Selecting the Optimal Focus Measure for Autofocusing and Depth-From-Focus

Murali Subbarao and Jenn-Kwei Tyan

Abstract—A method is described for selecting the optimal focus measure with respect to gray-level noise from a given set of focus measures in passive autofocusing and depth-from-focus applications. The method is based on two new metrics that have been defined for estimating the noise-sensitivity of different focus measures. The first metric—the *Autofocusing Uncertainty Measure* (AUM)—is useful in understanding the relation between gray-level noise and the resulting error in lens position for autofocusing. The second metric—*Autofocusing Root-Mean-Square Error* (ARMS error)—is an improved metric closely related to AUM. AUM and ARMS error metrics are based on a theoretical noise sensitivity analysis of focus measures, and they are related by a monotonic expression. The theoretical results are validated by actual and simulation experiments. For a given camera, the optimally accurate focus measure may change from one object to the other depending on their focused images. Therefore, selecting the optimal focus measure from a given set involves computing all focus measures in the set.

Index Terms—Focus measure, focusing, autofocusing, depth-from-focus, focus analysis.

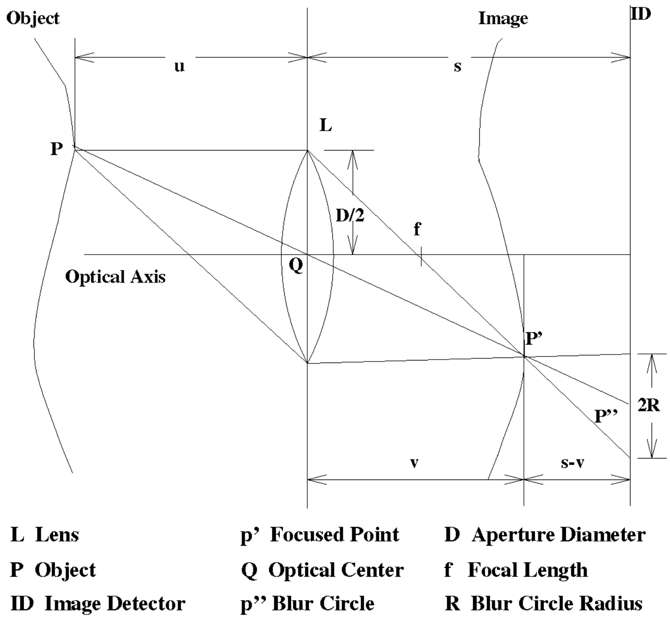

Fig. 1. Image formation in a convex lens.

1 Introduction

ELECTRONIC cameras can be autofocused by searching for the lens position that gives the best focused image [4], [5], [6]. In this approach, typically, a focus measure is computed for images acquired at several different lens positions, and the lens is moved to that position where the focus measure of the image is a maximum. The focused lens position v (see Fig. 1) depends on the distance u of the object to be focused and the focal length f of the lens. They are related by the lens formula

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

This same relation is used in depth-from-focus (DFF) methods to compute the object distance u from the focused lens position v [4], [5], [6].

Experimental evaluations of different focus measures have been reported in [1], [2], [3]. So far, there has not been any theoretical treatment of the noise sensitivity of focus measures. In the existing literature, all known work have been 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 image. This makes it difficult to arrive at general conclusions from experiments alone.

For a given camera and object, the most accurate focus measure can be selected from a given set through experiments as follows. For each focus measure, the object is autofocused several times, say 10, starting with an arbitrary default lens position. The mean of the 10 focused positions and their standard deviation are an estimate of the correct focused position and root-mean-square (RMS) error, respectively. The focus measure with the minimum estimate of RMS error is taken to be the optimal. In practical applications such as consumer video cameras or digital still cameras, it is desirable to find the best focus measure from a given set by autofocusing only once. It is quite undesirable to repeat 10 or several trials.

If one has detailed and accurate information on the focused image of the object to be focused and the camera characteristics such as its OTF, noise behavior, and camera parameters, then it would be possible to estimate the RMS error theoretically with only one trial. However such information is rarely available in practical applications.

In the absence of such detailed and accurate information, we propose two new metrics named *Autofocusing Uncertainty Measure* (AUM) and *Autofocusing Root-Mean-Square Error* (ARMS error) both of which can be computed with only one trial of autofocusing. In DFF applications, AUM and ARMS error can both be easily translated into uncertainties in depth using (1). The key assumption underlying the definition of AUM and ARMS error is that the mean value of focus measures are locally smooth with respect to lens position (e.g., quadratic near the peak). AUM and ARMS error metrics are general and applicable to any focus measure satisfying the local smoothness assumption. The analysis here deals with focusing errors caused only by gray-level noise and not other factors such as nonfront parallel surfaces. 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.

2 Model of Focus Measures

A detailed discussion of this topic can be found in several papers including [3]. Here, we summarize some relevant results based on geometric optics.

When a point object P is blurred on the image detector ID (see Fig. 1), it is imaged as a blur circle P” of radius R. This image $h(x, y)$ is the point spread function (PSF) of the camera. In a small image

• *The authors are with the Department of Electrical Engineering, State University of New York, Stony Brook, NY 11794-2350.*
E-mail: murali@sbee.sunysb.edu.

Manuscript received 12 Aug. 1996; revised 19 June 1998. Recommended for acceptance by S. Nayar.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 107052.

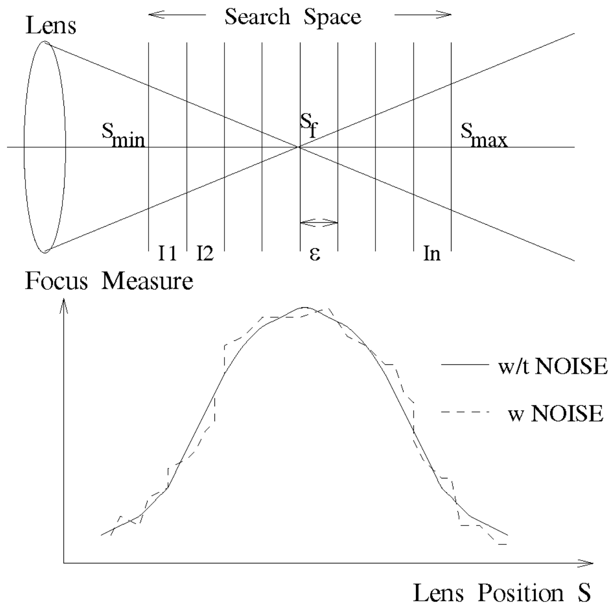

Fig. 2. Autofocusing algorithm.

region, if the imaged object surface is (approximately) a plane normal to the optical axis, then the PSF is the same for all points on the plane. Then the blurred image $g(x, y)$ in the small image region on the image detector ID is equal to the convolution of the focused image $f(x, y)$ and the PSF $h(x, y)$. Therefore, if G, F, and H are the Fourier transforms of g, f, and h, respectively, then $G = HF$. The OTF $H(\omega, \nu)$ has characteristics of a low-pass filter. As the blur increases, the higher frequencies are attenuated even more.

A general focus measure is modeled as follows. First, the image 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). Then, the energy (sum of squared values) of the filtered image is computed. This energy is the focus measure (see Section 6 for more details). Most FMFs correspond to filters that emphasize (or amplify) high frequencies. This seems appropriate since blurring has the opposite effect, i.e., high-frequency attenuation.

The focus measures modeled here cover most of the focus measures that have been used by researchers so far [3], except those based on sum of absolute values of image derivatives [1], [2]. Although AUM and ARMS error metrics are applicable to these focus measures also (which are based on sum of absolute values of image derivatives), we have not carried out a complete analysis of them here since they have been proved to be unsound [4], [5], [6] based on the effect of the OTF main lobe. These unsound focus measures may be optimal for some scenes, but for some other scenes they can give incorrect results even in the absence of all noise.

3 Autofocusing Algorithm

Fig. 2 shows a typical plot of a focus measure as a function of lens position. The problem is to find position s_f where the focus measure is maximum. Due to limited depth of field of the camera, we assume that the change in the best focused image is indistinguishable by the image detector when the lens is moved by an amount of up to $\pm\epsilon/2$ from s_f.

We propose the following algorithm for autofocusing. First the focus measure is computed at the current lens position and the lens is moved by about 10ϵ 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 used to narrow the search interval to about 10ϵ. The coarse search may use a binary or Fibonacci or a sequential search. See [4], [5], [6] for details. In this interval of size 10ϵ containing s_f, a focus measure is computed at three positions which are about 5ϵ apart. Then a quadratic or a Gaussian is fitted to these three (or more if desired) points. The position where the fitted curve has a maximum is taken to be the focused position s_f. Note that, according to geometric optics, the focus measure curve will be symmetric about the focus position s_f. Also, shifting the focus position s_f will shift the curve by the same amount with only small change in its shape.

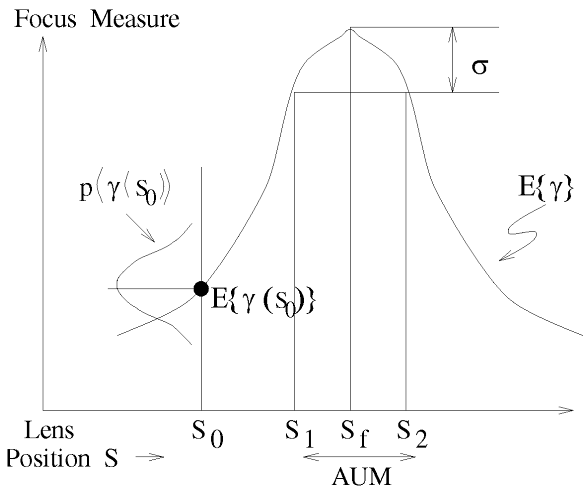

Fig. 3. Definition of AUM at the focused position S_f.

4 Autofocusing Uncertainty Measure (AUM)

First, we introduce AUM as a metric for focus measures to illustrate some underlying concepts. Later, we introduce the 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 $\mathrm{std}\{\gamma(s_0)\}$. 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 three to five 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. This assumption will be removed later in defining the ARMS error.

Referring to Fig. 3, the AUM at the maximum of the focus measure $\bar{\gamma}(s_f)$ is defined as follows:

$$AUM = s_2 - s_1, \tag{2}$$

where $s_1 < s_f < s_2$, $\left|\bar{\gamma}(s_f) - \bar{\gamma}(s_1)\right| = \left|\bar{\gamma}(s_f) - \bar{\gamma}(s_2)\right| = \sigma$, where σ is the standard deviation of the focus measure. In order to compute

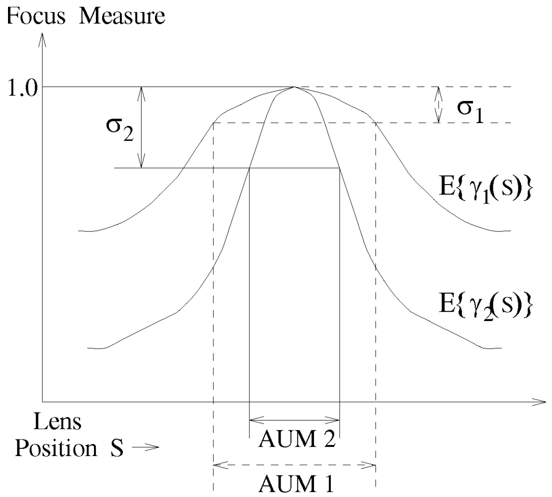

Fig. 4. Comparison of two focus measures γ_1 and γ_2 at the focused position.

AUM, we need to know σ. In Section 6, we derive a general formula that can be used to estimate σ as a function of the image and its noise level. Further we need to know the shape of the curve $\bar{\gamma}(s)$ near the peak. As discussed earlier, the position of $\bar{\gamma}_f$ and the function $\bar{\gamma}(s)$ near $\bar{\gamma}_f$ are estimated by fitting a curve (quadratic or Gaussian) to a few points (at least three) near the maximum. Intuitively, AUM is a measure similar to the RMS error in lens position that can be determined through repeated trials.

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.

For a position far away from the focused lens position s_f, see [4], [5], [6] for a definition of AUM.

5 ARMS Error

Now we derive an explicit expression for the *Autofocusing Root-Mean Square Error* (ARMS error). 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 is locally smooth with respect to lens position*. We model this local smoothness by a quadratic polynomial, but the analysis here can be extended to other models (e.g., cubic or Gaussian). However, such extensions do not appear to offer significant advantages compared to the quadratic model in practical applications.

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

$$\gamma(s) = as^2 + bs + c. \tag{3}$$

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 $\Gamma_- = \gamma(s_-)$, $\Gamma_0 = \gamma(s_0)$, and $\Gamma_+ = \gamma(s_+)$.

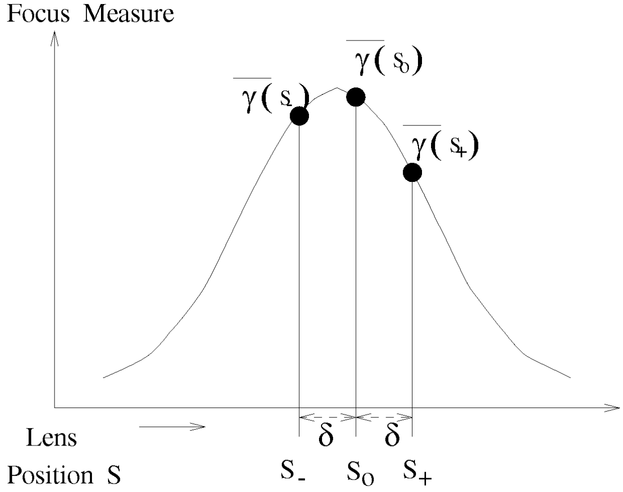

Fig. 5. Quadratic polynomial interpolation.

Near the focused position, $\Gamma_0 > \Gamma_-$ and $\Gamma_0 > \Gamma_+$. Solving for the coefficients of the quadratic expression, we obtain

$$a = \frac{\Gamma_+ + \Gamma_- - 2\Gamma_0}{2\delta^2},\ b = \frac{\Gamma_+ - \Gamma_-}{2\delta},\ c = \Gamma_0. \tag{4}$$

Let s_f be the lens position where the focus measure becomes the maximum and $\Gamma_f = \gamma(s_f)$. At s_f, the derivative of Γ vanishes. Therefore, we obtain

$$s_f = \frac{-b}{2a} = \frac{\delta}{2}\frac{\left(\Gamma_+ - \Gamma_-\right)}{\left(2\Gamma_0 - \Gamma_+ - \Gamma_-\right)}. \tag{5}$$

Substituting the above equation in (3), we obtain

$$\Gamma_f = -\frac{b^2 - 4ac}{4a}. \tag{6}$$

We are interested in the RMS value of s_f. 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:

$$\Gamma_i = \overline{\Gamma}_i + n_i \qquad \text{for } i = -,\ 0,\ +. \tag{7}$$

Therefore, (5) is expanded as

$$s_f = \frac{\delta}{2}\left(\frac{\overline{\Gamma}_+ - \overline{\Gamma}_-}{2\overline{\Gamma}_0 - \overline{\Gamma}_+ - \overline{\Gamma}_-}\right)\left(1 + \frac{n_+ - n_-}{\overline{\Gamma}_+ - \overline{\Gamma}_-}\right)\left(1 + \frac{2n_0 - n_+ - n_-}{2\overline{\Gamma}_0 - \overline{\Gamma}_+ - \overline{\Gamma}_-}\right)^{-1}. \tag{8}$$

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

$$\left|2\overline{\Gamma}_0 - \overline{\Gamma}_+ - \overline{\Gamma}_-\right| >> \left|2n_0 - n_+ - n_-\right|. \tag{9}$$

We obtain $s_f \approx s_f'$ where

$$s_f' = \bar{s}_f\left(1 + \frac{n_+ - n_-}{\overline{\Gamma}_+ - \overline{\Gamma}_-}\right). \tag{10}$$

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

$$s_f' = \bar{s}_f + \frac{\delta}{2}\left(\frac{n_+ - n_-}{2\overline{\Gamma}_0 - \overline{\Gamma}_+ - \overline{\Gamma}_-}\right). \tag{11}$$

Now the ARMS error is defined as the standard deviation of s_f', i.e.,

$$ARMS \text{ error} = \frac{\delta}{2}\frac{\sqrt{\sigma_+^2+\sigma_-^2}}{\left(2\overline{\Gamma}_0-\overline{\Gamma}_+-\overline{\Gamma}_-\right)}, \tag{12}$$

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

For a lens position away from the maximum focused position, we find that $\Gamma_- < \Gamma_{-0} < \Gamma_+$. 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 defined 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. 6). The linear model yields the expression

$$\frac{s-s_-}{s_+-s_-}=\frac{\Gamma-\Gamma_-}{\Gamma_+-\Gamma_-}. \tag{13}$$

The above equation can be rewritten as:

$$s=\delta\left(\frac{\Gamma-\Gamma_-}{\Gamma_+-\Gamma_-}\right)-\frac{\delta}{2}. \tag{14}$$

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 and assuming $\left|\overline{\Gamma}_+-\overline{\Gamma}_-\right| >> \left|n_+-n_-\right|$ and $\left|\overline{\Gamma}_+-\overline{\Gamma}_-\right| >> \left|2n_-\right|$, we obtain:

$$s' \approx \frac{\delta}{2}\left(\frac{n_+ + n_-}{\overline{\Gamma}_- - \overline{\Gamma}_+}\right). \tag{15}$$

Hence, the ARMS error would be

$$ARMS \text{ error} = std(s') = \frac{\delta\sqrt{\sigma_+^2+\sigma_-^2}}{2\left|\overline{\Gamma}_+-\overline{\Gamma}_-\right|}. \tag{16}$$

It is shown in [4], [5] that for points near the focused position, square of AUM is proportional to ARMS error $\left(AUM^2=\left(8\sqrt{2}\delta\right)ARMS\right)$, and for points away from the focused position, AUM and ARMS error are proportional $\left(AUM=2\sqrt{2}ARMS\right)$.

6 Mean and Variance of Focus Measures

In this section, we derive expressions for the *expected value* (mean) and *variance* of the focus measures modeled in Section 2. 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

$$f_\eta(m, n) = f(m, n) + \eta(m, n). \tag{17}$$

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. This σ_n can be easily estimated for a camera by imaging a uniformly bright object and computing the standard deviation of the gray level distribution. The images are assumed to be of size $(2N + 1) \times (2N + 1)$ and focus measure filter (FMF) $a(i, j)$ of size $(2M + 1) \times (2M + 1)$. Without loss of generality, the filtering operation will be represented by the *moving weighted sum* (MWS) operator instead of the usual *convolution* operator. MWS is correlation and is equivalent to convolution if, for example, the FMF is

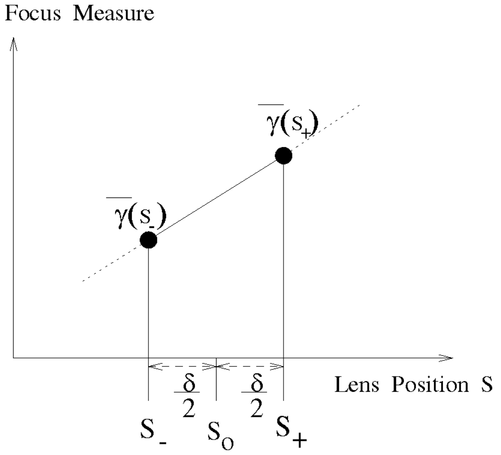

Fig. 6. Linear interpolation.

rotated by 180 degrees about its center by assigning $a(-i, -j)$ to $a(i, j)$. Denoting the MWS operator by $*$ it is defined by

$$a(i,j)*f_\eta(m,n)=\sum_{i,j}^{M}a(i,j)f_\eta(m+i,n+j), \tag{18}$$

where a double summation is abbreviated with a single summation to simplify notation.

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

$$g(m, n) = a(i, j) * f_\eta(m, n) = F(m, n) + \mathcal{N}(m, n), \tag{19}$$

where

$$F(m, n) = a(i, j) * f(m, n),\ \mathcal{N}(m, n) = a(i, j) * \eta(m, n). \tag{20}$$

The focus measure γ is defined as

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

$$=\gamma_{signal}+\gamma_{noise}+\frac{2}{(2N+1)^2}\sum_{m,n}^{N}F(m,n)\mathcal{N}(m,n), \tag{21}$$

where γ_{signal} and γ_{noise} are defined by:

$$\gamma_{signal}=\frac{1}{(2N+1)^2}\sum_{m,n}^{N}F^2(m,n),\ \gamma_{noise}=\frac{1}{(2N+1)^2}\sum_{m,n}^{N}\mathcal{N}^2(m,n). \tag{22}$$

Now, the expected value of the focus measure $E\{\gamma\}$ is (note that the expectation operator E is linear and commutes with summation):

$$E\{\gamma\}=\gamma_{signal}+E\{\gamma_{noise}\}+\frac{2}{(2N+1)^2}\sum_{m,n}^{N}F(m,n)E\{\mathcal{N}(m,n)\}$$

$$=\gamma_{signal}+A_n\sigma_n^2, \tag{23}$$

where

$$A_n=\sum_{i,j}^{M}a^2(i,j). \tag{24}$$

The above equation is a fundamental result. It shows that the expected value of the focus measure is a sum of two components– one due to signal alone and another due to noise alone. Therefore, if a focus measure is computed on a set of images for autofocusing, the effect of noise is to increase the computed focus measure by the same value on average for all images. The reason for this is that while the image signal changes in blur level with lens position, the noise characteristics of the camera remains the same. Therefore, the average increase in focus measure due to noise does not change the location of the focus measure peak. It is the variance of the focus measure that changes the location of the focus measure peak and therefore introduces error in autofocusing.

Now consider the variance of the focus measure:

$$Var\{\gamma\} = E\{\gamma^2\} - (E\{\gamma\})^2, \tag{25}$$

we obtain:

$$Var\{\gamma\} = \frac{A_n^2 E\{\eta^4\}}{(2N+1)^2} - \frac{A_n^2\sigma_n^4}{(2N+1)^2} + \frac{\sigma_n^4}{(2N+1)^2}\times$$

$$\sum_{i_1,j_1}^{M}\sum_{i_2,j_2}^{M}\sum_{i_3,j_3}^{M}\sum_{i_4,j_4}^{M} Q\cdot\left(\prod_{k=1}^{4} a(i_k,j_k)\right) + \frac{4\sigma_n^2}{(2N+1)^2}\gamma'_{signal}, \tag{26}$$

where

$$A(i,j) = a(i,j) * a(-i,-j), \quad F'(m,n) = A(i,j) * f(m,n), \tag{27}$$

$$\gamma'_{signal} = \frac{1}{(2N+1)^2}\sum_{m,n}^{M+N} F'^2(m,n), \tag{28}$$

and Q a Boolean variable with value 1 if the following condition is true and zero otherwise:

$$Q :: (C_5 \text{ OR } C_6) \,\&\, \text{NOT } C_1, \tag{29}$$

where

$$C_1 : (i_1 = i_2) \,\&\, (j_1 = j_2) \,\&\, (i_3 = i_4) \,\&\, (j_3 = j_4), \tag{30}$$

$$C_5 : (i_1 - i_3 = i_2 - i_4) \,\&\, (j_1 - j_3 = j_2 - j_4), \tag{31}$$

$$C_6 : (i_1 - i_4 = i_2 - i_3) \,\&\, (j_1 - j_4 = j_2 - j_3). \tag{32}$$

The equation above shows that the variance of a focus measure depends on the image signal in addition to noise level. The first three terms do not depend on the image signal. They can be computed and prestored. Among these three terms, the first two can be computed manually, but the third term may need a small computer program to evaluate. The last term in the above equation depends on the image being processed. Exact computation of this term requires knowledge of the noise-free image which is not possible. However the value of the term can be approximated using the noisy image $g(m, n)$ in place of $f(m, n)$. The approximation is valid for high signal to noise ratio [4], [5], [6].

The formula presented above can be applied directly in practical applications. Now we consider three examples to illustrate the application of the formula. In these examples, the noise will be modeled as Gaussian. For a zero mean Gaussian random variable η with standard deviation σ_n we have $E\{\eta^4\} = 3\sigma_n^4$. This result will be used in the following examples.

6.1 Gray Level Variance

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

$$a(i,j) = \begin{cases} 1 & \text{if } i = j = 0 \\ 0 & \text{otherwise} \end{cases}. \tag{33}$$

Using (26) for variance we obtain

$$Var\{\gamma\} = \frac{2\sigma_n^4}{(2N+1)^2} + \frac{4\sigma_n^2}{(2N+1)^4}\sum_{m,n}^{N} f^2(m,n). \tag{34}$$

6.2 Gradient Magnitude Squared

For gradient squared along x-axis and y-axis, respectively

$$a_x(i,j) = [-1 \;\; 1], \; a_y(i,j) = [-1 \;\; 1]^T. \tag{35}$$

Substituting $a_x(i, j)$ and $a_y(i, j)$ above in (26) for variance, we obtain:

$$Var\{\gamma\} = \frac{24\sigma_n^4}{(2N+1)^2} + \frac{4\sigma_n^2}{(2N+1)^4}\sum_{m,n}^{M+N}\left[A_x(i,j)*f(m,n) + A_y(i,j)*f(m,n)\right]^2, \tag{36}$$

where

$$A_x(i,j) = [-1 \;\; 2 \;\; -1], \; A_y(i,j) = [-1 \;\; 2 \;\; -1]^T. \tag{37}$$

6.3 Laplacian

The discrete Laplacian is approximated by

$$a(i,j) = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \tag{38}$$

Substituting this $a(i, j)$ into (26) for variance, we obtain

$$Var\{\gamma\} = \frac{1{,}352\sigma_n^4}{(2N+1)^2} + \frac{4\sigma_n^2}{(2N+1)^4}\sum_{m,n}^{M+N}\left[A(i,j)*f(m,n)\right]^2, \tag{39}$$

where

$$A(i,j) = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & -8 & 2 & 0 \\ 1 & -8 & 20 & -8 & 1 \\ 0 & 2 & -8 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}. \tag{40}$$

7 Experiments

In the first set of experiments, (26) 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) [3]. 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 gray level variance of the recorded image. Three objects labeled A, B, and C (see Fig. 7) were used in the experiments.

An object was placed in front of the camera, and for some fixed lens position, 10 images of size 32×32 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 (26). 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 (26) was obtained by averaging four 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 (26).

In the next experiment, the objects A, B, and C were autofocused using the algorithm described in Section 3. In each case, the experimental and theoretical ARMS error were computed (the unit is lens steps). Near the focus position, images were recorded at three positions s_-, s_0, and s_+ which were five steps apart. At each position, 10 images were recorded, and using these, the mean and the standard deviation of the focus measure there were computed. Then the theoretically estimated ARMS error was computed using (12). The same data was used to compute 10 experimental focus positions using (5). 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,

(a) (b) (c)

Fig. 7. Texture image. (a) Object A. (b) Object B. (c) Object C.

TABLE 1
EXPERIMENTALLY COMPUTED AND THEORETICALLY ESTIMATED STANDARD DEVIATIONS OF DIFFERENT FOCUS MEASURES

FMF	OBJ	SNR (dB)	std of FMF		ARMS	
			Theoretical	Experimental	Theoretical	Experimental
LAP	A	35	18.92	17.27	0.020	0.018
	B	28	3.71	4.05	0.044	0.043
	C	20	1.67	1.37	0.090	0.100
GRD	A	35	5.87	6.31	0.023	0.024
	B	28	1.06	1.25	0.048	0.049
	C	20	0.32	0.46	0.060	0.070
VAR	A	35	1.82	2.13	0.025	0.028
	B	28	0.85	1.02	0.100	0.110
	C	20	N/A	N/A	N/A	N/A

LAP: Laplacian, GRD: Gradient Magnitude Squared, VAR: Variance, OBJ: Object

and Laplacian squared. The measured noise standard deviation was 0.95 (gray-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, the 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. Autofocusing of object C was not possible using gray-level variance due to the absence of a well defined peak. Experiments similar to the ones above were also carried out on simulated image data (see [4], [5], [6]).

8 CONCLUSION

ARMS error has been defined as a metric for selecting the optimal focus measure for autofocusing with respect to gray-level noise from a given set of focus measures. It is based on the assumption of local smoothness of focus measures with respect to lens position. ARMS error can be applied to any focus measure whose variance can be expressed explicitly as a function of gray-level noise variance. Such an expression has been derived for a large class of focus measures that can be modeled as the energy of filtered images. Equations (23) and (26) for the mean and variance, respectively, of a focus measure along with (12) and (16) for ARMS error completely specify the dependence of autofocusing error on both gray-level noise and image content. These equations can be used to estimate the autofocusing accuracy of different focus measures, and the one with minimum error can be selected for application. In applications where computation needs to be minimized by computing only one focus measure, we recommend the use of the Laplacian as the focus measure filter. Laplacian has some desirable properties such as simplicity, rotational symmetry, elimination of unnecessary information, and retaining of necessary information [4], [5], [6].

This work can be extended in several ways. First, explicit expressions for the variance of other focus measures such as sum of absolute values of image derivatives could be derived so that ARMS error can be used to estimate their autofocusing accuracy. Second, in the definition of ARMS error, the local smoothness of focus measures could be modeled differently than here. Third, deriving an optimal focus measure filter for a given image and noise level remains to be investigated.

REFERENCES

[1] E. Krotkov, “Focusing,” *Int’l J. Computer Vision*, vol. 1, pp. 223-237, 1987.

[2] S.K. Nayar, “Shape From Focus System” *Proc. IEEE CS Conf. on Computer Vision and Pattern Recognition*, pp. 302-308, Champaign, Ill., June 1992.

[3] M. Subbarao, T. Choi, and A. Nikzad, “Focusing Techniques,” *J. Optical Eng.*, vol. 32, no. 11, pp. 2,824-2,836, Nov. 1993.

[4] M. Subbarao, and J.K. Tyan, “The Optimal Focus Measure for Passive Autofocusing and Depth-from-Focus,” *Proc. SPIE Conf. Viedometrics IV*, vol. 2,598, pp. 89-99, Philadelphia, Oct. 1995.

[5] M. Subbarao, and J.K. Tyan, “Root-Mean Square Error in Passive Autofocusing and 3D Shape Recovery,” *Proc. SPIE Conf., vol. 2,909*, Three-Dimensional Imaging and Laser-Based Systems for Metrology and Inspection II, pp. 162-177, Boston, Nov. 1996.

[6] J.K. Tyan, “Analysis and Application of Autofocusing and Three-Dimensional Shape Recovery Techniques Based on Image Focus and Defocus,” PhD thesis, Dept. of Electrical Eng., SUNY at Stony Brook, 1997.