IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 7, 2019, accepted January 31, 2019, date of publication February 5, 2019, date of current version February 22, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2897744

Improving Depth Computation From Robust Focus Approximation

SEONG-O SHIM[1], WAJID AZIZ[1,2], AMEEN BANJAR[1], ABDULLAH ALAMRI[1], AND MOHAMMED ALQARNI[1]

[1]College of Computer Sciences and Engineering, University of Jeddah, Jeddah 23890, Saudi Arabia
[2]Department of Computer Science, University of Azad Jammu & Kashmir, Muzafarabad 13100, Pakistan

Corresponding author: Seong-O Shim (sshim@uj.edu.sa)

This work was supported by the Deanship of Scientific Research (DSR), University of Jeddah, Jeddah, under Grant UJ-23-18-DR.

ABSTRACT Shape-from-focus (SFF) technique is to recover three-dimensional shape of an object from a sequence of two-dimensional images of the same object taken by gradually changing the focus settings of the imaging device. In SFF, for each pixel location (i, j), a focus measure operator computes sharpness (focus) value on a small neighborhood of the pixel (i, j) at each image along the z-axis (optical axis). The image sensor position z that has the maximum focus value and camera parameter information provide the distance information between the lens and the object point corresponding to the pixel (i, j). In traditional SFF methods, the final focus value of each pixel is determined by an average value of the initial focus values at the neighborhood of the pixel to remove the noise effects. However, it only reduces the noise effects instead of completely removing it. In this paper, we proposed a noise filtering technique that tries to remove noises while computing the focus values. First, an initial focus measure volume is computed by applying one of the focus measures on each pixel in the image sequence. Then, the focus value at each pixel is examined whether it is corrupted by noise or not by analyzing the neighboring focus values. The noisy focus value of the pixel is re-computed from noise-free focus values of neighboring pixels. The experimental results conducted on both the synthetic and real-world objects show the proposed method produces the better three dimensional shape in comparison to the existing methods.

INDEX TERMS Focus measure operator, focus value, shape-from-focus (SFF), three dimensional shape.

I. INTRODUCTION

Computing three-dimensional (3D) structure from multiple two-dimensional (2D) images is an important task in computer vision field. Many different types of 3D shape estimation techniques exist, and mainly they can be classified into contact, transmissive and reflective methods [1]. The contact method probes the subject through physical touch. This method is used mostly in manufacturing since it is very precise, but requires direct contact with the subject and might damage it. The transmissive method is used for the transparent object such as cell which cannot reflect light. The light travels through the object and the information of amplitude and phase changes of the light is used for the shape estimation. The reflective method analyzes the data based on wave particles reflected on the object, and it can be divided into optical and non-optical methods. The optical techniques can be further classified into passive and active methods. In active methods, we project light rays on the object and capture the reflected lights. However, in passive methods, the reflections of natural lights are collected.

The passive techniques can be categorized as shape from X, where X denotes the cue that is used to estimate the shape. The well-known examples of the cue include stereo, shading, structure from motion, texture, focus, defocus, etc. Shape from focus (SFF) recovers the shape from stack of images acquired by gradually changing the camera focus settings [2]–[14].

In SFF, at each position (i, j) in xy image plane, all the pixels lying along the z axis $(i, j, :)$ are collected and focus (sharpness) values on these pixels are compared. The z position of the maximally focused pixel provides depth information at (i, j). By collecting all the depth information at each (i, j), the final depth map of an object is made. The accuracy of the depth map is highly dependent on accurate measure of the focus value at each pixel. To measure the focus value at each pixel, a focus measure operator is applied on each pixel in the image stack and focus measure volume is constituted.

The associate editor coordinating the review of this manuscript and approving it for publication was Peng Liu.

Then, an averaging filter is applied on each pixel in the focus measure volume to reduce the noise effect.

In conventional SFF, the final focus value of each pixel in focus measure volume is determined by averaging the focus values of its neighborhood pixels lying on the same image plane to reduce the noise effect. However, averaging the neighboring focus values only suppresses the noise instead of eliminating it. If most of the neighboring focus values are corrupted by noise, averaging them cannot reduce the noise level of the center pixel's focus value.

In this paper, we try to remove the noise effect on computing the focus value of the pixel. After acquiring initial focus measure volume by applying one of existing focus measures, each focus value in the volume is examined whether it is noisy or noise-free. Then, the noisy focus values are re-computed from their neighboring noise-free focus values. Experimental results of SFF on a synthetic and real objects show that applying the proposed method on existing focus measures improve the quality of final recovered shape.

The rest of the paper is organized as follows: Section II provides a background on shape from focus. Section III presents a detailed description of the proposed method. Section IV evaluates the accuracy of the proposed methods on various objects. Finally, we conclude the paper in Section V.

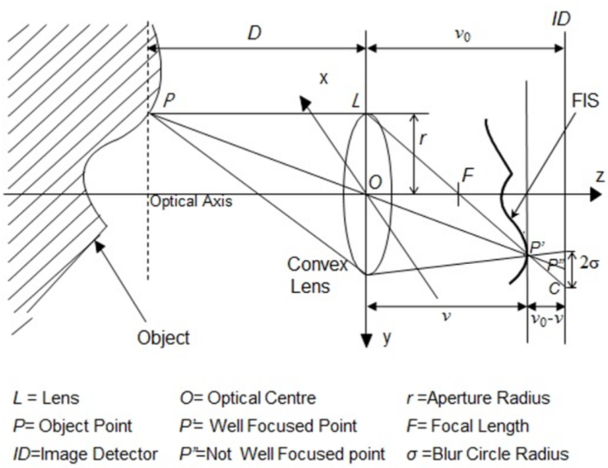

FIGURE 1. Image formation of 3D object point in convex lens.

II. SHAPE FROM FOCUS

The objective of SFF is to find the depth map of an object from multiple images acquired from the object by gradually changing the focus of a camera. The image formation of an object point with convex lens of focal length F is shown in Fig. 1. The aim of SFF is to estimate the depth D of each object point P. In our SFF system, image detector (ID) in Fig. 1 is placed near the lens and images are taken by gradually moving ID away from the lens. While the ID is moving away from the lens, the object point P is gradually focus on image space and attains the maximum focus at P' and again gradually defocused. When the point P is defocused, it is shown as a blurred circle with radius σ. The focus curve of the object point P along the optical axis can be shown as Fig. 2.

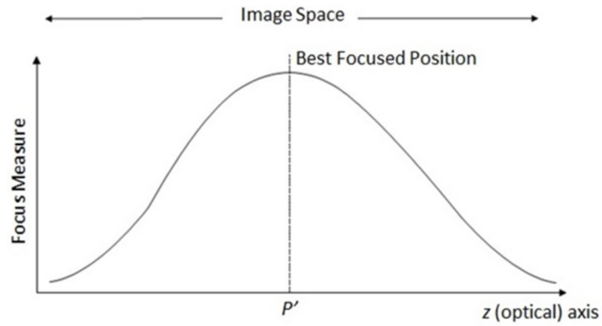

FIGURE 2. Focus curve of point P in image space along the optical axis.

In Fig. 1, we can derive the equation $r/v = \sigma/(v_0 - v)$ from two similar triangles ($\Delta P'OL$ and $\Delta P'P''C$). From thin lens equation, we get $1/v + 1/D = 1/F$. Depth D of the object point P can be obtained by combining these two equations as

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

where f is the f-number ($F/2r$) of the lens system. SFF techniques try to measure depth D by searching image detector position where blur circle radius σ is minimum ($\sigma \approx 0$). At $(i, j, \cdot)$ in image space corresponding to the point (x, y, D) in object space, the position v_0 of image detector where $\sigma \approx 0$ is found by searching the image position along z (optical) axis which produces maximum focus (sharpness) as

$$v_0(x, y) \leftarrow \max \arg_k FM(i, j, k) \tag{2}$$

where k is relative image detector position from the lens along optical axis and $FM(i, j, k)$ is the measure of focus (sharpness) at pixel (i, j, k) obtained by applying one of focus measure operators [15]–[25]. Then, the depth of the point (x, y) is computed as

$$D(x, y) = \frac{F \cdot v_0(x, y)}{v_0(x, y) - F} \tag{3}$$

III. THE PROPOSED METHOD

First, the total of L images of an object is captured by gradually changing the camera focus setting and a focus measure operator is applied on each pixel (i, j, k) to compute the initial focus measure volume $F_0(i, j, k)$. Then, each focus value in $F_0(i, j, k)$ is examined whether it is noise involved or not. First, at each (i, j, k), small three dimensional neighborhood $N_{(i,j,k)}(3 \times 3 \times 3)$ is selected from $F_0(i, j, k)$. Then, the number of pixels n_{ijk} in $N_{(i,j,k)}$ which has similar focus value of the center pixel (i, j, k) is counted as

$$n_{ijk} = \sum_{(i',j',k') \in N_{(i,j,k)}} C(i', j', k')$$

$$C(i', j', k') = \begin{cases} 1 & \text{if } \left|F_0(i, j, k) - F_0(i', j', k')\right| \le T_f \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

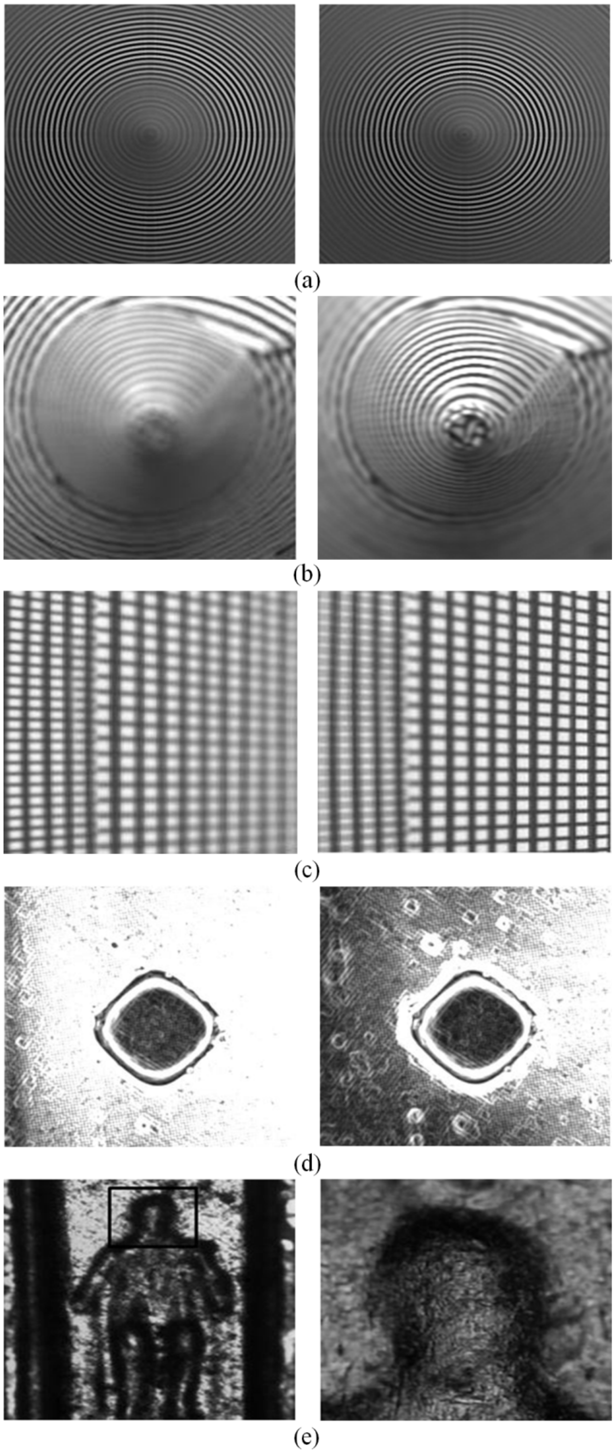

FIGURE 3. Example images from five objects. (a) Simulated cone at different lens steps. (b) Real cone at different lens steps. (c) Planer object at different lens steps. (d) Protrusion of TFT-LCD color filter at different lens steps. (e) Lincoln head part of US one cent coin.

where T_f is the threshold to decide whether two focus values are similar. When the surface of an object is assumed to be smooth, the focus values of neighboring pixels should be similar. If n_{ijk} is small, it means the focus value at (i, j, k) is not similar to the focus values of its neighboring pixels, and it

TABLE 1. RMSE and correlation on simulated cone object based on different focus measures.

	Focus measure	Original	Modified
RMSE	SML	8.0941	**7.8384**
	GLV	8.0886	**7.7316**
	TEN	8.1232	**7.8357**
	DCT	8.0878	**7.7581**
	PCA	8.0482	**9.8192**
Correlation	SML	0.9285	**0.9512**
	GLV	0.9317	**0.9538**
	TEN	0.9314	**0.9532**
	DCT	0.9274	**0.9515**
	PCA	0.9249	**0.9542**

is flagged as noisy. We define the noise factor $\alpha(i, j, k)$ that has value 1 if the focus value at (i, j, k) is noisy and 0 if it is noise-free.

$$\alpha(i, j, k) = \begin{cases} 1 & \text{if } n_{ijk} \le T_n \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

The final focus measure volume $F(i, j, k)$ is determined by re-computing the noisy focus values from the noise-free neighboring focus values as

$$F(i,j,k) = \begin{cases} F_0(i,j,k) & \mathit{if}\ \alpha(i,j,k) = 0 \\ \dfrac{\sum_{(i',j',k')\in N_{(i,j,k)}} \alpha(i',j',k') \cdot F_0(i',j',k')}{\sum_{(i',j',k')\in N_{(i,j,k)}} \alpha(i',j',k')}, & \mathit{if}\ \alpha(i,j,k) = 1 \end{cases} \tag{6}$$

For each pixel (i, j), the image sensor position $v_0(i, j)$ is determined by searching z position where $F(i, j, k)$ is maximized as

$$v_0(i,j) = \arg\max_k F(i,j,k) \tag{7}$$

Finally, the depth of the object point corresponding the pixel location (i, j) is computed as

$$D(i,j) = \frac{Fv_0}{v_0 - F} \tag{8}$$

The collection of the depth values at all the pixels' locations constitutes the depth map of the object.

IV. EXPERIMENTAL EVALUATIONS

One synthetic and four real objects were used for experiments [6]. The synthetic object is a simulated cone whose images were created by the camera simulation software. The program created a sequence of images of the simulated cone corresponding to different lens positions. Two of the real objects are cone and slanted planar objects whose images were generated from a CCD camera system by changing the distance between the lens and the object. Last two objects are microscopic images – protrusion of thin film transistor liquid crystal display (TFT-LCD) color filter and the Lincoln statue on US one cent coin whose images were obtained from

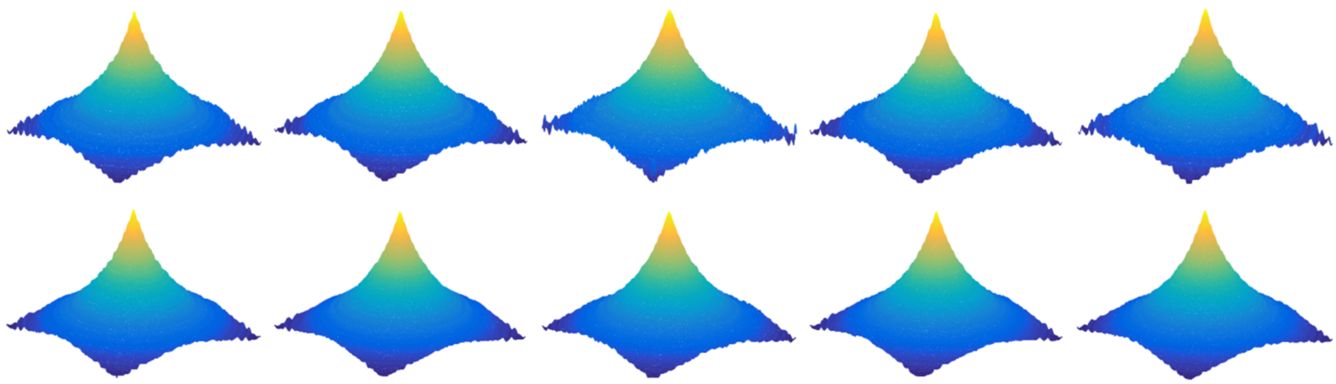

FIGURE 4. Reconstructed shapes of simulated cone object from SML, GLV, TEN, DCT, PCA focus measures from left to right respectively. First row presents the results from the original focus measures and second row presents the results from their modified focus measures using proposed method.

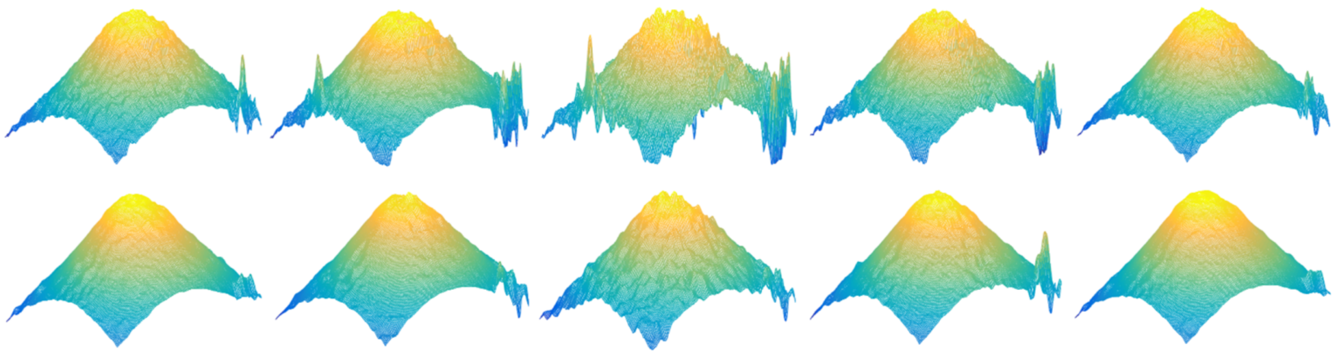

FIGURE 5. Reconstructed shapes of real cone object from SML, GLV, TEN, DCT, PCA focus measures from left to right respectively. First row presents the results from the original focus measures and second row presents the results from their modified focus measures using proposed method.

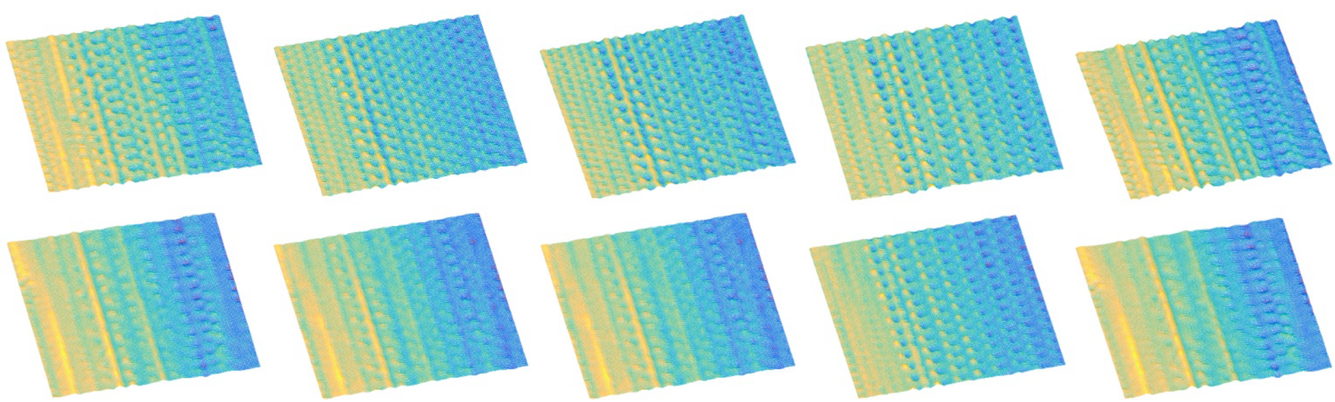

FIGURE 6. Reconstructed shapes of plane object from SML, GLV, TEN, DCT, PCA focus measures from left to right respectively. First row presents the results from the original focus measures and second row presents the results from their modified focus measures using proposed method.

microscope control system by changing the lens position through a step motor installed to the system. Example images of all the above mentioned objects are shown in Fig. 3.

In Table 1, the RMSE and correlation values between the actual depth map of the simulated cone object and the computed depth map based on five widely used focus measures – sum of modified Laplacian (SML) [15], gray level variance (GLV) [17], Tenengrad (TEN) [18], discrete cosine transform (DCT) [19], and principal component analysis (PCA) [20] - were compared. Results show that applying modified focus measures from proposed method could produce lower RMSE values and higher correlation values in comparison to applying original focus measures.

In Fig. 4, the SFF results on simulated cone object were compared. Shapes from the original focus measures have noises at bottom side of the cone and have rough surfaces. However, the modified focus measures have suppressed these noises and recovered smoother surfaces.

In Fig. 5, the SFF results on real cone object were compared. The results from the original focus measures show many noises in overall surface and the right side were badly corrupted by noise. However, the results from their

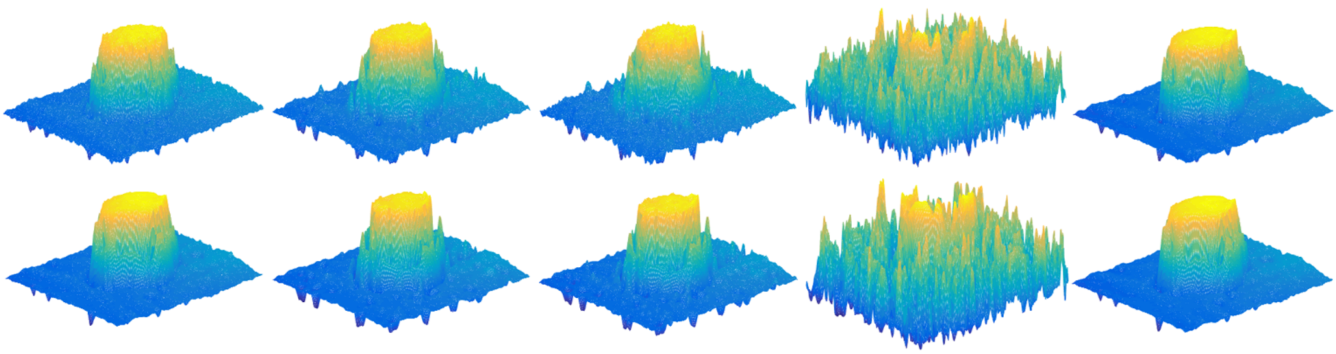

FIGURE 7. Reconstructed shapes of TFT LCD color filter object from SML, GLV, TEN, DCT, PCA focus measures from left to right respectively. First row presents the results from the original focus measures and second row presents the results from their modified focus measures using proposed method.

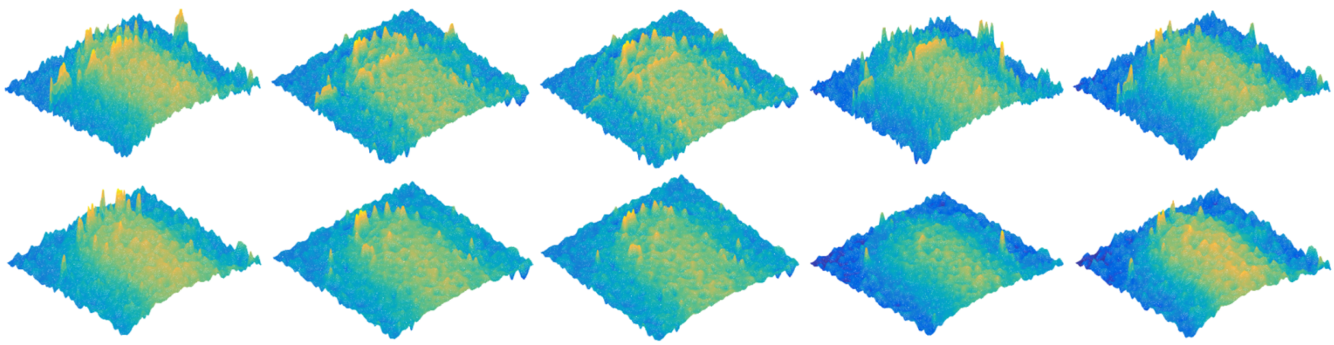

FIGURE 8. Reconstructed shapes of coin object from SML, GLV, TEN, DCT, PCA focus measures from left to right respectively. First row presents the results from the original focus measures and second row presents the results from their modified focus measures using proposed method.

corresponding modified focus measures generate smoother surface, and the shape corruption in the right side was dramatically restored. The shape improvements are more conspicuous when the shapes from the original focus measures contain more noises such as GLV and TEN focus measures.

In Fig. 6, the SFF results on planar object were compared. The reconstructed shapes on planar object from the proposed modified focus measures produced less fluctuations on overall surface in comparison to the shapes from the original focus measures.

The SFF results on microscopic objects - TFT-LCD color filter and Lincoln statue on US one cent coin - were compared in Fig. 7 and Fig. 8. On the results of TFT-LCD color filter and coin objects, the modified focus measures could generate less noisy surface and take the better shapes in comparison to the original focus measures. The shape reconstruction of TFT-LCD color filter was failed both in the original DCT focus measure and the modified DCT focus measure. However, the modified DCT focus measure could build the side part of the shape better than the original focus measure.

V. CONCLUSION

To improve the accuracy in focus value computation in SFF, we propose a technique to exclude noise while computing focus value of each image pixel. After getting initial focus measure volume, each focus value is inspected and labeled as noisy or noise-free. For the pixels whose focus value were classified as noisy, their final focus values were re-computed from the neighboring noise-free focus values. For quantitative analysis, RMSE and correlation between the computed depth map and the actual depth map of the synthetic object were compared. For qualitative analysis, the final reconstructed shapes from the existing focus measures and the proposed modified focus measures were compared on four different types of real objects. Results show that incorporating the proposed method into the existing focus measures produces less RMSE and higher correlation values on synthetic object and better reconstructed shapes in real objects compared to the original focus measures.

CONFLICT OF INTERESTS

The authors declare no conflict of interests regarding the publication of this article.

REFERENCES

[1] A. S. Malik and T.-S. Choi, "Effect of noise and source illumination on 3D shape recovery," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 22, no. 5, pp. 945–958, 2008.

[2] C. Ribal, N. Lermé, and S. Le Hégarat-Mascle, "Efficient graph cut optimization for shape from focus," *J. Vis. Commun. Image Represent.*, vol. 55, pp. 529–539, Aug. 2018.

[3] J. Humayun and A. S. Malik, "Real-time processing for shape-from-focus techniques," *J. Real-Time Image Process.*, vol. 11, no. 1, pp. 49–62, 2016.

[4] M. B. Ahmad and T.-S. Choi, "A heuristic approach for finding best focused shape," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 4, pp. 566–574, Apr. 2005.

[5] P. G. Kumar and R. R. Sahay, "Accurate structure recovery via weighted nuclear norm: A low rank approach to shape-from-focus," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Oct. 2017, pp. 563–574.

[6] S.-O. Shim and T.-S. Choi, "A novel iterative shape from focus algorithm based on combinatorial optimization," *Pattern Recognit.*, vol. 43, no. 10, pp. 3338–3347, 2010.

[7] T. Fan and H. Yu, "A novel shape from focus method based on 3D steerable filters for improved performance on treating textureless region," *Opt. Commun.*, vol. 410, pp. 254–261, Mar. 2018.

[8] R. Minhas, A. A. Mohammed, and Q. J. Wu, "Shape from focus using fast discrete curvelet transform," *Pattern Recognit.*, vol. 44, no. 4, pp. 839–853, 2011.

[9] M. S. Muhammad and T.-S. Choi, "Sampling for shape from focus in optical microscopy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 3, pp. 564–573, Mar. 2012.

[10] R. Hariharan and A. N. Rajagopalan, "Shape-from-focus by tensor voting," *IEEE Trans. Image Process.*, vol. 21, no. 7, pp. 3323–3328, Jun. 2012.

[11] S. Pertuz, D. Puig, and M. Garcia, "Reliability measure for shape-from-focus," *Image Vis. Comput.*, vol. 31, no. 10, pp. 725–734, 2013.

[12] 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," *Opt. Lasers Eng.*, vol. 51, no. 5, pp. 520–526, 2013.

[13] I. Lee, M. T. Mahmood, and T.-S. Choi, "Adaptive window selection for 3D shape recovery from image focus," *Opt. Laser Technol.*, vol. 45, pp. 21–31, Feb. 2013.

[14] I.-H. Lee, M. T. Mahmood, S.-O. Shim, and T.-S. Choi, "Optimizing image focus for 3D shape recovery through genetic algorithm," *Multimedia Tools Appl.*, vol. 71, no. 1, pp. 247–262, 2014.

[15] S. K. Nayar and Y. Nakagawa, "Shape from focus," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 16, no. 8, pp. 824–831, Aug. 1994.

[16] M. Subbarao and T.-S. Choi, "Accurate recovery of three-dimensional shape from image focus," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 3, pp. 266–274, Mar. 1995.

[17] T. Yeo, S. O. Jayasooriah, and R. Sinniah, "Autofocusing for tissue microscopy," *Image Vis. Comput.*, vol. 11, no. 10, pp. 629–639, 1993.

[18] J. M. Tenenbaum, "Accommodation in computer vision," Ph.D. dissertation, Dept. Comput. Sci., Stanford Univ., Standford, CA, USA, 1970.

[19] X. Zhang, H. Wu, and Y. Ma, "A new auto-focus measure based on medium frequency discrete cosine transform filtering and discrete cosine transform," *Appl. Comput. Harmon. Anal.*, vol. 40, no. 2, pp. 430–437, 2016.

[20] M. T. Mahmood, S.-O. Shim, and T.-S. Choi, "Shape from focus using principal component analysis in discrete wavelet transform," *Opt. Eng.*, vol. 48, no. 5, pp. 057203-1–057203-9, 2009.

[21] H.-J. Kim, M. T. Mahmood, and T.-S. Choi, "An efficient neural network for shape from focus with weight passing method," *Appl. Sci.*, vol. 8, no. 9, p. 1648, 2018.

[22] S.-Y. Lee, Y. Kumar, J.-M. Cho, S.-W. Lee, and S.-W. Kim, "Enhanced autofocus algorithm using robust focus measure and fuzzy reasoning," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, no. 9, pp. 1237–1246, Sep. 2008.

[23] I. De and B. Chanda, "Multi-focus image fusion using a morphology-based focus measure in a quad-tree structure," *Inf. Fusion*, vol. 14, no. 2, pp. 136–146, 2013.

[24] P. Sakurikar and P. J. Narayanan, "Composite focus measure for high quality depth maps," in *Proc. IEEE Conf. Comput. Vis.*, Oct. 2017, pp. 1623–1631.

[25] Y. Zhang, X. Bai, and T. Wang, "Boundary finding based multi-focus image fusion through multi-scale morphological focus-measure," *Inf. Fusion*, vol. 35, pp. 81–101, May 2017.

SEONG-O SHIM received the B.S. degree in electronics engineering from Ajou University, Suwon, South Korea, in 1999, and the M.S. degree in mechatronics and the Ph.D. degree in information and mechatronics from the Gwangju Institute of Science and Technology, Gwangju, South Korea, in 2001 and 2011, respectively.

He was with LG Electronics DTV Labs, Seoul, South Korea, where he worked on research and development of digital TV, from 2003 to 2007. He is currently an Associate Professor with the College of Computer Sciences and Engineering, University of Jeddah, Jeddah, Saudi Arabia. His research interests include computer vision, image processing, 3D shape recovery, and medical imaging.

WAJID AZIZ received the B.Sc. and M.Sc. degrees from the University of Azad Jammu & Kashmir University (UAJ&K), Muzaffarabad, Pakistan, and the Ph.D. degree from the Pakistan Institute of Engineering and Applied Sciences, Islamabad, Pakistan.

He started his career with the UAJ&K, in 1998, as a Lecturer. He is currently a Professor with the College of Computer Sciences and Engineering, University of Jeddah, Jeddah, Saudi Arabia. He has published three books and more than 45 research articles in the reputed national and International journals and conference proceeding. His research interests include biomedical information systems, biomedical signal processing, time series analysis, and intelligent data analytics. Based on his academic and research contributions, he was a recipient of the HEC University Best Teacher Award, from 2012 to 2013, by HEC Pakistan, in 2014, and the University Best Teacher Award by UAJ&K, in 2013.

AMEEN BANJAR received the B.S. degree in computer science from Taibah University, Medina, Saudi Arabia, in 2004, the M.S. degree in information and communication technology from the University of Wollongong, Wollongong, Australia, in 2010, and the Ph.D. degree in software engineering from the University of Technology Sydney, Sydney, Australia, in 2016.

He is currently an Assistant Professor with the College of Computer Sciences and Engineering, University of Jeddah, Jeddah, Saudi Arabia. His research interests include data science, software defined networks, and the Internet of Things.

ABDULLAH ALAMRI received the B.S. degree in computer science from King Khalid University, Abha, Saudi Arabia, in 2007, the M.S. degree in information technology from La Trobe University, Melbourne, Australia, in 2009, and the Ph.D. degree in computer science from RMIT University, Melbourne, Australia, in 2014.

He is currently an Assistant Professor with the College of Computer Sciences and Engineering, University of Jeddah, Jeddah, Saudi Arabia. His research interests include big data, the Internet of Things, database systems, and semantic Web.

MOHAMMED ALQARNI received the B.S. degree in computer science from King Khalid University, Abha, Saudi Arabia, in 2008, the M.S. degree in computational sciences from Laurentian University, Sudbury, Canada, in 2012, and the Ph.D. degree in computer science from McMaster University, Hamilton, Canada, in 2016.

He is currently an Assistant Professor and the Dean of the College of Computer Sciences and Engineering, University of Jeddah, Jeddah, Saudi Arabia. His research interests include networking, concurrent systems modeling, computational medicine, and computer vision.