International Journal of Computer and Information Technology (ISSN: 2279 – 0764)
Volume 04 – Issue 05, September 2015

Accurate Depth Map Computation from Image Focus by Equifocal Plane Approximation

Seong-O Shim
Computer Science Department, Faculty of Computing and IT
University of Jeddah
Jeddah, Saudi Arabia
Email: sshim [AT] uj.edu.sa

Abstract—**In this paper, a new depth map computation method from image focus based on equifocal plane approximation is proposed. Shape from focus (SFF) is a technique to compute depth map from multiple images of same scene obtained from different focus settings of camera. In SFF, measuring accurate focus value of each pixel in images is most crucial part in final quality of depth map. Traditional SFF techniques compute focus value of a pixel from two-dimensional local neighborhood. However, in optics with small depth of field, local neighborhood pixels have different level of focus and it lead to inaccurate focus value computation. To overcome this problem, we try to search local equifocal plane on each pixel where the neighboring pixels on this plane have same level of focus. Experimental results show the proposed method produces more accurate depth map in comparison to previous SFF methods.**

Keywords-component; depth map; shape from focus (SFF); image focus; equifocal plane

I. INTRODUCTION

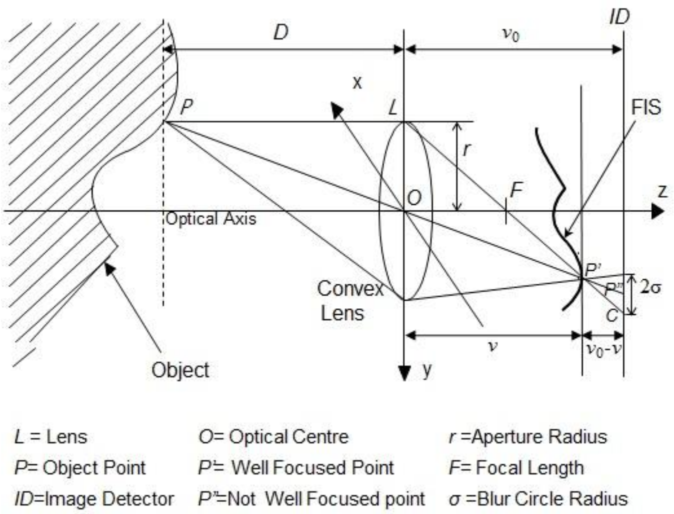


Figure 1. Image formation of 3D object point in convex lens.

Computing depth map of an object from multiple two-dimensional (2D) images is an important task in computer vision field. Among the variety of depth map estimation methods, shape from focus (SFF) [1][2][3][4][5][6][7] and shape from defocus (SFD) [8][9][10][11] try to recover depth map from multiple images obtained by different focus settings of the camera.

Fig. 1 shows the image formation geometry for an optical system with convex lens of focal length F . Our aim is to compute the depth D of each object point P from the lens. If image detector (ID) is located at P' , object point P will be appeared in the image as sharply focused point, and as image detector moves away from the point P' , it will be appeared in the image as a blurred circle with radius σ . The relationship between the point P in object space and its focused point P' in image space follows the famous thin lens law as $1/v+1/D=1/F$. And from the geometry of two similar triangles ($\Delta P'OL$ and $\Delta P'P''C$) in Fig. 1, we can obtain the equation $r/v=\sigma/(v_0-v)$. By combining these two equations, we can obtain:

$$D=\frac{Fv_0}{v_0-F-2\sigma f} \tag{1}$$

where f is the f-number ($F/2r$) of the lens system. In SFD, the depth D is computed by measuring the blur circle radius σ (which is the only unknown parameter in (1)) from one or two images. Whereas, SFF techniques estimate the depth D by searching image detector position v_0 where the blur circle radius approaches to zero ($\sigma\approx 0$). Then, the depth D is determined by already known parameters by putting the value of σ to zero in (1) as $D=Fv_0/(v_0-F)$.

In SFF, a stack of images (3D image volume) of an object is obtained by capturing images at each discrete position of image detector along z axis (optical axis). If magnification effect is assumed to be minimized, each point $(x,y,\cdot)$ in image space is gradually focused until it reaches maximum and then gradually defocused along optical axis. To measure the degree of focus, a focus measure operator [12][13][14][15][16][17] is applied on small regions of every pixel in the image volume and focus measure volume $FM(x,y,z)$ is made. Then, at each point (x, y), the depth $D(x, y)$ is determined by searching z position where $FM(x, y, z)$ is maximized as follows:

$$D(x,y) = \frac{F \cdot v_0^{(x,y)}}{v_0^{(x,y)} - F}, \quad v_0^{(x,y)} = \arg\max_z FM(x,y,z) \quad (2)$$

The traditional SFF techniques apply a focus measure operator on small two-dimensional window (local neighborhood) at each pixel. However, small neighboring pixels, taken with small depth of focus setting, have different level of focus. In this case, the computed focus value does not represent accurate focus level of the center pixel. To overcome this issue, Subbarao and Choi computed focus value from neighboring pixels lying on Focused Image Surface (FIS) [1] that is the image surface formed from best-focused points in image space as depicted in Fig. 1. Later, Ahmad and Choi improved the accuracy by searching optimal FIS shape based on dynamic programming [4]. In this paper, we propose a new SFF technique based on equifocal plane approximation. To avoid heavy computation in FIS based methods, we search tangential plane at each center pixel which is a good local approximation of a plane where the pixels are equally focused (equifocal plane). Computing the focus value of a pixel on neighborhood pixels lying on the searched plane could produce depth that is more accurate.

II. SFF BASED ON EQUIFOCAL APPROXIMATION

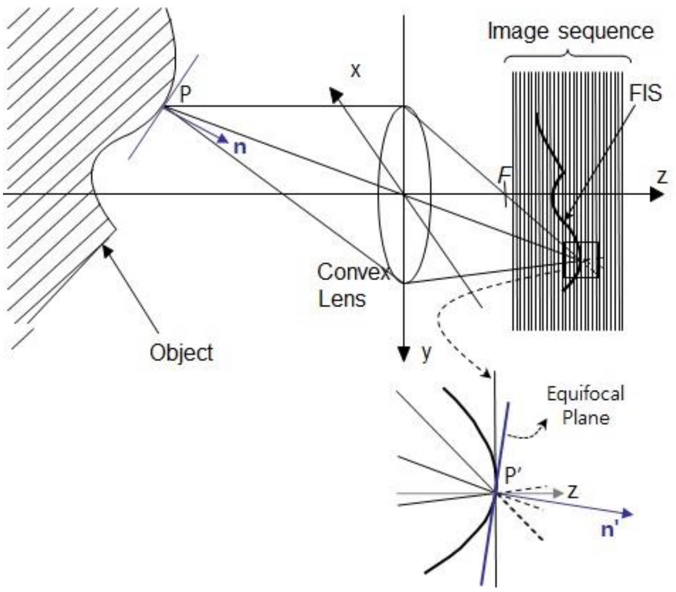


Figure 2. Equifocal plane of an object point P in image space.

A. Motivation

In FIS based method [1], first the initial FIS is formed by applying conventional SFF method. Then, at each pixel on FIS, the pixels on 3D neighborhood which produce maximum sum of focus values are searched by varying the slope of local initial FIS plane along each x, y, and z axes. However, this search process needs tremendous amount of computation and it is dependent on inaccurate neighboring focus values computed on original 2D image. In this paper, we approximate the local equifocal plane by searching tangential plane at each pixel in 3D image volume. In Fig. 2, the tangential plane on object point P has corresponding plane in image space. At each pixel, the small neighborhood pixels lying on this plane represent more equal level of focus among each other. Therefore, applying focus measure operator on these new neighborhood pixels can produce more accurate focus value of a center pixel. To find the tangential plane of each pixel, at each pixel $(x, y, \cdot)$, principal axis (which corresponds to its surface normal vector) is found from small neighborhood volume along which maximum focus changes occur.

B. Procedure

First, a small neighborhood $\mathrm{N}_{(x,y,z)}$ at each pixel (x, y, z) is defined by pixels whose distance to (x, y, z) is less than predefined value from the original image volume $I(x, y, z)$ as:

$$\mathrm{N}_{(x,y,z)} = \{I(p,q,r) \mid \operatorname{dist}((x,y,z),(p,q,r)) \le r\} \quad (3)$$

where the threshold value r is set to $\sqrt{2}$. Then, all the n elements of $\mathrm{N}_{(x,y,z)}$ are arranged into a column vector $\tilde{\mathrm{N}}_{(x,y,z)}$. By augmenting column vectors $\tilde{\mathrm{N}}_{(x,y,z)}$ at all the N image frames along z axis, we construct $n \times N$ matrix $\mathbf{M}_{(x,y)}$ as:

$$\begin{aligned} \mathbf{M}_{(x,y)} &= [m_{ij}]_{n \times N} = [\mathbf{m}_1 \ \mathbf{m}_2 \cdots \mathbf{m}_k \ \cdots \mathbf{m}_{N-1} \ \mathbf{m}_N] \\ \mathbf{m}_k &= \tilde{\mathrm{N}}_{(x,y,k)} \end{aligned} \quad (4)$$

Then, the covariance matrix $\mathbf{C_M}$ of $\mathbf{M}$ is defined as:

$$\begin{aligned} \mathbf{C_M} &= \frac{1}{N-1} \hat{\mathbf{M}} \hat{\mathbf{M}}^T, \\ \hat{\mathbf{M}} &= [\mathbf{m}_1 - \bar{\mathbf{m}} \cdots \mathbf{m}_k - \bar{\mathbf{m}} \cdots \mathbf{m}_N - \bar{\mathbf{m}}], \\ \bar{\mathbf{m}} &= \frac{1}{N} \sum_{k=1}^{N} \mathbf{m}_k \end{aligned} \quad (5)$$

By sorting all the eigenvectors $\mathbf{e}_i$ ($i = 1, 2, \ldots, n$) of $\mathbf{C_M}$ in a way that their corresponding eigenvalues $\lambda_i \ge \lambda_{i+1}$, eigenvector matrix $\mathbf{E}$ is defined as:

$$\mathbf{E} = [\mathbf{e}_1 \, \mathbf{e}_2 \ldots \mathbf{e}_{n-1} \, \mathbf{e}_n]^T \quad (6)$$

Then, the original data $\hat{\mathbf{M}}_{(x,y)}$ in (5) can be transformed to eigenspace data $\mathbf{F}$ as:

$$\mathbf{F}_{(x,y)} = [f_{ij}]_{n \times N} = \mathbf{E}_{(x,y)} \hat{\mathbf{M}}_{(x,y)} = [\mathbf{f}_1 \, \mathbf{f}_2 \cdots \mathbf{f}_k \cdots \mathbf{f}_{N-1} \, \mathbf{f}_N] \quad (7)$$

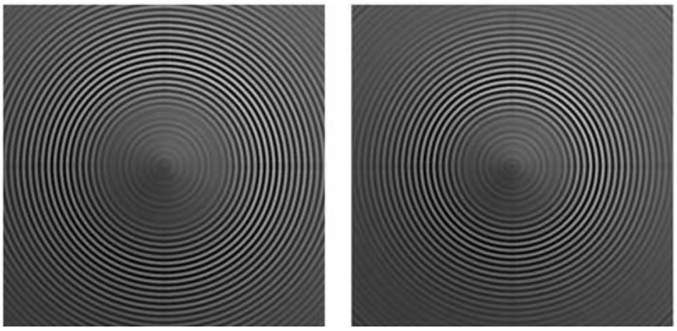
Each vector $\mathbf{m}_k$ in (4) is transformed to $\mathbf{f}_k$ in (7), and each row vector $\mathbf{e}_i$ of matrix $\mathbf{E}$ corresponds to the basis vector of i^{th} axis in eigenspace. Since, in eigenspace, the maximum variation of data occurs along the principal axis, its basis vector $\mathbf{e}_1$ can be considered as surface normal vector $\mathbf{n}'$ of equifocal plane in Fig. 2. Therefore, the first row in $\mathbf{F}_{(x,y)}$ can be assumed to be the data in eigenspace from the original data lying on equifocal plane at $(x, y, \cdot)$. The image detector position $v_0^{(x,y)}$ at (x, y) can be found from the position in first row in $\mathbf{F}_{(x,y)}$ that shows maximum value as:

$$v_0^{(x,y)} = \arg\max_k |f_{1k}|, \quad 1 \le k \le N \tag{8}$$

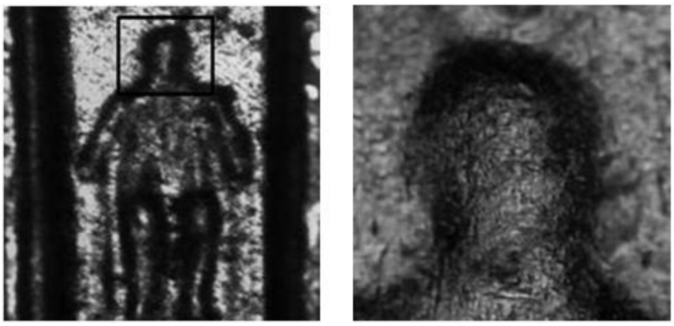
The final depth D at (x, y) is computed from $v_0^{(x,y)}$ as in (2).

III. Experiment and Discussion

A. Experimental Data



(a)



(b)

Figure 3. Test samples. (a) Simulated cone at different lens steps. (b) Lincoln head part of US 1 cent coin.

Experiments were conducted both on synthetic and real object. The sample images of these two objects are shown in Fig. 3. Synthetic object is a simulated cone object where true depth map information is already known. A total of 97 images of the simulated cone object was generated by the camera simulation software corresponding to 97 equally spaced lens positions. The real world object is the head part of the Lincoln statue on US one cent coin. Images of this real object were obtained at different lens positions from microscope system [4]. As shown in Fig. 4, the system consists of the optical microscope integrated with stepper motor control unit, frame grabber board with PC, and CCD camera mounted on ocular tube of the microscope. A software is developed to obtain images at each lens position by moving the stage through step motor driver having minimum 2.5 *nm* step length. A total of 60 images of the object was acquired by the system under $\times 100$ magnification with lens step size of 0.52 μm.

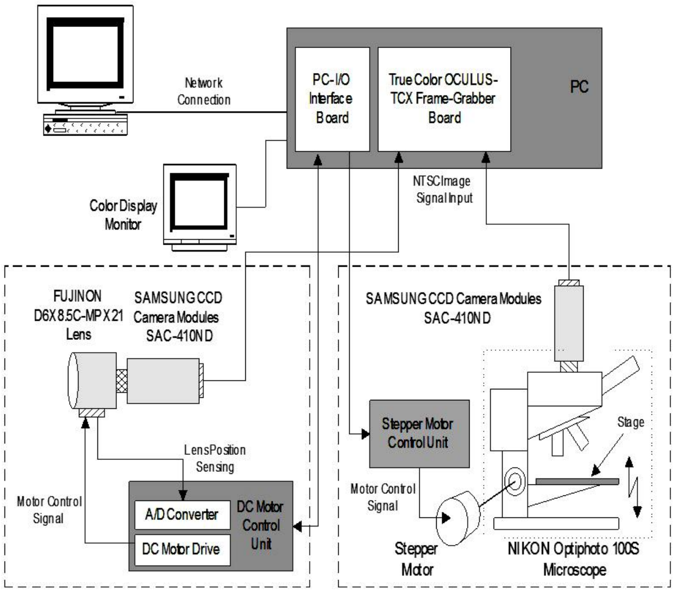


Figure 4. Automatic microscope imaging system.

B. Experiments and Analysis

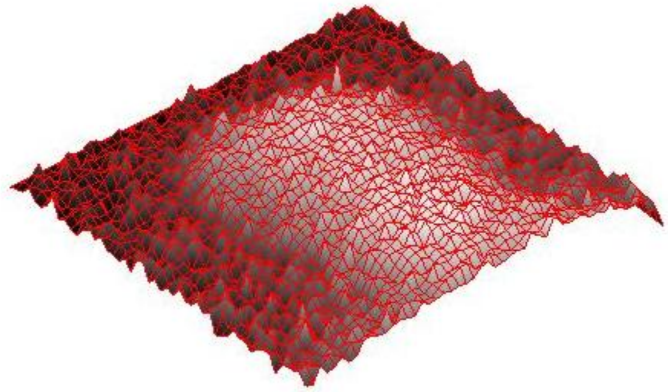
The proposed method was compared with previous SFF techniques based on FIS (SFF-FIS) [1], and based on dynamic programming (SFF-DP) [4] which were described in introduction section.

First, for quantitative analysis, RMSE value were compared between the true depth map of synthetic object and its computed depth maps obtained from SFF-FIS, SFF-DP, and the proposed method. The lower the RMSE value, the more the reconstructed depth map close to the actual depth map. As shown in Table I, the proposed method produced less RMSE values in comparison to SFF-FIS and SFF-DP.

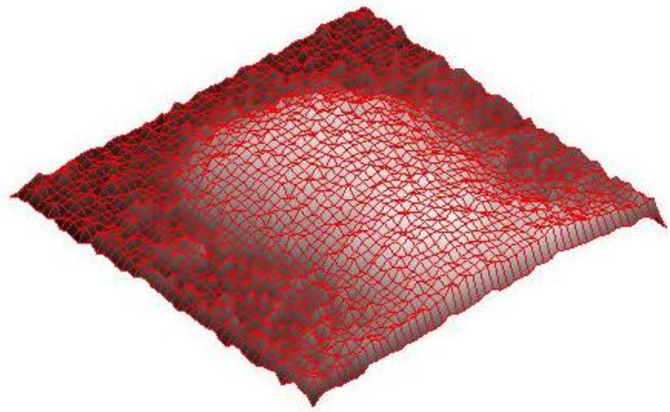
TABLE I. RMSE Values Between True Depth Map and Computed Depth Map on Simulated Cone

	SFF-FIS	SFF-DP	Proposed
RMSE	7.9132	7.7472	7.4495

In Fig. 5, for qualitative analysis, the shape recovery results from SFF-FIS and the proposed method were compared. The result from the proposed method generated smoother surface with less noise on overall surface area compared to the result from SFF-FIS.



(b)



(b)

Figure 5. Reconstructed shapes on Lincoln head part of US one cent coin. (a) SFF-FIS. (b) Proposed.

IV. CONCLUSION

In this paper, a new SFF method based on equifocal plane approximation was proposed. Contrary to computing focus value of a pixel on 2D local neighborhood, we try to find a equifocal plane at each pixel in image volume where the focus level of local neighborhood pixels lying on the plane is similar. To find the equifocal plane at each $(x, y, \cdot)$, a surface normal vector (basis vector corresponding to principal axis) is found in eigenspace. At each $(x, y, \cdot)$, the depth value is determined from the position along the principal axis in its transformed data into eigenspace that shows maximum value. Experimental results on synthetic object and real object showed that the proposed technique produces more accurate depth map compared to the existing methods.

REFERENCES

[1] M. Subbarao and T. S. Choi, “Accurate Recovery of Three-Dimensional Shape from Image Focus,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, no. 3, pp. 266-274, March 1995.

[2] E. Krotkov, “Focusing,” International Journal of Computer Vision, vol. 1, no. 3, pp. 223-237, 1987.

[3] S. K. Nayar and Y. Nakagawa, “Shape from focus,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no. 8, pp. 824-831, August 1994.

[4] M. B. Ahmad and T. S. Choi, “A Heuristic approach for finding best focused shape”, IEEE Transsactions on Circuits and Systems for Video Technology, vol. 15, no. 4, pp. 566-574, April 2005.

[5] S. O. Shim and T. S. Choi, “A novel iterative shape from focus algorithm based on combinatorial optimization,” Pattern Recognition, vol. 43, no. 10, pp. 3338-3347, October 2010.

[6] C. Y. Tseng and S. J. Wang, “Shape-From-Focus Depth Reconstruction With a Spatial Consistency Model,” IEEE Transsactions on Circuits and Systems for Video Technology, vol. 24, no. 12, pp. 2063-2076, September 2014.

[7] M. T. Mahmood, S. O. Shim, S. Alshomrani, and T. S. Choi, “Depth from image focus methods for micro-manufacturing,” The International Journal of Advanced Manufacturing Technology, vol. 6, no. 5, pp. 1701-1709, 2013.

[8] A. P. Pentland, “A new sense for depth of field,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 9, no. 4, pp. 523-531, July 1987

[9] A. N. Rajagopalan, S. Chaudhuri, and U. Mundenagudi, “Depth estimation and image restoration using defocused stereo pairs,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 26, no. 11, pp.1521-1525, November 2004.

[10] P. Favaro and S. Soatto, “A geometric approach to shape from defocus,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 3, pp. 406-417, March 2005.

[11] P. Favaro, S. Soatto, M. Burger, and S. J. Osher, “Shape from defocus via diffusion”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 3, pp. 518-531, March 2008

[12] Y. Sun, S. Duthaler, and B. J. Nelson, “Autofocusing in computer microscopy: Selecting the optical focus algorithm,” Microscopy Research and Technique, vol. 65, no. 3, pp. 139-149, 2004.

[13] J. M. Tenenbaum, “Accomodation in computer vision,” Ph. D. dissertation, Standford University, Stanford, CA, 1970.

[14] I. H. Lee, S. O. Shim, and T. S. Choi, “Improving focus measurement via variable window shape on surface radiance distribution for 3D shape reconstruction,” Optics and Lasers in Engineering, vol. 51, no. 5, pp. 520-526, 2013.

[15] S. Pertuz, D. Puig, M. A. Garcia, “Analysis of focus measure operators for shape-from-focus,” Pattern Recognition, vol. 46, no. 5, pp. 1415-1432, 2013.

[16] F. Zhong, Y. Ma, and H. Li, “Multifocus image fusion using focus measure of fractional differential and NSCT,” Pattern Recognition and Image Analysis, vol. 24, no. 2, pp. 234-242, June 2014.

[17] I. H. Lee, M. T. Mahmood, and T. S. Choi, “Robust Focus Measure Operator Using Adaptive Log-Polar Mapping for Three-Dimensional Shape Recovery,” Microscopy and Microanalysis, vol. 21, no. 2, pp. 1-17, 2015.