An extensive empirical evaluation of focus measures for digital photography

Hashim Mir, Peter Xu, and Peter van Beek

Cheriton School of Computer Science
University of Waterloo, Canada

ABSTRACT

Automatic focusing of a digital camera in live preview mode, where the camera's display screen is used as a viewfinder, is done through contrast detection. In focusing using contrast detection, a focus measure is used to map an image to a value that represents the degree of focus of the image. Many focus measures have been proposed and evaluated in the literature. However, previous studies on focus measures have either used a small number of benchmarks images in their evaluation, been directed at microscopy and not digital cameras, or have been based on *ad hoc* evaluation criteria. In this paper, we perform an extensive empirical evaluation of focus measures for digital photography and advocate using three standard statistical measures of performance—precision, recall, and mean absolute error—as evaluation criteria. Our experimental results indicate that (i) some popular focus measures perform poorly when applied to autofocusing in digital photography, and (ii) simple focus measures based on taking the first derivative of an image perform exceedingly well in digital photography.

Keywords: Passive autofocus, contrast-detection, focus measures, live preview, digital camera

1. INTRODUCTION

Digital cameras are equipped with one or more passive autofocus mechanisms where a lens is focused using only the camera's optical system and an algorithm for controlling the lens. Passive autofocus mechanisms come in two basic kinds: contrast-detection and phase-detection. Contrast-detection autofocus is the most common—being standard in a wide range of cameras from mobile phones cameras, to point-and-shoots, to high-end DSLRs—whereas currently only high-end DSLRs also come equipped with phase-detection autofocus. Phase-detection is faster and better able to track subject movement, whereas contrast-detection is less costly and can be more accurate.[1] Our concern here is with contrast-detection autofocus.

Contrast-detection autofocus makes use of a focus measure that maps an image to a value that represents the degree of focus of the image. A contrast-detection autofocus algorithm iteratively moves the lens searching for an image with the maximal focus according to the focus measure. Lenses are moved by step motors that can be positioned at discrete positions. The images are acquired from the same stream of images that is displayed in the camera's live preview mode. (In live preview mode, the camera displays images streamed from the sensor at video frame rates; e.g., 24 frames per second on many Canon cameras.) The speed and accuracy of the autofocus algorithm are crucial to user satisfaction, and the choice of focus measure is an important factor in the performance of autofocus.

Previous studies on focus measures have either used a small number of benchmarks images in their evaluation, been directed at microscopy and not digital cameras, or have been based on *ad hoc* evaluation criteria. In this paper, we perform an extensive empirical evaluation of focus measures for digital photography and advocate using three standard statistical measures of performance—precision, recall, and mean absolute error—as evaluation criteria (see below for definitions). Our experimental results indicate that some popular focus measures perform poorly when applied to autofocusing in digital photography or to more representative sets of images. As well, our experimental results indicate that simple focus measures based on taking the first derivative of an image perform exceedingly well in digital photography, although care must be taken in which direction the derivative is taken.

Correspondence to Peter van Beek: vanbeek@uwaterloo.ca, `https://cs.uwaterloo.ca/~vanbeek/`
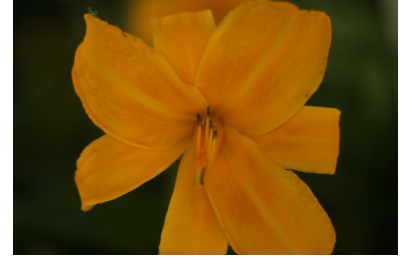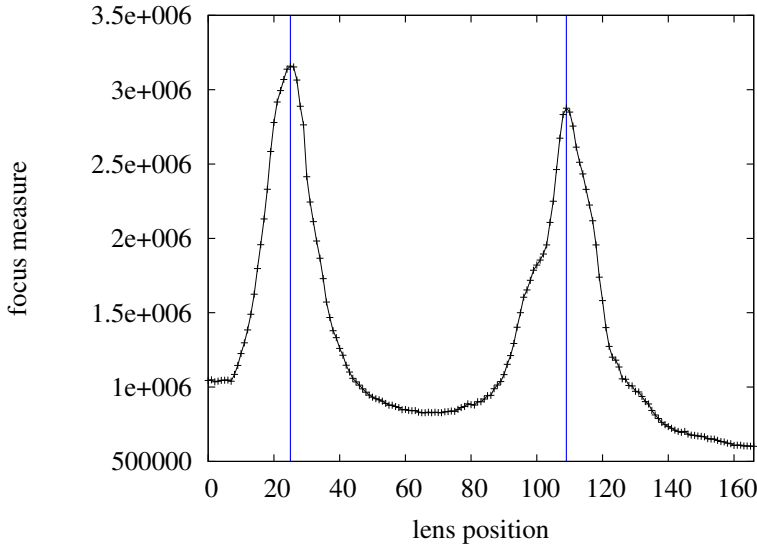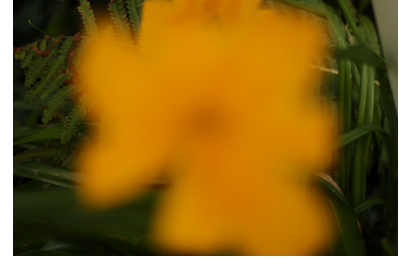
Figure 1. (a) Focus measures of images at each of the 167 lens positions (Canon 50 mm lens) for an example scene using the squared gradient focus measure. The two (blue) vertical bars refer to the two images that have objects that are in maximal focus: (b) flower in focus, and (c) fern and grasses in focus.

2. FOCUS MEASURES

Many focus measures have been proposed and evaluated in the literature (see, e.g.,[2–8]). Let $f(x, y)$ be the luminance or grayscale at pixel (x, y) in an image of size $M \times N$ pixels. As an example, the value $\phi(p)$ of the squared gradient focus measure for an image acquired when the lens is at position p is given by,

$$\phi(p) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-2} (f(x, y+1) - f(x, y))^2.$$

Figure 1a shows the squared gradient focus measure for the images acquired at all of the possible lens positions (Canon 50 mm lens) for an example scene. We assume that the region of interest (ROI) is the entire image. In practice, a user can either (i) specify the ROI by moving a rectangle over the desired part of the image when the camera is in live preview mode, or (ii) have the camera automatically determine the object or region of interest to bring into focus (e.g., using face or object recognition[9, 10]). Our evaluation methodology is easily adapted to the case where the ROI is an arbitrary sub-area of an image.

We review below the various focus measures that have been proposed in the literature and that we evaluated in our study. The focus measures are classified under six types: first derivative, second derivative, image histogram, image statistics, correlation, and data compression.

2.1 Focus measures based on first-order differentiation

Many focus measures based on first-order differentiation, or on smoothing followed by first-order differentiation, have been defined in the autofocusing literature or can be adapted from the image processing literature.

Brenner et al.[11] proposed a focus measure based on summing the squares of the horizontal first-derivative,

$$\sum_{x=0}^{M-1} \sum_{y=0}^{N-3} (f(x, y+2) - f(x, y))^2. \qquad \text{(Brenner)}$$

Brenner's focus measure has also been stated and evaluated in comparative studies using the gradient in the vertical direction,[3, 12]

$$\sum_{x=0}^{M-3} \sum_{y=0}^{N-1} (f(x+2, y) - f(x, y))^2. \qquad \text{(vertical Brenner)}$$

Variations on summing the squares of the gradients in the horizontal,[2, 7] and vertical direction,[12] have also been proposed and evaluated,

$$\sum_{x=0}^{M-1} \sum_{y=0}^{N-2} (f(x, y+1) - f(x, y))^2, \text{ and} \qquad \text{(squared gradient)}$$

$$\sum_{x=0}^{M-2} \sum_{y=0}^{N-1} (f(x+1, y) - f(x, y))^2. \qquad \text{(vertical squared gradient)}$$

In addition to using just the horizontal gradient (as in the Brenner and the squared gradient measures) or using just the vertical gradient (as in the vertical Brenner and the vertical squared gradient measures), the horizontal and vertical gradients can also be combined in various ways: (a) the sum of the squares of the horizontal and vertical gradients, (b) the sum of the absolute values of the horizontal and vertical gradients, (c) the maximum of the absolute values of the gradients, (d) the maximum of the squares of the gradients, (e) the absolute value of the horizontal gradient, and (f) the absolute value of the vertical gradient, For example, Santos et al.[4] use the maximum of the horizontal and vertical squared gradients. (In our experiments, we evaluated each of these alternatives and found that the versions based on just the horizontal gradient alone performed better than all other alternatives.)

Table 1. First-order derivative operators for constructing focus measures.

Operator	Horizontal	Vertical
3×3 difference	$\begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & +1 \\ 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & +1 & 0 \end{bmatrix}$
3×3 Sobel	$\begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix}$	$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ +1 & +2 & +1 \end{bmatrix}$
3×3 Scharr	$\begin{bmatrix} -3 & 0 & +3 \\ -10 & 0 & +10 \\ -3 & 0 & +3 \end{bmatrix}$	$\begin{bmatrix} -3 & -10 & -3 \\ 0 & 0 & 0 \\ +3 & +10 & +3 \end{bmatrix}$
3×3 Roberts	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & +1 & 0 \\ -1 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & +1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$
3×3 Prewitt	$\begin{bmatrix} -1 & 0 & +1 \\ -1 & 0 & +1 \\ -1 & 0 & +1 \end{bmatrix}$	$\begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ +1 & +1 & +1 \end{bmatrix}$
5×5 Sobel	$\begin{bmatrix} -1 & -2 & 0 & +2 & +1 \\ -4 & -8 & 0 & +8 & +4 \\ -6 & -12 & 0 & +12 & +6 \\ -4 & -8 & 0 & +8 & +4 \\ -1 & -2 & 0 & +2 & +1 \end{bmatrix}$	$\begin{bmatrix} -1 & -4 & -6 & -4 & -1 \\ -2 & -8 & -12 & -8 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ +2 & +8 & +12 & +8 & +2 \\ +1 & +4 & +6 & +4 & +1 \end{bmatrix}$

Focus measures have also been proposed that are based on first-order operators adapted from the image processing literature. These operators can be viewed as first performing smoothing, to reduce noise, and then differentiating. Table 1 summarizes the operators we considered in our study (not shown are 7×7 Gaussian derivative operators). Conceptually, a focus measure can be constructed from a pair of horizontal and vertical operators by convolving the operators with the image (see, for example, Gonzalez and Woods,[13] Nixon and Aguado[14]) and then summing the squares of these values. As noted above, the horizontal and vertical gradients can be combined in various ways. However, summing the squares of the gradients was found to perform the best of all alternatives. For example, consider the pair of 3×3 difference operators. The corresponding focus measure is given by,

$$\sum_{x=1}^{M-2} \sum_{y=1}^{N-2} (f(x+1,y) - f(x-1,y))^2 + (f(x,y+1) - f(x,y-1))^2. \qquad (3 \times 3 \text{ difference})$$

Geusebroek et al.[15] introduced a focus measure based on the Gaussian derivative operator. Krotkov[16] credits Tenenbaum with proposing the sum of squared 3×3 Sobel focus measure. The Sobel operator has been evaluated many times in the literature as a focus measure.[12, 16–19] Other comparative studies have included the Roberts operator,[17] the Prewitt operator,[18] and the Gaussian derivative operator.[12]

Table 2. Second-order derivative operators for constructing focus measures.

Operator	3×3	5×5
Laplacian	$\begin{bmatrix} -1 & -1 & -1 \\ -1 & +8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & -3 & -4 & -3 & -1 \\ -3 & 0 & +6 & 0 & -3 \\ -4 & +6 & +20 & +6 & -4 \\ -3 & 0 & +6 & 0 & -3 \\ -1 & -3 & -4 & -3 & -1 \end{bmatrix}$
Horizontal Sobel	$\begin{bmatrix} +1 & +2 & +1 \\ -2 & -4 & -2 \\ +1 & +2 & +1 \end{bmatrix}$	$\begin{bmatrix} +1 & +4 & +6 & +4 & +1 \\ 0 & 0 & 0 & 0 & 0 \\ -2 & -8 & -12 & -8 & -2 \\ 0 & 0 & 0 & 0 & 0 \\ +1 & +4 & +6 & +4 & +1 \end{bmatrix}$
Vertical Sobel	$\begin{bmatrix} +1 & -2 & +1 \\ +2 & -4 & +2 \\ +1 & -2 & +1 \end{bmatrix}$	$\begin{bmatrix} +1 & 0 & -2 & 0 & +1 \\ +4 & 0 & -8 & 0 & +4 \\ +6 & 0 & -12 & 0 & +6 \\ +4 & 0 & -8 & 0 & +4 \\ +1 & 0 & -2 & 0 & +1 \end{bmatrix}$
Cross Sobel	$\begin{bmatrix} -1 & 0 & +1 \\ 0 & 0 & 0 \\ +1 & 0 & -1 \end{bmatrix}$	$\begin{bmatrix} -1 & -2 & 0 & +2 & +1 \\ -2 & -4 & 0 & +4 & +2 \\ 0 & 0 & 0 & 0 & 0 \\ +2 & +4 & 0 & -4 & -2 \\ +1 & +2 & 0 & -2 & -1 \end{bmatrix}$

2.2 Focus measures based on second-order differentiation

Similarly to first-order differentiation, focus measures based on second-order differentiation have been defined in the autofocusing literature or can be adapted from the image processing literature. Table 2 summarizes the operators we considered in our study (not shown is a 9×9 Laplacian of Gaussian operator). Conceptually, a focus measure can be constructed from either a 3×3 or a 5×5 operator by convolving the operator with the image and then summing the squares of these values. The Laplacian operator has been evaluated in many comparative studies[2, 7, 17, 19, 20] as well as the square of the Laplacian.[21]

2.3 Focus measures based on image histogram

Several focus measures based on the image histogram have been defined. Firestone et al.[3] proposed a focus measure based on the range of the histogram,

$$\max\{h(k)\} - \min\{h(k)\}, \qquad \text{(range histogram)}$$

where $h(k)$ is the number of pixels with luminance value k. This focus measure has also been evaluated in other comparative studies.[4,12]

Firestone et al.[3] also proposed a focus measure based on the entropy,

$$-\sum_{k=0}^{255} p(k) \log_2 p(k), \qquad \text{(entropy histogram)}$$

where $p(k)$ is the relative frequency of luminance value k. Intuitively, entropy is a measure of the information content of the image. This focus measure has also been evaluated in other comparative studies.[4,7,12,19]

Mason and Green[22] proposed a focus measure based,

$$\sum_{k=\mu}^{255} (k - \theta_{MG}) \cdot h(k), \qquad \text{(M \& G histogram)}$$

where θ_{MG} is a threshold value and $h(k)$ is the number of pixels with luminance value k (see Mason and Green[22] for specification of θ_{MG}). This focus measure was also evaluated by Firestone et al.[3]

Mendelsohn and Mayall[23] proposed a focus measure,

$$\sum_{k=\mu}^{255} k \cdot h(k), \qquad \text{(M \& M histogram)}$$

where μ is the mean luminance of the image and $h(k)$ is the number of pixels with luminance value k. This focus measure was also evaluated by Firestone et al.[3]

2.4 Focus measures based on image statistics

Several focus measures based on various statistics of the distribution of luminance or grayscale values in the image have been defined. Focus measures based on the variance and normalized variance have been evaluated in many comparative studies,[2–4,7,12,17–19]

$$\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (f(x,y) - \mu)^2, \text{ and} \qquad \text{(variance)}$$

$$\frac{1}{MN\mu} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (f(x,y) - \mu)^2, \qquad \text{(normalized variance)}$$

where μ is the mean luminance of the image.

Threshold pixel count[2,4,12] is a focus measure based on the number of pixels below a threshold θ,

$$\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} i(f(x,y) < \theta), \qquad \text{(threshold pixel count)}$$

where i is an indicator function that returns 1 if its argument is true; 0 otherwise. In our experiments, we found that a threshold of 150 gave among the best results.

Threshold content[2, 4, 23] sums the luminance values above a threshold θ,

$$\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} s(f(x,y),\theta), \qquad \text{(threshold content)}$$

where s is a function that returns $f(x,y)$ if $f(x,y) \geq \theta$; 0 otherwise. In our experiments, we found that a threshold of 150 gave among the best results.

Power[3, 4, 12, 17] sums the square of the luminance values

$$\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y)^2. \qquad \text{(power)}$$

Thresholded versions of the power focus measure, where the sum is only over measure values $f(x,y)$ that exceed a threshold θ, have also been proposed. However, we found in our experiments that the best threshold was 0 and hence the threshold is omitted.

2.5 Focus measures based on correlation

Vollath[24, 25] proposed two focus measures, called F_4 and F_5, based on autocorrelation,

$$\sum_{x=0}^{M-2} \sum_{y=0}^{N-1} f(x,y)f(x+1,y) - \sum_{x=0}^{M-3} \sum_{y=0}^{N-1} f(x,y)f(x+2,y), \text{ and} \qquad \text{(Vollath's } F_4\text{)}$$

$$\sum_{x=0}^{M-2} \sum_{y=0}^{N-1} f(x,y)f(x+1,y) - MN\mu^2, \qquad \text{(Vollath's } F_5\text{)}$$

where μ is the mean luminance of the image. These focus measures have also been evaluated in other comparative studies.[4, 7, 12]

A focus measure based on a standard definition of autocorrelation from signal processing can also be defined,

$$(MN - k)\sigma^2 - \left(\sum_{x=0}^{M-1} \sum_{y=0}^{N-k-1} (f(x,y)-\mu)(f(x,y+k)-\mu) \right), \qquad \text{(autocorrelation)}$$

where μ and σ^2 are the mean and variance of the luminance of the image, respectively, and k is a fixed constant. In our experiments, we found that $k = 2$ produced the best results. Intuitively, the summation term of the autocorrelation focus measure computes the similarity between pixels k horizontal steps apart in the image.

2.6 Focus measures based on data compression

We also used focus measures based on data compression (for background on data compression see, for example, Witten et al.[26]). The intuition is that an in-focus image contains more information than an out-of-focus image, and that the relative amount of information corresponds to the file size when the image is compressed. For data compression, we used JPEG compression, which is designed for compressing images, and bzip2 and gzip, which are designed more for compressing text documents.

3. RELATED WORK

The problem of evaluating focus measures for passive contrast-based autofocusing has been well-studied in the literature. There have been several extensive comparative studies of focus measures, as well as the proposal of new focus measures followed by an experimental comparison to earlier proposals. The work can be categorized as focus measures for microscopy applications and focus measures for digital photography. Table 3 summarizes previous comparative studies.

Table 3. Summary of previous comparative studies.

setting	citation	conclusions
microscopy applications	Groen, Young, and Ligthart[2]	squared gradient
	Firestone et al.[3]	vertical Brenner, variance
	Santos et al.[4]	Vollath's F_4, F_5
	Liu, Wang, and Sun[12]	variance
digital photography	Subbarao, and Tyan[5]	Laplacian
	Ng, Neow, and Ang[19]	first-order 3×3 Sobel, variance
	Shih[18]	first-order 3×3 Prewitt and Sobel
	Yousefi, Rahman, and Kehtarnavaz[7]	Vollath's F_4

3.1 Focus measures for microscopy applications

Groen, Young, and Ligthart[2] perform a limited comparative study of focus measures for autofocusing in microscopy. The evaluation is on a small number of images with the result that the squared gradient is deemed the best focus measure for their application.

Firestone et al.[3] perform an extended comparative study of focus measures for autofocusing in microscopy. Their conclusion is that the vertical Brenner and the variance focus measures are the best focus measure for their application.

Santos et al.[4] perform an extended comparative study of focus measures for autofocusing in molecular cytogenetic analysis, a specific application that arises in microscopy. Their conclusion is that Vollath's F_4 and F_5 are the best focus measures for their application.

Liu, Wang, and Sun[12] perform an extended comparative study of focus measures for autofocusing in microscopic analysis of blood smears and pap smears. Liu et al. use the Fibonacci algorithm as part of their evaluation criteria. The Fibonacci algorithm is a search algorithm that finds the maximum of a function over an interval. The Fibonacci algorithm is, in general, unsuitable for digital photography because it assumes the function is unimodal (has a single peak over the interval). Liu et al.'s conclusion is that variance is the best focus measure for their application.

3.2 Focus measures for digital photography

Subbarao, and Tyan[5] theoretically study the noise-sensitivity of various focus measures. Their analysis can be used estimate the accuracy of different focus measures to select which focus measure to use given the noise properties of an image. If one must go with a single focus measure, rather than apply a portfolio of focus measures and select based on their analysis, their recommendation is the Laplacian.

Ng, Neow, and Ang[19] perform a limited evaluation of focus measures for machine vision applications and use accuracy and computation time as the evaluation criteria. (We argue that computation time may be a suitable tie-breaking criteria, but that accuracy is a much more important criteria in digital photography as an out-of-focus image can be disappointing at best to a photographer and worthless in the worst-case.) Their conclusion is that the first-order 3×3 Sobel and variance focus measures perform the best.

Shih[18] performs a comparative evaluation on real scenes that arise in digital photography and uses accuracy and unimodality as the evaluation criteria. (We argue that unimodality is an unsuitable criteria in digital photography as real images often do not give rise to unimodal focus measure functions; see, for example, Figures 1 & 2). Shih recommends the first-order 3×3 Sobel and Prewitt focus measures, but does not consider the Brenner or squared gradient focus measures in the comparison.

Yousefi, Rahman, and Kehtarnavaz[7]—in addition to introducing a new focus measure, denoted SOD, that is based on the derivative of a Gaussian—perform a comparative evaluation on real and simulated scenes. They use accuracy and computation time as their evaluation criteria. Yousefi, Rahman, and Kehtarnavaz[7] found that Vollath's F_4 was slightly more accurate, but that their proposed SOD focus measure was faster to compute.

4. EXPERIMENTAL METHODOLOGY

In this section, we describe the methodology we followed to perform a comparative evaluation of focus measures for digital photography. The experimental methodology consisted of two stages. All implementations were in C++ running under Windows 7*.

In the first stage, we implemented a camera remote control application whereby a camera is tethered to a computer via a USB cable and controlled by the software running on the computer. Our remote control application makes use of the Canon SDK (Version 2.11) and can control and replicate the basic functionality of the camera such as setting the aperture, displaying the live preview stream, and controlling the focus position of the lens. Using the remote control software, we gathered 25 sets of benchmark images that covered a range of common photography settings including landscapes, closeups, interiors, still lifes, and so on. Each of the sets of benchmark images contains either 167 (when using a 50 mm lens) or 231 (when using a 200 mm lens) jpeg images, one for each focus position of the lens. The jpeg images are captured from the live preview stream once the lens is moved from one position to the next. There were a total of 4,303 images (23 of the benchmarks have 167 images; two of the benchmarks have 231 images). The camera used in our experiments was a Canon EOS 550D/Rebel T2i.

In the second stage, we implemented more than 30 focus measures that we found in an extensive survey of the literature (see Section 2). Given a benchmark set of images and a focus measure, the focus measure was applied to each jpeg image in the benchmark. The peaks in the graph of a focus measure correspond to the lens position (or image) where the focus measure predicts the corresponding object will be in maximum focus. The predictions of the focus measure were then compared to the ground truth, where the ground truth is the true lens position (or image) where the corresponding object was in maximum focus. The ground truth was determined by one of the authors viewing the images but there was little or no subjectivity involved as the image in sharpest focus was generally clear. In some cases two or three images were considered equally in focus and in those cases we considered all of these as true or correct. Given the predictions of each focus measure and the ground truth, we then calculated the precision and recall of each focus measure.

Table 4. Notation for defining precision and recall.

	ground truth	
predicted	tp (true positive)	fp (false positive)
	fn (false negative)	tn (true negative)

Consider the notation shown in Table 4. For example, tp means that the image corresponding to a peak in the focus measure was in maximal focus (i.e., the image that the focus measure predicted was in maximal focus was correct) and fn means that an image that was in maximal focus did not correspond to a peak in the focus measure (i.e., the focus measure failed to predict that the image was in maximal focus). Given this notation, the standard definitions of precision and recall can be stated,

$$\text{precision} = \frac{tp}{tp + fp}, \qquad \text{recall} = \frac{tp}{tp + fn}.$$

We also calculated the mean absolute error, $|x_t - \hat{x}_t|$, where x_t is the lens position of the actual image that is in maximal focus and $\hat{x}_t$ is the lens position of the image that a focus measure predicts is in maximal focus (see Figure 2 for an example). The mean absolute error was chosen over the mean squared error as the former prefers prediction errors that are occasionally large but small on average, while the latter prefers prediction errors that avoid large errors while still possibly being quite unsatisfactory overall. In other words, if focus fails, it should fail big. There is little that is more frustrating than to return from a shoot to discover that, while an image looked sharp in preview mode on the camera, one has *just* failed to capture a subject, such as the eyes in a portrait shot, in sharp focus.

---

*The software and data are available at: `https://cs.uwaterloo.ca/~vanbeek/research`.
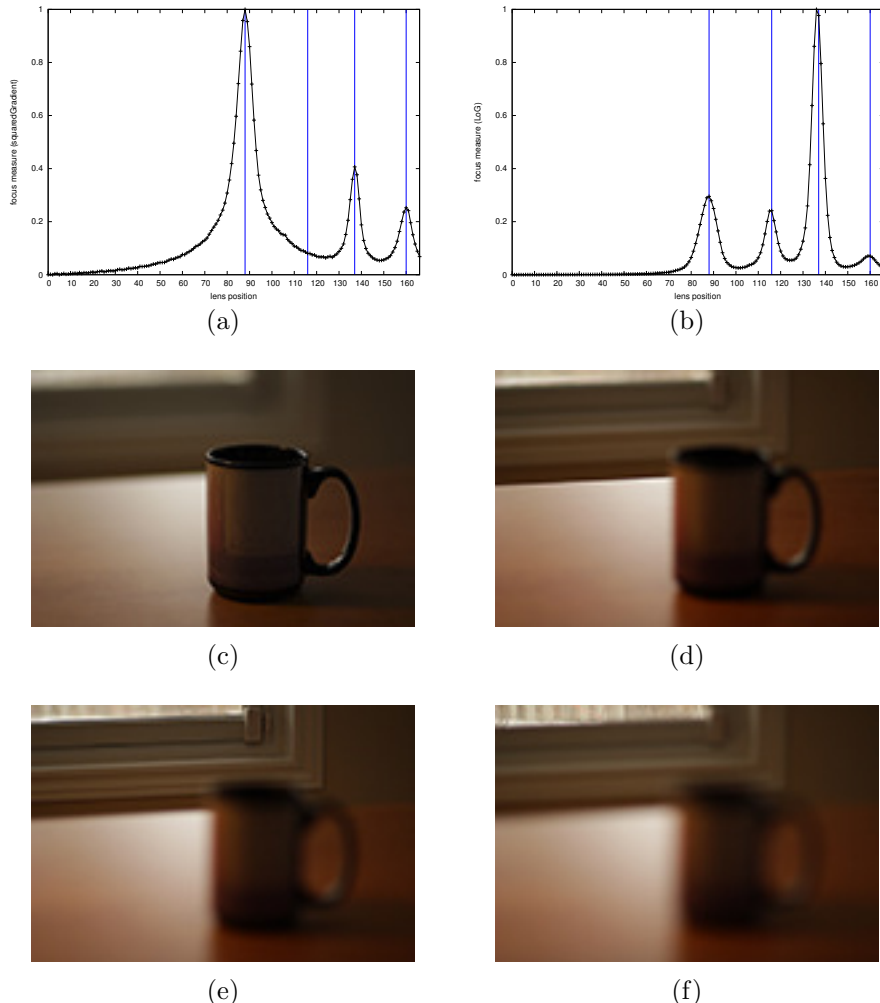
Figure 2. Focus measures of images at each of the 167 lens positions (50 mm lens) for an example scene using (a) the squared gradient focus measure, and (b) the Laplacian of Gaussian focus measure. The four vertical bars refer to the four images that have objects that are in maximal focus: (c) cup, (d) edge of desk, (e) window sill, (f) and fence in focus. On this benchmark, the squared gradient has precision 3/3 and recall 3/4, and the Laplacian of Gaussian has precision 4/4 and recall 4/4.

5. EXPERIMENTAL RESULTS AND DISCUSSION

Table 5 shows a summary of the results of our experiments. Several focus measures stand out from the rest. In terms of precision, the Brenner and the squared gradient measures—simple and fast to compute measures that are based on taking the first derivative of the image—are perfectly precise. In terms of recall, the Brenner and the squared gradient also perform well, and are matched only by some of the larger second-derivative convolution operators, such as the Laplacian of Gaussian and the 5×5 vertical Sobel. The Brenner and the squared gradient also have the lowest (perfect) mean squared error. Of particular note is how our results contrast with previous experimental studies, especially those for digital photography (see Table 3). Our experimental results indicate that some popular focus measures performed poorly when applied to autofocusing in digital photography or to more representative sets of images. As well, our experimental results indicate that care must be taken in which direction the derivative is taken. Previous work has not explicitly noted that the Brenner and the squared gradient can be formulated in both the horizontal and vertical directions and compared the various formulations. For microscopy, the nature of the images ensures that the images are isotropic—or similar in all directions—and therefore the focus measures are rotationally invariant. Such is not the case in photography.

Table 5. Precision, recall, and mean absolute error (MAE) of various focus measures on the benchmarks.

type	focus measure	precision	recall	MAE
first derivative	Brenner	100.00	99.00	0.00
	squared gradient	100.00	99.00	0.00
	3×3 difference	98.00	97.00	0.02
	3×3 Sobel	98.00	97.00	0.02
	3×3 Scharr	98.00	97.00	0.02
	3×3 Roberts	98.00	97.00	0.02
	3×3 Prewitt	98.00	97.00	0.02
	Gaussian	95.00	92.00	0.05
	vertical squared gradient	91.33	92.00	0.55
	vertical Brenner	91.33	90.00	0.23
	5×5 Sobel	85.67	85.00	0.14
second derivative	Laplacian of Gaussian	98.00	99.00	0.32
	5×5 vertical Sobel	96.67	99.00	0.65
	5×5 Laplacian	96.67	99.00	0.46
	5×5 cross Sobel	90.33	93.00	0.66
	3×3 cross Sobel	89.67	93.00	1.18
	3×3 vertical Sobel	88.73	96.00	3.94
	3×3 Laplacian	85.00	91.00	1.41
	5×5 horizontal Sobel	83.33	88.00	1.08
	3×3 horizontal Sobel	73.27	78.00	2.95
image histogram	range histogram	20.46	41.00	44.27
	entropy histogram	16.86	32.00	48.63
	M & G histogram	15.24	24.00	53.48
	M & M histogram	1.03	4.00	42.01
image statistics	normalized variance	43.57	53.00	26.36
	variance	32.33	35.00	20.80
	threshold pixel count	4.24	18.00	52.38
	threshold content	2.13	4.00	36.31
	power	0.53	2.00	43.80
correlation	autocorrelation	80.40	99.00	26.87
	Vollath's F_4	61.74	87.00	24.60
	Vollath's F_5	33.00	35.00	19.00
compression	jpeg	66.38	78.00	10.77
	bzip2	50.64	70.00	34.28
	gzip	40.44	53.00	26.98

6. CONCLUSION

The speed and accuracy of a digital camera's contrast-based autofocus algorithm are crucial to user satisfaction. Contrast-detection autofocus makes use of a focus measure that maps an image to a value that represents the degree of focus of the image and the choice of focus measure is an important factor in the performance of autofocus. Previous studies on focus measures have either used a small number of benchmarks images in their evaluation, been directed at microscopy and not digital cameras, or have been based on *ad hoc* evaluation criteria. In this paper, we perform an extensive empirical evaluation of focus measures. To compare the focus measures, we gathered an extensive set of benchmark images that covered a range of common photography settings. To evaluate the focus measures, we advocate using three standard statistical measures of performance—precision, recall, and mean absolute error—as evaluation criteria (see below for definitions). Our experimental results indicate that two simple focus measures based on taking the first derivative of an image perform exceedingly well in digital photography.

REFERENCES

[1] Cicala, R., "Autofocus reality," (July 2012). Available at: http://www.lensrentals.com/blog/2012/07.

[2] Groen, F., Young, I., and Ligthart, G., "A comparison of different focus functions for use in autofocus algorithms," *Cytometry* **6**, 81–91 (1985).

[3] Firestone, L., Cook, K., Culp, K., Talsania, N., and Preston, K., "Comparison of autofocus methods for automated microscopy," *Cytometry* **12**, 195–206 (1991).

[4] Santos, A., Ortiz de Solórzano, C., Vaquero, J. J., Peña, J. M., Malpica, N., and del Pozo, F., "Evaluation of autofocus functions in molecular cytogenetic analysis," *Journal of Microscopy* **188**, 264–272 (1997).

[5] Subbarao, M. and Tyan, J.-K., "Selecting the optimal focus measure for autofocusing and depth-from-focus," *IEEE Trans. Pattern Anal. Mach. Intell.* **20**, 864–870 (1998).

[6] Jeon, J., Lee, J., and Paik, J., "Robust focus measure for unsupervised auto-focusing based on optimum discrete cosine transform coefficients," *IEEE Trans. Consum. Electron.* **57**, 1–5 (2011).

[7] Yousefi, S., Rahman, M., and Kehtarnavaz, N., "A new auto-focus sharpness function for digital and smartphone cameras," *IEEE Trans. Consum. Electron.* **57**, 1003–1009 (2011).

[8] Tsai, D. C. and Chen, H. H., "Reciprocal focus profile," *IEEE Trans. Image Processing* **21**, 459–468 (2012).

[9] Lee, S. Y., Kumar, Y., Cho, J. M., Lee, S. W., and Kim, S. W., "Enhanced autofocus algorithm using robust focus measure and fuzzy reasoning," *IEEE Trans. Circuits Syst. Video Tech.* **18**, 1237–1246 (2008).

[10] Rahman, M. and Kehtarnavaz, N., "Real-time face-priority auto focus for digital and cell-phone cameras," *IEEE Trans. Consum. Electron.* **54**, 1506–1513 (2008).

[11] Brenner, J., Dew, B., Horton, J., King, J., Neirath, P., and Sellers, W., "An automated microscope for cytologic research," *J. Histochem. Cytochem.* **24**, 100–111 (1971).

[12] Liu, X. Y., Wang, W. H., and Sun, Y., "Dynamic evaluation of autofocusing for automated microscopic analysis of blood smear and pap smear," *Journal of Microscopy* **227**, 15–23 (2007).

[13] Gonzalez, R. C. and Woods, R. E., [*Digital Image Processing, 3rd Edition*], Prentice Hall (2008).

[14] Nixon, M. and Aguado, A., [*Feature Extraction & Image Processing, 2nd Edition*], Academic Press (2008).

[15] Geusebroek, J.-M., Cornelissen, F., Smeulders, A. W. M., and Geerts, H., "Robust autofocusing in microscopy," *Cytometry* **39**, 1–9 (2000).

[16] Krotkov, E., "Focusing," *Int. J. Comput. Vis.* **1**, 223–237 (1987).

[17] Lee, J.-H., Kim, K.-S., Nam, B.-D., Lee, J.-C., Kwon, Y.-M., and Kim, H.-G., "Implementation of a passive automatic focusing algorithm for digital still camera," *IEEE Trans. Consum. Electron.* **41**, 449–454 (1995).

[18] Shih, L., "Autofocus survey: A comparison of algorithms," in [*Proc. SPIE 6502, Digital Photography III*], (2007).

[19] Ng, N. K. C., Neow, P. A., and Ang Jr., M. H., "Practical issues in pixel-based autofocusing for machine vision," in [*Proc. of the IEEE International Conference on Robotics and Automation*], 2791–2796 (2001).

[20] Nayar, S. and Nakagawa, Y., "Shape from focus," *IEEE Trans. Pattern Anal. Mach. Intell.* **16**, 824–831 (1994).

[21] Subbarao, M. and Wei, T.-C., "Depth from defocus and rapid autofocusing: A practical approach," in [*Proc. of IEEE Conference on Computer Vision and Pattern Recognition*], 773–776 (1992).

[22] Mason, D. and Green, D., "Automatic focusing of a computer controlled microscope," *IEEE Trans. Biomed. Eng.* **22**, 312–317 (1975).

[23] Mendelsohn, M. and Mayall, B., "Computer-oriented analysis of human chromosomes," *Comput. Biol. Med.* **2**, 137–150 (1972).

[24] Vollath, D., "Automatic focusing by correlative methods," *Journal of Microscopy* **147**, 279–288 (1987).

[25] Vollath, D., "The influence of the scene parameters and of noise on the behaviour of automatic focusing algorithms," *Journal of Microscopy* **151**, 133–146 (1988).

[26] Witten, I. H., Moffat, A., and Bell, T. C., [*Managing Gigabytes: Compressing and Indexing Documents and Images*], Morgan Kaufmann, 2nd ed. (1999).