Removal of Non-Gaussian Jitter Noise for Shape From Focus Through Improved Maximum Correntropy Criterion Kalman Filter

HOON-SEOK JANG$^{\dagger}$, MANNAN SAEED MUHAMMAD$^2$, AND MIN-KOO KANG$^3$

$^1$IT Application Research Center, Jeonbuk Regional Branch, Korea Electronics Technology Institute, Jeonju 54853, South Korea

$^2$College of Information and Communication, Sungkyunkwan University, Suwon 16419, South Korea

$^3$Center for Imaging Media Research, Korea Institute of Science and Technology, Seoul 02792, South Korea

Corresponding author: Min-Koo Kang (minkoo@kist.re.kr)

This work was supported in part by the Civil-Military Technology Cooperation Program of the Institute of Civil-Military Technology Cooperation (ICMTC) under Project 17-CM-DP-29, and in part by the Korea Institute of Science and Technology (KIST) Institutional Program under Project 2E29450.

ABSTRACT

Three-dimensional (3D) shape reconstruction from one or multiple observations is a primary problem of computer vision. Shape from Focus (SFF) is a passive optical method that uses multiple two-dimensional (2D) images with different focus levels. When obtaining 2D images in each step along the optical axis, mechanical vibrations, referred as jitter noise, occur. SFF techniques are vulnerable to jitter noise that can vary focus values in 2D images. In this paper, new filtering method, which provides high accuracy of 3D shape reconstruction and low computational cost, is proposed. First, jitter noise is modeled as Lévy distribution. This assumption makes it possible to show the influence of proposed filtering method in real environment with non-Gaussian jitter noise. Second, focus curves are modeled as Gaussian function to compare the performance of proposed filtering method with those of the conventional filtering methods. Finally, improved maximum correntropy criterion Kalman filter (IMCC-KF) is designed as a post-processing step, and is applied to the modeled focus curves. The experiments are performed on real and synthetic objects and the results demonstrate the effectiveness of proposed method.

INDEX TERMS

Focus curve, improved maximum correntropy criterion Kalman filter (IMCC-KF), jitter noise, non-Gaussian distribution, shape from focus (SFF).

I. INTRODUCTION

Obtaining the 3D shape of an object from 2D images is a fundamental purpose of research in computer vision. Many 3D shape reconstruction methods, using various cues, have been proposed, which include focus, [1], [2], defocus, [3], [4], texture, [5], [6], stereo, [7], [8], and motion, [9], [10]. Three dimensional shape recovery methods based on focus are important due to their low computational cost and easy implementation. Particularly, these methods have many advantages over other 3D shape recovery methods utilizing other cues, as those encounter correspondence problems.

In Shape from Focus (SFF), the object placed on a translational stage, is moved at a constant step size along the optical axis, for capturing a 2D image sequence with different focus settings, [11]. This is followed by a focus measure (FM) operator applied to each 2D image, for computing focus quality of each pixel in acquired image sequence. Some famous FMs include modified Laplacian (ML), sum of modified Laplacian (SML), gray level variance (GLV), and Tenenbaum (TEN), [12]. Sparse depth map is obtained by finding the maximum focus value and corresponding frame number, of each pixel, along the optical axis. This depth map may have many errors due to information loss between consecutive frames. To minimize these errors, many approximation methods have been proposed by using a variety of optimization techniques, such as dynamic programming (DP), [13], and Bezier surface approximation (BS), [14]. Multiple SFF techniques have been applied to different fields including depth estimation, robot navigation, camera auto-focusing, 3D camera, comparison of polymers, and LCD/TFT displays manufacturing, [15]–[17]. However, a fundamental problem of mechanical vibrations, referred as jitter noise, occurring in each image acquisition step, affects the accuracy of SFF.
methods, as shown in Fig. 1. This jitter noise is different from image noise, and it changes focus values along the optical axis due to random vibrations, thus, degrading the performance of 3D shape recovery. Filtering methods, such as Kalman filter, [18], Bayes filter, [19], particle filter, [20], modified Kalman filter, [21], and adaptive neural network filter, [22], have been proposed for removing this type of jitter noise. Since the conventional filters use minimum mean squared error (MMSE) criterion as a cost function, they only capture the second-order statistics of the error signal. This results in poor performance of state estimation in the presence of non-Gaussian jitter noise with large outliers, [23]. Correntropy criterion has received increasing interest in the field of machine learning, pattern recognition, and signal processing over the years, [24]–[26]. Correntropy criterion is defined as a generalized similarity measure between observed output and estimated output. It captures higher-order statistics of the error signal than the MMSE criterion, and provides significant performance improvement for state estimation in the field of non-Gaussian noise signal processing, [23]. Also, maximum correntropy criterion Kalman filter (MCC-KF) removes the non-Gaussian noise effectively by maximizing the correntropy, [27]. However, MCC-KF has problems of numerical instability and computational complexity in the covariance matrix and Kalman gain, which provides high estimation quality and low computational cost.

II. RELATED WORK

A. SHAPE FROM FOCUS

The goal of SFF is to compute the depth of all object points in the scene, using focus cues in the images sequence, [6]. To get a depth map, this sequence of images is obtained by moving an unknown object on the translational stage, towards or away from the optical lens, at a constant step size. For SFF, it is required to have constant magnification and the depth of field of the image system to be shallow, [29]. The basic geometry of image construction is shown in Fig. 2.

The light rays from object point \( P \), pass through the optical lens, and are converged to a specific image point. If the image point \( P' \) is on the image detector, a defocused image is obtained, which is a blur circle with radius \( R \). If the image point \( P' \) lies on the image detector, a focused image is formed. When light rays from object point are converged to the focused image point, Gaussian lens formula can be employed to get the depth of object:

\[
\frac{1}{f} = \frac{1}{u} + \frac{1}{v},
\]

where, \( f \) is the focal length, \( u \) is the object distance from the optical lens, and \( v \) is the image distance from the optical lens. Optimal 3D shape of the object with the best focus quality is acquired by computing the object distance \( u \) for all object points. To find the best focused image points for all pixels in obtained image sequence, an FM operator needs to be applied to each 2D image. In next section, some of the famous FM operators are discussed.

B. FOCUS MEASURE (FM) OPERATORS

A focus measure (FM) operator is used to find the focus quality of each pixel in an image. Many FM operators have been proposed in literature, [12], [30]–[33], and can be distinguished into six groups according to their working principle—gradient based operators, Laplacian based operators, wavelet based operators, discrete cosine transform (DCT) based operators, statistics based operators, and
miscellaneous operators. Gradient based operators, such as Tenenbaum (TEN), compute the focus quality by using first derivative. Laplacian based operators, such as modified Laplacian (ML), measure the degree of focus by utilizing second derivatives. To recover weak-textured surfaces, sum modified Laplacian (SML) can be employed by summing ML values in a local window. Wavelet based and DCT based operators estimate the amount of edges by using discrete wavelet transform and discrete cosine transform, respectively. Statistics based operators get the focus values through several image statistics, e.g. gray level variance (GLV) utilizes the fact that the variance of gray-level in a sharp image is higher than that in a blurred image. FM operators, which do not belong to any of previous groups, are classified into miscellaneous operators [34].

After the application of one of these FM operators, sparse depth map is obtained, which has errors due to discrete-ness between consecutive frames or several types of noise. In next section, some approximation methods for reducing these errors are introduced.

C. APPROXIMATION METHODS

To improve the accuracy of initial depth map, many approximation methods have been proposed in literature. Subbarao and Choi, proposed a concept of focused image surface (FIS) to overcome constant approximation, [35]. FIS, based on planar approximation, is defined as the surface formed by the set of points at which the object points are focused by a camera lens. However, this method has high computational complexity. Bilal and Choi proposed a SFF method based on DP to reduce the computational time of FIS, [13]. Muhammad and Choi proposed the use of Bezier surface to interpolate initial depth map over a small patch on the object surface, [14]. Muhammad and Choi also proposed a focus curve fitting method based on Lorentzian-Cauchy function to improve the performance of conventional focus curve fitting methods, such as Gaussian interpolation method, [29]. Jang et al. proposed the use of Kalman filter to remove Gaussian jitter noise, [18]. To reflect the real environment of SFF, Jang et al. proposed utilization of modified Kalman filter to improve the performance of conventional Kalman filter in the presence of non-Gaussian jitter noise, [21]. In order to provide better performance than modified Kalman filter, in terms of accuracy of state estimation, Lee et al. proposed a new filtering method based on adaptive neural network, [22]. Chen et al. proposed a 3D shape reconstruction method using maximum correntropy Kalman filter, [27], where a maximum correntropy Kalman filter is utilized for removing the noise related to the surface texture of the object. In next section, jitter noise is modeled as non-Gaussian function for applying improved maximum correntropy Kalman filter as proposed method.

III. NOISE MODELING

In SFF, when 2D images are obtained, mechanical vibrations occur in each step, also referred as jitter noise. The jitter noise is modeled as Lévy symmetric stable function for reflecting the real environment of SFF, [21], [36]–[38]. Lévy symmetric stable distribution has two parameters, one of which is stability index \( \alpha \) and the other is scale parameter \( c \). These parameters have the range of values as \( 0 < \alpha \leq 2 \), and \( c > 0 \). The Fourier transform relating the probability density function is given by:

\[
p(\Delta z_n) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp \left(-ix\Delta z_n - |cx|^\alpha \right) \, dx, \tag{2}
\]

where, \( \Delta z_n \) represents the degree of change of the position of each image frame due to jitter noise, modeled by Lévy symmetric stable function. The graphical representation of (2) is shown in Fig. 3. In this manuscript, \( \alpha \) and \( c \) are taken as 1.3 and 10, respectively.

In next section, focus curves are modeled as Gaussian function, together with the Lévy noise. The performance of proposed method with conventional methods is then compared.
IV. FOCUS CURVE MODELING

To consider the effect of Lévy noise, the focus curve is modeled as Gaussian function with mean $\bar{z}_f$ and standard deviation $\sigma_f$, shown in Fig. 4.

Gaussian modeled focus curve is represented as:

$$F(z) = F_p \cdot \exp \left( -\frac{1}{2} \left( \frac{z - \bar{z}_f}{\sigma_f} \right)^2 \right),$$  \hspace{1cm} (3)

where, $z$ is the position of each image frame without jitter noise, $F(z)$ is focus value at $z$, and $F_p$ is amplitude of Gaussian focus curve. $\bar{z}_f$, $\sigma_f$, and $F_p$ need to be computed for finding out $F(z)$. By applying natural logarithm to (3), the following is obtained:

$$\ln(F(z)) = \ln(F_p) - \frac{1}{2} \left( \frac{z - \bar{z}_f}{\sigma_f} \right)^2.$$  \hspace{1cm} (4)

After applying FM (like SML), conventional best focused position $z_{cf}$ is obtained. The positions $z_{cf-1}$ and $z_{cf+1}$ are on both sides of $z_{cf}$. Their corresponding focus values are $F_{cf-1}$, and $F_{cf+1}$, and are obtained as follows:

$$\ln(F_{cf}) - \ln(F_{cf-1}) = -\frac{1}{2} \left( \frac{(z_{cf} - \bar{z}_f)^2 - (z_{cf-1} - \bar{z}_f)^2}{\sigma_f^2} \right),$$  \hspace{1cm} (5)

$$\ln(F_{cf}) - \ln(F_{cf+1}) = -\frac{1}{2} \left( \frac{(z_{cf} - \bar{z}_f)^2 - (z_{cf+1} - \bar{z}_f)^2}{\sigma_f^2} \right).$$  \hspace{1cm} (6)

Using (6), (7) is acquired as:

$$\frac{1}{\sigma_f^2} = \frac{\ln(F_{cf}) - \ln(F_{cf+1})}{\ln(F_{cf}) - \ln(F_{cf-1})}.$$  \hspace{1cm} (7)

Applying (7) to (5), (8) is obtained:

$$\ln \left( \frac{(z_{cf} - \bar{z}_f)^2 - (z_{cf-1} - \bar{z}_f)^2}{(z_{cf} - \bar{z}_f)^2 - (z_{cf+1} - \bar{z}_f)^2} \right).$$  \hspace{1cm} (8)
Assuming step size is equal to 1 as $z_{cf+1} - z_{cf} = z_{cf} - z_{cf-1} = 1$ and using (8):

Utilizing (7) and (9), as shown at the bottom of next page, (10) is obtained as:

$$
\sigma_f^2 = -\frac{1}{2} \left( \frac{2}{(\ln (F_{cf})-\ln (F_{cf-1}))) + (\ln (F_{cf})-\ln (F_{cf+1}))} \right).
$$

Using (3), (9), and (10), $F_p$ is acquired as follows:

$$
F_p = \frac{F_{cf}}{\exp \left( \frac{1}{2} \left( \frac{z_f - z_f^2}{\sigma_f^2} \right)^2 \right)}.
$$

By substituting (9), (10), (11) into (3), Gaussian focus curve without jitter (Lévy) noise is obtained, [19]. Previously modeled jitter noise is added to (3) as:

$$
F_n (z) = F_p \cdot e^{\frac{1}{2} \left( \frac{z+\Delta z_n - z_f}{\sigma_f} \right)^2},
$$

where, $\Delta z_n$ is jitter noise. In next section, improved maximum correntropy Kalman filter, as proposed method, is designed and applied to modeled focus curve for removing the jitter noise.

V. PROPOSED METHOD

Various filtering techniques have been utilized for removing jitter noise, [18]–[22]. Conventional filters are linear filtering methods using minimum mean square error (MMSE) criterion. They capture only the second order statistics of the error signal such as jitter noise; hence, resulting in poor performance of state estimation in the presence of non-Gaussian jitter noise with large outliers, [23]. Maximum correntropy criterion has received much attention in many fields such as machine learning and signal processing, [24]–[26]. Maximum correntropy criterion maximizes a generalized similarity measure between observed output and estimated output. It can capture second and higher order statistics of the error signal and has provided significant
H.-S. Jang et al.: Removal of Non-Gaussian Jitter Noise for SFF Through IMCC-KF

FIGURE 8. Comparisons of Gaussian focus curves optimized by various filters: (a) Simulated cone (50, 50), (b) LCD-TFT filter (100, 100), (c) Coin (150, 150), (d) Letter-I (200, 200).

performance improvement for state estimation in the field of non-Gaussian noise signal processing, [23]. It has been demonstrated that maximum correntropy criterion Kalman filter (MCC-KF) is effective for removing non-Gaussian noise with heavy tails, [27]. However, MCC-KF has the problems of the numerical instability and computational complexity in the covariance matrix and Kalman gain. To solve the problems, improved maximum correntropy Kalman filter (IMCC-KF) is proposed in this manuscript. In order to apply proposed filter, the system state is defined as position of each image frame in image sequence, and can be changed by the jitter noise. In IMCC-KF algorithm, jitter noise is taken as measurement noise, and is removed through following three steps. First, there is initialization for state vector and covariance as:

$$X_{ik} = Z_{ik},$$  
$$P_{ik} = R_{ik},$$  

where, $X_{ik}$ is “state vector estimate”, $Z_{ik}$ is “observation vector”, $P_{ik}$ is “covariance of state vector estimate”, and $R_{ik}$ is “covariance of measurement noise”. For SFF system,

$$Z_f = \frac{(\ln(F_{cf}) - \ln(F_{cf+1})) \left( z_{cf}^2 - z_{cf-1}^2 \right) - (\ln(F_{cf}) - \ln(F_{cf-1})) \left( z_{cf}^2 - z_{cf+1}^2 \right)}{2 \left( (\ln(F_{cf}) - \ln(F_{cf-1})) + (\ln(F_{cf}) - \ln(F_{cf+1})) \right)}$$  

(9)
$X_{ik}$ represents estimated position of each image frame in image sequence by IMCC-KF, $Z_{ik}$ represents position of each image frame before filtering, and $R_{ik}$ represents variance of jitter noise. In this paper, $R_{ik}$ is set as $10^4$ as an alternative to infinite value for Lévy distribution.

Second, there is time update for state vector and covariance as:

$$X_{ik+1} = A_{ik} \cdot X_{ik},$$
$$P_{ik+1} = A_{ik} \cdot P_{ik} \cdot A_{ik}^T + Q_{ik},$$

where, $A_{ik}$ is “state transition matrix” and $Q_{ik}$ is “covariance of process noise”. In this manuscript, $A_{ik}$ is set to 1 by default. Since only jitter noise is considered in this manuscript, $Q_{ik}$ is set as 0.

Finally, there is measurement update for state vector and covariance. The Kalman gain needs to be computed as:

$$\lambda_{ik} = G_{i}(\parallel Z_{ik} - B_{ik} \cdot X_{ik} \parallel_{R_{ik}}^{-1})$$
$$= \exp \left( \frac{\left( \parallel Z_{ik} - B_{ik} \cdot X_{ik} \parallel_{R_{ik}}^{-1} \right)^2}{2 \cdot \sigma^2} \right),$$
$$K_{ik} = \lambda_{ik} \cdot P_{ik} \cdot B_{ik}^T \cdot \left( \lambda_{ik} \cdot B_{ik} \cdot P_{ik} \cdot B_{ik}^T + R_{ik} \right)^{-1},$$

where, $I$ is “identity matrix”. IMCC-KF algorithm is repeated $N$ times, which is set to 100. After implementing IMCC-KF algorithm, for each image frame in the image sequence, a filtered image sequence is obtained. After that, optimal depth map is acquired by applying a FM and some approximation methods to the filtered image sequence.

Table 1 provides the summary of MCC-KF algorithm, whereas, the summary of IMCC-KF algorithm is shown in Table 2. Computing Kalman gain and updating the variance of positions of image frames in Table 1 require four matrices’ inversions and are therefore, computationally expensive process. This implementation makes maximum correntropy criterion Kalman filter (MCC-KF) impractical due to the numerical instability induced by round-off errors in matrices’ inversions, and computational complexity. However, IMCC-KF requires only one inversion of the matrix in computing Kalman gain, and uses computationally simpler expression for updating the variance of positions of image frames.
This modification results in numerical stability and computational simplicity of IMCC-KF. A list of frequently used symbols and notations is shown in Table 3.

VI. RESULTS AND DISCUSSION

A. DATA SETS

In order to demonstrate the effectiveness of proposed filter, four objects are used for experiments, as shown in Fig. 5. One is simulated cone, a synthetic object, shown in Fig 5(a). The image sequence of simulated cone has 97 images, each with the dimensions of 360×360 pixels. These images are generated by camera simulation software, [39]. Fig 5(b) shows a liquid crystal display-thin film transistor (LCD-TFT) filter with microscopic images of an LCD color filter. This image sequence has 60 images, each with the dimensions of 300×300 pixels. Fig 5(c) shows the coin sequence, with magnified images of Lincoln’s head from the back of a US penny with 80 images in the sequence, each with the dimensions of 300×300 pixels. Final one is letter-I engraved in the metallic surface, shown in Fig 5(d). This image sequence has 60 images, each with the dimensions of 300×300 pixels. LCD-TFT filter, coin, and letter-I, as real objects, are acquired through a microscopic control system (MCS), [40]. MCS consists of a personal computer integrated with a frame grabber board (Matrox Meteor-II) and a CCD camera (SAMSUNG CAMERA SCC-341) mounted on a microscope (NIKON OPTIPHOT-100S), [29]. Computer software acquires the real object images by controlling the object position through a stepper motor driver (MAC 5,000), possessing a 2.5nm minimum step length. LCD-TFT filter images are obtained under 50× magnification, while coin and letter-I images are acquired under 10× magnification.

B. COMPARATIVE ANALYSIS

For experiments, various parameters need to be set. Since jitter noise with Lévy distribution is used in this manuscript, the standard deviation of jitter noise is assumed to be 10^4 as an alternative to infinite value. In order to obtain focus values
TABLE 1. Maximum correntropy criterion Kalman filter algorithm.

| Algorithm 2 | Computing filtered position of each image frame and remaining jitter noise |
|-------------|--------------------------------------------------------------------------------|
| 1: procedure Filtered position $X_{mk}$ & remaining jitter noise $\zeta$ |
| 2: $X_{mk} \leftarrow Z_{mk}$ $\triangleright$ Set initial position of image frame with unfiltered position |
| 3: $P_{mk} \leftarrow R_{mk}$ $\triangleright$ Set initial variance of position of image frame with variance of jitter noise |
| 4: for $i = 1 \rightarrow N$ do $\triangleright$ Total number of iterations of MCC-KF |
| 5: $\lambda_{mk} \leftarrow G_{\phi} \left( \|Z_{mk} - B_{mk} \cdot X_{mk}\|_{z_{mk}} \right)^{-1}$ $\triangleright$ Compute scalar term for Kalman gain |
| 6: $K_{ik} \leftarrow \lambda_{mk} \cdot (P_{mk}^{-1} + \lambda_{mk} \cdot B_{ik} \cdot R_{ik} \cdot B_{ik}^{T})^{-1} \cdot B_{ik}^{T} \cdot R_{mk}^{-1}$ $\triangleright$ Compute Kalman gain |
| 7: $P_{mk+1} \leftarrow (I - K_{ik} \cdot B_{ik}) \cdot P_{mk} \cdot (I - K_{ik} \cdot B_{ik})^{T} + K_{ik} \cdot B_{ik} \cdot K_{ik}^{T}$ $\triangleright$ Update variance of position of image frame |
| 8: $X_{mk+1} \leftarrow X_{mk} + K_{ik} \cdot (Z_{mk} - B_{mk} \cdot X_{mk})$ $\triangleright$ Update position of image frame |
| 9: $\zeta \leftarrow |X_{mk+1} - z|$ $\triangleright$ Compute remaining jitter noise |
| 10: end for |
| 11: end procedure |

TABLE 2. Improved maximum correntropy criterion Kalman filter algorithm.

| Algorithm 1 | Computing filtered position of each image frame and remaining jitter noise |
|-------------|--------------------------------------------------------------------------------|
| 1: procedure Filtered position $X_{ik}$ & remaining jitter noise $\zeta$ |
| 2: $X_{ik} \leftarrow Z_{ik}$ $\triangleright$ Set initial position of image frame with unfiltered position |
| 3: $P_{ik} \leftarrow R_{ik}$ $\triangleright$ Set initial variance of position of image frame with variance of jitter noise |
| 4: for $i = 1 \rightarrow N$ do $\triangleright$ Total number of iterations of IMCC-KF |
| 5: $\lambda_{ik} \leftarrow G_{\phi} \left( \|Z_{ik} - B_{ik} \cdot X_{ik}\|_{z_{ik}} \right)^{-1}$ $\triangleright$ Compute scalar term for Kalman gain |
| 6: $K_{ik} \leftarrow \lambda_{ik} \cdot P_{ik} \cdot B_{ik}^{T} \cdot (\lambda_{ik} \cdot B_{ik} \cdot P_{ik} \cdot B_{ik}^{T} + R_{ik})^{-1}$ $\triangleright$ Compute Kalman gain |
| 7: $P_{ik+1} \leftarrow (I - K_{ik} \cdot B_{ik}) \cdot P_{ik}$ $\triangleright$ Update variance of position of image frame |
| 8: $X_{ik+1} \leftarrow X_{ik} + K_{ik} \cdot (Z_{ik} - B_{ik} \cdot X_{ik})$ $\triangleright$ Update position of image frame |
| 9: $\zeta \leftarrow |X_{ik+1} - z|$ $\triangleright$ Compute remaining jitter noise |
| 10: end for |
| 11: end procedure |

in each pixel of 2D image frames, SML with $7 \times 7$ window is utilized. For performance comparison, Kalman filter (KF), modified Kalman filter (MKF), adaptive neural network filter (ANNF), and maximum correntropy criterion Kalman filter (MCC-KF) are employed.

The initial values of the corresponding variables in Table 1 and 2 are same. The number of iterations of the filters including IMCC-KF is set to 100. Fig. 6 shows the performance of various filters in the 80th frame of simulated cone for various iterations. Fig. 7 shows the performance of various filters in the specific frames of simulated cone for 100th iteration. True position is the position which is not affected by the jitter noise and KF, MKF, ANNF, MCC-KF, and IMCC-KF are the filtered positions using KF, MKF, ANNF, MCC-KF, and IMCC-KF, respectively. It is clear from Fig. 6 and 7 that IMCC-KF converges better to True position than KF, MKF, ANNF, and MCC-KF. Fig. 8 provides the comparisons of Gaussian focus curves which are optimized by these filters in specific image positions. Without noise is ideal Gaussian focus curve which is not affected by jitter noise and KF, MKF, ANNF, MCC-KF, and IMCC-KF are Gaussian focus curves which are optimized by utilizing KF, MKF, ANNF, MCC-KF, and IMCC-KF, respectively. It is evident from Fig. 8 that optimal position with the highest focus values in IMCC-KF is closer to the optimal position in Without noise than optimal positions in KF, MKF, ANNF, and MCC-KF. For performance comparison of 3D shape recovery of simulated cone, true-depth-map is shown in Fig. 9. Table 4, 5, and 6 provide the quantitative comparison of 3D shape recovery of simulated cone by utilizing SML, GLV, and TEN, before and after filtering in terms of Root Mean Square Error (RMSE), correlation, and Peak-Signal-to-Noise-Ratio (PSNR), respectively. SML, GLV, and TEN are the most frequently used focus measure operators. RMSE, correlation, and PSNR are commonly utilized performance metrics, [41].

To obtain RMSE, Mean Square Error (MSE) needs to be computed. MSE represents the average of square of the error between the true and the estimated depth map. RMSE can be
TABLE 5. Quantitative comparison of 3D shape recovery of simulated cone by utilizing focus measure operators before and after filtering in terms of correlation.

|                | SML   | GLV   | TEN   |
|----------------|-------|-------|-------|
| Before filtering | 0.2940 | 0.3999 | 0.6372 |
| KF             | 0.8489 | 0.8262 | 0.8047 |
| MKF            | 0.8642 | 0.8756 | 0.8727 |
| ANNF           | 0.9545 | 0.9486 | 0.9267 |
| MCC-KF         | 0.9556 | 0.9491 | 0.9279 |
| IMCC-KF        | 0.9679 | 0.9718 | 0.9703 |

TABLE 6. Quantitative comparison of 3D shape recovery of simulated cone by utilizing focus measure operators before and after filtering in terms of PSNR.

|                | SML    | GLV    | TEN    |
|----------------|--------|--------|--------|
| Before filtering | 11.9293 | 12.9479 | 15.0102 |
| KF             | 18.3044 | 18.4356 | 20.1438 |
| MKF            | 21.0202 | 20.8249 | 20.3557 |
| ANNF           | 21.6528 | 21.5180 | 20.9587 |
| MCC-KF         | 22.0833 | 21.7202 | 21.1388 |
| IMCC-KF        | 22.2247 | 22.6630 | 22.6272 |

computed by adding the root operation as:

$$\text{RMSE} = \sqrt{\frac{1}{I \times J} \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \left( d(i,j) - \hat{d}(i,j) \right)^2}, \quad (22)$$

where, $I \times J$ is size of each 2D image in image sequence and $d(i,j)$ and $\hat{d}(i,j)$ are true and estimated depth maps, respectively. The lower the RMSE, better the performance. Correlation represents the similarity between true and estimated depth map as (23), shown at the bottom of this page, where, $\bar{d}(i,j)$ and $\bar{\hat{d}}(i,j)$ are the means of true and estimated depth maps, respectively. Higher correlation corresponds to better performance. PSNR represents the ratio between the maximum possible power of a signal and the power of corrupting noise. PSNR can be represented in terms of the logarithmic decibel scale as:

$$\text{PSNR} = 10 \log_{10} \left( \frac{d_{\text{max}}^2}{\text{MSE}} \right), \quad (24)$$

where, $d_{\text{max}}$ is maximum depth value in the depth map. Higher value of PSNR signifies better performance. Table 4, 5, and 6, represent the 3D shape recovery results of simulated cone before and after filtering using KF, MKF, ANNF, MCC-KF, and IMCC-KF, respectively. It is shown in these tables that SML performance is the best, followed by GLV, and finally TEN. The performance of the FMs before filtering is difficult to distinguish due to unremoved jitter noise. It is also evident from these tables that IMCC-KF performance is the best, followed by MCC-KF, ANNF, MKF, and finally KF. Fig. 10, 11, 12, and 13 show the qualitative comparison of 3D shape recovery of simulated and real objects using SML, GLV, and TEN after various filtering techniques. It is clear from these figures that surfaces of 3D shapes of various objects, optimized by IMCC-KF, are smoother than those optimized by other filtering techniques. Table 7 provides the computation time of proposed and conventional filtering techniques in the last frame of various objects using SML. The last frame of simulated cone, LCD-TFT filter, coin, and letter-I is 97th frame, 60th frame, 68th frame, and 60th frame, in the respective image sequences. The table represents the time taken to filter the jitter noise in the last frame.
of various objects by utilizing $KF$, $MKF$, $ANNF$, $MCC$-$KF$, and $IMCC$-$KF$, respectively. The unit of computation time is micro-seconds. It is evident from Table 7 that $IMCC$-$KF$ is the fastest, followed by $MCC$-$KF$, $MKF$, and the slowest is $ANNF$.

VII. CONCLUSION

In SFF, when an object on a translational stage is moved at predetermined step along the optical axis, mechanical vibrations occur, which are referred as jitter noise. In this manuscript, jitter noise is modeled as Lévy symmetric stable function to reflect the real environment of SFF. Next, focus curves, obtained by using one of focus measure operators, are modeled as Gaussian function for performance comparison of various filters. Finally, IMCC-$KF$, which is a new filtering method for removing jitter noise, is designed and applied. Through various experiments with synthetic and real objects, it is demonstrated that proposed filtering method provides more accurate and faster performance results than conventional filtering methods due to the numerical stability and computational simplicity in the covariance matrix and Kalman gain.

REFERENCES

[1] 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.

[2] 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, Jan. 2013.

[3] P. Favarro, S. Soatto, M. Burger, and S. J. Osher, “Shape from defocus via diffusion,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 3, pp. 518–531, Mar. 2008.

[4] J. Lin, X. Ji, W. Xu, and Q. Dai, “Absolute depth estimation from a single defocused image,” IEEE Trans. Image Process., vol. 22, no. 11, pp. 4545–4550, Nov. 2013.

[5] B. J. Super and A. C. Bovik, “Shape from texture using local spectral moments,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 17, no. 4, pp. 333–343, Apr. 1995.

[6] A. Blake, H. H. Bulthoff, and D. Sheinberg, “Shape from texture: Ideal observers and human psychophysics,” Vis. Res., vol. 33, no. 12, pp. 1723–1737, Aug. 1993.

[7] S. C. de Vries, A. M. L. Kappers, and J. I. Koenderink, “Shape from stereo: A systematic approach using quadratic surfaces,” Perception Psychophys., vol. 53, no. 1, pp. 71–80, Jan. 1993.

[8] F. Lafarge, R. Keriven, M. Bredif, and H.-H. Vu, “A hybrid multiview stereo algorithm for modeling urban scenes,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 1, pp. 5–17, Jan. 2013.

[9] S. Parashar, D. Pizarro, and A. Bartoli, “Isometric non-rigid shape-from-motion with Riemannian geometry solved in linear time,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 10, pp. 2442–2454, Oct. 2018.

[10] A. Chhatkuli, D. Pizarro, T. Collins, and A. Bartoli, “Inextensible non-rigid structure-from-motion by second-order cone programming,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 40, no. 10, pp. 2428–2441, Oct. 2018.
[35] M. Subbarao and T. 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.
[36] Y. Liang and W. Chen, “A survey on computing Lévy stable distributions and a new MATLAB toolbox,” Signal Process., vol. 93, no. 1, pp. 242–251, Jan. 2013.
[37] A. Potrykus and S. Adhikari, “Dynamical response of damped structural systems driven by jump processes,” Probabilistic Eng. Mech., vol. 25, no. 3, pp. 305–314, Jul. 2010.
[38] G. Zak, A. Wylomanska, and R. Zimroz, “Application of alpha-stable distribution approach for local damage detection in rotating machines,” J. Vibroeng., vol. 17, no. 6, pp. 2987–3002, Sep. 2015.
[39] M. Subbarao and M.-C. Lu, “Image sensing model and computer simulation for CCD camera systems,” Mach. Vis. Appl., vol. 7, no. 4, pp. 277–289, Dec. 1994.
[40] M. T. Mahmood, W.-J. Choi, and T.-S. Choi, “PCA-based method for 3D shape recovery of microscopic objects from image focus using discrete cosine transform,” Microsc. Res. Techn., vol. 71, no. 12, pp. 897–907, Dec. 2008.
[41] M. T. Mahmood, A. Majid, and T.-S. Choi, “Optimal depth estimation by combining focus measures using genetic programming,” Inf. Sci., vol. 181, no. 7, pp. 1249–1263, Apr. 2011.

HOON-SEOK JANG received the B.S. degree in electronic engineering from the Kumoh National Institute of Technology, Gumi, South Korea, in 2013, and the M.S. and Ph.D. degrees in mechatronic engineering from the Gwangju Institute of Science and Technology, Gwangju, South Korea, in 2014 and 2019, respectively. He spent one year as a Postdoctor at the Center for Imaging Media Research, Korea Institute of Science and Technology (KIST), Seoul, South Korea. He is currently a Senior Researcher with the Korea Electronics Technology Institute (KETI), Icheon, South Korea. His research interests include image/signal processing and machine/deep learning.

MANNAN SAEED MUHAMMAD received the B.S. degree in mechatronic engineering from the National University of Science and Technology, Rawalpindi, Pakistan, in 2003, and the M.S. and Ph.D. degrees in mechatronic engineering from the Gwangju Institute of Science and Technology, Gwangju, South Korea, in 2006 and 2012, respectively. He is currently an Assistant Professor with Sungkyunkwan University, Suwon, South Korea. His research interests include robot vision, image processing, and robots in healthcare.

MIN-KOO KANG was born in Ulsan, South Korea, in 1984. He received the B.S. degree in electronic engineering from Inha University, South Korea, in 2008, and the M.S. and Ph.D. degrees in electrical engineering from the Gwangju Institute of Science and Technology, South Korea, in 2010 and 2015, respectively. He spent three years as a Postdoctor at the Center for Imaging Media Research, Korea Institute of Science and Technology (KIST), Seoul, South Korea, where he has been a Senior Researcher, since 2018. His current research interests include super-multiview video and high-density multiview video signal processing, 3-D human factor, and viewer-interactive video systems.