Root-Mean Square Error in Passive Autofocusing and 3D Shape Recovery

Murali Subbarao *Jenn-Kwei Tyan*
Department of Electrical Engineering,
State University of New York
Stony Brook, NY 11794-2350
email: murali@sbee.sunysb.edu , jktyan@sbee.sunysb.edu

ABSTRACT

Image focus analysis is an important technique for passive autofocusing and three-dimensional shape measurement. Electronic noise in digital images introduces errors in this technique. It is therefore important to derive robust focus measures that minimize error. In our earlier research, we have developed a method for noise sensitivity analysis of focus measures. In this paper we derive explicit expressions for the root-mean square (RMS) error in autofocusing based on image focus analysis. This is motivated by the Autofocusing Uncertainty Measure (AUM) defined earlier by us as a metric for comparing the noise sensitivity of different focus measures in autofocusing and 3D shape-from-focus. The RMS error we derive is shown to be proportional to the square of the AUM. The expression for RMS error derived by us has the same advantage as AUM in that it can be computed in only one trial of autofocusing. We validate our theory on RMS error and AUM through experiments. It is shown that the theoretically estimated and experimentally measured values of the standard deviation of a set of focus measures are in agreement. Our results are based on a theoretical noise sensitivity analysis of focus measures, and they show that for a given camera the optimally accurate focus measure may change from one object to the other depending on their focused images.

Keywords: focus measures, autofocusing, shape-from-focus, noise sensitivity.

1 INTRODUCTION

Image Focus Anaylsis[1,3,4,6] is an important technique for passive autofocusing and three-dimensional (3D) shape recovery of objects. These techniques find applications in machine vision, consumer video cameras, and video microscopy. In image focus analysis, a sequence of images of a 3D scene are acquired by a camera with different degrees of blur or defocus. The change in the level of defocus is obtained by changing either the lens position or the focal length of the lens in the camera. A focus measure is computed for each of the images in the image sequence in small image regions. The value of the focus measure increases as the image sharpness or contrast increases and it attains the maximum for the sharpest focused image. Thus the sharpest focused image regions can be detected and extracted. This facilitates autofocusing of small image regions by adjusting the camera parameters (lens position and/or focal length) so that a focus measure attains a maximum for that image region. Also, such focused image regions can be synthesized to obtain a large image where all image regions are in focus. Further, the distance or depth of

object surface patches that correspond to small image regions can be obtained from a knowledge of the lens position and the focal length that result in sharpest focused images of the surface patches. This is done using the lens formula:

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v} \tag{1}$$

where f is the focal length, u is the object distance, and v is the distance of the focused image (see Fig. 1). The three-dimensional shape or depth-map of the scene can be obtained by synthesizing the depth of small surface patches in the scene.

The accuracy of autofocusing and 3D shape measurement using the image focus analysis technique depends on the particular focus measure that is used. Experimental evaluations of different focus measures have been reported in.[1–4,8] Until recently, there was no theoretical treatment of the noise sensitivity of focus measures. The published literature consisted of a combination of experimental observations and subjective judgement. The noise sensitivity of a focus measure depends not only on the noise characteristics but also on the image itself. The optimally accurate focus measure for a given noise characteristics may change from one object to the other depending on its focused image. This makes it difficult to arrive at general conclusions from experiments alone.

A rigorous and general noise sensitivity analysis for a large class of focus measures was provided in.[6] A new metric named *Autofocusing Uncertainty Measure* (AUM) was proposed. Here we propose another new metric named *Autofocusing Root-Mean-Square Error* (ARMS error). This metric corresponds to the traditional Root-Mean Square (RMS) error that is widely used to specify the noise-sensitivity of estimates. ARMS error is similar to AUM and shares the same advantages. Also, they are related by a monotonic relation. In depth-from-focus (DFF) applications, AUM and ARMS error can both be easily translated into uncertainties in depth using Eq. (1).

The analysis here shows that the autofocusing noise sensitivity of a focus measure depends on the image of the object to be autofocused in addition to the camera characteristics. For an object with unknown focused image, finding the optimally accurate focus measure involves computing all the candidate focus measures at a set of lens positions and computing AUM/ARMS error for each of the lens positions. Then the lens is moved to the focused position estimated by the optimal focus measure (which has minimum AUM/ARMS error). Usually the number of candidate focus measures that should be considered for good performance is only a few (about 3 to 5). Also, almost all focus measures require only a modest amount of computing. Therefore selecting the optimal focus measure from a candidate set comes at a small computational cost. However, if it is necessary to use minimal computing in autofocusing by using the same focus measure for all objects, then it has been argued[6] that the the energy of the image Laplacian is a good focus measure to use. This focus measure has been shown to have some important desirable characteristics based on a spatial domain analysis.

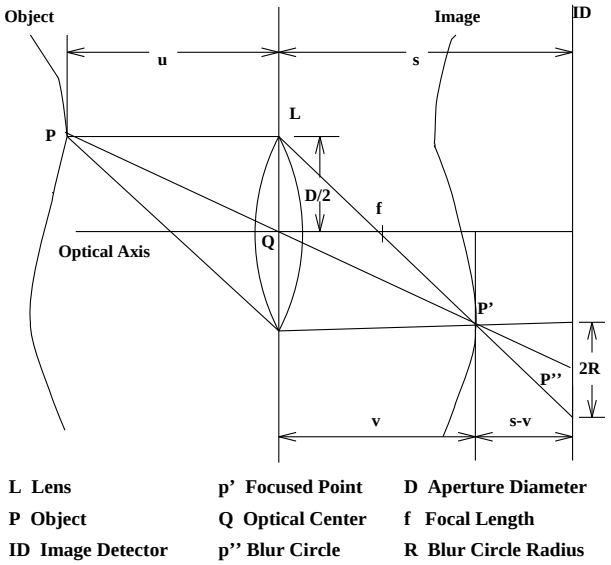


Figure 1: Image Formation in a Convex Lens

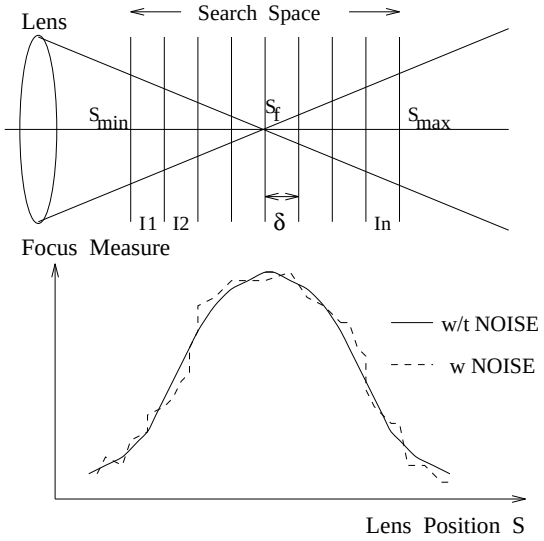


Figure 2: Autofocusing Algorithm

2 FOCUS MEASURES AND AUTOFOCUSING

A detailed discussion of the image sensing model can be found in several papers including.[4] Here we simply note some relevant results based on geometric optics. In Fig. 1, a blurred image $g(x, y)$ in a small image region on the image detector is equal to the convolution of the focused image $f(x, y)$ and the PSF $h(x, y)$. Assuming additive noise $\eta(x, y)$, this model of image sensing is summarized by

$$g(x, y) \;=\; h(x, y) * f(x, y) + \eta(x, y). \tag{2}$$

A more detailed model of image sensing is described in.[5]

A general focus measure is modeled as follows. First the image $g(x, y)$ for which the focus measure needs to be computed is normalized for brightness by dividing the image by its mean brightness. Then it is convolved with a focus measure filter (FMF) $a(x, y)$. Then the energy (sum of squared values) of the filtered image is computed. This energy is the focus measure (see Appendix).

2.1 Autofocusing Algorithm

Detailed autofocusing algorithms can be found elsewhere.[6] Their general structure can be summarized as follows. First the focus measure is computed at the current lens position and the lens is moved by a small distance to another position. The focus measure is again computed. The sign of the change in the two focus measures is used to determine the direction in which the lens should be moved. Then a coarse search is begun by moving the lens systematically (e.g. sequential and/or binary and/or Fibonacci search) and a rough location is found where the focus measure is a maximum. Then the lens is moved to several (about 3 to 5) positions that are close apart near the rough location of the focused position. At each of these positions the focus measure is computed. Then a curve (e.g. quadratic or a Gaussian) is fitted to these focus

measures. The actual focused position is estimated as the position where the fitted curve attains the maximum (see Fig. 2).

3 AUM

First we review AUM as a metric for focus measures to illustrate some underlying concepts. Next we introduce ARMS error which is based on weaker assumptions than AUM. At any lens position s_0 (see Fig. 3), each focus measure γ is associated with a probability density function $p(\gamma(s_0))$, an expected value (mean) $E\{\gamma(s_0)\}$, and a standard deviation std$\{\gamma(s_0)\}$. However, the focus measure with the minimum standard deviation is not necessarily the best because we are not interested in the accuracy of the focus measure itself, but in the corresponding mean lens position and its standard deviation. Estimating the standard deviation of the lens position requires a knowledge of the function that relates the expected value of the focus measure to the lens position (see Fig. 3). This function depends on the camera PSF as a function of camera parameters and the focused image of the object. In the absence of accurate information about the camera PSF and the object, the function is estimated in a desired interval through sampling and interpolation. For example, near the maximum, the focus measure may be computed at 3 to 5 nearby lens positions and a smooth function such as a quadratic polynomial or a Gaussian is fitted. The assumption is that the computed values of the focus measure are (nearly) the expected values of the focus measure.

Referring to Fig. 3, the AUM at the maximum of the focus measure γ is defined as follows:

$$\gamma_{max} = \gamma(s_f) \quad , \qquad \sigma = \gamma(s_f) - \gamma(s_1) = \gamma(s_f) - \gamma(s_2) \ , \qquad \text{AUM} = s_2 - s_1 \tag{3}$$

where σ is the standard deviation of the focus measure. In order to compute AUM, we need to know σ. In the Appendix section we provide a general formula that can be used to estimate σ as a function of the image and its noise level. We need to know the shape of the curve $\gamma(s)$ near the peak. As discussed earlier, the position of γ_{max} and the function $\gamma(s)$ near γ_{max} are estimated by fitting a curve (quadratic or Gaussian) to a few points (at least 3) near the maximum. AUM is a measure similar to the RMS error in lens position.

Fig. 4 shows a typical comparison of two focus measures. The maximum values of the two focus measures have been normalized to be the same. We see that although $\sigma_2 > \sigma_1$, $AUM_2 < AUM_1$, implying that γ_2 is better than γ_1.

Referring to Fig. 5, focus measure $\overline{\gamma}$ is modeled to be locally quadratic in a small interval of size 2δ with respect to lens position near the focused position:

$$\overline{\gamma}(s) = as^2 + bs + c \tag{4}$$

Let the focus measure be given at three arbitrary positions which are δ apart. Without loss of generality, let the three positions be $s_- = -\delta$, $s_0 = 0$, and $s_+ = +\delta$. Let $\overline{\Gamma}_- = \overline{\gamma}(s_-)$, $\overline{\Gamma}_0 = \overline{\gamma}(s_0)$ and $\overline{\Gamma}_+ = \overline{\gamma}(s_+)$. Near the focused position, $\overline{\Gamma}_0 > \overline{\Gamma}_-$ and $\overline{\Gamma}_0 > \overline{\Gamma}_+$. Solving for the coefficients of the quadratic expression, we obtain

$$a \ = \ \frac{\overline{\Gamma}_+ + \overline{\Gamma}_- - 2\overline{\Gamma}_0}{2\delta^2} \ , \quad b \ = \ \frac{\overline{\Gamma}_+ - \overline{\Gamma}_-}{2\delta} \ , \quad c \ = \ \overline{\Gamma}_0 \tag{5}$$

Let s_f be the lens position where the focus measure becomes the maximum and $\overline{\Gamma}_f = \overline{\gamma}(s_f)$. At s_f, the derivative of $\overline{\Gamma}$ vanishes. Therefore we obtain

$$s_f = \frac{-b}{2a} = \frac{\delta}{2}\frac{\left(\overline{\Gamma}_+ - \overline{\Gamma}_-\right)}{\left(2\overline{\Gamma}_0 - \overline{\Gamma}_+ - \overline{\Gamma}_-\right)} \tag{6}$$

Substituting the above equation in (4) we obtain

$$\overline{\Gamma}_f = -\frac{b^2 - 4ac}{4a} \tag{7}$$

Given that $AUM = s_2 - s_1$, we obtain s_1 and s_2 as the roots of the equation

$$\overline{\Gamma}(s) = \overline{\Gamma}_f - \sigma = as^2 + bs + c \tag{8}$$

Solving the above equation, we obtain

$$\text{AUM} = \left|\frac{\sqrt{b^2 - 4a\left(c + \sigma - \overline{\Gamma}_f\right)}}{a}\right| \tag{9}$$

Substituting Eq (5) and (7) in the above equation yields

$$\text{AUM} = 2\delta\left(\frac{2\sigma}{2\overline{\Gamma}_0 - \overline{\Gamma}_+ - \overline{\Gamma}_-}\right)^{\frac{1}{2}} \tag{10}$$

At a position far away from the focused lens position s_f, AUM is defined as in Fig. 6. This is somewhat similar to that near the peak:

$$\sigma = (\overline{\gamma}(s_2) - \overline{\gamma}(s_1))/2\ , \qquad \text{AUM} = s_2 - s_1 \tag{11}$$

Once again, σ is computed from the known noise characteristics and the image. The shape of the focus measure curve is estimated by a linear (2 points) interpolation using the values of the focus measure at s_- and s_+ that are δ apart. Without loss of generality, let $s_- = -\delta/2$ and $s_+ = +\delta/2$ and the focus measures at these points be $\overline{\Gamma}_-$ and $\overline{\Gamma}_+$ respectively (see Fig. 8). The linear model yields the expression

$$\frac{s - s_-}{s_+ - s_-} = \frac{\overline{\Gamma} - \overline{\Gamma}_-}{\overline{\Gamma}_+ - \overline{\Gamma}_-} \tag{12}$$

The above equation can be rewritten as:

$$s = \delta\left(\frac{\overline{\Gamma} - \overline{\Gamma}_-}{\overline{\Gamma}_+ - \overline{\Gamma}_-}\right) - \frac{\delta}{2} \tag{13}$$

We obtain s_1 and s_2 by solving

$$\overline{\Gamma}(s) = \frac{\overline{\Gamma}_+ + \overline{\Gamma}_-}{2} \pm \sigma \tag{14}$$

where σ is the standard deviation of the focus measure. Using equation (13) and solving for AUM, we obtain

$$\text{AUM} = |s_1 - s_2| = \frac{2\delta\sigma}{|\overline{\Gamma}_+ - \overline{\Gamma}_-|} \tag{15}$$

Fig. 7 shows a comparison of two focus measures far away from the focused position. Once again we see that although $\sigma_2 > \sigma_1$, $AUM_2 < AUM_1$, implying that γ_2 is better than γ_1.

4 ARMS ERROR

In this section, an explicit expression for the *Autofocusing Root-Mean Square Error* (ARMS error) is derived, which is based on weaker assumptions than AUM. An exact expression for the RMS error depends on the Optical Transfer Function (OTF) of the camera and the Fourier spectrum of the focused image. Deriving such an exact expression is complicated because of the nature of the camera's OTF and the variability of the Fourier spectrum of the focused image for different objects. Further, usefulness of such an expression in practical applications is limited since all the information necessary to evaluate the expression (e.g. OTF and camera parameters) may not be available. However, an approximate expression that is very useful in practical applications can be derived under some weak assumptions. The assumption we use is that the expected value of the focus measure can be modeled to be quadratic locally with respect to the lens position. The analysis here can be extended to other models (e.g. cubic or Gaussian) but such extensions do not appear to be useful in practical applications at present.

We are interested in the RMS value of s_{max}. For this reason, the focus measure Γ_i will be expressed as the summation of their expected value $\overline{\Gamma}_i$ and their noise component n_i:

$$\Gamma_i = \overline{\Gamma}_i + n_i \qquad for \;\; i = -, 0, +. \tag{16}$$

In this case we obtain a set of equations similar to Eqs (4) to (7) with the difference that $\overline{\Gamma}_i$ are replaced by Γ_i , therefore we obtain

$$\begin{aligned} s_{max} &= \frac{\delta}{2}\left(\frac{\Gamma_+ - \Gamma_-}{2\Gamma_0 - \Gamma_+ - \Gamma_-}\right) \\ &= \frac{\delta}{2}\left(\frac{\overline{\Gamma}_+ - \overline{\Gamma}_- + n_+ - n_-}{2\overline{\Gamma}_0 - \overline{\Gamma}_+ - \overline{\Gamma}_- + 2n_0 - n_+ - n_-}\right) \\ &= \frac{\delta}{2}\left(\frac{\overline{\Gamma}_+ - \overline{\Gamma}_-}{2\overline{\Gamma}_0 - \overline{\Gamma}_+ - \overline{\Gamma}_-}\right)\left(1 + \frac{n_+ - n_-}{\overline{\Gamma}_+ - \overline{\Gamma}_-}\right)\left(1 + \frac{2n_0 - n_+ - n_-}{2\overline{\Gamma}_0 - \overline{\Gamma}_+ - \overline{\Gamma}_-}\right)^{-1} \end{aligned} \tag{17}$$

Near the focused position we have $\overline{\Gamma}_0 > \overline{\Gamma}_+$ and $\overline{\Gamma}_0 > \overline{\Gamma}_-$. Therefore, if the signal to noise ratio is sufficiently large, we have

$$|2\overline{\Gamma}_0 - \overline{\Gamma}_+ - \overline{\Gamma}_-| \gg |2n_0 - n_+ - n_-| \tag{18}$$

We obtain $s_{max} \approx s'_{max}$ where

$$s'_{max} = \overline{s}_{max} \left(1 + \frac{n_+ - n_-}{\overline{\Gamma}_+ - \overline{\Gamma}_-} \right) \tag{19}$$

Note: we cannot assume that $|\overline{\Gamma}_+ - \overline{\Gamma}_-| \gg |n_+ - n_-|$ because, near the focused position, $\overline{\Gamma}_+$ and $\overline{\Gamma}_-$ may be nearly equal. Simplifying the expression for s'_{max} we obtain

$$s'_{max} = \overline{s}_{max} + \frac{\delta}{2} \left(\frac{n_+ - n_-}{2\overline{\Gamma}_0 - \overline{\Gamma}_+ - \overline{\Gamma}_-} \right) \tag{20}$$

Now the ARMS error is defined as the standard deviation of s'_{max}, i.e.

$$\begin{aligned} \text{ARMS } error &= \frac{\delta}{2} \frac{1}{\left(2\overline{\Gamma}_0 - \overline{\Gamma}_+ - \overline{\Gamma}_- \right)} \cdot std\left(n_+ - n_- \right) \\ &= \frac{\delta}{2} \frac{1}{\left(2\overline{\Gamma}_0 - \overline{\Gamma}_+ - \overline{\Gamma}_- \right)} \cdot \left(\sigma_+^2 + \sigma_-^2 \right)^{\frac{1}{2}} \end{aligned} \tag{21}$$

where σ_+ and σ_- are the standard deviations of the focus measures Γ_+ and Γ_- respectively.

For a lens position far from the maximum focused position, the above expression for ARMS error will not be valid since the assumption in Equation (18) will not be valid. In this case, the local linear model for the focus measure will be better than the local quadratic model. The ARMS error for this case is based on focus measures at only two lens positions (rather than three) that are δ apart. Without loss of generality, let the two positions be $s_- = -\delta/2$ and $s_+ = +\delta/2$ and the focus measures at these points be Γ_- and Γ_+ respectively (similar to Fig. 8). The linear model yields the expression

$$\frac{s - s_-}{s_+ - s_-} = \frac{\Gamma - \Gamma_-}{\Gamma_+ - \Gamma_-} \tag{22}$$

The above equation can be rewritten as:

$$s = \delta \left(\frac{\Gamma - \Gamma_-}{\Gamma_+ - \Gamma_-} \right) - \frac{\delta}{2} \tag{23}$$

Once again, we express Γ_+ and Γ_- as $\Gamma_+ = \overline{\Gamma}_+ + n_+$ and $\Gamma_- = \overline{\Gamma}_- + n_-$ where $\overline{\Gamma}_+$ and $\overline{\Gamma}_-$ are the expected values and n_+ and n_- are the noise components.

Now the ARMS error is defined as the standard deviation of s' where s' is the solution of $\Gamma(s) = \frac{\overline{\Gamma}_+ + \overline{\Gamma}_-}{2}$. Solving this equation we obtain

$$\begin{aligned} s' &= \frac{\delta}{2} \left(\frac{\overline{\Gamma}_+ - \overline{\Gamma}_- - 2n_-}{\overline{\Gamma}_+ - \overline{\Gamma}_- + n_+ - n_-} \right) - \frac{\delta}{2} \\ &= \frac{\delta}{2} \left[\left(1 - \frac{2n_-}{\overline{\Gamma}_+ - \overline{\Gamma}_-} \right) \left(1 + \frac{n_+ - n_-}{\overline{\Gamma}_+ - \overline{\Gamma}_-} \right)^{-1} - 1 \right] \end{aligned} \tag{24}$$

Assuming $|\overline{\Gamma}_+ - \overline{\Gamma}_-| \gg |n_+ - n_-|$ and $|\overline{\Gamma}_+ - \overline{\Gamma}_-| \gg |2n_-|$, we obtain

$$\begin{aligned} s' &\approx \frac{\delta}{2}\left[1 - \frac{2n_-}{\overline{\Gamma}_+ - \overline{\Gamma}_-} - \frac{n_+ - n_-}{\overline{\Gamma}_+ - \overline{\Gamma}_-} - 1\right] \\ &\approx \frac{\delta}{2}\left(\frac{n_+ + n_-}{\overline{\Gamma}_- - \overline{\Gamma}_+}\right) \end{aligned} \tag{25}$$

Hence, the ARMS error would be

$$\text{ARMS error} = \; std(s') \; = \frac{\delta\,\left(\sigma_+^2 + \sigma_-^2\right)^{\frac{1}{2}}}{2|\overline{\Gamma}_+ - \overline{\Gamma}_-|} \tag{26}$$

4.1 Relation between AUM and ARMS error

Comparing the expressions for AUM and ARMS error from equations (10) and (21) we find

$$\begin{aligned} \frac{\text{AUM}^2}{\text{ARMS}} &= 16\delta\frac{\sigma}{\sqrt{\sigma_+^2 + \sigma_-^2}} \\ &\approx 8\sqrt{2}\delta \qquad \textbf{if} \quad \sigma_+ \approx \sigma_- \approx \sigma \end{aligned} \tag{27}$$

The ratio of the square of AUM and ARMS error is a constant. Therefore AUM and ARMS error are monotonically related. If we redefine AUM so that instead of using equation (8) we obtain s_1 and s_2 by solving

$$\Gamma(s) = \Gamma_{max} - \sigma^2 \tag{28}$$

then we find that AUM and ARMS error are linearly related for a given focus measure and focused image.

For a lens position far away from the focused position, comparing the expressions (15) and (26) for AUM and ARMS error yields

$$\begin{aligned} \frac{\text{AUM}}{\text{ARMS}} &= \frac{4\sigma}{\sqrt{\sigma_+^2 + \sigma_-^2}} \\ &\approx 2\sqrt{2} \qquad \textbf{if} \quad \sigma_+ \approx \sigma_- \approx \sigma \end{aligned} \tag{29}$$

For this case, they are linearly related.

5 EXPERIMENTS

In the first set of experiments, Eq. (36) in Appendix for the variance of focus measures was verified as follows. The autofocusing algorithm described earlier was implemented on a system named Stony Brook Passive Autofocusing and Ranging Camera System (SPARCS).[4] In SPARCS, a 35 mm focal length lens is used. The lens is driven by a stepper motor that can move the

lens to 97 different step positions. The standard deviation of the camera noise was estimated by imaging a flat and uniformly bright object and then computing the grey level variance of the recorded image. Three objects labeled A,B, and C (see Fig. 9) were used in the experiments.

An object was placed in front of the camera, and for some fixed lens position, 10 images of size 32 x 32 of the object were recorded. These images slightly differed from each other due to electronic noise. A given focus measure was computed for each of the 10 images. The standard deviation of the resulting 10 focus measures was then computed. This was the experimentally determined standard deviation of the focus measure. The theoretical estimation of the standard deviation of the focus measure was computed using Eq. (36). For this purpose, the standard deviation of the noise was obtained as mentioned earlier using a flat uniformly bright object. The noise-free image needed in Eq. (36) was obtained by averaging 4 noisy images of the object. Table 1 shows the experimentally computed and theoretically estimated standard deviations of different focus measures. We see that the two values are close thus verifying Equation (36).

In the second experiment, the objects A, B, and C, were autofocused using the algorithm described in Section 2. In each case, the experimental and theoretical ARMS error were computed (the units is lens steps). Near the focus position, images were recorded at 3 positions s_-, s_0 and s_+ which were 5 steps apart. At each position, 10 images were recorded. At each position, using the 10 recorded images at that position, the mean and the standard deviation of the focus measure there were computed. Then the theoretically estimated ARMS error was computed using Eq. (21). The same data was used to compute 10 experimental focus positions using Eq. (17). The standard deviation of these 10 positions was the experimental ARMS error. The resulting values are shown in the last two columns of Table 1. We see that they are very close. These values also indicate the relative autofocusing accuracy of the three focus measure filters– gray level variance, gradient magnitude squared and Laplacian squared. The measured noise standard deviation was 0.95 (grey level units) for the camera, and the SNR for the three objects were 35 dB, 28 dB and 20 dB respectively.

Three main conclusions can be drawn from the experimental results. First, for a given object (i.e. fixed image content), ARMS error decreases with increasing signal-to-noise ratio (SNR). This implies that low contrast objects and noisy cameras have more autofocusing error. Second, the focus measure with minimum standard deviation is not necessarily the focus measure that gives minimum error in autofocusing. Third, best focus measure could be different for different objects depending on both image content and noise characteristics; SNR alone cannot be used to determine the best focus measure. For example, the best focus measure for the objects with SNR 35 dB and SNR 28 dB are the Laplacian squared, but for the object with SNR 20 dB, the best focus measure is gradient magnitude squared. The gray level variance performed very poorly for object C and the autofocusing was totally unreliable. This is indicated by the N/A entries in the table.

5.1 Computer simulation experiments

Experiments similar to the ones above were carried out on simulated image data. The purpose of these experiments was to further verify our theoretical results. In these experiments, unlike

in the previous experiments, the noise-free image data and precise characteristics of noise were known accurately. Therefore we expected a closer agreement than previous experiments between theoretically estimated and experimentally determined values of the standard deviation of focus measures. This expectation was satisfied thus verifying our theory more accurately. In addition, unlike the previous experiments, the simulation experiments were carried out at many different levels of noise rather than at only one level of noise. The theory was verified to be correct at all noise levels. The test object shown in Fig. 10.a was added with various levels of zero-mean Gaussian random noise to get a set of noisy images. At each noise level, the mean and the standard deviation of the focus measure were computed using 10 noisy images. Then the standard deviation of the focus measures were estimated theoretically using Eqs. (39), (42) and (46). The plots in Figs. 10b to 10d show that the experimental and theoretical standard deviation are in close agreement at all noise levels for all three focus measures.

Another experiment similar to the second experiment for real data described earlier was conducted on simulation data as follows. The focused image of a planar object normal to the optical axis was used as input to a program that models image sensing in a CCD video camera. The program was a modified version of the Image Defocus Simulator (IDS) developed by Lu.[5] The IDS program was modified to improve the accuracy of blurred images computed for small degrees of blur. The improved accuracy was achieved by increasing the sampling rate and by using a wave optics model[9] of the camera's PSF. A sequence of blurred images were generated corresponding to different lens positions in the SPARCS camera system in our laboratory mentioned earlier.

Three images near the focused position were selected from the image sequence generated above and a specified level of zero-mean Gaussian random noise was added to these. Then the focused position was computed using Eq. (17). The above step was repeated 10 times, and the standard deviation of the resulting 10 values of the focused positions was calculated to obtain experimental value of the ARMS error. Then the ARMS error was estimated theoretically using Eq. (21). The process above was repeated for various noise levels and three different focus measures. The results are plotted in Fig. 11. We see that the two ARMS error are in good agreement. In a similar manner, a plot of the two AUMs for various noise levels are shown in Fig. 12. The monotonic relation between ARMS error and AUM are also demonstrated from those two plots.

6 APPENDIX

In this section we present expressions for the *expected value* (mean) and *variance* of a focus measure. Detailed derivations of these expressions can be found in our previous work.[6] These are useful in computing the *standard deviation* σ of the focus measure and its AUM/ARMS error.

Let $f(m,n)$ be the blurred noise free discrete image and $\eta(m,n)$ be the additive noise. The noisy blurred digital image recorded by the camera is

$$f_\eta(m,n) = f(m,n) + \eta(m,n) \tag{30}$$

The noise $\eta(m,n)$ at different pixels are assumed to be independent, identically distributed random variables with zero mean and standard deviation σ_n. Let $g(m,n)$ be the image obtained

by filtering the noisy blurred image $f_\eta(m,n)$ with the FMF $a(i,j)$:

$$\begin{aligned} g(m,n) &= a(i,j) \star f_\eta(m,n) \\ &= a(i,j) \star f(m,n) + a(i,j) \star \eta(m,n) \end{aligned} \tag{31}$$

where $\star$ means the moving weighted sum (MWS) operator that is defined by

$$a(i,j) \star f_\eta(m,n) = \sum_{i,j}^{M} a(i,j) f_\eta(m+i, n+j) \tag{32}$$

The focus measure γ is defined as

$$\gamma = \frac{1}{(2N+1)^2} \sum_{m,n}^{N} g^2(m,n) \tag{33}$$

The following expressions can be derived for the mean and variance of the focus measure:

$$E\{\gamma\} = \frac{1}{(2N+1)^2} \sum_{m,n}^{N} [a(i,j) \star f(m,n)]^2 + A_n \sigma_n^2 \tag{34}$$

$$A_n = \sum_{i,j}^{M} a^2(i,j) \tag{35}$$

$$\begin{aligned} Var\{\gamma\} &= \frac{A_n^2 E\{\eta^4\}}{(2N+1)^2} - \frac{A_n^2 \sigma_n^4}{(2N+1)^2} + \frac{\sigma_n^4}{(2N+1)^2} \sum_{i_1,j_1}^{M} \sum_{i_2,j_2}^{M} \sum_{i_3,j_3}^{M} \sum_{i_4,j_4}^{M} Q \cdot \left(\prod_{k=1}^{4} a(i_k, j_k) \right) \\ &\quad + \frac{4\sigma_n^2}{(2N+1)^2} [a(i,j) * a(-i,-j) * f(m,n)]^2 \end{aligned} \tag{36}$$

where $*$ represents the convolution operator and Q is a boolean variable with value 1 if the following condition is true and zero otherwise:

$$\begin{aligned} Q \;::\; & (((i_1 - i_3 = i_2 - i_4) \& (j_1 - j_3 = j_2 - j_4)) \text{ OR} \\ & ((i_1 - i_4 = i_2 - i_3) \& (j_1 - j_4 = j_2 - j_3))) \text{ \&NOT} \\ & ((i_1 = i_2) \& (j_1 = j_2) \& (i_3 = i_4) \& (j_3 = j_4)) \end{aligned} \tag{37}$$

We consider three examples to illustrate the application of the above formula. If the noise is modeled as Gaussian, the variable η with standard deviation σ_n we have[10] $E\{\eta^4\} = 3\sigma_n^4$. This result will be used in the following examples.

1. Gray Level Variance

The image is normalized by subtracting the mean grey value from the grey level of each pixel. The focus measure filter in this case is

$$a(i,j) = \begin{cases} 1 & \text{if } i = j = 0 \\ 0 & \text{otherwise} \end{cases} \tag{38}$$

Using the formula (36) we obtain

$$Var\{\gamma\} = \frac{2\sigma_n^4}{(2N+1)^2} + \frac{4\sigma_n^2}{(2N+1)^4}\sum_{m,n}^{N} f^2(m,n) \tag{39}$$

2. Gradient Magnitude Squared

There are two components: gradient squared along x-axis

$$a_x(i,j) = [-1 \quad 1] \tag{40}$$

and gradient squared along y-axis

$$a_y(i,j) = [-1 \quad 1]^T \tag{41}$$

We obtain

$$Var\{\gamma\} = \frac{24\sigma_n^4}{(2N+1)^2} + \frac{4\sigma_n^2}{(2N+1)^4}\sum_{m,n}^{M+N} [A_x(i,j) * f(m,n) + A_y(i,j) * f(m,n)]^2 \tag{42}$$

where

$$A_x(i,j) = a_x(i,j) * a_x(-i,-j) \tag{43}$$
$$A_y(i,j) = a_y(i,j) * a_y(-i,-j) \tag{44}$$

Note that a cross item of focus measure is generated by the effect of noise on x direction and y direction which are not independent.[7]

3. Laplacian

The discrete Laplacian is approximated by

$$a(i,j) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{45}$$

Substituting this $a(i,j)$ into formula (36) for variance we obtain

$$Var\{\gamma\} = \frac{1352\sigma_n^4}{(2N+1)^2} + \frac{4\sigma_n^2}{(2N+1)^4}\sum_{m,n}^{M+N} [A(i,j) * f(m,n)]^2 \tag{46}$$

where

$$A(i,j) = a(i,j) * a(-i,-j) \tag{47}$$

7 REFERENCES

[1] E. Krotkov, "Focusing", *International Journal of Computer Vision*, 1, 223-237, 1987.

[2] G. Ligthart and F. Groen, "A Comparison of Different Autofocus Algorithms", *International Conference on Pattern Recognition*, pp. 597-600,1982.

[3] Nayar, S.K., "Shape from Focus System" *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Champaign, Illinois, pp. 302-308 (June 1992).

[4] M. Subbarao, T. Choi, and A. Nikzad, "Focusing Techniques", *Journal of Optical Engineering*, Vol. 32 No. 11, pp. 2824-2836, November 1993.

[5] M. Subbarao, and M. C. Lu, "Computer Modeling and Simulation of Camera Defocus", *Machine Vision and Applications*, (1994) 7, pp. 277-289.

[6] M. Subbarao, and J.K. Tyan, "The Optimal Focus Measure for Passive Autofocusing and Depth-from-Focus", *Proceedings of SPIE conference on Viedometrics IV* , Philadelphia, Oct 1995.

[7] M. Subbarao, and J.K. Tyan, "The Optimal Focus Measure for Passive Autofocusing and Depth-from-Focus", *Tech. Report No. 96.07.01*, Computer Vision Laboratory, Dept. of Electrical Engineering, SUNY, Stony Brook (1996).

[8] Schlag, J.F., Sanderson, A.C., Neumann, C.P., and Wimberly, F.C., "Implementation of Automatic Focusing Algorithms for a Computer Vision System with Camera Control", *Tech. Report CMU-RI-TR-83-14*, Carnegie Mellon University (August 1983).

[9] H. H. Hopkins, "The frequence response of a defocused optical system" *Proceeding Royal Society of London*, A 231, 1955, pp 91-103.

[10] A. Papoulis, "*Probability, Random Variables, and Stochastic Processes*", McGraw-Hill Book Company, the 3rd, 1991.

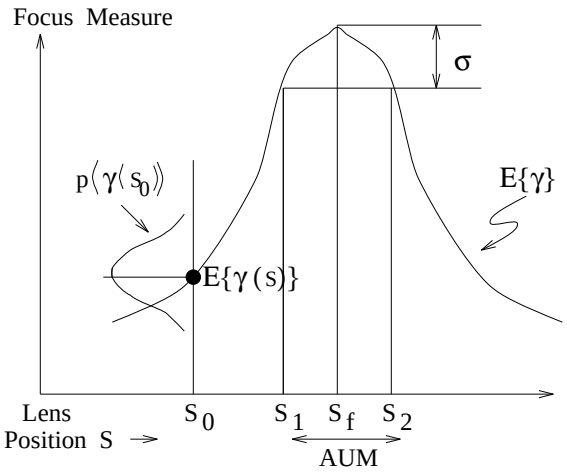


Figure 3: Definition of AUM at the focused position S_f.

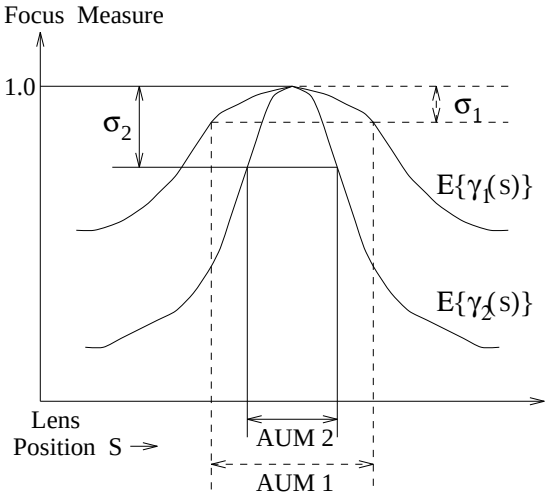


Figure 4: Comparison of two focus measures γ_1 and γ_2 at the focused position

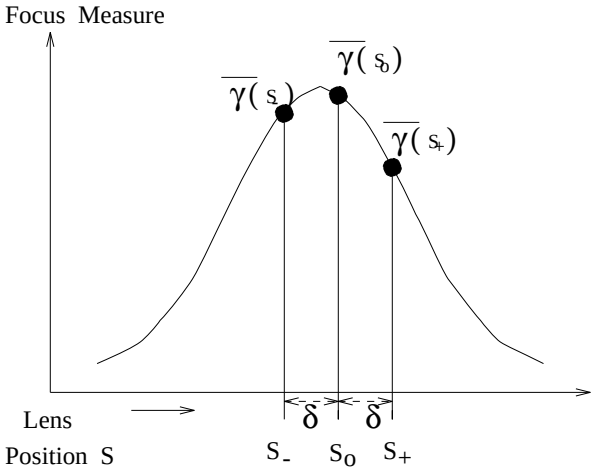


Figure 5: quadratic polynomial interpolation

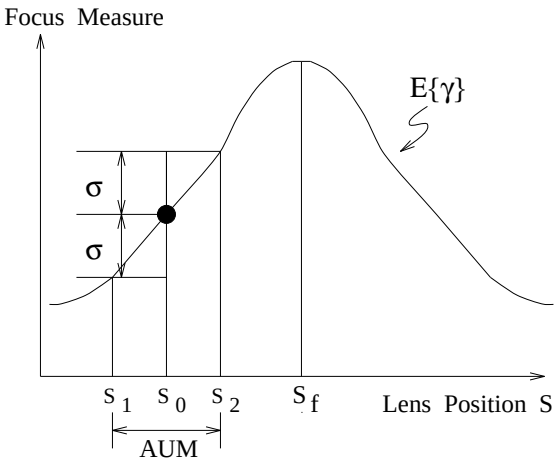


Figure 6: Definition of AUM at a position S_0 far from the focused position S_f.

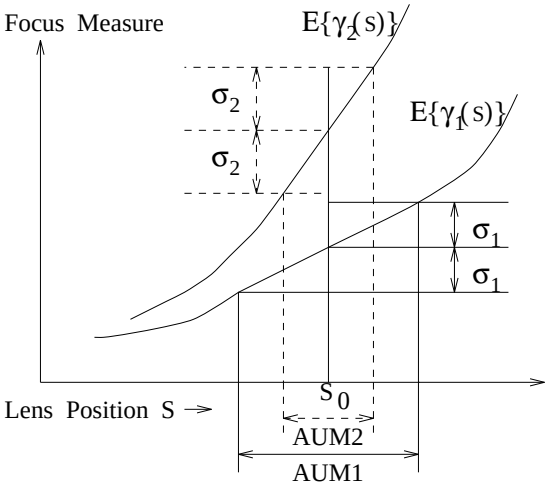


Figure 7: Comparison of two focus measures γ_1 and γ_2 at a position S_0 far from the focused position S_f. Note: $\sigma_2 > \sigma_1$ but $AUM2 < AUM1$, therefore γ_2 is better than γ_1.

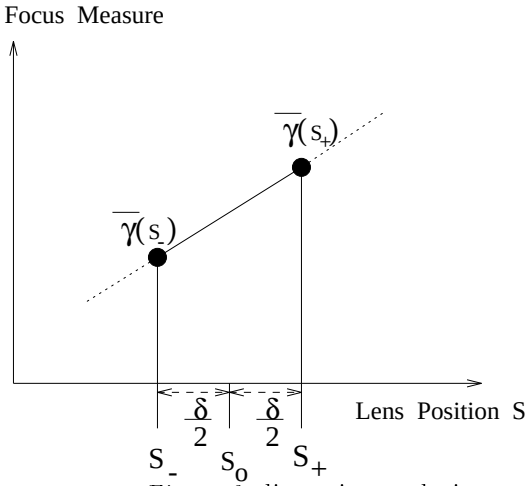
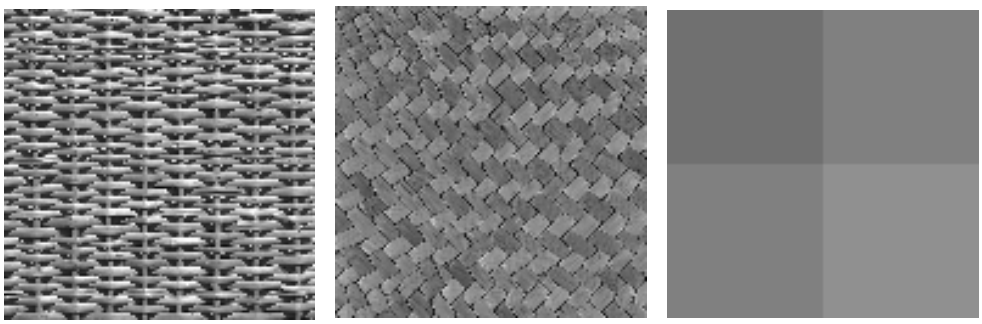


Figure 8: linear interpolation



Object A Object B Object C

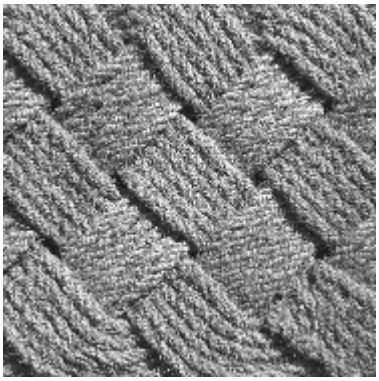
Figure 9: Texture image

Object A SNR: 35dB	Theoretical std of FM	Experiment std of FM	Theoretical ARMS	Experiment ARMS
Laplacian	18.92	17.27	0.020	0.018
Gradient	5.87	6.31	0.023	0.024
Variance	1.82	2.13	0.025	0.028

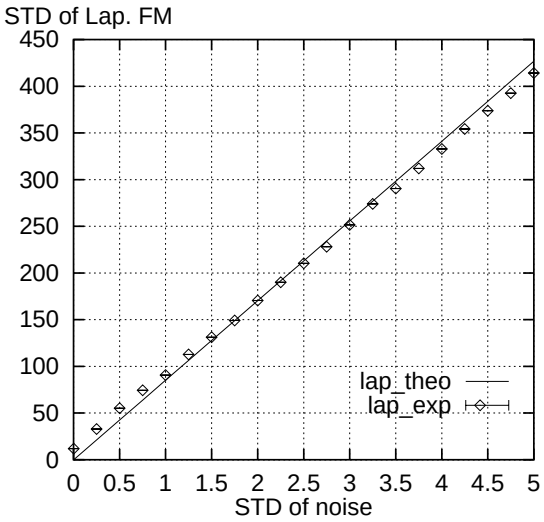
Object B SNR: 28dB	Theoretical std of FM	Experiment std of FM	Theoretical ARMS	Experiment ARMS
Laplacian	3.71	4.05	0.044	0.043
Gradient	1.06	1.25	0.048	0.049
Variance	0.85	1.02	0.10	0.11

Object C SNR: 20dB	Theoretical std of FM	Experiment std of FM	Theoretical ARMS	Experiment ARMS
Laplacian	1.67	1.37	0.09	0.10
Gradient	0.32	0.46	0.06	0.07
Variance	N/A	N/A	N/A	N/A

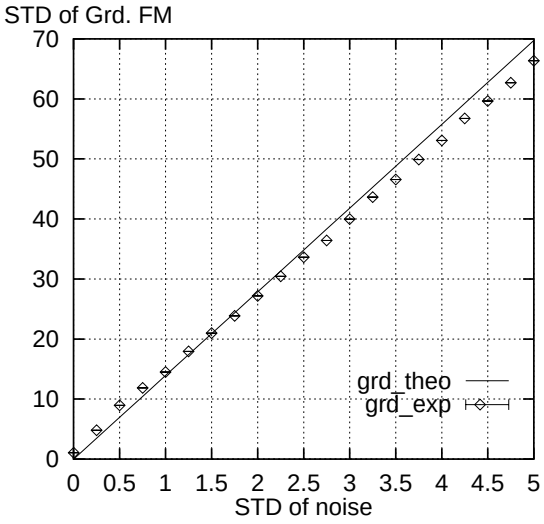
Table 1: Experimental records



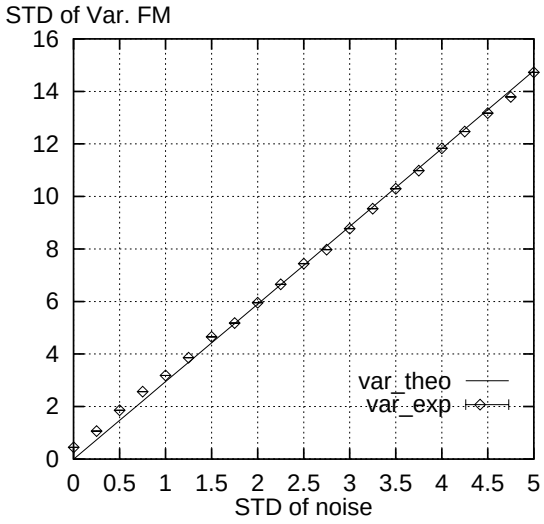
Test object
(a)



(b)



(c)



(d)

Figure 10: Simulation records

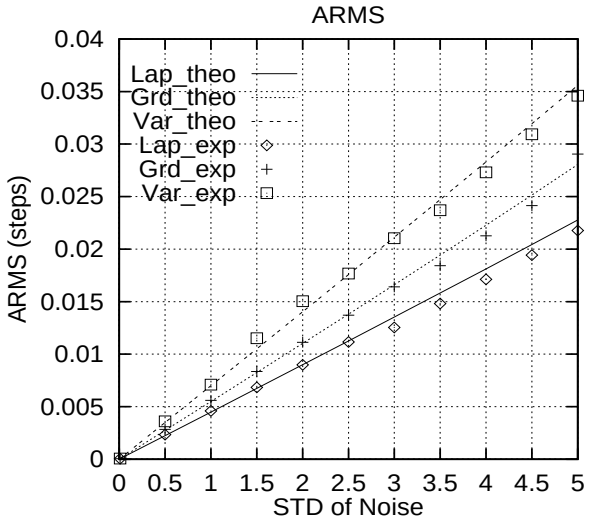


Figure 11: ARMS vs Noise

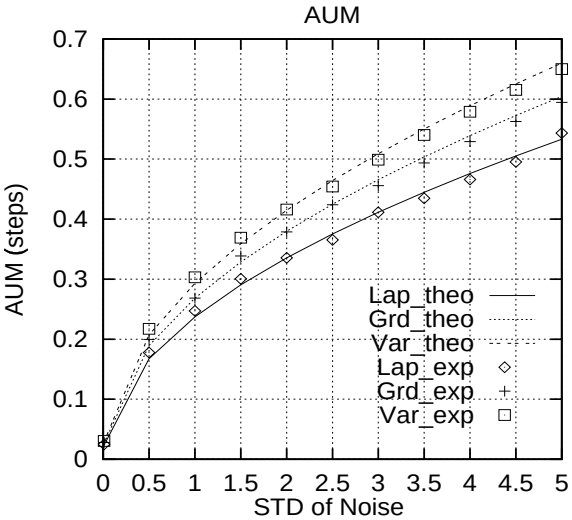


Figure 12: AUM vs Noise