,y) - I(x+step,y)\right| + \left|2I(x,y) - I(x,y-step) - I(x,y+step)\right| \quad \ldots(5)$$

Finally, the depth map or the focus measure at a point (x,y) was computed as the sum of ML values, in a small window around (x,y), that are greater than a threshold value T:

$$F(x,y) = \sum_{i=x-N}^{i=x+N} \sum_{j=y-N}^{j=y+N} \nabla^2_{ML} I(i,j) \quad \ldots(6)$$

$$for \quad \nabla^2_{ML} I(i,j) \geq T$$

The parameter N determines the window size used to compute the focus measure. Nayar referred the above focus measure as the Sum-Modified-Laplacian (SML).

2.1.3. Grey level variance. Intuitively, high grey-level variance is associated with sharp image structure while low variance is associated with blurring, which reduces the amount of grey-level fluctuation. The variance of an N x N image is defined as

$$\sigma^2 = \frac{1}{N^2} \sum_{x=1}^{N} \sum_{y=1}^{N} [I(x,y) - \mu]^2 \quad \ldots(7)$$

where μ is the mean of the grey-level distribution within the image. The criterion is to maximize σ^2, which indirectly corresponds to maximizing the integral of the power spectrum of the intensity distribution, since the sum of the squared intensities is related by Parseval's Theorem to the energy in the power spectrum. One problem with the variance measure is that it depends only upon the probability distribution of the grey-levels, and not on its spatial distribution.

3. SFF using 3D window

We discuss few focus measure operators from the literature in section 2. From equations (1) and (6), we see that focus measures are summed in small 2D windows. In general, the SFF methods in the literature do not yield accurate shape or depth-map of objects. The main reason for this is that the focus measures in these methods are defined and computed over image frames. The focus measures at each pixel in the image frame are computed using a small 2D window around the pixel.

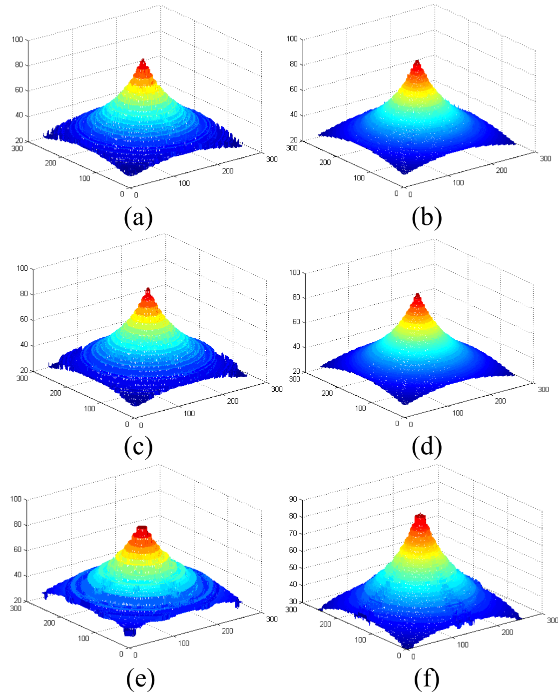


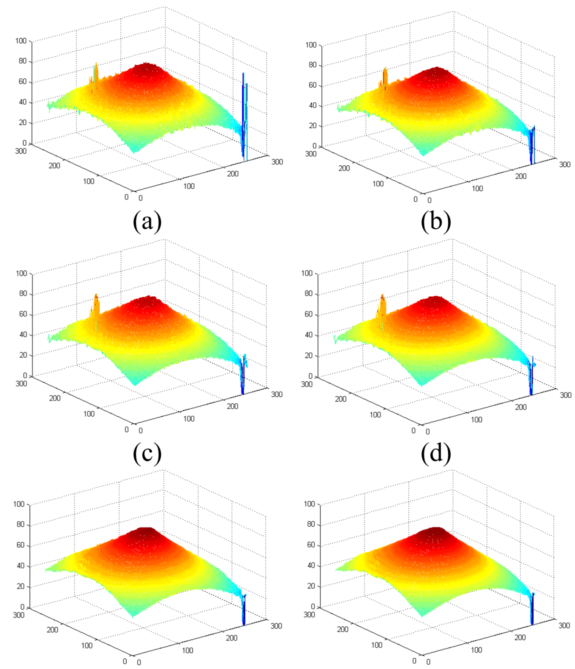
Figure 5. 3D shape recovery of the Simulated cone using SFF with Focus Measure (a) Gradient Magnitude Maximization with 2D window (b) with 3D window (c) Modified Laplacian with 2D window (d) with 3D window (e) Gray Level Variance with 2D window (f) with 3D window

At each pixel, the image frame among the image sequence that gives a maximum sharpness measure is determined. This corresponds to a piecewise constant approximation of the object shape in the window. Because of this approximation, the focused image reconstructed from the image sequence is an approximation of the actual focused image. The depth maps can be improved if 3D windows around the pixels are used.

In this paper, we propose to use the sum defined in equations (1) and (6) over 3D windows instead of 2D windows. As we have sequence of images with small step distance among the frames, 3D windows can be easily created. The three axes of the 3D windows are: rows, columns and image frames. Equations (1) and (6) are modified and expressed as:

$$F_{x,y} = \frac{1}{IXY} \sum_{p=i-N}^{p=i+N} \sum_{m=x-N}^{m=x+N} \sum_{n=y-N}^{x=y+N} O_{p,m,n} \quad for \quad O_{p,m,n} \geq T \quad \ldots(8)$$

where T represents the thereshold. The total number of rows, columns, and images are represented by X, Y, and I, respectively. Previously, sums were taken in single image frames, but now, sums are taken in small number of image frames around the pixel.



(a) (b) (c) (d) (e) (f)

Figure 6. 3D shape recovery of the Real cone using SFF with Focus Measure (a) Gradient Magnitude Maximization with 2D window (b) with 3D window (c) Modified Laplacian with 2D window (d) with 3D window (e) Gray Level Variance with 2D window (f) with 3D window

The surface of simulation and real cones should be very smooth and the tips should be very sharp. The results of SFF algorithms using 2D and 3D windows are shown in Fig. 5 and 6. We see clearly from the Fig. 5 and 6 (b) (d) and (f), that SFF using 3D window gives smoother surface and sharper tips of the cones than using 2D windows Fig. 5 and 6 (a) (c) and (e), respectively.

4. Conclusion

In this paper, 3D shape recovery from image focus is discussed. Previous algorithms in the literature use 2D windows for focus measure. We propose the use of 3D windows for focus measure for better depth maps.

5. Acknowledgement

This study was supported by Ministry of Culture & tourism and Culture & Content Agency in Republic of Korea.

6. References

[1] B.K.P. Horn, "Focusing," MIT Artificial Intelligence Laboratory, Memo no 160, May 1968.

[2] E. Krotkov, "Focusing," International Journal of Computer Vision, Vol. 1 pp. 223-237, 1987.

[3] J.F. Schlag, A.C. Sanderson, C.P. Neumann, F.C. Wimberly, "Implementation of Automatic Focusing Algorithms for a Computer Vision System with Camera Control," Carnegie Mellon University, CMU-RI-TR-83-14, Aug. 1983.

[4] J.M. Tenenbaum, "Accommodation in Computer Vision," Ph.D. dissertation, Standford University, 1970.

[5] R.A. Jarvis, "A Perspective on Range Finding Techniques for Computer Vision," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 5, no. 2, Mar. 1983.

[6] J. Ens, and P. Lawrence, "An Investigation of Methods for Determining Depth from Focus," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 15, no. 2, pp. 97-108, Feb. 1993.

[7] Shree K. Nayar and Yasuo Nakagawa, "Shape from focus," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 16, No. 8, August 1994.

[8] M. Subbarao and J.K. Tyan, "Selecting the Optimal Focus Measure for Autofocusing and Depth-from-Focus," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 20, no. 8, pp. 864-870, Aug. 1998.

[9] T. Darrell, and K. Wohn, "Pyramid Based Depth from Focus," Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition, pp. 504-509, Jun. 1988.

[10] M. Subbarao, and G. Surya, "Depth from Defocus: A Spatial Domain Approach," J. Computer Vision, vol. 13, no. 3, pp. 271-294, Dec. 1994.

[11] A. Pentland, "A New Sense for Depth of Field," IEEE Trans. Pattern Analysis and Machine In-telligence, vol. 9, no. 4, pp. 523-531, Jul. 1987.

[12] Y.Y. Schechner, and N. Kiryati, "Depth from Defocus vs. Stereo: How Difference Really are They?," J. Computer Vision, vol. 39, no. 2, pp. 141-162, Sep. 2000.

[13] P. Favaro, S. Soatto, "A Geometric Approach to Shape from Defocus," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 27, no. 3, pp. 406-417, Mar. 2005.

[14] Y. Zhang, Ying Zhang and C. Wen, "A New Focus Measure Method using Moments," Image and Vision Computing, vol. 18, no. 12, pp. 959-965, Sep. 2000.

[15] P.T. Yap, and P. Raveendran, "Image Focus Measure Based on Chebyshev Moments," IEE Proc. Vision, Image and Signal Processing, vol. 151, no. 2, pp. 128-136, Apr. 2004.

[16] J. Kautsky, J. Flusser, B. Zitova and S. Simberova, “A New Wavelet-Based Measure of Image Focus,” Pattern Recognition Letters, vol. 23, no. 14, pp. 1785-1794, Dec. 2002.