IJCTA, 8(3), 2015, pp. 847-855

Evaluation of Statistical Focus Measures in a Parallax Affected SFF-Inspired Approach

[1]C.R. Srinivasan, [2]R. Senthilnathan*, [3]P. Subhasree, [3]R. Sivaramakrishnan, and [1]R. Srividya

Abstract: Shape From Focus (SFF) is a popular technique in the field of computer vision for scene reconstruction. The SFF technique is based on the fact that the focus levels of the pixels of the image preserves depth information. The usage of telecentric lenses for conventional SFF limits its application for only small objects so as to preserve magnification constancy. In the current research work a new SFF-inspired algorithm is developed which utilizes a wide angle lens in place of a telecentric lens. This extends the range of object that the system can deal with, though severe magnification changes occur when a stack of images are acquired with respect to the scene. This problem is addressed using a variable window approach when focus measures are computed. This paper is a segment of the larger research work which aims at evaluation of 15 different focus measures. The paper presents significant results of performance evaluation of three different focus measures most commonly used in SFF and auto-focus algorithms. The evaluation is carried out based on two different performance evaluation criteria namely root mean square error and computation time. The analysis of focus measures are carried out under various operating conditions such as different spatial resolution, window size, contrast changes, gray level saturation and camera noise.

Keywords: Shape from Focus, Variable Magnification, Variable Window Size, Statistical Focus Measures

1. INTRODUCTION

Computer vision is one of the most researched field and has many application in many fields like computer aided inspection, robotics, entertainment and other scientific and industrial applications. Scene reconstruction is the most important problem addressed in the field of 3-D Computer Vision. Out of these methods available for scene reconstruction Shape from Focus (SFF) [1] and Shape from Defocus (SFD) use multiple images of the scene taken with varying focus levels. The difference between the methods come the in the form that the SFD [2] generally require one sharp image of the foreground and one sharp image of the background. The distance of all the points that lie between background and foreground is interpolated by a sharpness measure. Unlike the SFD method, SFF generally requires more number of images acquired along different focal distances. This generally makes the SFF method time consuming and computationally tedious compared to SFD, though the precision of reconstruction in SFF is better than SFD. The SFF techniques have been successfully used in medical diagnostics, computer aided inspection involving microscopic imaging or imaging of small objects [3]. The SFF techniques demand images of the scene to be acquired from different focal distances which can be obtained by translating the camera or the object or by changing the focus setting of the lens. One of the basic limitations of the SFF method is that it is highly sensitive to magnifications changes. Many authors in the recent in years have presented methods to improve the applicability of SFF techniques by means of novel image processing procedures [4]. The attempts made for increasing the applicability of the SFF technique is purely for complete scene

1 Manipal Institute of Technology, Manipal University
2 Department of Mechatronics Engineering, SRM University, Kattankulathur
3 Division of Mechatronics, Department of Production Technology, MIT Campus, Anna University
* e-mail: senthilnathan.r@ktr.srmuniv.ac.in

reconstruction rather than using some approximate information from SFF with an imaging system prone to magnification changes across the focal stack. This paper reports a small part of the research work which deals with development of an algorithm inspired by conventional SFF which may be used for scenario involving parallax. Previous work from the authors presented results of evaluation of focus measures based on image gradient. This paper presents results of evaluation of three other focus measures commonly used in autofocus and SFF procedures. Since the current paper is only part of a larger research work most of the common introductory contents are carried over from the previous articles [5, 6].

2. PROPOSED SFF-INSPIRED ALGORITHM

As mentioned in the previous section of the paper the setup required for SFF demands a precise translating mechanism. The photograph of the experimental setup is shown in Fig. 1. The images are taken on the go using time-synchronization approach and hence a simple DC motor is used for the purpose.

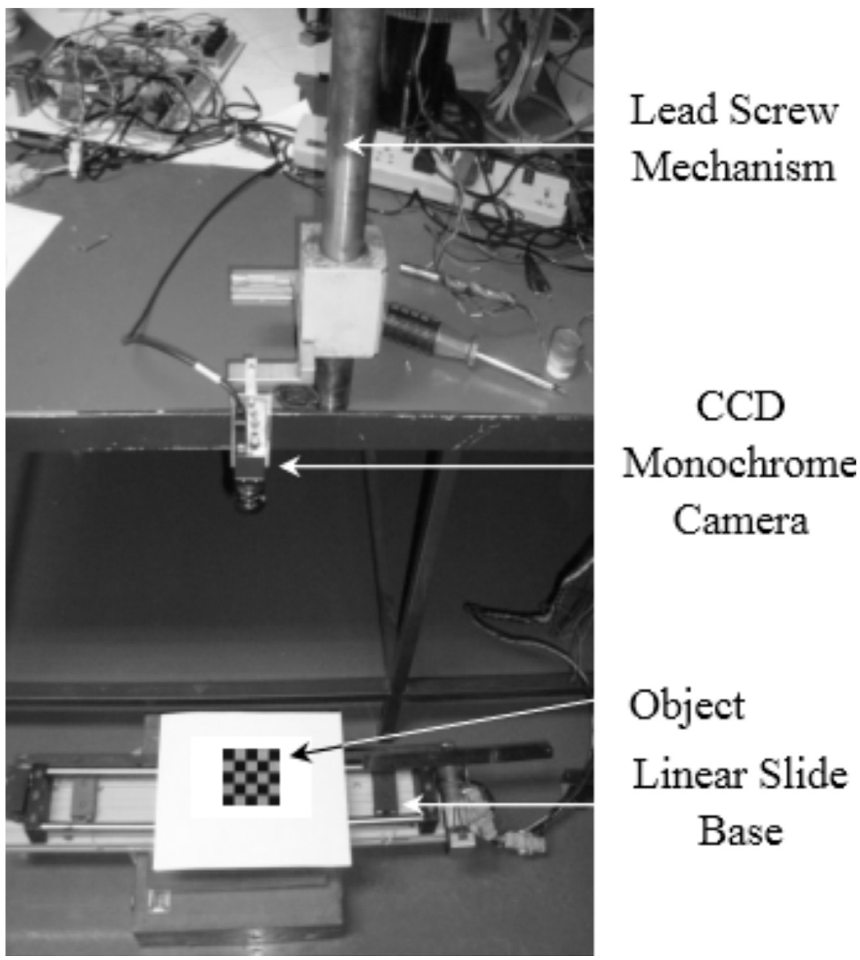


Figure 1: Experimental Setup

The various specifications of the imaging system are shown in Table 1. A focus measure is basically a mathematical function, which gives a measure of the focus of the image indirectly, by measuring the contrast of the image. It is generally computed in a small square window around the pixels in the image. A high value for the focus measure indicates a sharply focussed region in the image, and a low value indicates blurred regions. Most focus measures presented in SFF and autofocus algorithms are based on either image derivative or statistical information of the image content. The following section presents three important measures used in the current study that do not fall under any of the previously mentioned three categories.

Table 1

System Specifications

Parameter	*Specification*	*Parameter*	*Specification*
Light Source	White LED area light	Aspect Ratio	4:3
Illuminance	160 lux	Sensor	Sony ICX424
Lighting Technique	Partially diffused bright field incident lighting	Sensor Type	CCD Progressive, Monochrome
Lens Type	Fixed focal length prime lens	Operating Frame Rate	45 Frames Per Second
Focal Length & f#	16 mm and 1.3	Mode of operation	Mono8 mode
Camera Make and Model	Allied Vision Technologies Guppy F033b	Trigger Type	Software trigger
Interface	IEEE 1394a – 400Mb/s, 1 port	Image Acquisition Time	180ms per image
Computer Interface	PCI – IEEE 1394a	Processor	Intel Core i5, 2.5GHz
Resolution	640 × 480	Memory	4 GB, 1300 MHz

Statistical information in the image related to local grey level differences contributing to texture in the image may be used as a focus measure. This again is based on the same idea of exploiting the contrast change that is experienced because of image defocus. The variance of the grey levels is one of the most commonly used methods to measure the image focus. It has been popular in both autofocus [7] and SFF [8]. The grey level variance is given by:

$$F_{GLV} = \sum_{(i,j)\in\Omega(x,y)} (I(i,j)-\mu)^2 \tag{1}$$

where μ is the mean grey level of the pixels in the neighbourhood Ω. The grey level variance can be compensated for the differences in the average image brightness among different images by normalizing the value of F_{GLV} in Equation (1) by the mean grey value μ [9]. The definition of the normalized grey level variance is as follows:

$$F_{NGLV} = \frac{F_{GLV}}{\mu} \tag{2}$$

Many autofocus literatures [7] have used the range of the histogram in the local neighbourhood Ω as a focus measure which may be used in SFF. The range of the histogram is defined as

$$F_{HR} = \max(k \mid H > 0) - \min(k \mid H > 0) \tag{3}$$

The following section of the paper presents the proposed SFF-inspired algorithm, which forms the central theme of the larger research work, a portion of which is presented in this paper. First a set of feature points present across the stack of images is detected using Speeded-Up Robust Features (SURF) feature detector [10]. The stack suffers from combined variations in focus and magnification because of the relative motion between the camera and the scene. The focus measure of only those particular pixels is computed. This is different from the conventional SFF route, where the focus measure is computed for all the pixels in all the images in the focal stack. In the current research, since a wide angle lens is used with a higher DOF, in order to achieve a complete Gaussian distribution, large camera motion would be required. Extremely low magnification causes the spatial resolution of the image to become too poor for any measurements possible from the images. Because of these reasons a coarse method of depth estimate is adopted for this research. The algorithm may be summarized as follows:

i. The initial location of the camera from the measurement plane is known a priori as s_m, where m = 1 for the initial location of the camera.

ii. Accumulate the image sequences acquired at each step m where the stand-off distance (s_m) increases in steps of Δd.

iii. Measure focus, F_m for each of the SURF feature points, across the stack at each step whose correspondences are matched using the SSD metric.

iv. Find the step number m in which the focus measure is the maximum for a point (x, y), such that $F_m = F_{max}$, where F_{max} is the maximum value of the focus measure for a particular pixel.

v. Assign the value of the distance of the camera motion as the height of the object corresponding to the particular pixel, such that the height of the scene point $\bar{h} = m\Delta d$.

vi. Once the height of a point is computed, the depth of the point ${}^{C}z$ may be computed as ${}^{C}z = s_1 - \bar{h}$.

This algorithm, as may be observed, gives only a rough estimate of the depth. The performance of the algorithm is directly dependent on the selection of Δd. Lower values of Δd give better accuracy, albeit there is always a non-zero resolution error. Interestingly, at times the estimated depth becomes equal to the actual depth, depending on the particular scene point under consideration. In other words the depth error may be zero, although the system as a whole suffers from a non-zero resolution error. The size of the window about which the focus measure is computed, is a vital parameter in the SFF method. Generally, the window size must be as small as possible to obtain accurate results. When the size of the window is large, a large neighbourhood is included to compute the focus. If the depth of the scene corresponding to different points in the window varies, it may lead to averaging of different focus levels caused by different depths of the scene points. In the current research, since the images suffer from magnification changes, a variable window size approach is developed. According to this method, the window size applied to a particular frame is scaled by the magnification factor corresponding to that frame. This means that in the current scenario, the window sizes would be reducing, starting from the first frame to the fifteenth image in the focal stack. Larger window sizes offer better results, but lead to averaging errors. It is justified, since the current work which uses the SFF-inspired algorithm only to obtain a sparse and coarse depth estimate. This issue may be considered as a drawback of the proposed method, as it inherently suffers from slightly higher averaging errors compared to the conventional SFF.

3. EVALUATION OF FOCUS MEASURES

The focus measures are evaluated under different operating conditions namely different spatial resolutions, camera noise, gray level saturation and contrast. Fig. 3 shows the stacks of three different spatial resolutions and the first image in each stack. The distances between the object and the measurement plane in the three cases are, 30 mm, 50 mm and 84 mm respectively. The three cases are indicated as spatial resolution 1, 2 and 3 respectively.

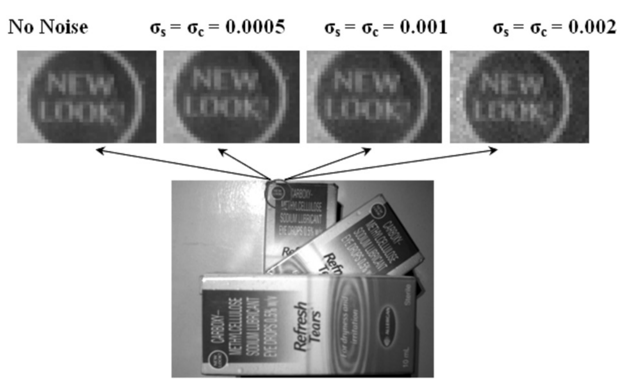


Figure 2: Different Noise levels

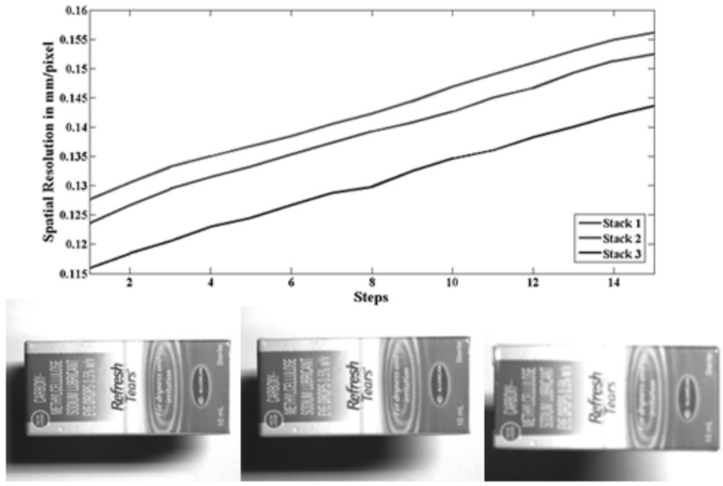


Figure 3: Different Spatial Resolution

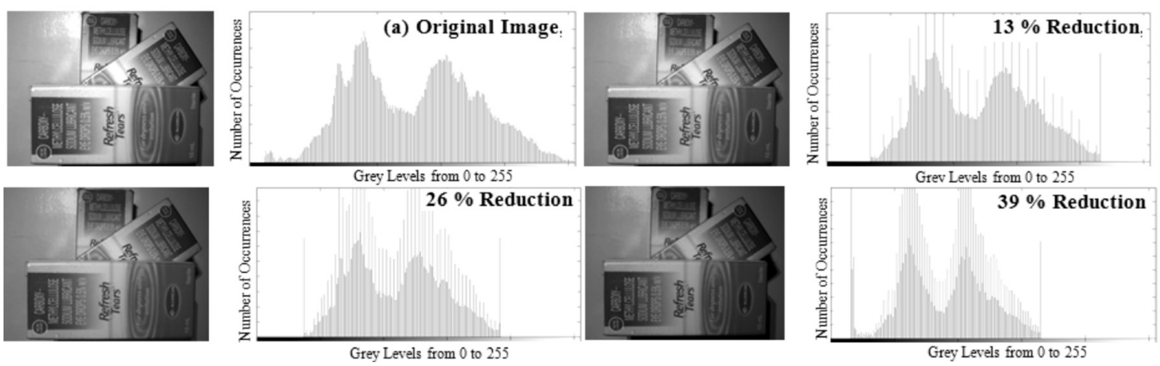


Figure 4: Contrast Reduction

Figure 5: Saturation Levels

Out of various camera noises that may corrupt an image acquired from a CCD sensor, the significant noise sources may be grouped as irradiance-dependent and irradiance-independent sources [11]. The noises may be modelled as follows:

$$I_{noise} = F(I + n_s + n_c) + n_q \tag{8}$$

where I_{noise} is the image that is obtained after adding the noise components to the original image I. The parameter F in the above equation is the camera response function, n_s is the irradiance-dependent noise component, n_c is the irradiance-independent noise component, and n_q is the quantization and amplification noise. The noise components are basically Gaussian white noise with zero mean, and the variances for n_s and n_c are $Var(n_s) = I \cdot \sigma_s^2$ and $Var(n_c) = \sigma_c^2$ respectively. The feature detectors and match metrics are evaluated for three different levels of noise, namely, for $\sigma_s = \sigma_c = 0.0005$, $\sigma_s = \sigma_c = 0.001$ and $\sigma_s = \sigma_c = 0.002$. Fig. 2 shows a magnified view of a small portion in the image (highlighted in red colour) which is subjected to different levels of noise corruption.

Image contrast is an important factor that affects the performance of the algorithms, related to feature detection and matching. In the current study, the contrast of the original images is reduced to different levels to analyse the robustness of the detectors in such a scenario. The contrasts of the images are reduced by compressing the histogram of the respective images. Three different set of points are chosen, to achieve 13 %, 26 % and 39 % contrast reduction, with reference to the original images. Fig. 4 shows the first image of the focal stack for different contrast levels and their corresponding grey level histogram.

In this research, image saturation has been evaluated by adding a constant offset to the original image, as presented in the equation below:

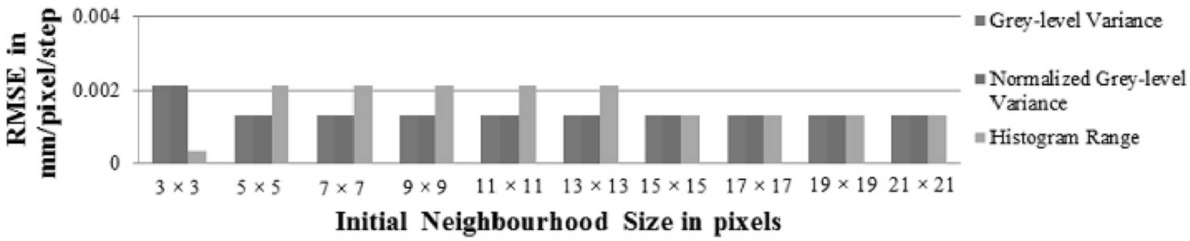
$$I_{sat} = I + S \tag{9}$$

where I_{sat} is the saturated image obtained by adding a constant offset S to the original image I. In the current study, three different offsets, namely, 25, 51, 77 are added to the images, which may be considered as 10%, 20% and 30% saturation respectively, for an 8-bit dynamic range. Fig. 5 shows the sample images subjected to various levels of saturation considered for evaluation. In the current research, two criteria are used for the evaluation of the focus measures under different operating conditions, namely, the execution time of the focus measures, and the Root Mean Square Error (RMSE). The RMSE is normalized by the number of pixels in the image and the number of steps of image acquisition. The normalized RMSE is defined as follows:

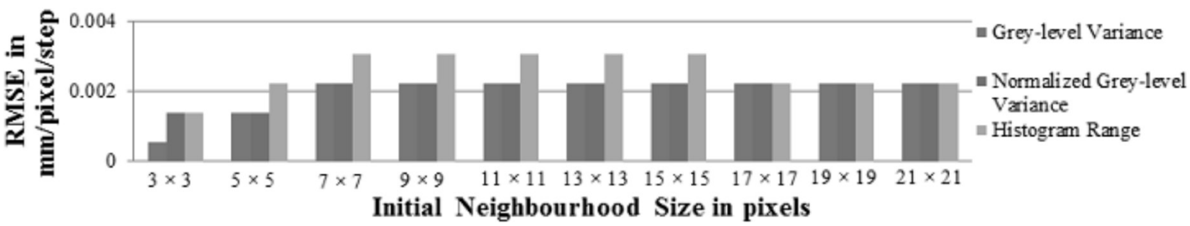
$$RMSE = \frac{\sqrt{\frac{1}{MN}\sum_{(i,j)}(G_T(i,j) - Z(i,j))^2}}{No.of\ Steps} \tag{10}$$

In the above equation N and M are the number of pixels in the horizontal and vertical dimensions of the image. $G_T(i, j)$ is the ground truth information about the actual depth of a point (i, j) in the scene, and $Z(i, j)$ is the estimated depth of the scene obtained from a particular focus measure at that point. The values of $G_T(i, j)$ are found, based on the physical measurements of the object's dimensions, with an uncertainty of 1mm. The execution time is calculated by software means in MATLAB, which gives an approximate estimate of the time taken for the execution of a set of functions. In order to reduce error, the execution time is averaged from 20 trials for the same function's execution, and also all other application software is prevented from running.

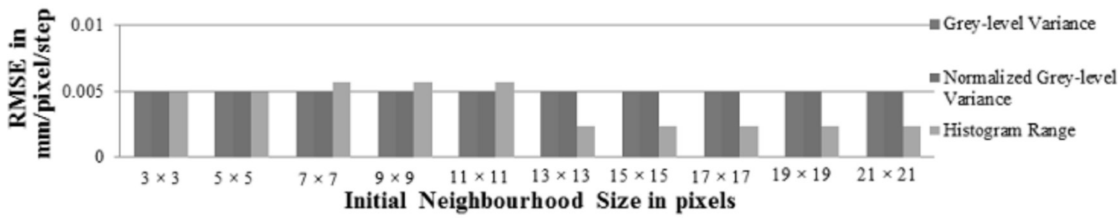
The following plots in Fig. 6 show the results of RMSE for various operating conditions.



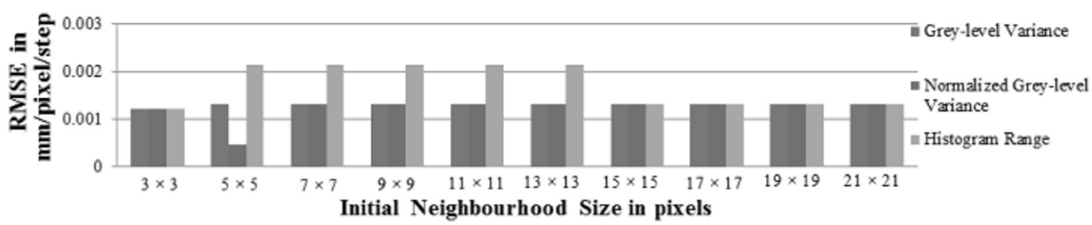
(i) Root Mean Square Error for Spatial Resolution 1



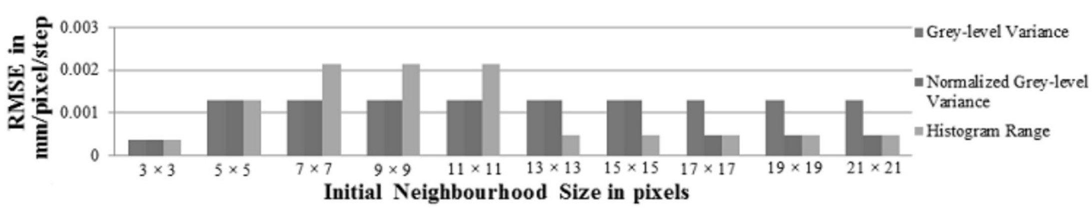
(ii) Root Mean Square Error for Spatial Resolution 2



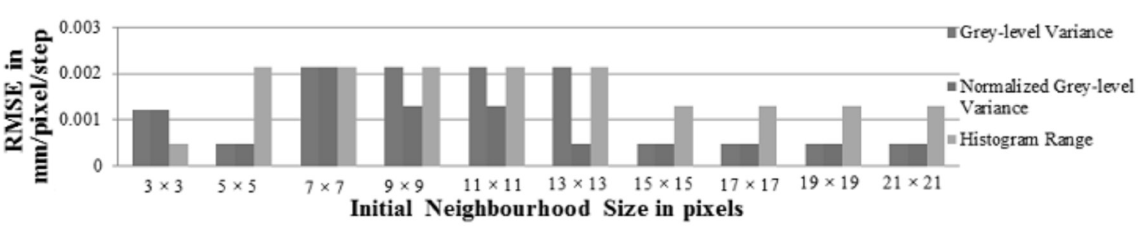
(iii) Root Mean Square Error for Spatial Resolution 3



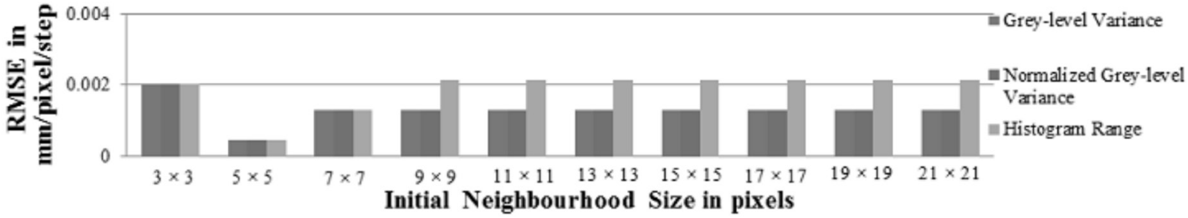
(iv) Root Mean Square Error for Camera Noise Level 1



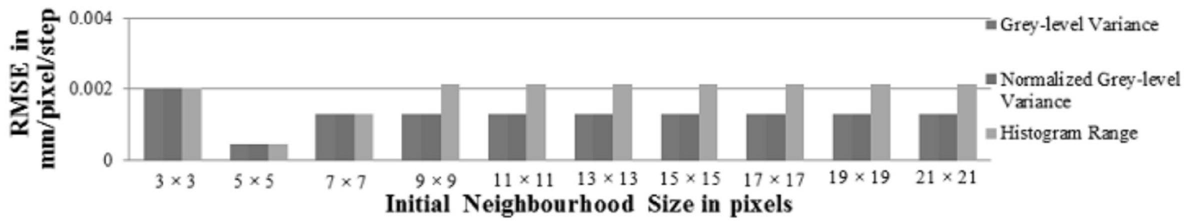
(v) Root Mean Square Error for Camera Noise Level 2



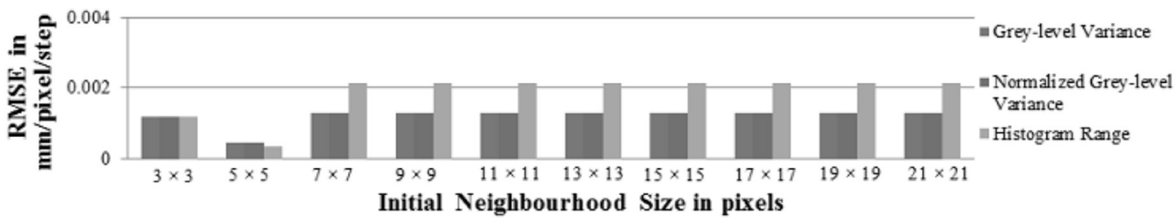
(vi) Root Mean Square Error for Camera Noise Level 3



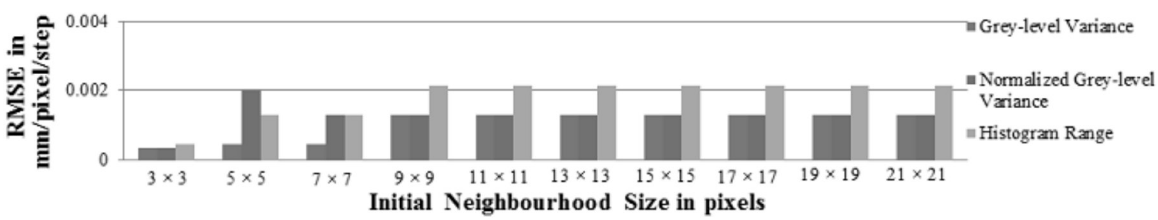
(vii) Root Mean Square Error for 13% Contrast Reduction



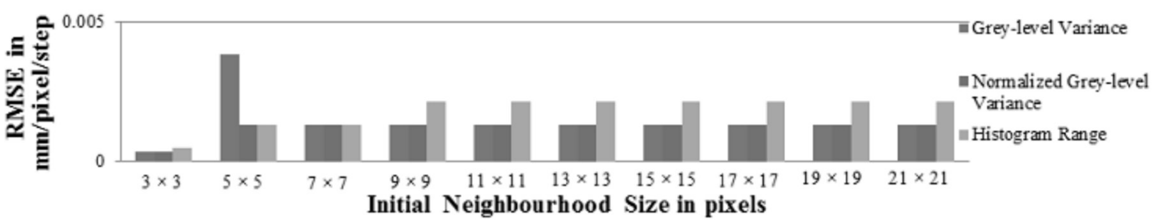
(viii) Root Mean Square Error for 26% Contrast Reduction



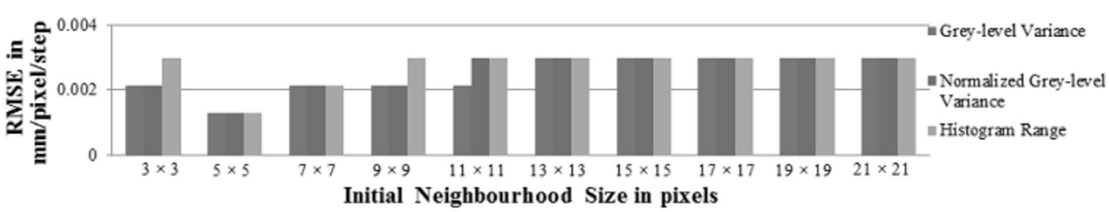
(ix) Root Mean Square Error for 39% Contrast Reduction



(x) Root Mean Square Error for 10% Saturation



(xi) Root Mean Square Error for 20% Saturation



(xii) Root Mean Square Error for 30% Saturation

Figure 7: Results of RMSE for Various Operating Conditions

Table 2 shows the average execution time for the various focus measures for a single point computed around a window, whose initial size is 21 × 21, since the execution time for a larger window takes more time, compared to a smaller window; hence, evaluating in the worst case scenario.

Table 2
Average Computation Time

Focus Measure	*Average Computation Time (ms)*
Grey-level Variance	0.461
Normalized Grey Level	0.303
Histogram range	0.15

All the statistical measures exhibited an almost similar performance in terms of the RMSE. Since estimating the range of the histogram is the cheapest method in terms of the algorithm, it took the least time for execution. Higher window sizes returned lower RMSE. This is mainly due to the larger neighbourhood of focus measure which makes the measurement robust (at the cost of higher computational time) compared to the noise measurements obtained from smaller masks. Selection of a unique size is subjected to the trade off and a window size of 15 × 15 is chosen to be the candidate window size for all the focus measurement purpose. The results presented in this paper may be used to select the right statistical focus measure for the proposed SFF-inspired algorithm.

References

1. S. K. Nayar, and Y. Nakagawa, "Shape from focus: An Effective Approach for Rough Surfaces," *CRA90*, pp. 218-225, 1990.
2. Y. Xiong and S. A. Shafer, "Depth from focusing and defocusing," *IEEE Computer Vision Pattern Recognition*, pp. 68-73, 1993.
3. F. S. Helmli and S. Scherer, "Adaptive shape from focus with an error estimation in light microscopy," *Proc. of 2nd International Symposium on Image and Signal Processing and Analysis*, 2001.
4. R. R. Sahay and A. N. Rajagopalan, "Dealing with Parallax In Shape-From-Focus," *IEEE Transactions on Image Processing*, Vol. 20 No. 2, pp. 558-569, 2011.
5. R. Senthilnathan and R. Sivaramakrishnan, "Estimation of relative depth in the scene using SFF-inspired focus cue," *Proc. of the IEEE International Conference on Advanced Communication Control and Computing Technologies ICACCCT*, pp. 332-337, 2012.
6. R. Senthilnathan, P. Subhasree and R. Sivaramakrishnan, "Performance Analysis of Gradient-Based Focus Measures in a Parallax Affected SFF Scenario," *International Journal of Computer Aided Manufacturing*, Vol. 1 No. 1, pp. 1-12, 2015.
7. L. Firestone, K. Cook, K. Culp, N. Talsania and K. Preston Jr., "Comparison of autofocus methods for automated microscopy", *Cytometry*, Vol. 12, pp. 195-206, 1991.
8. Y. An, G. Kang, J.J. Kim, H.S. Chung, and J. Park, "Shape from focus through Laplacian using 3D window", *Proc. of International Conference on Future Generation Communication and Networking*, pp. 46–50, 2008.
9. S. Y. Lee, J. T. Yoo, Y. Kumar, and S. W. Kim, "Reduced energy–ratio measure for robust autofocusing in digital camera", *IEEE Signal Processing Letters*, Vol. 16, pp. 133-136, 2009.
10. H. Bay, T. Tuytelaars and L. Van Gool, "SURF: Speeded up robust features," *Proc. of European Conference on Computer Vision*," pp. 404-417, 2006.
11. S. Pertuz, D. Puig and M. A. Garcia, "Analysis of focus measure operators for shape–from–focus," *Pattern Recognition*, Vol. 46, pp. 1415-1432, 2013.