

# A Bayes-Spectral-Entropy-Based Measure of Camera Focus Using a Discrete Cosine Transform

Matej Kristan<sup>\*</sup>, Janez Perš, Matej Perše, Stanislav Kovačič

*Faculty of Electrical Engineering, University of Ljubljana, Tržaška 25, 1001 Ljubljana, Slovenia*

---

## Abstract

In this paper we present a novel measure of camera focus based on the Bayes spectral entropy of an image spectrum. In order to estimate the degree of focus, the image is divided into non-overlapping sub-images of 8 by 8 pixels. Next, sharpness values are calculated separately for each sub-image and their mean is taken as a measure of the overall focus. The sub-image spectra are obtained by an  $8 \times 8$  discrete cosine transform (DCT). Comparisons were made against four well-known measures that were chosen as reference, on images captured with a standard visible-light camera and a thermal camera. The proposed measure outperformed the reference measures by exhibiting a wider working range and a smaller failure rate. To assess its robustness to noise, additional tests were conducted with noisy images.

*Key words:* Camera focusing, Spectral methods, DCT, Bayes entropy

---

## 1. Introduction

Proper camera focus has a profound effect on image quality and therefore plays an important role in many computer vision applications. Focus can be achieved manually, or automatically using an active or passive approach. Active methods are based on emitting a sound wave or an infrared signal in order to estimate the distance from an object, and then using the measured distance, they calculate the appropriate lens position. On the other hand, passive methods rely solely on the captured image of an object being focused, by adjusting the lens parameters with respect to some predefined criteria of camera focus. Some passive systems use specialized image

---

<sup>\*</sup> Corresponding author.

Tel: ++ 386 1 4768 876

Fax: ++ 386 1 4768 279

*Email address:* [matej.kristan@fe.uni-lj.si](mailto:matej.kristan@fe.uni-lj.si)

(Matej Kristan).

sensors, while others maximize the image sharpness to obtain the optimum focus value. The latter rely solely on the quality of the acquired image, they do not estimate the distance from an object and thus do not require any special-purpose hardware in addition to a CCD sensor. Active systems are fast, but restricted by the distance from the object of interest, while computationally more costly passive systems rely on the image of the photographed object and are therefore not subjected to these restrictions. Passive focusing based on image quality is a common tool in general-purpose CCD cameras and consists of three major parts: the region being focused (the focusing window); a measure of focus, which yields a maximum at the focused position; and an optimization algorithm, which searches for a global maximum of the focus measure. Many types of focus measure have been investigated in microscopy (Firestone et al., 1991; Yeo et al., 1993; Santos et al., 1997; Geusebroek et al., 2000), digital holography (Gillespie et al., 1989) and digital photography (Krotkov, 1987; Subbarao and Tyan, 1998; Widjaja and Jutamulia, 1998; Choi et al., 1999; Chern et al., 2001; Lin et al., 2004; Yap and Raveendran, 2004). Usually, these measures are calculated using certain norms of the image gradients or derivatives (Krotkov, 1987; Subbarao and Tyan, 1998; Chern et al., 2001), manipulation of gray level histograms (Krotkov, 1987; Chern et al., 2001) or operations on the longest lines in the images (Lin et al., 2004).

In this paper we present a new measure of focus based on estimating the uniformity of the image spectrum. We assume that when a typical image is out of focus its amplitude spectrum is less uniform than when it is focused. To assess this uniformity we made use of the Bayes entropy. The image is transformed to its spectral representation via a discrete cosine transform (DCT), which has a low computational complexity and is most famous for being a part of the JPEG compression standard (Pennebaker and Mitchell, 1993). Since most consumer digital cameras already incorporate JPEG standard compression (and are hence capable of calculating the DCT) we define our measure of focus in a way that makes it compatible with this standard.

The remainder of this paper is organized as follows: We first describe the 2D-DCT in Section 2, and the normalized DCT in Section 3. Section 4 defines the basis of our proposed measure, the implementation of which is then described in Section 5. The experiments are described and commented on in Section 6, and finally, the conclusions are drawn in Section 7.

## 2. Discrete cosine transform of an image

Mathematically, the DCT is a one-to-one mapping between the spatial and spectral domains. An image is transformed to its spectral representation by projection onto a set of orthogonal 2-D basis functions. The amplitudes of these projections are called the *DCT coefficients*, and they are the output of the transform. The spectral representation of an image function  $f(x, y)$  defined on  $x \in \{0, \dots, N - 1\}$ ,  $y \in \{0, \dots, M - 1\}$  is thus obtained via a DCT operator  $\mathcal{F}_C$  as a set of real-valued DCT coefficients:

$$\mathcal{F}_C [f(x, y)] = \{B_{\omega\nu}(f(x, y))\} \tag{1}$$

defined as

$$B_{\omega\nu}(f(x, y)) = \sum_{(x, y)} f(x, y) C_{\omega\nu}(x, y; M, N), \quad (2)$$

where  $C_{\omega\nu}$  are the basis functions of the transform and are defined by the following expressions:

$$\begin{aligned} C_{\omega\nu}(x, y; M, N) &= c_\omega(x; M) c_\nu(y; N) \\ c_\varphi(z; A) &= \alpha(\varphi; A) \cos\left(\frac{\pi\varphi(2z+1)}{2A}\right) \\ \alpha(\varphi, A) &= \begin{cases} \frac{1}{\sqrt{A}} & ; \varphi = 0 \\ \sqrt{\frac{2}{A}} & ; otherwise \end{cases} \end{aligned} \quad (3)$$

Low-order basis functions represent low spatial frequencies, while those of higher orders represent the higher spatial frequencies. Thus, the coefficients of lower orders depict slow spatial variations in image intensity, while those of higher orders depict rapid spatial variations. In a typical image the intensity varies slowly from point to point across the image, which means that its spectrum obtained with the DCT is dominated by the coefficients of lower orders.

It is well known that the DCT is closely related to the discrete Fourier transform (DFT), which is a standard tool in signal analysis and has previously been reported by authors like [Krotkov \(1987\)](#), [Gillespie et al. \(1989\)](#) and [Chern et al. \(2001\)](#) as the preferred transform for spectral-based focusing algorithms. However, the DCT has a greater energy-compaction property than the DFT, meaning that most of the image information tends to be concentrated in a few low-frequency DCT coefficients, approaching the Karhunen-Loève transform (which is the optimum in the decorrelation sense) for signals based on certain limits of Markov processes. As a consequence, when the DCT is used, a smaller number of coefficients have to be calculated to achieve a good spectral representation of an image. This is also the primary reason why the JPEG standard compression scheme is based on the DCT – see, for example, [Pennebaker and Mitchell \(1993\)](#). Another advantage of using the DCT to obtain a spectral representation of an image over using the DFT, is that its basis functions are composed entirely of real-valued components. And since the DFT's basis functions are complex, this entails a larger computational load than when using the DCT. Furthermore, [Krotkov \(1987\)](#) and [Yeo et al. \(1993\)](#) suggest that the spectrum of the DFT contains some information that is superfluous to the task of focusing (e.g., the phase information). Based on these arguments, we chose the DCT to be the basis of our focus measure in preference to the DFT.

### 3. Normalized DCT

Let  $f(x, y)$  be a grayscale image defined on  $x \in \{0, \dots, N-1\}$ ,  $y \in \{0, \dots, M-1\}$ . The normalized DCT (nDCT) of an image  $f(x, y)$  is defined as

$$\tilde{\mathcal{F}}_C[f](\omega, \nu) = \frac{|\mathcal{F}_C[f](\omega, \nu)|}{\sum_{(\omega, \nu) \in \mathcal{D}} |\mathcal{F}_C[f](\omega, \nu)|}, \quad (4)$$

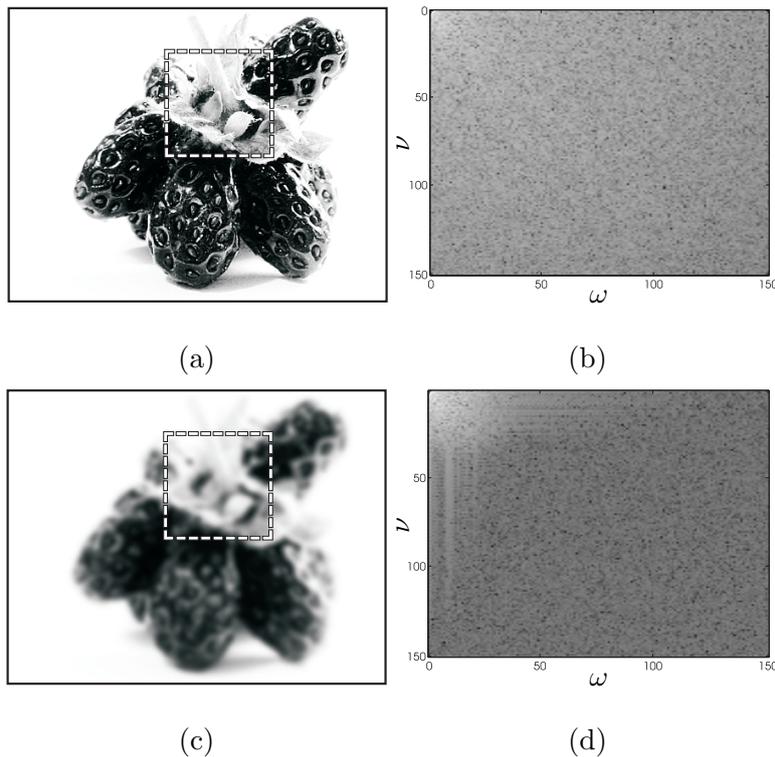


Fig. 1. Illustrative example of the relation between focus and the normalized DCT. The left-hand column shows a sharp image (a) and the same image blurred with a Gaussian filter (c). Normalized DCTs were calculated within the dashed rectangles and for the purposes of visualization,  $\log$ s of the transforms are shown in the right-hand column. The coefficients with higher values are shown in light-gray and those with lower values are shown in dark-gray. By visual comparison of (b) and (d) it can be observed that the focused image produces a much more uniform spectrum than its blurred counterpart.

where  $\mathcal{D}$  denotes the spectral domain, and  $|\cdot|$  is the  $L_1$ -norm. A convenient property of this transform is its invariance to changes in the contrast of an image. The latter can be shown by writing  $f'(x, y) = \alpha f(x, y)$ , where  $\alpha$  is a nonzero scaling factor. Due to the linearity of a DCT operator, the nDCT of a modified image can be written as:

$$\begin{aligned} \tilde{\mathcal{F}}_{\mathcal{C}} [f'] (\omega, \nu) &= \frac{\alpha \left| \tilde{\mathcal{F}}_{\mathcal{C}} [f] (\omega, \nu) \right|}{\sum_{(\omega, \nu) \in \mathcal{D}} \alpha \left| \tilde{\mathcal{F}}_{\mathcal{C}} [f] (\omega, \nu) \right|} \\ &= \tilde{\mathcal{F}}_{\mathcal{C}} [f] (\omega, \nu) \end{aligned} \quad (5)$$

Equation (5) implies that the normalized DCT is contrast invariant, and hence any measure defined on this transform is contrast invariant as well.

The relationship between the image focus and the image spectrum is illustrated in Fig. 1, where the normalized DCTs are shown for blurred and sharp images. For a better visualization the  $\log$  of the transform was taken, and it can be clearly seen that the normalized DCT of a blurred image contains a large mode at low frequencies, whereas the sharp image produces a very uniform spectrum.

#### 4. Entropy of a normalized spectrum

A normalized DCT spectrum of an image  $f(x, y)$  is a real-valued positive function, defined on the spectral domain  $\mathcal{D}$ , whose values sum to one. This means that it can be thought of as a probability distribution function, and its dependency on the focus can be viewed as follows: When an image is out of focus its normalized spectrum exhibits a pronounced mode at very low frequencies. As the image approaches the optimum focus, higher spectral components appear and the mode becomes less pronounced. Therefore, we assume that when an image becomes focused its normalized DCT (or at least the nonzero part of it) tends to a uniform distribution.

A common measure that yields a maximum when the underlying distribution is uniform is the entropy function. Let us assume that an image contains noise that is independent and identically distributed (i.i.d.). The power spectrum of such noise is uniform. This means that an i.i.d. noise affects the image spectrum uniformly, and we can assume that the coefficients corresponding only to the noise have smaller values than those corresponding to the rest of the image content. We therefore need an entropy measure that weights the lower-valued coefficients less than the higher-valued ones. This is a property of a Bayes entropy function (Devijver, 1974), defined as

$$H^2 = 2\left(1 - \sum_{i \in \mathcal{I}} p_i^2\right), \quad (6)$$

where  $p_i$  is a distribution function defined on domain  $\mathcal{I}$ . In our case,  $p_i$  is simply the normalized DCT of an image, and the measure of the image sharpness is

$$\begin{aligned} M'_{Be}(f) &= 1 - \sum_{\omega=0}^{M-1} \sum_{\nu=0}^{N-1} \tilde{F}_C(\omega, \nu)^2 \\ &= 1 - \frac{\sum_{\omega=0}^{M-1} \sum_{\nu=0}^{N-1} |F_C(\omega, \nu)|^2}{\left(\sum_{\omega=0}^{M-1} \sum_{\nu=0}^{N-1} |F_C(\omega, \nu)|\right)^2}, \end{aligned} \quad (7)$$

where  $F_C(\omega, \nu)$  is a DCT of an image and  $\tilde{F}_C(\omega, \nu)$  is its normalized DCT spectrum. The scaling factor in Eq. (6) was ignored, since it does not affect the shape of the sharpness measure. Note that this measure is defined on a normalized DCT spectrum, which means it is independent of the image contrast, and yet for its actual calculation no normalization of the spectrum is needed.

The normalized DCT spectrum of a typical noiseless image is dominated by the coefficients of lower orders, so the coefficients of higher orders should be very close to zero. However, in the presence of noise, these high-order coefficients take some nonzero values, and since they provide very little or no information about the camera focus, they should be ignored. We thus define a measure of image sharpness that ignores the coefficients of an order higher than some predefined value  $t$  as

$$M'_{Be}(f, t) = 1 - \frac{\sum_{\omega+\nu \leq t} |F_C(\omega, \nu)|^2}{\left( \sum_{\omega+\nu \leq t} |F_C(\omega, \nu)| \right)^2}. \quad (8)$$

## 5. Implementation

The main weakness of spectral methods is that they are computationally expensive and require special-purpose hardware for real-time applications (Yeo et al., 1993). We therefore based our measure of focus on an existing solution, i.e., the DCT, which operates on images of  $8 \times 8$  pixels, also referred to as  $8 \times 8$  DCT. This particular transform was chosen because many efficient schemes for its calculation exist (Wallace, 1992) and hardware implementations are commonly available (Fernandez et al., 2000).

According to the JPEG compression standard (Pennebaker and Mitchell, 1993), which is present in nearly every consumer digital camera, an image is firstly divided into sub-images of  $8 \times 8$  pixels that are separately transformed by  $8 \times 8$  DCT and then compressed. In order to comply with this scheme, our measure of focus is calculated as follows: (i) The analyzed part of the image (i.e., the focusing window) is divided into non-overlapping sub-images of size  $8 \times 8$  pixels. (ii) Each sub-image is transformed with  $8 \times 8$  DCT, and its value of sharpness is calculated. (iii) An overall image-focus value is estimated by taking the mean of all the sub-image sharpness values.

Let  $f(x, y)$  be an image-focusing window of size  $M \times N$ ,  $x \in \{0, \dots, M-1\}$ ,  $y \in \{0, \dots, N-1\}$ , where  $M$  and  $N$  are multiples of 8. Such an image comprises  $\frac{M}{8} \times \frac{N}{8}$  non-overlapping sub-images of  $8 \times 8$  pixels  $s_{ij}^8(x, y)$ , defined as

$$s_{ij}^8(x, y) = \{f(x, y)\}_{x=8i, y=8j}^{x=8i+7, y=8j+7}, \quad (9)$$

where  $i = 0, \dots, \frac{M}{8} - 1$ ,  $j = 0, \dots, \frac{N}{8} - 1$ . The Bayes-spectral-entropy-based focus measure of the image  $f(x, y)$  is then defined as

$$M_{Be}^8(f) = \underset{i,j}{mean} [\{M'_{Be}(s_{ij}^8(x, y), t)\}], \quad (10)$$

where the sharpness measure  $M'_{Be}$  is defined by Eq. (8). The parameter  $t$  in Eq. (10) was determined experimentally and is set to  $t = 6$ .

## 6. Experimental study

To assess the proposed focus measure  $M_{Be}^8$ , experiments were conducted using a standard visible-light camera and a thermal camera. The proposed measure was compared against four reference measures by observing their behavior on images with different spatial and noise contents.

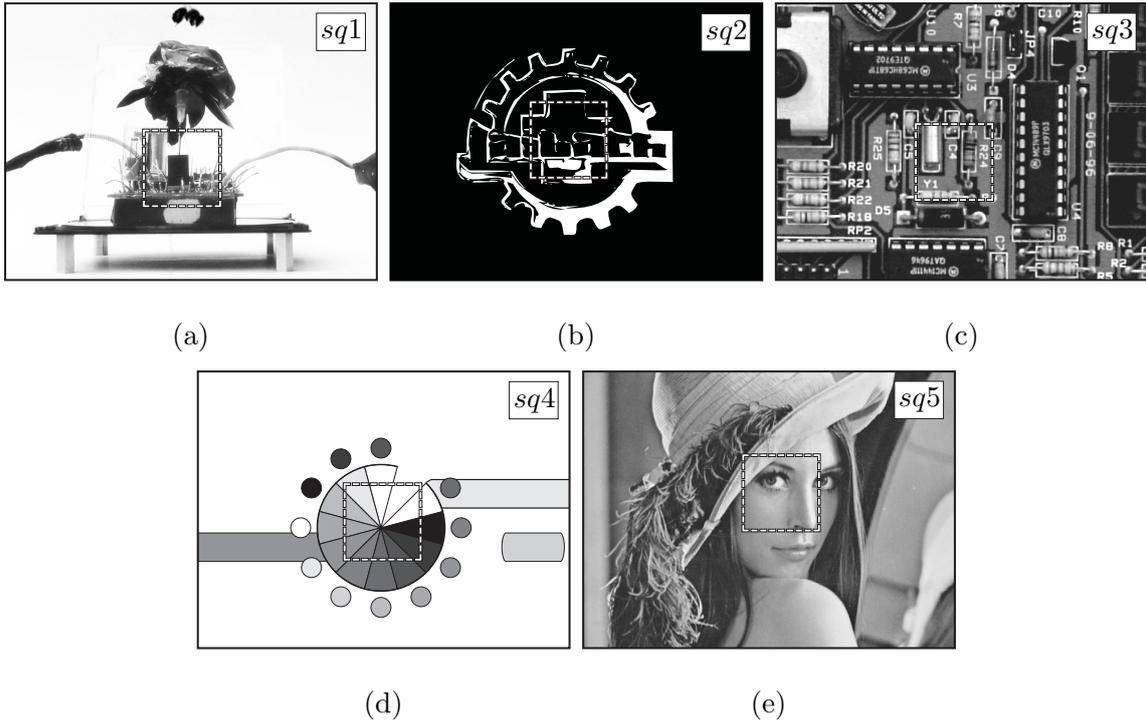


Fig. 2. Five images (a, b, c, d, e) with different content were used as targets in the experiment with the visible-light camera. For each of the five targets a sequence of differently focused images was recorded. These sequences are denoted as  $sq1$ ,  $sq2$ ,  $sq3$ ,  $sq4$  and  $sq5$ , respectively. The dashed rectangles in the images depict the regions used for focusing (i.e., focusing windows).

### 6.1. Acquisition of test images

A *DFW-X700 Sony* digital camera with a *Cosmicar 75mm* television lens was used to record the visible-light images. Computer control of the focus was facilitated via a stepper motor with which 36 equidistant positions of focus could be achieved, numbered from 0 to 35. Five planar targets (Fig. 2) were placed in front of the camera at a distance of 1.6m, with the optimum focus achieved at the 26<sup>th</sup> position. For each target a sequence of 36 differently focused grayscale images of size  $640 \times 480$  pixels was recorded, each image corresponding to a different focus position. Measures were calculated on focusing windows of  $128 \times 128$  pixels.

The thermal images were recorded using a *FLIR ThermoCAM<sup>TM</sup> E25* thermal camera with a standard  $25^\circ$  lens. A stepper motor was used to control the focus with 19 positions, numbered from 0 to 18. Three objects (Fig. 3a,b,c) were placed in front of the camera at a distance of 1.4m, with the optimum focus achieved at the 13<sup>th</sup> position. The first two objects were parts of a powered electrical circuit, and the third object was wrinkled aluminium foil. For each of the three objects a series of differently focused grayscale images (each  $160 \times 120$  pixels) was recorded while gradually changing the focus from 0 to 18. The measures were calculated on focusing windows of  $64 \times 64$  pixels.

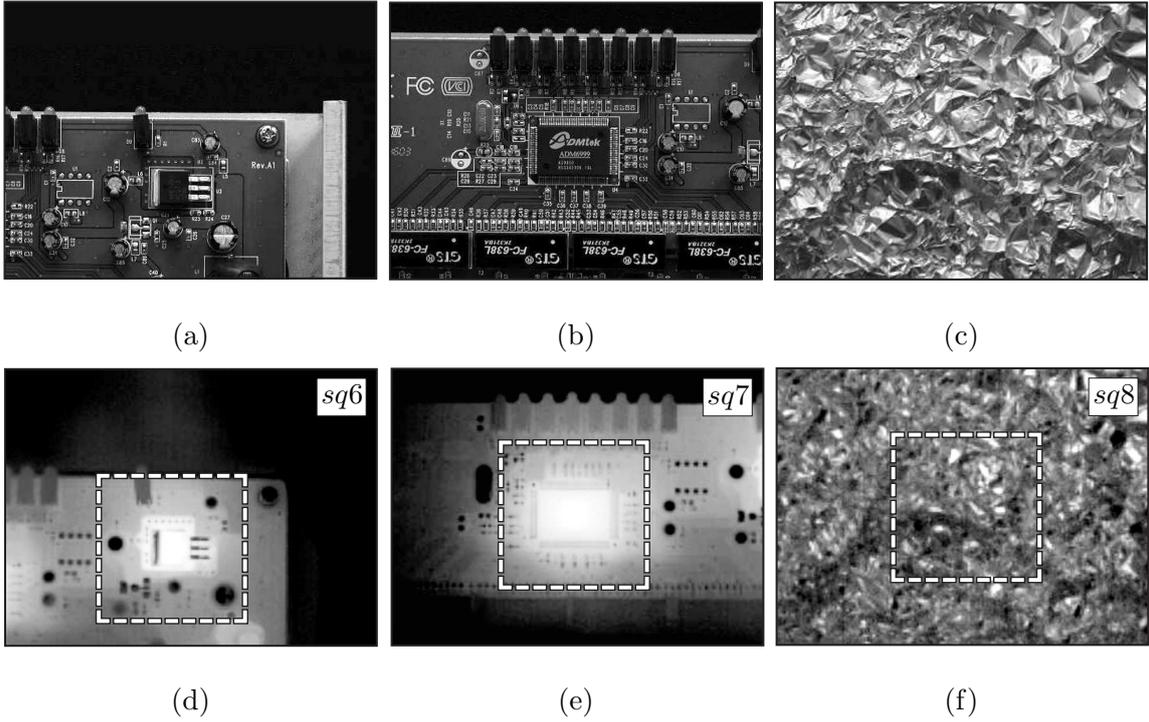


Fig. 3. Images of objects (a, b, c) used in tests with the thermal camera and their thermal images (d, e, f). For each of the three objects a sequence of differently focused images was recorded. These three sequences are denoted as *sq6*, *sq7* and *sq8*. The focusing windows for all the sequences are depicted by the dashed rectangles.

## 6.2. Noise of image sequences

The signal-to-noise ratio (SNR) was estimated for each sequence of images in the following way. For each series of recorded images, the image corresponding to the last step was recorded twice. We denote these last two images as  $f_1(x, y)$  and  $f_2(x, y)$ . Assuming that image noise is an i.i.d. process, its variance can be estimated as one half of the expected squared difference between the images  $f_1(x, y)$  and  $f_2(x, y)$ ,

$$\sigma_n^2 = \frac{1}{2} \langle (f_1 - f_2)^2 \rangle \quad (11)$$

and the SNR for a particular series can therefore be estimated from the ratio between the range of image intensity and the standard deviation of the image noise (Eq. 12). The SNR values for all the sequences are presented in Table 1.

$$SNR = 20 \log_{10} \left( \frac{\max(f_1) - \min(f_1)}{\sigma_n} \right) \quad (12)$$

Table 1  
Signal-to-noise ratios of the test sequences.

	<i>visible-light camera</i>					<i>thermal camera</i>		
sequence no.	1	2	3	4	5	6	7	8
SNR[dB]	39	38	37	34	38	42	40	27

Sequence numbers refer to the sequences depicted in Figs. 2 and 3.

### 6.3. Reference measures of focus

Numerous focus measures have been presented in the literature (Krotkov, 1987; Gillespie et al., 1989; Firestone et al., 1991; Yeo et al., 1993; Santos et al., 1997; Choi et al., 1999; Subbarao and Tyan, 1998; Widjaja and Jutamulia, 1998; Geusebroek et al., 2000; Chern et al., 2001; Lin et al., 2004; Yap and Raveendran, 2004). From among these we selected four well-known measures to compare against our measure. The first measure was the energy of the Laplacian image, which was selected by Subbarao and Tyan (1998) as the optimum focus measure in terms of the error, and is defined as

$$M_{Le}(f) = \sum_{(x,y)} L(x,y)^2, \quad (13)$$

where  $L(x,y)$  is the convolution of the focusing window with a Laplace operator.

The second measure was the energy of an image convoluted with Sobel operators, also known as Tenengrad measure (Krotkov, 1987), and defined as

$$M_{Tg} = \sum_{(x,y)} (S_x(x,y)^2 + S_y(x,y)^2), \quad (14)$$

where  $S_x(x,y)$  and  $S_y(x,y)$  are the convolutions of the image with Sobel horizontal and vertical operators, respectively.

The third measure was based on evaluating the image-intensity differences two pixels apart, also known as the Brenner measure. This measure was reported by Firestone et al. (1991) to be the best performing among a number of other measures, and is defined as

$$M_{Br} = \sum_{(x,y)} (f(x,y) - f(x+2,y))^2, \quad (15)$$

where  $f(x,y)$  is the image function.

The fourth measure, known as Vollath's  $F_4$ , was reported by Santos et al. (1997) to achieve good focusing, even in low-information-content images, and is defined as

$$M_{Vo} = \sum_{(x,y)} f(x,y) (f(x+1,y) - f(x+2,y)), \quad (16)$$

where  $f(x,y)$  is, again, the image function.

These four measures were selected to be the reference measures for our tests, because they all operate on the spectral content of an image and have been reported to perform very well.

### 6.4. Experiments with noise-free images

Experiments with noise-free images were conducted to evaluate the effectiveness of the proposed measure. For each sequence, focus measures were calculated within the focusing windows. The recorded graphs of the measures  $M_{Be}^g$ ,  $M_{Tg}$ ,  $M_{Le}$ ,  $M_{Br}$  and  $M_{Vo}$  are shown in Figs. 4 and 5.

The results for the visible-light camera and the thermal camera are shown in the left- and right-hand columns, respectively. In all graphs, the dashed vertical lines depict the optimum focus positions. At this point it should be noted that the image sequences were not really noise-free, rather they had a high SNR (Table 1). Nevertheless, the sequences were considered as noise-free.

From a visual inspection of Fig. 4(a,b), we conclude that  $M_{Be}^8$  produces a similar response for different sequences. The graphs of  $M_{Be}^8$  consistently exhibit a well-expressed global maximum at the correct position and they gradually decrease when moving away from that position. The measures  $M_{Le}$  (Fig. 4e,f) and  $M_{Br}$  (Fig. 5a,b) also exhibit the correct global optimum, but generally do not decrease as gradually as  $M_{Be}^8$ . This is clearly seen in the graph of  $M_{Br}$  (Fig. 5b, *sq6*), where the value of the measure at focus position 1 is almost as high as that at the optimum position. A similar behavior in responses to different focus positions can be seen in Fig. 4(c,d) and Fig. 5(c,d) for the measures  $M_{Tg}$  and  $M_{Vo}$ , except for the case of sequence *sq5*, where both measures exhibit an erroneous global maximum.

For an objective evaluation we calculated two features that delineate the behavioral properties of a focus measure: (a) the average range of focus positions, where the measure decreased monotonically from its exhibited optimum focus position (range of measure) and (b) the portion of all cases where the measure failed to exhibit an optimum at the correct focus position (failure rate). The results for all the measures are presented in the first two columns of Table 2, and as previously visually observed,  $M_{Be}^8$  achieves the largest average range and zero failure rate, followed by  $M_{Br}$ ,  $M_{Le}$ ,  $M_{Tg}$  and  $M_{Vo}$ .

### 6.5. Experiments with noisy images

In order to evaluate the robustness of the proposed measure to the presence of noise, all the sequences of recorded images were degraded with an artificially added noise, and as before, average ranges and failure rates were calculated for each measure. Tests were made with four different noise variances and each experiment was conducted on all the sequences 30 times. The average ranges and failure rates of all the focus measures are shown in Table 2 by the columns under the SNR values from 33dB to 17dB, where for each measure the marginal values on the average ranges and failure rates for artificially degraded sequences are also shown (the last two columns of Table 2).

Again, the measure  $M_{Be}^8$  marginally outperformed all the other measures, and was followed by  $M_{Tg}$ ,  $M_{Le}$ ,  $M_{Vo}$  and  $M_{Br}$ . In all four experiments with different noise variances,  $M_{Be}^8$  achieved the largest average range. In terms of failure rates,  $M_{Be}^8$  generally performed better or as well as the other measures, except in the case of a severe noise of 17dB, where  $M_{Tg}$  exhibited a smaller failure rate. A detailed analysis of the results of the experiments with a 17dB SNR revealed that  $M_{Be}^8$  always produced the correct optimum for the visible-light images, but had difficulties with the thermal-light sequences *sq6* and *sq7* (Fig. 3d,e). In these two sequences the thermal camera did not produce a completely sharp image, due to the dispersion of heat. The image content was overwhelmed by the substantial amount of added noise throughout the sequences, which consequently resulted in a failure of the measure. For reference, if a noise of 17dB was applied to

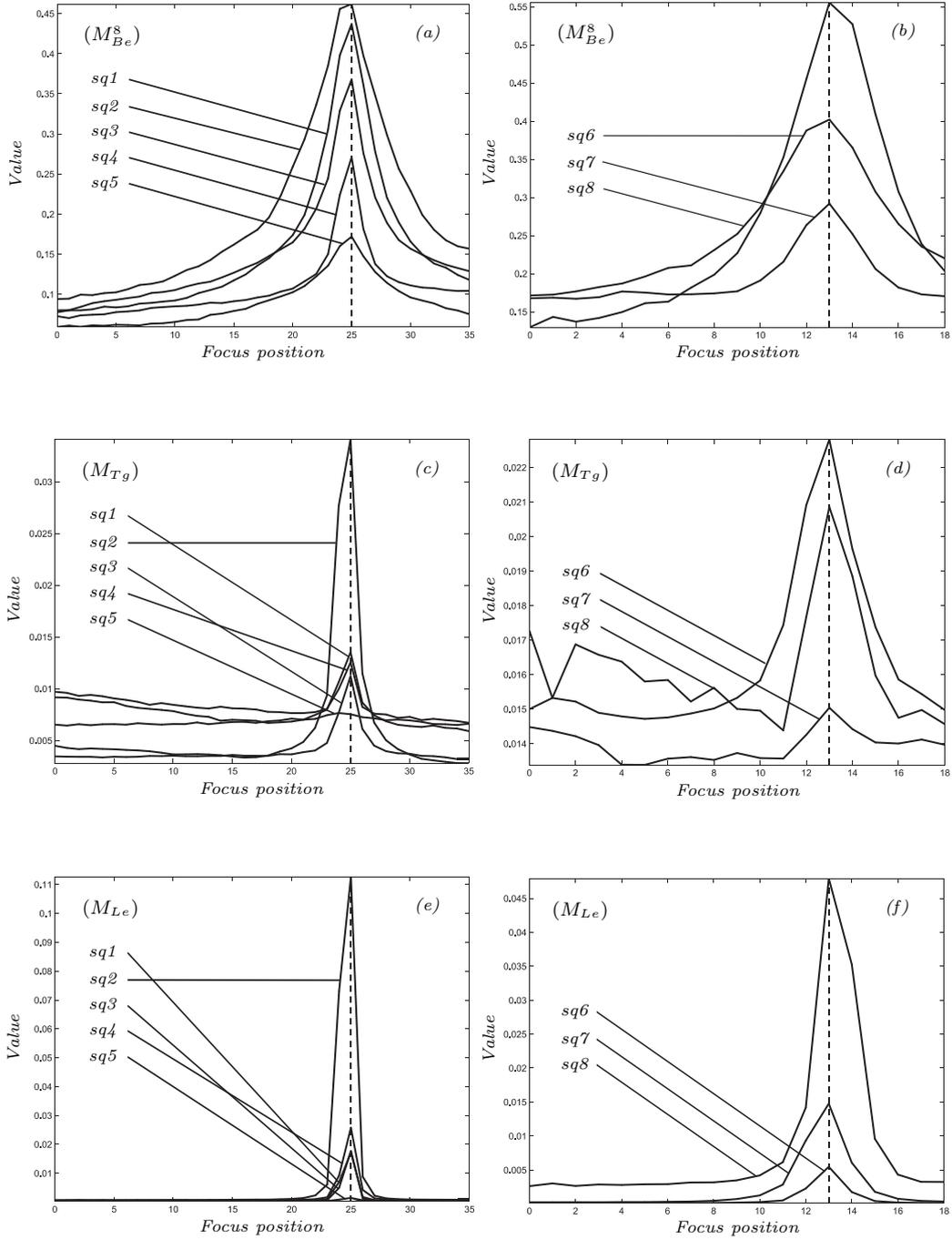


Fig. 4. Graphs of focus measures  $M_{Be}^s$ ,  $M_{Tg}$  and  $M_{Le}$  for noise-free sequences of recorded images. The left-hand column shows the results for the visible-light camera sequences, while the results from the thermal camera sequences are shown in the right-hand column. In all the graphs the correct focus position is depicted by a dashed vertical line. The labels  $sq(1-8)$  refer to the sequences depicted in Figs. 2 and 3.

an image with an intensity range of 256 gray levels, then the standard deviation of such a noise would be approximately 51 gray levels.

We must emphasize, however, that the measure did not fail consistently throughout all 30 experiments with the 17dB SNR on sequences  $sq6$  and  $sq7$ , and that these sequences caused

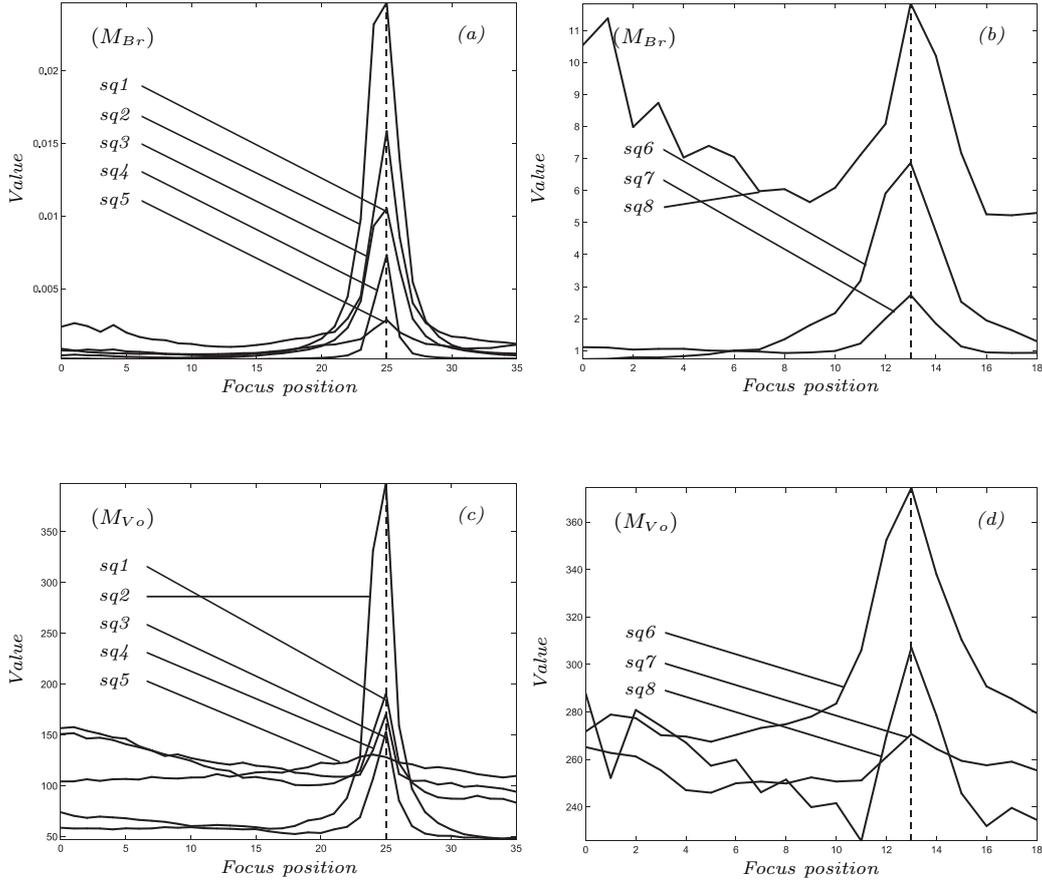


Fig. 5. Graphs of focus measures  $M_{Br}^8$  and  $M_{Vo}$  for noise-free sequences of recorded images. The left-hand column shows the results for the visible-light camera sequences, while the results from the thermal camera are shown in the right-hand column. In all the graphs the correct focus position is depicted by a dashed vertical line. The labels  $sq(1-8)$  refer to the sequences depicted in Figs. 2 and 3.

Table 2

Average ranges ( $d_r$ ) and failure rates ( $f_r$ ) calculated on sequences with different amounts of added noise.

$\sigma_n$ (SNR)	0	0.03	0.09	0.14	0.20	marginal						
	( $\infty$ dB)	(33dB)	(24dB)	(20dB)	(17dB)	(33dB-17dB)						
	$d_r$	$f_r$	$d_r$	$f_r$	$d_r$	$f_r$	$d_r$	$f_r$				
$M_{Be}^8$	0.84	0.0	0.73	0.0	0.55	0.1	0.48	0.1	0.43	0.2	0.55	0.1
$M_{Tg}$	0.4	0.1	0.39	0.1	0.4	0.1	0.39	0.1	0.36	0.1	0.38	0.1
$M_{Le}$	0.54	0.0	0.42	0.0	0.36	0.2	0.34	0.2	0.32	0.2	0.36	0.15
$M_{Br}$	0.63	0.0	0.31	0.3	0.18	0.6	0.19	0.6	0.22	0.3	0.23	0.45
$M_{Vo}$	0.38	0.1	0.36	0.1	0.31	0.2	0.26	0.3	0.22	0.4	0.29	0.25

Standard deviations of added noise ( $\sigma_n$ ) are given for each sequence in portions of image intensity range, and for reference, their respective SNRs are also given. The columns under  $\infty$ dB show the results for experiment with noise-free sequences, while the columns under 33dB to 17dB show the results for the experiments with sequences degraded with artificial noise. The last two columns show the marginal average range and failure rate for the experiments with noisy sequences.

problems for other measures, even with smaller amounts of added noise. We therefore conclude that the measure  $M_{Be}^8$  performs well even in the presence of severe noise.

## 7. Conclusion

A novel measure of camera focus  $M_{Be}^8$ , based on the Bayes entropy of an image spectrum, is presented. The focus is usually estimated only on a small part of the image (i.e., the focusing window), but calculating the spectrum of such a region still requires a lot of computational power. The measure was therefore calculated via smaller sub-images, using  $8 \times 8$  DCT to obtain their spectra. Since the spectra of typical images are dominated by the low-order coefficients, our measure ignores the coefficients of an order higher than some predefined threshold. This threshold was determined experimentally, but note that when no prior knowledge is available on the types of images, the threshold can be set to its maximum value (e.g., in the case of  $8 \times 8$  DCT the maximum value of the threshold would be 16).

The proposed measure  $M_{Be}^8$  belongs to a class of measures used by passive systems that rely on image content. Focusing using these measures depends heavily on the information available from the CCD sensor. Low-lit or dark scenes might therefore present a considerable difficulty to  $M_{Be}^8$ , since in these situations the object being focused is often not clearly seen by the CCD until the flash is fired. This is a drawback that is characteristic of passive systems.

The image information that is used by  $M_{Be}^8$  for estimating the focus is the image spectrum. This means that changing the image focus must result in a change of its spectrum, and usually this is true. However when this assumption does not hold, focusing will fail. For example, when focusing a region of smoothly varying intensity,  $M_{Be}^8$  might prove to be unreliable if the noise in the CCD is very high. In these situations, the noise can overwhelm the image spectrum to such an extent that the change in focus would not significantly affect the spectrum in a way that  $M_{Be}^8$  can detect.

However, when  $M_{Be}^8$  was compared with four reference measures of focus, it exhibited a marginally larger range, a smaller failure rate and a robustness to noise.

One strength of this measure is that standard solutions can be used for its hardware implementation (i.e., fast  $8 \times 8$  DCT). Indeed, since nearly every consumer digital camera incorporates JPEG compression and is therefore capable of performing  $8 \times 8$  DCT, we believe that measure  $M_{Be}^8$  could be appropriate for such systems. Furthermore, the calculation of  $M_{Be}^8$  does not impose very rigid restrictions on the shape of a focusing window, so this measure can easily be extended to multi-point focus systems or systems for small-depth-of-field focusing.

### References

- Chern N.K., Neow P.A., Ang Jr., M.H., Practical Issues in Pixel-Based Autofocusing for Machine Vision, 2001. Int. Conf. On Robotics and Automation., 2791-2796
- Choi, K.S., Lee, J.S., Ko, S.J., 1999. New Autofocusing Technique Using the Frequency Selective Weighted Median Filter for Video Cameras, IEEE Trans. on Consumer Electronics. 45 (3), 820-827

- Devijver, P.A., 1974. On a new class of bounds on Bayes risk in multihypothesis pattern recognition, *IEEE Trans. Comput. C* (23), 70-80
- Fernandez, P.G., Garcia, A., Ramirez, J., Parrilla, L., Lloris, A. , 2000. Fast RNS-Based DCT Computation with Fewer Multiplication Stages, *Proc. of XV DCIS*, 276-281
- Firestone, L., Cook, K., Culp, K., Talsania, N., Preston Jr, K., 1991. Comparison of Autofocus Methods for Automated Microscopy, *Cytometry*. 12 (3), 195-206
- Gillespie, J., King, R.A., 1989. The use of self-entropy as a focus measure in digital holography, *Pattern Recognition Letters*. 9 (1), 19-25
- Geusebroek, J.M., Cornelissen F., Smeulders A.W., Geerts H., 2000. Robust autofocusing in microscopy, *Cytometry*. 36 (1), 1-9
- Krotkov, E.P., 1987. Focusing, *Int. Jour. of Computer Vision*. 1 (3), 223-237
- Lin, J., Zhang, C., Shi, Q., 2004. Estimating the amount of defocus through a wavelet transform approach, *Pattern Recognition Letters* 25 (4), 407-411
- Pennebaker, W. B., Mitchell, J. L., 1993. *JPEG: Still Image Data Compression Standard*, Tomson Publishing Inc., New York.
- Santos, A., Ortiz de Solorzano, C., Vaquero, J.J., Pena J.M., Malpica N., del Pozo, F., Evaluation of autofocus functions in molecular cytogenetic analysis, 1997. *Journal of Microscopy*. 188 (3), 264-72
- Subbarao, M., Tyan, J.K., Selecting the Optimal Focus Measure for Autofocusing and Depth-From-Focus, 1998. *IEEE Trans. on Pattern Analysis and Machine Intelligence* 20 (8), 864-870
- Yap, P.T., Raveendran, P., 2004. Focus measure based on Chebyshev moments, *VISP*. 151 (2), 128-136
- Yeo, T.T.E., Ong, S.H., Jayasooriah, Sinniah, R., Autofocusing for tissue microscopy, 1993. *Image and Vision Computing*. 11 (10), 629-639
- Wallace, G.K., *The JPEG Still Picture Compression Standard*, 1992. *IEEE Trans. Consum. Electronics*. 38 (1), 18-34
- Widjaja, J., Jutamulia S., 1998. Use of wavelet analysis for improving autofocusing capability, *Opt. Commun*. 151, 1214