



A novel iterative shape from focus algorithm based on combinatorial optimization

Seong-O Shim, Tae-Sun Choi*

Gwangju Institute of Science and Technology, 261 Cheomdan Gwagiro, Buk-Gu, Gwangju 500-712, Republic of Korea

ARTICLE INFO

Article history:

Received 9 February 2009

Received in revised form

26 November 2009

Accepted 21 May 2010

Keywords:

Shape from focus (SFF)

3D shape

Depth map

Focus measure

Combinatorial optimization

ABSTRACT

Shape from focus (SFF) is a technique to estimate the depth and 3D shape of an object from a sequence of images obtained at different focus settings. In this paper, the SFF is presented as a combinatorial optimization problem. The proposed algorithm tries to find the combination of pixel frames which produces maximum focus measure computed over pixels lying on those frames. To reduce the high computational complexity, a local search method is proposed. After the estimate of the initial depth map solution of an object, the neighborhood is defined, and an intermediate image volume is generated from the neighborhood. The updated depth map solution is found from the intermediate image volume. This update process of the depth map solution continues until the amount of improvement is negligible. The results of the proposed SFF algorithm have shown significant improvements in both the accuracy of the depth map estimation and the computational complexity, with respect to the existing SFF methods.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

In computer vision, the techniques to recover the three-dimensional (3D) geometry of a scene or an object from a collection of images are known as shape-from-X. Where, X denotes the accommodation cue to infer the shape, such as stereo, motion, shading, texture, etc. Shape-from-focus/defocus (SFF/SFD) deals with the recovery of shape from multiple images of the same scene that are captured with different geometries of the imaging device. SFD techniques [1–4] estimate the depth of a point by measuring the amount of blur (error in focus) from one or two images. However, SFF requires searching the best focus setting that gives the best focus at each point [5–7]. Therefore, in SFF, every point of an object needs to be well focused in a particular image frame in the collection of images. SFF achieves better quality of shape in comparison to SFD with the cost of higher computational complexity. In this paper, we focus the discussion on SFF.

In SFF, a sequence of images of an unknown object is acquired by changing the level of the object focus. The change in the level of focus is made either by varying the focus value of the camera, or varying the distance of the object from the lens. The acquired image sequence is a three dimensional image volume where the row and the column of each image frame are the first and second dimension and the image frames along the optical (or temporal)

axis the third dimension. In the image space, each object point is gradually focused until it attains maximum focus and then gradually blurred along the temporal axis. Then, a focus measure operator is applied to compute the focus value on the small regions of every pixel in the image volume. The focus value increases as the image sharpness or contrast increases and it attains maximum value at the best focused point. For each point (ith row, jth column) which corresponds to each object point, the image frame that exhibits maximum sharpness value along the axis of third dimension is determined. This image frame has the information about the distance v between the lens and the image detector when it is taken. Then, by thin lens formula, the distance D of an object point from the lens is given as

$$D = \frac{vf}{v-f} \quad (1)$$

where f is the focal length of the lens. The collection of depth information at each location (ith row, jth column) constitutes the depth map of an object. In addition, the collection of the gray level (or color) value of the best focused image frame at each point (ith row, jth column) constitutes all-in-focus image. Hence, in SFF, the measure of focus or sharpness is the most crucial part for the quality of the final depth estimation. Traditionally, the focus measure is applied on each 2D image frame of the image volume. However, for an object with complex geometry, the images acquired from the lens with limited depth of field have different focus level. Hence, the focus value of a focus measure operator on 2D image frame does not represent accurate focus level at the pixel.

* Corresponding author.

E-mail addresses: seongo@gist.ac.kr (S.-O. Shim), tschoi@gist.ac.kr (T.-S. Choi).

In this paper, the SFF is modeled as combinatorial optimization problem. The optimal depth map solution of an object is considered as the combination of pixel frame numbers that gives the maximum focus measure. Trying all combinations of the pixel frames requires high computational time. To reduce the computational complexity, a local search algorithm is proposed. First the initial depth map solution is obtained by applying a focus measure and then finding the frame number that maximizes the focus value along the optical axis. At each point (i th row, j th column), the neighborhood is defined from the initial solution by taking several preceding and following frames with respect to the initial depth. The intermediate image volume is obtained by collecting the pixels values of neighborhood at each point. The updated solution is found from the intermediate image volume. This update process continues until the convergence criterion is met. The process to obtain temporary image volume has the effect of aligning the curved object patch, corresponding to the focused image surface, perpendicular to the optical axis. Therefore, applying the focus measure on the intermediate image volume gives more accurate focus level at each pixel.

This article is organized as follows. In Section 2, we give brief overview of the SFF techniques. In Section 3, the proposed SFF algorithm is described. Comparison with the previous SFF methods is done in Section 4. Finally, Section 5 concludes our work.

2. Previous works

A focus measure is defined as a quantity to locally evaluate the sharpness of a pixel. It takes small local neighborhood and computes the sharpness of a chosen center pixel. Since each object point has different surface characteristic and geometry, the focus measure values of the same object point from different optical settings are compared. A variety of focus measures have been proposed in the spatial domain and the transformed domains [5,8–10]. Among them, Sum modified Laplacian (SML), Tenenbaum gradient (TEN), and gray level variance (GLV) are most commonly used. The SML is based on the Laplacian operator which is suitable for measuring sharpness magnitude. In the case of Laplacian, the second derivatives in x and y directions can cancel each other. To avoid this problem, Nayar and Nakagawa [6] proposed modified Laplacian (ML), which is defined as

$$\text{ML}(x,y) = \left| \frac{\partial^2 g(x,y)}{\partial x^2} \right| + \left| \frac{\partial^2 g(x,y)}{\partial y^2} \right| \quad (2)$$

where $g(x,y)$ is the gray value in an image at the coordinate (x,y) . In order to increase the robustness against noise, the focus measure at (x,y) is presented as sum of ML values in a local window.

$$\text{SML}(x_0,y_0) = \sum_{p(x,y) \in U(x_0,y_0)} \text{ML}(x,y) \quad (3)$$

Tenenbaum [11] proposed a focus measure based on Sobel gradient operator. It uses sum of squared responses of the horizontal and vertical Sobel masks for gradient magnitude maximization. For robustness, it is also summed in a local window

$$\text{TEN}(x_0,y_0) = \sum_{p(x,y) \in U(x_0,y_0)} [G_x(x,y)^2 + G_y(x,y)^2] \quad (4)$$

The GLV focus measure is based on variations in gray level values. The variance of the gray level values in a sharp image region is higher than that in a blurred image region. GLV focus measure is obtained by taking the variance of the gray level values

of pixels within a local window as

$$\text{GLV}(x_0,y_0) = \frac{1}{N-1} \sum_{p(x,y) \in U(x_0,y_0)} [g(x,y) - \mu_{U(x_0,y_0)}]^2 \quad (5)$$

where $\mu_{U(x_0,y_0)}$ is the mean of the gray values in the neighborhood $U(x_0,y_0)$.

The conventional SFF methods assume the object shape is piecewise constant in a small region around the each pixel position. Therefore, the focus measure is applied over pixels from individual planar image frames. However, in most practical situations, this assumption is not valid due to the complex geometry of the object surface. The conventional SFF methods thus fail to produce accurate shape estimation for the objects with complex geometry. For an accurate estimation of the object shape, Subbarao and Choi [7] proposed a technique based on the concept named the focused image surface (FIS). The FIS of an object is defined as the surface formed by a set of points at which the object points are focused by the lens. According to paraxial geometric optics, there is one-to-one correspondence between the shape of an object and the shape of its FIS. Therefore, the problem of SFF can be described as the problem of determining the shape of FIS. After the initial estimate of FIS using a traditional SFF method, the initial FIS is refined by searching the planar surface which maximizes a focus measure computed over pixels lying on the FIS. Whereas the conventional SFF method is based on the piecewise constant approximation of the FIS, the SFF method based on FIS has shown improved results by exhaustive search of the FIS shape using planar surface approximation. Later, Choi and Yun [12] proposed the approximation of FIS by a piecewise curved surface rather than piecewise planar approximation to improve the accuracy for the objects with complex geometry. The piecewise curved surface was estimated by interpolation using second order Lagrange polynomial. Asif and Choi [13] used neural networks to learn the shape of FIS. Due to their nonlinear characteristics, neural networks can be used to approximate arbitrary shape. Furthermore, Ahmad and Choi [14] proposed the use of dynamic programming to search the optimal FIS shape. However, all those methods compute focus measure on 2D image frame.

3. Theory

3.1. Motivation

The SFF attempts to search for the surface in an image volume that gives the best focus measure. Fig. 1 illustrates small 3D image volume of an object and its corresponding depth map. The combination of pixel frame numbers of the depth map in Fig. 1 produces maximum focus measure computed over pixels on these image frames. Therefore, the SFF problem can be thought of as the one of choosing the combination of the pixel frames in an image volume so that the focus measure is maximized, which is the combinatorial optimization problem [15–17]. Combinatorial optimization problem consists of a collection of instances (F, c) , where F is the set of feasible solutions and c is the cost function. The goal is to find a globally optimal solution, i.e., a feasible solution S^* such that $c(S^*) \leq c(S)$ for all $S \in F$. Simply trying all combinations of pixel frame numbers and searching for the combination (surface) that gives the maximum contrast value globally requires tremendous amount of computations. The number of combinations for an $I \times J \times K$ image volume (I rows, J columns, and K frames) is given by $K^{I \times J}$. To reduce the computational complexity, we proposed a local search algorithm.

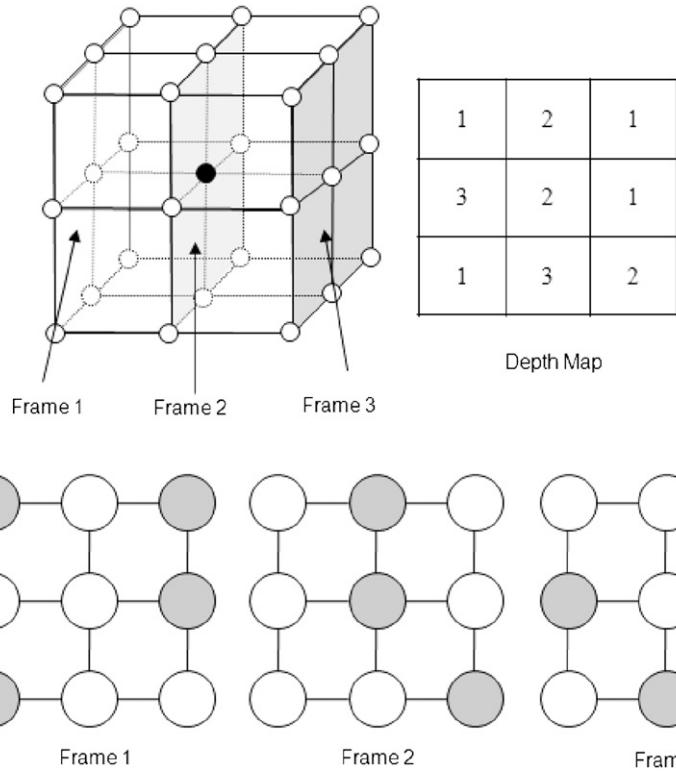


Fig. 1. Small 3D image volume of an object and its depth map. (Gray color circles in each image frame represent the pixels that show the maximum focus measure at that position.)

3.2. Proposed local search algorithm

Local search [18–20] is an effective means of solving difficult combinatorial optimization problems. Instead of a global solution, it tries to find a locally optimal solution. A local search algorithm starts with an initial feasible solution and then repeatedly searches neighborhoods to find better solutions until it reaches a near optimal solution. In this paper, the solution of the SFF problem is defined as the depth map which is the set of depth solutions at each object point. For convenience of notation, we refer the point (i th row, j th column) in the image detector as point (i,j) , and the pixel (i th row, j th column) in the image frame k as pixel (i,j,k) .

First, the initial depth map of the object is estimated by applying one of the traditional focus measures. For each image frame k from the total of K image frames, the focus measure $FM(i,j,k)$ is computed at each pixel (i,j,k) in the acquired image volume $g(i,j,k)$ by applying a focus measure on a small window centered at (i,j,k) . Then, for the point (i,j) , the depth map $D^{[0]}(i,j)$ is computed by taking the frame number that produces the maximum focus measure from $FM(i,j,k)$

$$D^{[0]}(i,j) = \operatorname{argmax}_k FM(i,j,k), \quad 1 \leq i \leq I, \quad 1 \leq j \leq J, \quad 1 \leq k \leq K \quad (6)$$

where I and J are the width and the height of the image frame, and K is the total number of image frames. We consider the computed depth map $D^{[0]}(i,j)$ as initial solution at each point (i,j) , and the focus measure function $FM(i,j,k)$ as cost function. The neighborhood of the estimated solution is found as follows. For each point (i,j) , the neighborhood $N^{[0]}(i,j)$ of the solution $D^{[0]}(i,j)$ is defined by taking few preceding and following frames with respect to the frame $D^{[0]}(i,j)$ as

$$N^{[0]}(i,j) = \{k | D^{[0]}(i,j) - b \leq k \leq D^{[0]}(i,j) + b\} \quad (7)$$

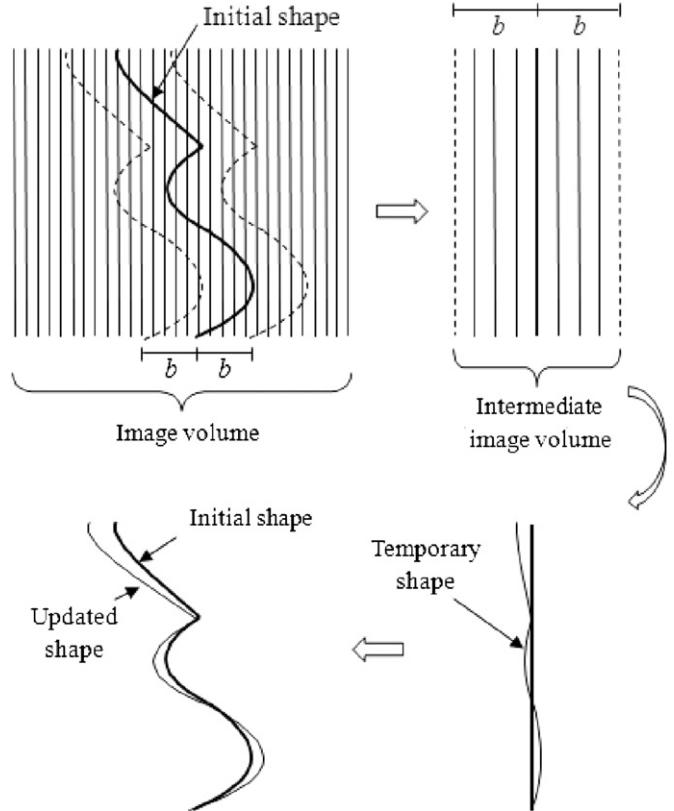


Fig. 2. 3D shape estimation process of the proposed algorithm.

where b is the radius of the neighborhood $N^{[0]}(i,j)$ centered at $D^{[0]}(i,j)$. We assume that $N^{[0]}(i,j)$ is greater than 1 (first image frame) and smaller than K (last image frame). Then an

intermediate image volume $g_{temp}(i,j,k')$ consisting of $K'(1 < K' < K)$ frames is obtained from the original image volume $g(i,j,k)$ as

$$g_{temp}(i,j,k') = g(i,j,N^{[0]}(i,j)), \quad 1 \leq i \leq I, \quad 1 \leq j \leq J, \quad 1 \leq k' \leq (2b+1) \quad (8)$$

The intermediate image volume $g_{temp}(i,j,k')$ has some special properties. The pixels corresponding to the initial depths, $(i,j,D^{[0]}(i,j))$, are re-located to the frame $k'=1+b$, namely, $(i,j,1+b)$. The image frame $g_{temp}(i,j,1+b)$ corresponds to the all-in-focus image. For a given frame number k' , $g_{temp}(i,j,k')$ has the approximately same depth for all (i,j) as shown in Fig. 2. Therefore, a focus measure value obtained by applying a focus measure operator on each image frame in $g_{temp}(i,j,k')$ reflects more accurate focus level for the pixel (i,j,k') . After obtaining a new focus measure volume $FM_{temp}(i,j,k')$ from $g_{temp}(i,j,k')$, a temporary solution $D_{temp}(i,j)$ is computed as

$$D_{temp}(i,j) = \underset{k'}{\operatorname{argmax}} FM_{temp}(i,j,k'), \quad 1 \leq i \leq I, \quad 1 \leq j \leq J, \quad 1 \leq k' \leq (2b+1) \quad (9)$$

Based on $D_{temp}(i,j)$, the updated solution $D^{[1]}(i,j)$ is computed from

$$D^{[1]}(i,j) = D^{[0]}(i,j) + (D_{temp}(i,j) - (1+b)) \quad (10)$$

where $(D_{temp}(i,j) - (1+b))$ is the number of frames that needs to be adjusted to the initial depth map $D^{[0]}(i,j)$. Finally, averaging filter

Table 1
Summary of the proposed SFF algorithm.

Algorithm:

- (1) Compute the focus measure volume $FM(i,j,k)$ from the image volume $g(i,j,k)$ by applying one of the focus measure operators.
- (2) Compute the initial depth map solution $D^{[l]}(i,j)$, $l=0, 1, 2, \dots$, of an object from Eq. (6).
- (3) Apply averaging filter on $D^{[l]}(i,j)$, $l=0, 1, 2, \dots$.
- (4) Define the neighborhood $N^{[l]}(i,j)$ of the solution $D^{[l]}(i,j)$ from Eq. (7).
- (5) Generate an intermediate image volume $g_{temp}(i,j,k')$ from an original image volume from Eq. (8).
- (6) Apply the focus measure operator on the intermediate image volume $g_{temp}(i,j,k')$ to obtain its corresponding focus measure volume $FM_{temp}(i,j,k')$.
- (7) Computes an updated depth map solution $D^{[l+1]}(i,j)$ from Eq. (10).
- (8) Apply averaging filter on $D^{[l+1]}(i,j)$.
- (9) Replace $D^{[l]}(i,j)$ with $D^{[l+1]}(i,j)$ and return to step 4 until the difference between $D^{[l]}(i,j)$ and $D^{[l+1]}(i,j)$ is negligible.

with 3×3 window size is applied on $D^{[1]}(i,j)$ to remove possible spikes caused by error of the initial depth estimates. This update process of the depth map solution is repeated several times until the amount of the difference between the final updated solution and the previous solution is negligible. The number of iteration is investigated heuristically in the following section.

The proposed algorithm refines the depth map based on the first estimate of the depth map. Therefore, if the initial depth value of a point is deviated from the real depth value, this error might propagate in the final depth value. In addition, the shape details might be smeared at depth discontinuities such as object edges and occlusions. This problem can be prevented by introducing the slope constraint. If we assume the maximum slope of the object is α_{max} and the distance between the image frames is ζ and the pixel size is $d \times d$, then the depth difference $|\delta_{ij}(x,y)|$ in (11) between the point (i,j) and its adjacent point $(i+x, j+y)$ should not exceed $d \alpha_{max}$. Therefore, at each iteration of the algorithm, we impose a slope constraint on $D^{[l]}(i,j)$ as

$$\delta_{ij}(x,y) = D^{[l]}(i+x, j+y) - D^{[l]}(i,j), \quad 0 < |x|, |y| \leq 1 \quad (11)$$

$$D^{[l+1]}(i,j) = D^{[l]}(i,j) \quad \text{if } \exists (x,y) \text{ s.t. } |\delta_{ij}(x,y)| > d\alpha_{max}/\zeta \quad (12)$$

The complete procedure of the proposed SFF algorithm is summarized in Table 1, and illustrated in Fig. 2.

4. Results and discussion

4.1. Experimental setup

In this section, the proposed algorithm was analyzed and compared with the previous SFF methods. The proposed SFF method based on local search (SFF-LS), searches the optimal depth map using the initial estimates of shape (FIS). Therefore, the proposed technique was compared with other well known methods, i.e., SFF based on FIS (SFF-FIS) [7] and SFF based on dynamic programming (SFF-DP) [14], which use the FIS concept in their algorithms. In addition, the proposed SFF algorithm was compared with the recent SFD algorithm based on diffusion (SFD-DFU) [4].

Experiments were conducted on six different types of objects. First object is a simulated cone (synthetic object) whose images are generated from the camera simulation software. A sequence of 97 images of the simulation cone was generated corresponding to 97 lens positions. The second and third objects are real cone and

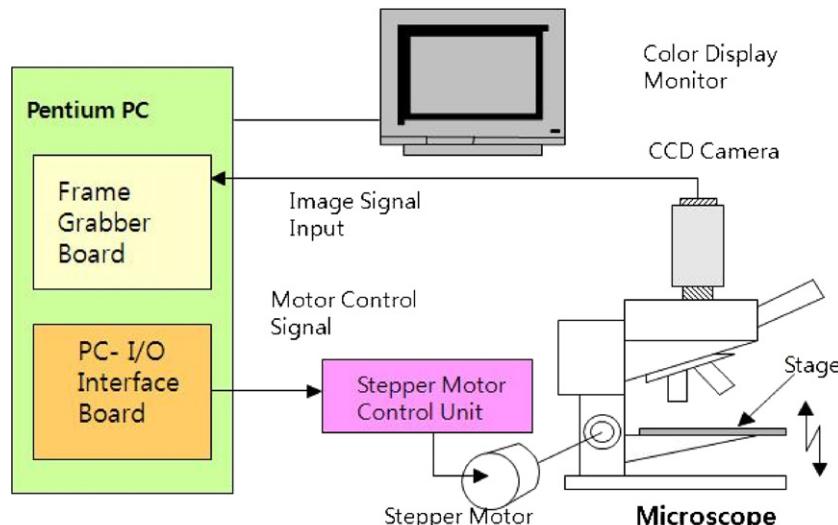


Fig. 3. Microscope control system (MCS).

slanted planar objects whose images were taken using a CCD camera system. The real cone object was made of hardboard with black and white stripes drawn on the surface so that a dense

texture of ring patterns is viewed in images. For the real cone and planar object, the displacement between two consecutive frames was about 0.03 mm. The length of the real cone object was about

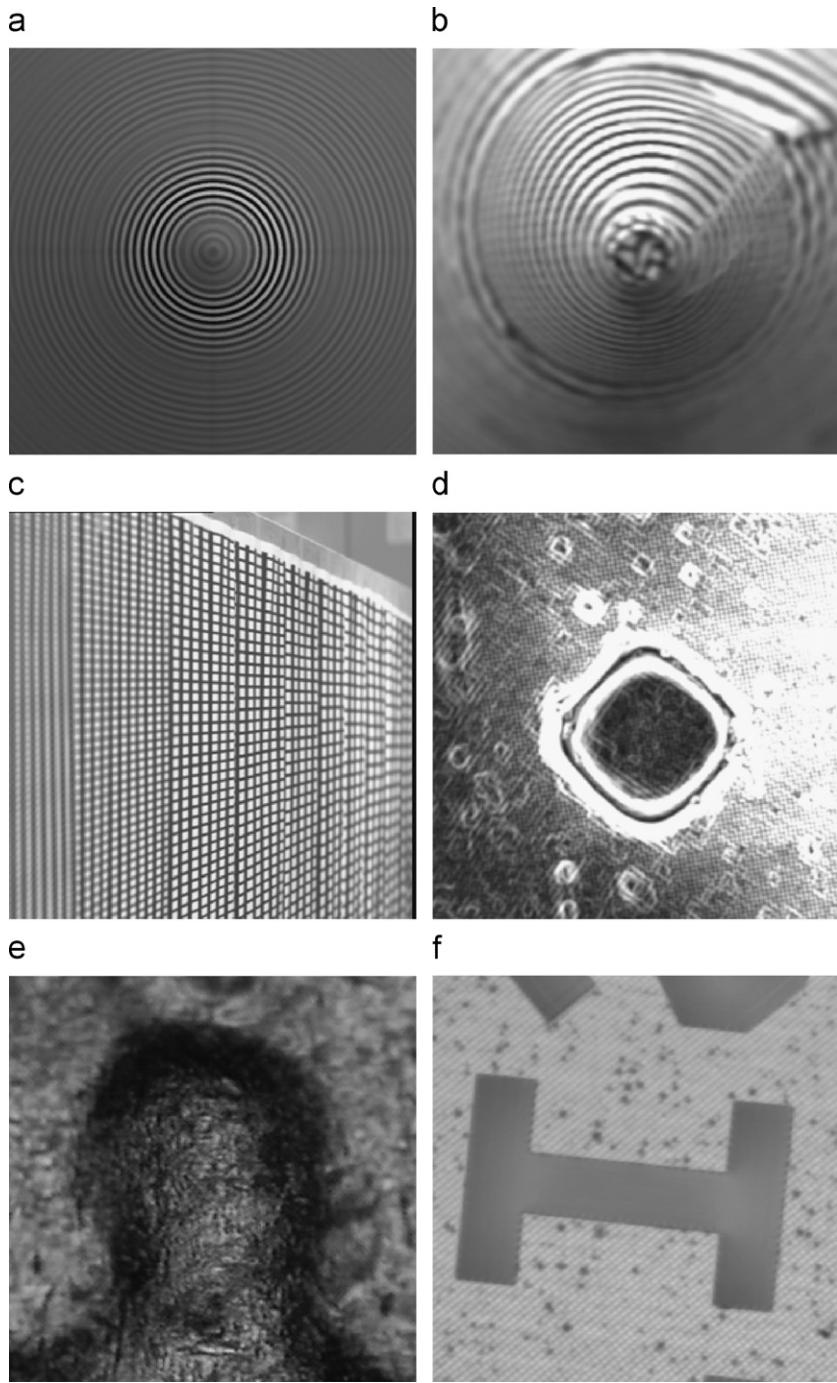


Fig. 4. Test 3D objects: (a) simulated cone, (b) real cone, (c) planar object, (d) TFT-LCD color filter, (e) head part of Lincoln statue on US one cent coin, and (f) letter 'H'.

Table 2

Depth changes at each stage of iterations (focus measure=SML).

| Point | Initial depth | 1st update | 2nd update | 3rd update | 4th update | Real depth |
|-----------|---------------|------------|------------|------------|------------|------------|
| (70,70) | 34 | 35 | 36 | 37 | 37 | 37 |
| (90,90) | 38 | 39 | 40 | 40 | 40 | 41 |
| (120,120) | 46 | 48 | 48 | 49 | 50 | 53 |
| (140,140) | 56 | 58 | 58 | 59 | 59 | 61 |
| (170,170) | 81 | 82 | 82 | 83 | 83 | 83 |

Table 3

Comparison of RMSE and correlation values on simulated cone at each stage of iterations.

| Point | Focus measure | Initial depth | 1st update | 2nd update | 3rd update | 4th update |
|-------------|---------------|---------------|------------|------------|------------|------------|
| RMSE | SML | 8.0941 | 7.4822 | 7.3490 | 7.2803 | 7.2404 |
| | TEN | 8.1232 | 7.3414 | 7.1824 | 7.0709 | 6.9863 |
| | GLV | 8.0886 | 7.3706 | 7.2066 | 7.0947 | 7.0270 |
| Correlation | SML | 0.9285 | 0.9479 | 0.9495 | 0.9499 | 0.9510 |
| | TEN | 0.9314 | 0.9518 | 0.9528 | 0.9532 | 0.9534 |
| | GLV | 0.9317 | 0.9508 | 0.9520 | 0.9526 | 0.9528 |

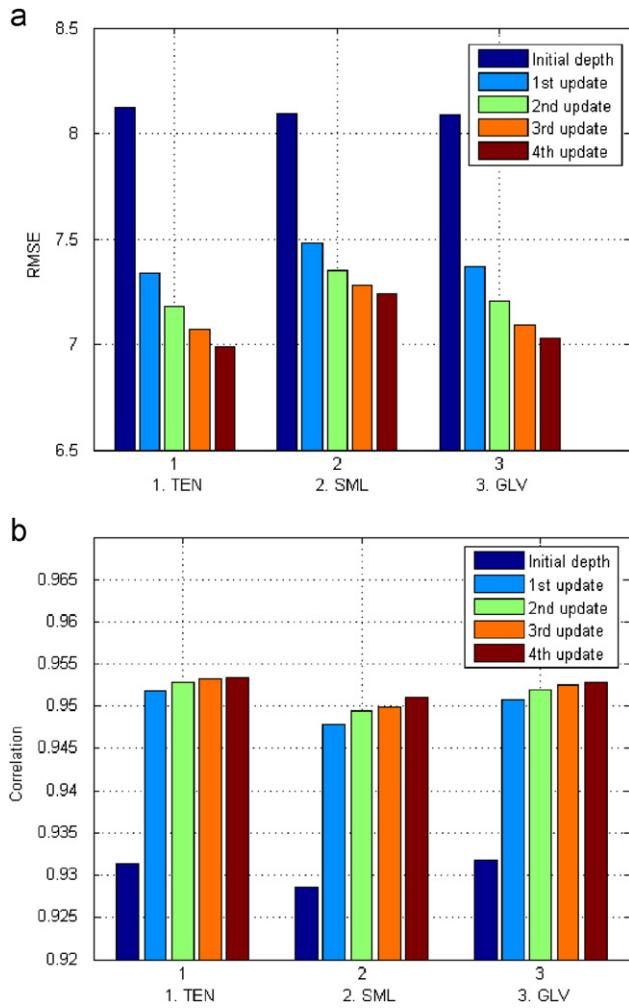


Fig. 5. Comparison of RMSE and correlation at each stage of iterations (simulated cone): (a) RMSE and (b) correlation.

Table 4

Comparison of RMSE and correlation values on simulated cone.

| Point | SFF-FIS | | | SFF-DP | | | SFF-LS | | |
|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|
| | 1st update | 2nd update | 3rd update | 1st update | 2nd update | 3rd update | 1st update | 2nd update | 3rd update |
| RMSE | 7.9132 | 7.7472 | 7.3414 | 7.1824 | 7.0709 | 6.9863 | | | |
| Correlation | 0.9473 | 0.9436 | 0.9518 | 0.9528 | 0.9532 | 0.9534 | | | |

200 cm and its base diameter was 38 cm. Step 1 corresponded to focusing an object at infinity and step 97 corresponded to focusing at about 50 cm from the lens. The planar object had only horizontal slope and equally spaced strips to make a poor

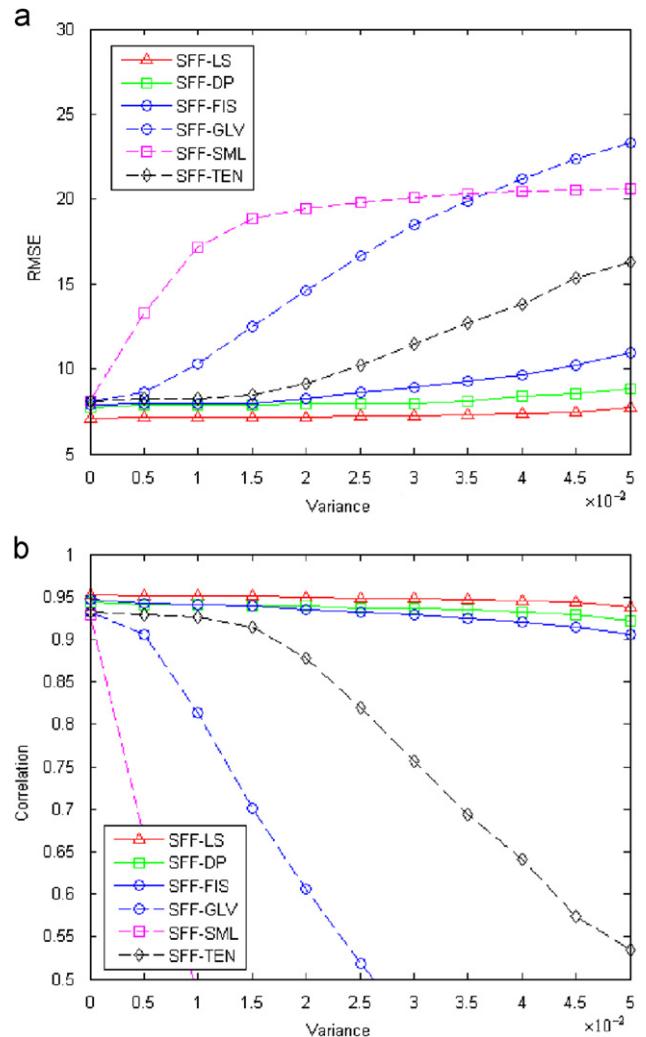


Fig. 6. RMSE and correlation comparisons on simulated cone in the presence of Gaussian noise: (a) RMSE and (b) correlation.

texture. The fourth, fifth, and sixth objects are the microscopic objects—thin film transistor liquid crystal display (TFT-LCD) color filter, the head part of the Lincoln statue on US one cent coin, and the letter 'H' engraved by laser beam.

The images of these microscopic objects are taken at different lens steps from microscope control system (MCS) shown in Fig. 3. The system consists of a personal computer integrated with a frame grabber board (Matrox Meteor-II), stepper motor control unit, optical microscope (NIKON OPTIPHOT-100S), and CCD camera (SAMSUNG CAMERA SCC-341) mounted on ocular tube of the microscope. Software was developed to acquire images by controlling the lens position through a step motor driver

(MAC 5000) having a 2.5 nm step length. A sequence of 60 images of TFT-LCD object was obtained under $500 \times$ magnification with 0.1 μm lens step size. And a sequence of 60 images of the coin was obtained under $100 \times$ magnification with 0.52 μm lens step size. For the letter ‘H’ object, a sequence of 60 images was obtained under $50 \times$ magnification with 1 μm lens step size. One of the frames of all the above mentioned images are shown in Fig. 4.

4.2. Quantitative analysis

The proposed algorithm was investigated on the synthetic object where the depth information is already known. For quantitative analysis, the root mean square error (RMSE) and the correlation were computed using the actual and the reconstructed depth maps of the simulated cone. If $f(x,y)$ is the actual depth map and $g(x,y)$ is the reconstructed depth map, the RMSE and the correlation r between $f(x,y)$ and $g(x,y)$ are given by

$$\text{RMSE} = \sqrt{\frac{1}{XY} \sum_{x=1}^X \sum_{y=1}^Y |f(x,y) - g(x,y)|^2} \quad (13)$$

$$r = \frac{\sum_{x=1}^X \sum_{y=1}^Y (f(x,y) - \bar{f})(g(x,y) - \bar{g})}{\sqrt{(\sum_{x=1}^X \sum_{y=1}^Y (f(x,y) - \bar{f})^2)(\sum_{x=1}^X \sum_{y=1}^Y (g(x,y) - \bar{g})^2)}} \quad (14)$$

If the reconstructed depth map is closer to the actual depth map, it has smaller RMSE and larger correlation values. Table 2 shows the changes of computed depth values at certain points. SML focus measure was applied to compute the depth map. It can be seen that the computed depth values approaches the real depth values as the number of iteration increases. At each stage of iterations, the RMSE and correlation values between the actual and the reconstructed depth maps of the simulated cone were computed in Table 3. Where, the SML, TEN, and GLV, which are the most commonly used focus measures described in Section 2, were investigated. For comparison purpose, 5×5 window size was applied while computing focus measure and averaging filter with 3×3 window size was applied on the final depth map for all

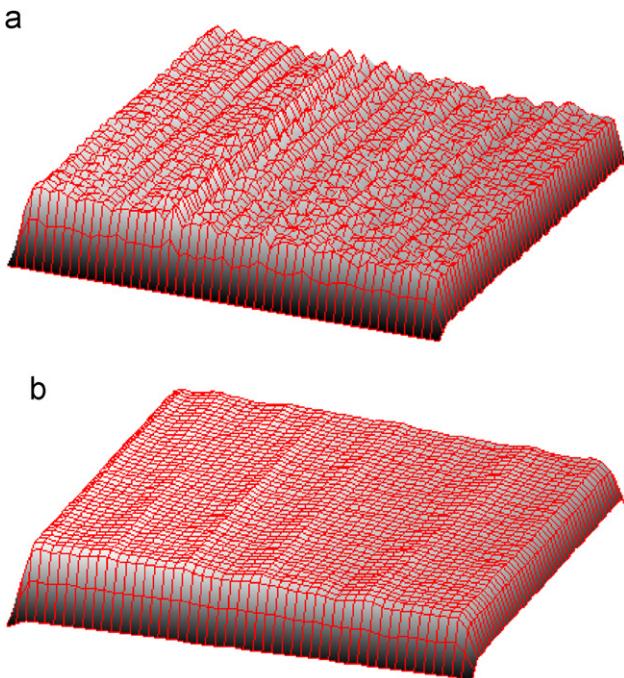


Fig. 7. SFF results on planar object based on TEN focus measure: (a) initial shape and (b) final shape of the algorithm.

SFF methods in experiments. Malik and Choi [21] concluded that the optimal window size is 3×3 for estimation of depth map.

From Table 3, it can be observed that the RMSE values get lower and the correlation values get higher as the number of iteration increases for all three focus measures. And, the improvements are most significant at first iteration and the amount of improvement gets smaller as the number of iteration increases. The improvements of the results are shown graphically in Fig. 5. It can be seen from Fig. 5 that the algorithm applied to TEN focus measure shows both the best results and highest improvements in RMSE and correlation values. If the number of iteration increases, the result improves at the cost of computational time. According to Table 3, three iterations can be considered as good tradeoff between the quality of the result and the computational time.

In Table 4, the proposed algorithm was compared with the previous SFF algorithms based on FIS (SFF-FIS) [7] and dynamic programming (SFF-DP) [14]. For all these SFF methods, TEN focus measure was used. In Table 4, we can see that even the first iteration of the proposed algorithm outperforms SFF-FIS and SFF-DP both in RMSE and correlation values. As the number of iteration increases, the amount of improvements decreases.

For robustness analysis, SFF algorithms were applied on noisy images, and their RMSE and correlation values on the

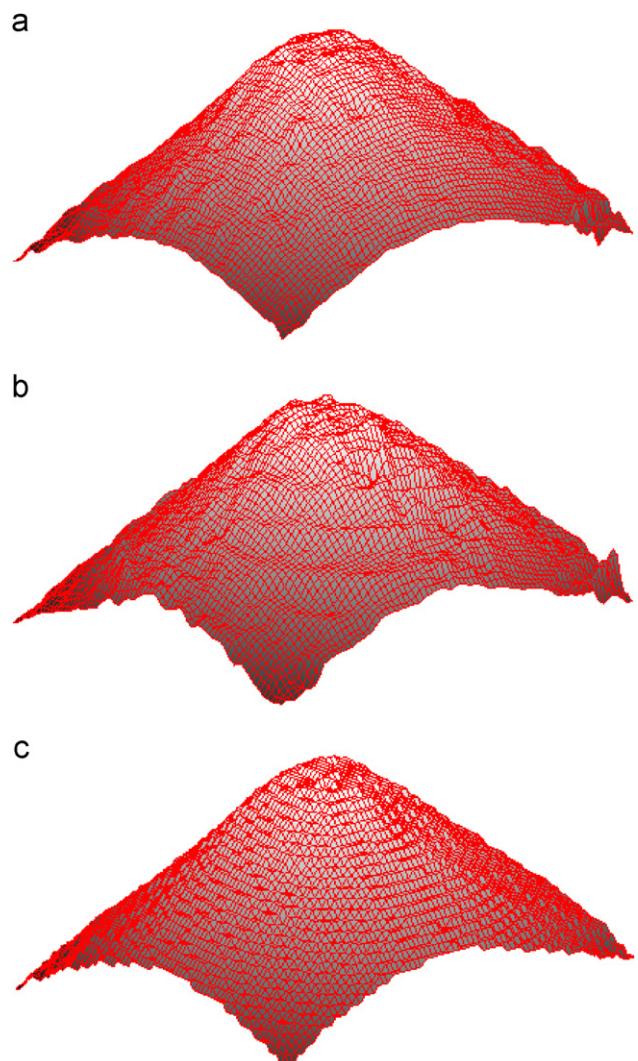


Fig. 8. Comparison of the SFF results on real cone: (a) SFF-FIS, (b) SFF-DP, and (c) SFF-LS.

reconstructed shapes were compared. Additive white Gaussian noise with zero mean and nonzero variance was imposed to the simulated cone image sequence. The image measurement systems including the microscopic systems are constrained by the CCD noise. Thermal noise is the main contributor to CCD noise and hence for accurate shape recovery, it needs to be accounted for. Thermal noise can be modeled using Gaussian probability density function. In addition, Gaussian noises arise in the images because of the factors such as the electronic circuit noise and the sensor noise caused by poor illumination and/or high temperature.

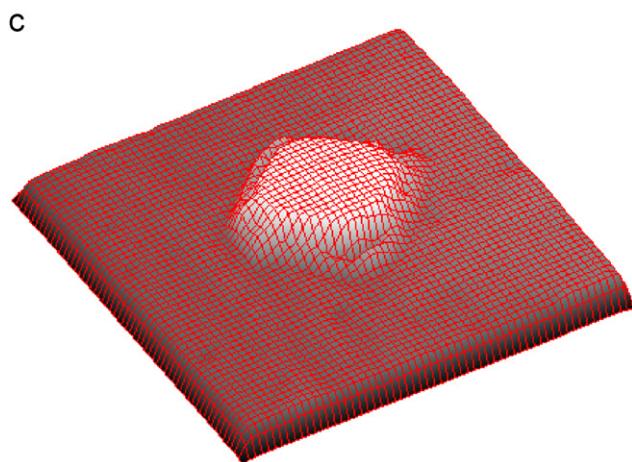
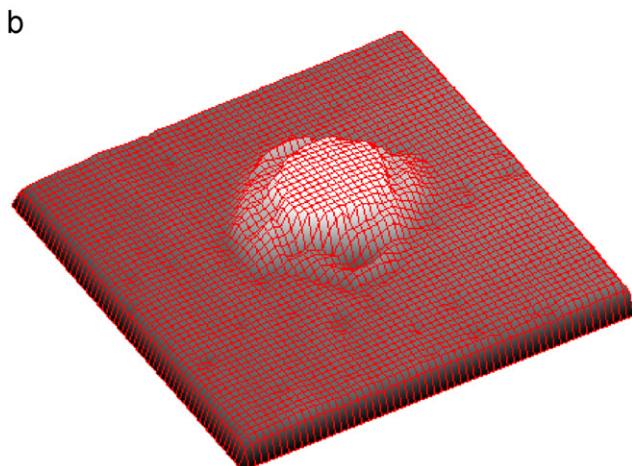
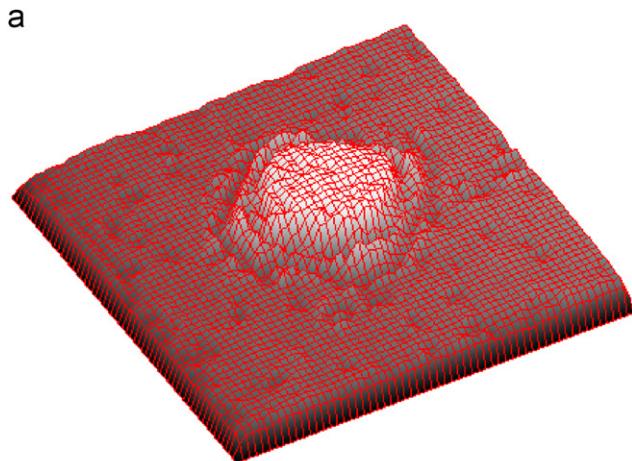


Fig. 9. Comparison of the SFF results on TFT-LCD color filter: (a) SFF-FIS, (b) SFF-DP, and (c) SFF-LS.

In Fig. 6, the RMSE and correlation of the previous SFF methods and the proposed SFF-LS were plotted. For each level of variance, the RMSE and Correlation values were measured repeatedly for 30 times, and their mean values were plotted. Among the traditional methods, SFF-TEN were most robust against noise as the noise level increased. Therefore, the TEN focus measure was used to estimate the initial shape in SFF-FIS, SFF-DP including the proposed SFF-LS. The proposed SFF-LS method showed best robustness against noise in terms of both RMSE and correlation. Between SFF-FIS and SFF-DP, even though SFF-FIS has higher correlation value without noise, SFF-DP outperformed SFF-FIS in terms of both the RMSE and correlation as the noise level increased.

Since many of the focus measures are based on the first derivative (for example, TEN focus measure) and the second derivative (for example, Laplacian, the modified Laplacian, and SML focus measures), their performance degrades as the noise increases. The same problem is observed with the focus measure based on the variance of the pixels values. On the other hand, our approach is quite robust to the noise as shown by the almost

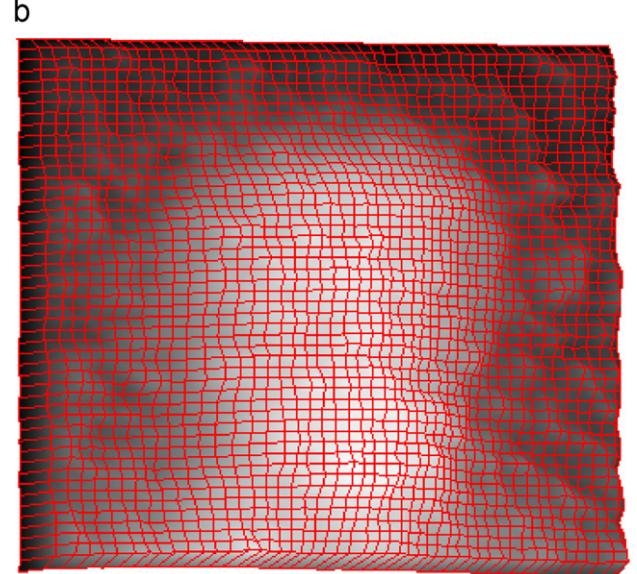
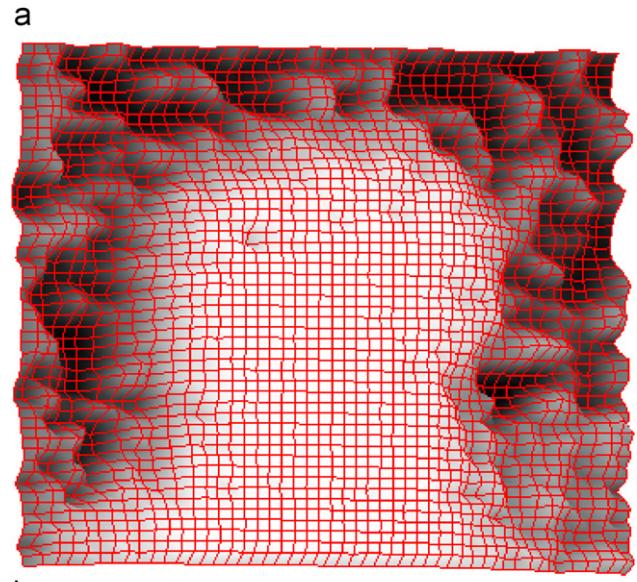


Fig. 10. Comparison between SFF-LS and SFD-DFU on Lincoln head part of US one cent coin: (a) SFD-DFU and (b) SFF-LS.

constant RMSE value in spite of increase in the variance of the noise. This is due to the fact that our approach incorporates various neighboring frames to calculate accurate focus level. The effect of the noise is reduced because of the iterative adjustments of the neighboring pixels values and the refinements of the center pixel value while computing focus measure.

4.3. Qualitative analysis

In Fig. 7, the initial shape and the final reconstructed shape of the proposed algorithm based on TEN focus measure was compared. The initial shape of the figure corresponds to the traditional SFF method. The number of iterations of the proposed algorithm was set to 3. The result shows that the proposed algorithm reduced the fluctuations or errors of the initial shape on overall area and is closer to the real shape.

The Comparison with other popular SFF methods which rely on first estimate of the shape, SFF-FIS and SFF-DP, were conducted for the real cone object and the TFT-LCD color filter, and their

results were shown in Figs. 8 and 9, respectively. For all these methods, TEN focus measure was applied in the initial estimates of the shape. For both objects, the proposed SFF-LS shows smoother surface in overall area and takes better shape in comparison to SFF-FIS and SFF-DP.

In Figs. 10 and 11, the proposed SFF algorithm was compared with the recent SFD algorithm based on diffusion (SFD-DFU) [4] on the embossed and engraved objects, respectively. It can be seen that the proposed SFF-LS produced more accurate shape with less fluctuations. While the SFD-DFU uses two images for the shape estimation, the SFF-LS uses multiple images. Therefore, the SFF-LS could generate more accurate shape and could express more fine details of the object. However, SFD algorithm can be more suitable for real time applications due to the advantages in both the image acquisition and computation times.

4.4. Computational complexity

SFF-FIS uses one of the traditional SFF methods (SFF-TR) for finding FIS which is the initial estimate for the shape of an object. For each point (i,j) , SFF-FIS searches every possible FIS by varying the position and slopes of the initial shape at that point. Therefore, if the number of arithmetic operations in SFF-TR is $O(TR)$, the number of operations in SFF-FIS is approximately $N_z N_p N_q N_w O(TR)$, where N_z is the number of steps in varying positions, N_p and N_q represent x - and y -slope, respectively, and N_w is the window size. This shows that SFF-FIS achieved improvement of quality of shape at a tremendous computational cost. SFF-DP reduced computation time to $3 \times O(TR)$ [14]. In the proposed algorithm, the computation to find intermediate image volume is insignificant since it does not involve any multiplication operation. Therefore, the number of operations in the proposed SFF-LS algorithm is approximately $O(TR)(1 + ((2b+1)/K)n)$, where $(2b+1)$ is the number of image frames in the intermediate image volume and n is the number of iterations. Therefore, the number of arithmetic operations in the proposed SFF-LS algorithm is lower than the previous SFF-FIS and SFF-DP methods.

5. Conclusions

In this paper, the SFF problem is presented as combinatorial optimization problem. To reduce the computational complexity, a local search algorithm is proposed. Based on the initial estimate of the object depth map, a temporary image volume is created by taking into account several image frames preceding and following the best focused image frames corresponding to the initial depth map. The depth map is updated by applying focus measure on the temporary image volume. Pixels on each image frames in the temporary image volume have approximately same depth. Therefore, applying focus measure operator on the newly formed image frames could produce more accurate focus level at each pixel. This update process of the depth map is repeated several times until the improvement of the depth map is negligible. Heuristically, it is found that three iterations of the algorithm is good tradeoff between the accuracy and the computational time.

The proposed algorithm was compared with the previous SFF methods based on FIS (SFF-FIS) and dynamic programming (SFF-DP). Moreover, the proposed algorithm was compared with the recent SFD algorithm based on diffusion (SFD-DFU). Experimental results show that the proposed technique achieves better quality of 3D reconstructed shape and lower computational time compared to SFF-FIS and SFF-DP. With respect to SFD-DFU, the proposed algorithm shows better performance in terms of accuracy with the disadvantage of requiring more image data.

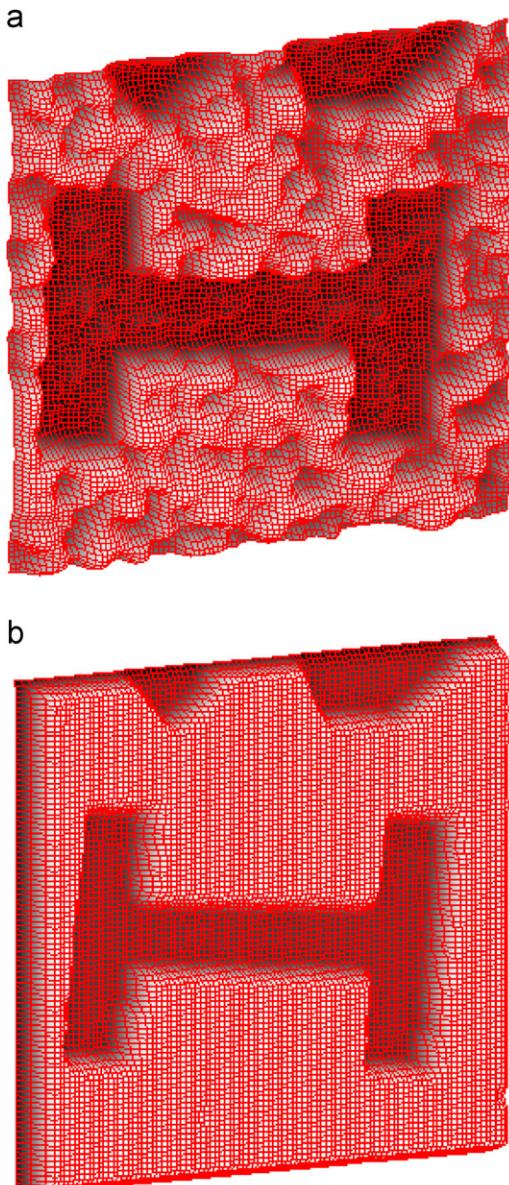


Fig. 11. Comparison between SFF-LS and SFD-DFU on letter 'H' object: (a) SFD-DFU and (b) SFF-LS.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MEST) (No. 2009-0083733).

References

- [1] A. Pentland, A new sense for depth of field, *IEEE Trans. Pattern Anal. Mach. Intell.* 9 (4) (1987) 523–531.
- [2] M. Subbarao, G. Surya, Depth from defocus: a spatial domain approach, *Int. J. Comput. Vision* 13 (3) (1994) 271–294.
- [3] P. Favaro, S. Soatto, A geometric approach to shape from defocus, *IEEE Trans. Pattern Anal. Mach. Intell.* 27 (3) (2005) 406–417.
- [4] P. Favaro, S. Soatto, M. Burger, S.J. Osher, Shape from defocus via diffusion, *IEEE Trans. Pattern Anal. Mach. Intell.* 30 (3) (2008) 518–531.
- [5] E.P. Krotkov, Focusing, *Int. J. Comput. Vision* 1 (3) (1987) 223–237.
- [6] S. Nayar, Y. Nakagawa, Shape from focus, *IEEE Trans. Pattern Anal. Mach. Intell.* 16 (8) (1994) 824–831.
- [7] M. Subbarao, T.-S. Choi, Accurate recovery of three dimensional shape from image focus, *IEEE Trans. Pattern Anal. Mach. Intell.* 17 (3) (1995) 266–274.
- [8] Y. Sun, S. Duthaler, B.J. Nelson, Autofocusing in computer microscopy: selecting the optimal focus algorithm, *Microsc. Res. Tech.* 65 (3) (2004) 139–149.
- [9] A.S. Malik, T.-S. Choi, A Novel algorithm for estimation of depth map using image focus for 3D shape recovery in the presence of noise, *Pattern Recognition* 41 (7) (2008) 2200–2225.
- [10] M. Subbarao, J.K. Tyan, Selecting the optimal focus measure for autofocusing and depth-from-focus, *IEEE Trans. Pattern Anal. Mach. Intell.* 20 (8) (1998) 266–274.
- [11] J.M. Tenenbaum, Accommodation in computer vision, Ph.D. Thesis, Stanford University, 1970.
- [12] T.S. Choi, J. Yun, Three-dimensional shape recovery from focused image surface, *Opt. Eng.* 39 (5) (2000) 1321–1326.
- [13] M. Asif, T.-S. Choi, Shape from focus using multilayer feedforward neural network, *IEEE Trans. Image Process.* 10 (11) (2001) 1670–1675.
- [14] M.B. Ahmad, T.-S. Choi, A heuristic approach for finding best focused shape, *IEEE Trans. Circuits Syst. Video Technol.* 15 (43) (2005) 566–574.
- [15] E. Lawler, in: *Combinatorial Optimization: Networks and Matroids*, Holt, Rinehart and Winston, New York, 1976.
- [16] C.H. Papadimitriou, K. Steiglitz, in: *Combinatorial Optimization: Algorithms and Complexity*, Prentice-Hall, Englewood Cliffs, NJ, 1982.
- [17] B. Korte, J. Vygen, in: *Combinatorial Optimization: Theory and Algorithms*, 2nd ed., Springer, Berlin, 2002.
- [18] B.W. Kernighan, S. Lin, An efficient heuristic procedure for partitioning graphs, *Bell Syst. Tech. J.* 49 (1972) 291–307.
- [19] S. Lin, B. Kernighan, An effective heuristic algorithm for the traveling salesman problem, *Oper. Res.* 21 (1973) 498–516.
- [20] R.K. Ahuja, O. Ergun, A.P. Punnen, A survey of very large-scale neighborhood search techniques, *Discrete Appl. Math.* 123 (2002) 75–102.
- [21] A.S. Malik, T.-S. Choi, Consideration of illumination effects and optimization of window size for accurate calculation of depth map for 3D shape recovery, *Pattern Recognition* 40 (1) (2007) 154–170.

About the Author—SEONG-O SHIM received the B.S. degree in electrical engineering from Ajou University, Suwon, Korea, in 1999, the M.S. degree in Mechatronics from Gwangju Institute of Science & Technology, Gwangju, Korea, in 2001.

From 2003 to 2007, he was with LG Electronics DTV Labs, Seoul, Korea, working on research and development of digital TV. He is currently doing research at Signal and Image Processing Lab in School of Information & Mechatronics at Gwangju Institute of Science and Technology, Korea. His research interests include image processing, 3D shape recovery, content based image retrieval (CBIR), medical imaging.

About the Author—TAE-SUN CHOI received the B.S. degree in electrical engineering from the Seoul National University, Seoul, Korea, in 1976, the M.S. degree in electrical engineering from the Korea Advanced Institute of Science and Technology, Seoul, Korea, in 1979, and the Ph.D. degree in electrical engineering from the State University of New York at Stony Brook, in 1993.

He is currently a Professor in the Department of Mechatronics at Gwangju Institute of Science and Technology, Gwangju, Korea. His research interests include image processing, machine/robot vision, and visual communications.