Development of New Optimized Sampling method for 3D Shape Recovery in the Presence of Noise

Hyeong-Geun Lee* , Hoon-Seok Jang**

Abstract Noise affects the accuracy of three-dimensional shape recovery. Its occurrence is unpredictable and depends on several mechanical, environmental, and other factors. When two-dimensional image sequences are obtained for shape from focus (SFF), mechanical vibration occurs in the translational stage, causing an error in the three-dimensional shape recovery. To address this issue, mechanical vibration is modeled using Newton’s second law and the principle of the rack and pinion gear. Then, an optimal sampling step size considering the mechanical vibration is suggested through theoretical demonstration. Experiments conducted with real objects verify the effectiveness of the proposed sampling step size. In this paper, in a realistic environment with noise, the potential of obtaining more accurate three-dimensional reconstruction results of the objects is explored by acquiring the optimal sampling step size, which improves the sampling step size relative to those reported in a previous study performed under similar conditions.

Key Words : Mechanical Vibration, Rack and Pinion Gear, Sampling Step Size, Shape from Focus.

1. Introduction

Shape from focus (SFF) is a three-dimensional (3D) shape recovery method. It obtains image sequences by moving the translational stage along the optical axis with a discrete sampling step size. The final depth map can then be obtained by finding the frame value with the best sharpness after applying a focus measure such as the sum of modified Laplacian (SML) for each pixel of obtained images [1-3]. However, this approach results in the loss of information between two contiguous frames. It is, therefore, difficult to find the exact depth values for few pixels.

Therefore, researchers have suggested various approximation methods to obtain a more improved depth map. The SFF method based on Gaussian interpolation provides the optimized depth value by performing Gaussian approximation of SML curve for each pixel [1]. Instead of conducting piecewise constant approximation on the surface of an object, shape from focus using focused image surface (SFF.FIS) based on piecewise planar approximation has been suggested [4]. To approximate the surface of the complex object, shape from focus using curved window (SFF.CW) has been proposed [5]. A 3D shape recovery method, which uses a multilayer feedforward neural network (SFF.NN) to obtain the optimized FIS, has been suggested [6]. This method is based on the description of 3D FIS with respect to neural network weights. This neural network is trained to learn the shape of the FIS, which
maximizes the focus measure. Because FIS requires heavy computation, shape from focus using principal component analysis (SFF.PCA) was recently proposed to replace FIS. It uses principal component analysis and transforms the data to the eigenspace \[7\]. In addition, proposed for this purpose were SFF.DP, which utilizes a dynamic programming method \[8\], SFF.BS, which approximates the object shape using the Bezier surface, \[9\], and SFF.LC, which approximates the SML curve by using a Lorentzian-Cauchy function \[8\].

When the 3D shape is recovered by the SFF methods, the methods require a long computation time on the account of oversampling, which minimizes the sampling step size, i.e., the gap between continuous image frames. To solve this problem, a method that provides an optimized sampling step size without considering mechanical vibration was proposed \[10\]. When a stack of two-dimensional (2D) images are obtained by moving the translational stage along the optical axis for the SFF methods, mechanical vibration occurs. Mechanical vibration is defined as random or periodic mechanical oscillations. This vibration changes the focus values of 2D image sequences and affects the accuracy of 3D shape recovery.

Therefore, in this study, the vibration that occurs when 2D image sequences are obtained by a microscope (hereafter, the referenced microscope is the Optiphot 100S) was modeled. The optimized sampling step size considering the modeled vibration is herein proposed.

2. Mechanical Vibration

Mechanical vibration is one of the main factors affecting the accuracy of 3D shape recovery. To closely analyze the mechanical vibration, it is necessary to model it. In SFF, the main assumption is that the points on the object plane and the corresponding pixels representing each point of the object in the image plane do not change. Thus, in this paper, we consider only the z-axis vibration.

In this section, we model the mechanical vibration that occurs when 2D image sequences are obtained by a microscope. The microscope follows the principle of the rack and pinion gear, which changes rotational motion to a straight-line motion. Thus, when torque is applied to pinion, the microscope stage moves with the rack. We therefore set equation and parameter values based on this principle for modeling.

First, the equation of motion is derived based on Newton’s second law, which states that “the sum of the external force acting on an object is equal to the product of mass and acceleration of the object.” This is expressed as \(1\).

\[
\sum_{i=1}^{n} F_i = m\alpha, \tag{1}
\]

where, \(F\) is the external force acting on the object, \(m\) is the mass, and \(\alpha\) is the acceleration factor. From Eq. (1), Eq. (2) is obtained as the equation of the mechanical vibration that occurs in the translation stage when 2D images are obtained by the microscope.
\[ F = m \ddot{z}(t) + c \dot{z}(t) + k z(t), \quad (2) \]

where \( z(t) \) is the location of the stage according to time, \( c \) is viscous damping properties, \( k \) is stiffness. Eq. (3) can be acquired when Eq. (2) is divided by mass, \( m \).

\[ \frac{F}{m} = \ddot{z}(t) + 2\zeta w_n \dot{z}(t) + w_n^2 z(t), \quad (3) \]

where, \( \zeta = \frac{c}{2mw_n} \) is the viscous damping factor and \( w_n = \sqrt{\frac{k}{m}} \) is the undamped natural frequency of oscillation [11] [12]. From \( w_n \), \( n \) denotes 'natural' of the undamped natural frequency.

To obtain the final equation of the mechanical vibration, parameters should be set according to the microscope.

First, assume stiffness \( k \) of an aluminum alloy, which is the stiffness of the rack and pinion. Stiffness of aluminum alloy has the value of 7.0 to 7.5 (\( \times 10^7 \text{kg/m} \)) [13]. Thus, it is herein assumed that \( k \) is 7.25 (\( \times 10^7 \text{kg/m} \)).

It is additionally assumed that mass \( m \) is 3kg, which is the sum of the stage, object, and rack, and the radius of the pinion is 35 mm.

Moreover, \( \zeta \) is the material damping; hence, its value is between 0.1 and 1. Thus, it is herein assumed that \( \zeta \) is 0.3.

The graph of the stage position according to various torques acting on the pinion in each step for experimented objects is represented in Figs. 1, 2, and 3. The settling time and peak time are defined to compute the amplitude of the mechanical vibration. The settling time is the amount of time in which the system reaches within 2% of the final value. The peak time is the amount of time in which the system reaches the maximum value. In Figs. 1, 2, and 3, the settling time and peak time of the system microscope used for this research are represented by using various sampling step sizes of experimented objects. The settling time represents the time in which the stage reaches the sampling step size. The difference between the stage position in the peak time and stage position in the settling time represents the mechanical vibration amplitude. Related equations are the same as Eqs. (4) and (5), [14].

\[ T_s = \frac{4}{\zeta w_n}, \quad (4) \]

\[ T_p = \frac{\pi}{w_n \sqrt{1 - \zeta^2}}, \quad (5) \]

where, \( T_s \) is the amount of time in which the system reaches within 2% of the final value and \( T_p \) is the amount of time in which the system reaches the maximum value. In this paper, the peak time is 0.21 s and the settling time is 0.86 s. From Figs. 1, 2, and 3, the mechanical vibration generally occurs within each sampling step size.

In the next section, we assume that 2D image sequences are obtained when the vibration amplitude is the greatest in the modeled vibration graph, and we obtain the optimal sampling step size for real objects.
3. Proposed Method

Previous research results obtained an optimal sampling step size without considering noise, especially mechanical vibration, as Eq. (6) [10].

\[
\Delta u_{\mu} = \frac{\lambda \sqrt{n^2 - NA^2}}{4\rho NA^2}, \tag{6}
\]

Here, \(n\) is the reflective index of the immersion medium, \(\lambda\) is the light wavelength, \(NA\) is the numerical aperture of the objective lens, and \(\rho\) is the constant, which determines the cut-off frequency.

In this section, we theoretically prove how the mechanical vibration modeled in the previous section affects the optimal sampling step size. For theoretical demonstration, the basic image formation geometry is shown in Fig. 4. In Fig. 4, \(u_p\) is the focused object distance, \(u_f\) is the farthest object distance as the defocused object distance, \(u_n\) is the nearest object distance as the defocused object distance, \(v_p\) is the focused image distance, \(v_f\) is the defocused image distance corresponding to \(u_f\), and \(v_n\) is the defocused image distance corresponding to \(u_n\). We make a depth map by calculating \(u_p\) for every
pixel. We can use the lens formula to calculate \( u_p \). The relationship between the object distance \( u_p \), lens focal distance \( f \), and focused image distance \( v_p \) is given by the Gaussian lens law as Eq. (7).

\[
\frac{1}{f} = \frac{1}{u_p} + \frac{1}{v_p},
\]

(7)

where, \( f \) is the focal length. To calculate the optimal sampling step size, we must know the depth of field (\( D_{\text{field}} \)), which is shown in Eq. (8).

\[
D_{\text{field}} = 2(m + 1)\left(\frac{Nc}{N^2c^2 - f^2m^2}\right)f^2,
\]

(8)

where, \( c \) is the blur circle, \( N \) is the lens f-number, and \( m \) is the magnification. The lens f-number can be approximated as Eq. (9) when the numerical aperture (\( NA \)) is small.

\[
N \approx \frac{1}{2NA}.
\]

(9)

Magnification is represented in terms of object distance \( u_p \) and focused image distance \( v_p \) as Eq. (10).

\[
m = \frac{v_p}{u_p}.
\]

(10)

In the previous paper showing the optimal sampling step size without considering vibration, the optimal sampling step size is represented in terms of the depth of field and \( \rho \) as Eq. (11).

\[
\Delta u = \frac{D_{\text{field}}}{4\rho}
\]

(11)

where \( \rho = -\log_\beta \). Here, we utilize \( \rho \) which is used in the previous paper (\( \rho = -3.388 \) for coin, \( \rho = -0.1774 \) for LCD-TFT, and \( \rho = -3.360 \) for letter-I) [8].

Because the parameters related to the camera are already known and is constant, the optimal sampling step size is only affected by blur circle \( c \) and magnification \( m \). If mechanical vibration occurs, object distance \( u_p \) is changed according to the amount of mechanical vibration. In the previous section, mechanical vibration occurs within 40% of each sampling step size. By using Eq. (10), magnification \( m \) affecting the optimal sampling step size is obtained. The object distance and image distance used in this experiment, as well as the maximal amplitude of mechanical vibration (approximately 40% of each sampling step size) in the case of the previously optimal sampling step size, are utilized in Eqs. (12) and (13).

\[
m = \frac{v_p}{u_p} = \frac{65}{6.5} = 10,
\]

(12)

\[
m = \frac{v_p}{u_p} = \frac{65}{6.5 - 0.3167 \times 10^{-3}} = 10.00048725
\]

(13)

For the coin object, working distance of the microscope used in this experiment is 6.5 mm. The stack of 2D images is obtained under 10 \( \times \) magnification and the maximal amplitude of mechanical vibration is \( 0.3167 \times 10^{-3} \) as Eqs. (12) and (13).

The circle of confusion (CoC), which is another factor affecting the optimal sampling step size, can be calculated by using the PhotoPills application because it is difficult.
to manually calculate CoC [15]. By entering the values affecting CoC into the CoC calculator, CoC can be acquired to show the effect of different viewing distances. Based on the parameters utilized in our experiments, CoC is 0.03 mm regardless of whether the vibration is considered.

By using the fact that the numerical aperture for the coin is 0.3 and Eq. (11), the new optimized sampling step size is 813.2691 nm compared to the previous optimized sampling step size of 813.3123 nm. Thus, 52 images are used for 3D shape recovery. For LCD-TFT and letter-I objects, the 2D image sequence of LCD-TFT is acquired under 50 × magnification and the 2D image sequence of letter-I is obtained under 10 × magnification. For 50 × magnification, the working distance \( u_w \) of our microscope used in this experiment is 0.54 mm. The numerical aperture for LCD-TFT is 0.8. Based on these values, new optimized sampling step sizes are acquired in the same way as described above. The new optimized sampling step size for LCD-TFT is 107.7987 nm compared to the previously optimized sampling step size of 107.8072 nm. For letter-I, the new optimized sampling step size is 820.0460 nm compared to the previously optimized sampling step size of 820.0899 nm. For 3D shape recovery, 61 and 12 frames are utilized in LCD-TFT and letter-I, respectively.

In Section 4, we compare the depth map obtained by using the previously optimized sampling step size with the depth map acquired by utilizing the new optimized sampling step size in the presence of noise. The focus curves obtained by using SML as one of the focus measure operators are modeled by quadratic approximation in the transformed domain.

4. Experimental Results

Various experiments were performed using real objects to prove the effectiveness of the optimal sampling step size. Here, 52, 61, and 12 frames were used in coin, LCD-TFT, and letter-I image sequences, respectively. The dimensions of these image sequences were 300 by 300 pixels. These images were obtained by the SFF system, which varied the object plane with a predetermined sampling step size for acquiring 2D image sequences with different focus values [7]. The image sequences of coin and letter-I were obtained under 10 × magnifications. The image sequence of LCD-TFT was acquired under 50 × magnification. Thus, the numerical aperture for the coin and letter-I was 0.3 and for LCD-TFT, and the numerical aperture was 0.8. The images for experiments are shown in Fig. 5.

![Fig. 5. 30th frame of experimented objects: (a) coin, (b) LCD-TFT, (c) letter-I.](image)

To show the improved quality of 3D shape recovery in the presence of mechanical vibration by using the new optimized sampling step size, we employed the root mean square error (RMSE). RMSE shows the error of the 3D shape recovered by SFF techniques by
Development of New Optimized Sampling method for 3D Shape Recovery in the Presence of Noise

comparing it to the original shape and correlation, which shows the similarity between the original shape and the recovered shape.

Figs. 7, 8 and 9 provide the 3D shape recovery results of real objects using various SFF techniques with the previously optimized sampling step size and the new optimized sampling step size.

It is clear from these figures and tables that recovering the 3D shape with the new optimized sampling step size improves the accuracy of 3D shape reconstruction in the presence of mechanical vibration.

Table 2 RMSE and correlation of 3D shape recovery results by using various SFF techniques for the LCD-TFT object

Table 3 RMSE and correlation of 3D shape recovery results by using various SFF techniques for the letter-I object

Since the original shape does not exist for real objects, 3D shapes recovered by SFF.LC were used as the original shapes, as shown in Fig. 6. SFF.LC provides the best 3D shape recovery results among the SFF techniques utilized in this study.

Tables 1, 2 and 3 show the RMSE and correlation of 3D shape recovery results in the presence of the mechanical vibration maximal amplitude by using SFF techniques.
Fig. 7 3D shape recovery results of coin using various SFF techniques in the presence of noise (maximal amplitude of vibration: 0.3167μm): SFF.TR (a), (b); SFF.BS (c), (d); SFF.DP (e), (f); previous optimized sampling step size (a), (c), (e); new optimized sampling step size (b), (d), (f).

Fig. 8 3D shape recovery results of LCD-TFT using various SFF techniques in the presence of noise (maximal amplitude of vibration: 0.0419μm): SFF.TR (a), (b); SFF.BS (c), (d); SFF.DP (e), (f); previous optimized sampling step size (a), (c), (e); new optimized sampling step size (b), (d), (f).

Fig. 9 3D shape recovery results of letter-I using various SFF techniques in the presence of noise (maximal amplitude of vibration: 0.319μm): SFF.TR (a), (b); SFF.BS (c), (d); SFF.DP (e), (f); previous optimized sampling step size (a), (c), (e); new optimized sampling step size (b), (d), (f).

5. Conclusion

In this study, we modeled the mechanical vibration that can occur in a microscope (Nikon Optiphot 100S) and acquired an optimal sampling step size in the presence of modeled mechanical vibration.

Experimental results demonstrated that the proposed sampling step size provided better performance of 3D shape recovery results compared to the previously optimized sampling step size in the presence of mechanical vibration in terms of RMSE and correlation, which are 3D shape recovery performance indicators. It was possible to obtain a new optimized and distinct sampling step size according to the proposed method of noise modeling.
REFERENCES


Author Biography

Hyeong-Geun Lee [Regular Member]

- 2015. 02: Social Welfare, Hansel University (B.S.)
- 2018. 10 ~ 2019.05: Seodaeman Disabled Family Support Center(Staff)
- 2019. 06 ~ Present: Lifelong Education Department, Nowon lifelong education center for people with developmental disabilities (Staff)

Interests: Machine Learning, Deep Learning, Signal & Image Processing
Hoon-Seok Jang  [Regular Member]

- 2014. 08: Mechatronics Engineering, Gwangju Institute of Science and Technology (M.S.)
- 2019. 02: Mechatronics Engineering, Gwangju Institute of Science and Technology (Ph.D.)
- 2020. 02: Center for Imaging Media Research, Korea Institute of Science and Technology (Post-Doc)
- 2020. 02 ~ Present: IT Application Research Center, Jeonbuk Regional Branch, Korea Electronics Technology Institute (Senior Researcher)

[Interests] Signal & Image Processing, 3D Shape Recovery, Machine Learning, Deep Learning