ELSEVIER

Pattern Recognition Letters 23 (2002) 1785–1794

Pattern Recognition Letters

www.elsevier.com/locate/patrec

A new wavelet-based measure of image focus

Jaroslav Kautsky [a], Jan Flusser [b,*], Barbara Zitová [b], Stanislava Šimberová [c]

[a] *Department of Mathematics and Statistics, Flinders University of South Australia, GPO Box 2100, Adelaide, SA 5001, Australia*
[b] *Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, 18208 Prague 8, Czech Republic*
[c] *Astronomical Institute, Academy of Sciences of the Czech Republic, 25165 Ondřejov, Czech Republic*

Received 6 November 2001; received in revised form 28 January 2002

Abstract

We present a new measure of image focus. It is based on wavelet transform of the image and is defined as a ratio of high-pass band and low-pass band norms. We show this measure is monotonic with respect to the degree of defocusation and sufficiently robust. We experimentally illustrate its performance on simulated as well as real data and compare it with existing focus measures (gray-level variance and energy of Laplacian). Finally, an application of the new measure in astronomical imaging is shown.

Keywords: Image blurring; Focus measure; Wavelet transform; Astronomical imaging

1. Introduction

The problem of selecting the best-focused image from a sequence of differently defocused/blurred images of the same scene often arises in many application areas. In computer vision, autofocusing algorithms evaluate the degree of defocus of the images taken with various settings of camera parameters. The image with minimum value of defocusation measure defines the parameters for camera autofocusation. This approach is applied in industrial visual inspection and in digital microscopy, among others. In remote sensing and in astronomy we face similar tasks. However, the image blurring is caused by atmospheric turbulence rather than by camera defocus. The goal is to select "good" images for visual interpretation or further computer analysis. Having a reliable and robust focus measure (or inversely, blur measure) is a key point to resolve these problems.

To be as general as possible, we do not restrict ourselves to any particular model of blurring. However, we assume the relationship between the original scene $f(x,y)$ and the acquired set of images $g_1(x,y),\dots,g_n(x,y)$ can be expressed by convolution

$$g_i(x,y) = (f * h_i)(x,y), \quad i = 1,\dots,n \tag{1}$$

* Corresponding author. Tel.: +420-2-6605-2357; fax: +420-2-8468-0730.

E-mail addresses: jarka@ist.flinders.edu.au (J. Kautsky), flusser@utia.cas.cz (J. Flusser), zitova@utia.cas.cz (B. Zitová), ssimbero@asu.cas.cz (S. Šimberová).

0167-8655/02/$ - see front matter
PII: S0167-8655(02)00152-6

where $h_i(x,y)$ is the point-spread function (PSF) of the blur in the ith observation. In the best possible case (not occurring in practice), $h_i(x,y) = \delta(x,y)$ and we get ideal image $g_i(x,y) = f(x,y)$. In practice all the $h_i(x,y)$ have a character of an unknown low-pass filter.

By the term "focus measure" or "blur measure" we understand any functional defined on the space of image functions which reflects the amount of blurring introduced by $h_i(x,y)$. Thus, having focus/blur measure M, we look for the "best" image $g_{i_0}(x,y)$ by maximizing/minimizing $M(g_i)$ over $i = 1,\dots,n$. Furthermore, we may also want to order the images according to their quality, which is equivalent to ordering the sequence $\{M(g_i)\}$, $i = 1,\dots,n$.

Any reasonable focus measure should satisfy some basic requirements. First, it should be content-independent, which means it must not be based on any particular structures in the image (e.g., on isolated bright points). Secondly, it should be monotonic with respect to blur. The more blurred the image, the less the focus measure should be. Finally, the measure should be robust to noise.

Our motivation for this work came from the area of astronomical imaging, particularly from the processing of images of solar atmosphere. In the ground-based observations, the short-exposure images from the telescope are corrupted by "seeing". This degradation leads to image blurring according to Eq. (1), where the actual PSF is a composition of the intrinsic PSF of the telescope (which is constant over the observation period) and of a random component describing the perturbations of the wavefronts in the earth's atmosphere. Different parts of the solar atmosphere are observed in different spectral bands. The lower part called photosphere is usually observed in visible light of $\lambda = 590$ nm while the medium part called chromosphere is best to observe in H_α ($\lambda = 656.3$ nm) wavelength. In visible light the effects of fluctuations in the refractive index of the air caused by temperature variations are more significant than in H_α. Since the atmospheric conditions may change very quickly, the acquired image sequence usually contains images of different quality from sharp to heavy blurred ones. Such sequence, which is a result of one observation session, typically consists of several hundreds of images. Automatic selection of a few "good" images which can be used for further investigation of astronomical phenomena is a real challenge.

In this paper, we propose a new focus measure defined by means of wavelet transform of the image. In Section 2, we briefly review existing focus measures. The new focus measure is introduced in Section 3 and its robustness is analysed in Section 4. In Section 5, its performance on simulated as well as real data is shown.

2. Existing focus measures

Several focus measures have been reported in the literature (Krotkov, 1987; Nayar and Nakagawa, 1994; Subbarao et al., 1993; Subbarao and Choi, 1995; Subbarao and Tyan, 1998; Zhang and Wen, 2000). Most of them are based on the idea to emphasize high frequencies of the image and measure their quantity. It corresponds with our intuitive expectation that the blurring suppresses high frequencies regardless of the particular PSF.

One of the simplest focus measures is variance of image gray levels (Subbarao et al., 1993, 1998)

$$M_1 = \iint (g_i(x,y) - m_i)^2 \mathrm{d}x\,\mathrm{d}y,$$

where m_i denotes the mean gray level value of $g_i(x,y)$. This measure was proved to be monotonic (Subbarao et al., 1993) and, for zero-mean images, equivalent to the image energy in Fourier domain

$$\iint |G_i(u,v)|^2 \mathrm{d}u\,\mathrm{d}v.$$

Several focus measures are defined by means of image derivatives. L_1 norm of image gradient

$$M_2 = \iint \left|\frac{\partial g_i(x,y)}{\partial x}\right| + \left|\frac{\partial g_i(x,y)}{\partial y}\right| \mathrm{d}x\,\mathrm{d}y,$$

L_2 norm of image gradient (sometimes called gradient energy)

$$M_3 = \iint \left(\frac{\partial g_i(x,y)}{\partial x}\right)^2 + \left(\frac{\partial g_i(x,y)}{\partial y}\right)^2 \mathrm{d}x\,\mathrm{d}y,$$

L_1 norm of second derivatives

$$M_4 = \iint \left| \frac{\partial^2 g_i(x,y)}{\partial x^2} \right| + \left| \frac{\partial^2 g_i(x,y)}{\partial y^2} \right| \mathrm{d}x\,\mathrm{d}y$$

and energy of image Laplacian

$$M_5 = \iint \left(\frac{\partial^2 g_i(x,y)}{\partial x^2} + \frac{\partial^2 g_i(x,y)}{\partial y^2} \right)^2 \mathrm{d}x\,\mathrm{d}y,$$

belong to the most popular ones. Subbarao et al. (1993) proved the monotonicity of M_3 and M_5 and showed they can be equivalently evaluated in Fourier domain as the energy of high-pass filtered image, where the magnitudes of the corresponding filters are $\sqrt{(u^2+v^2)}$ and (u^2+v^2), respectively. M_2 and M_4 are nonlinear measures, thus they cannot be expressed in Fourier domain. Nevertheless, they have been used successfully in the literature. Krotkov (1987) described a modification of M_2 in such a way that only points where the norm of image gradient exceeds a pre-defined threshold are used in the integrand. Nayar and Nakagawa (1994) did the same but used M_4 instead of M_2.

Subbarao and Tyan (1998) analysed the robustness of M_1, M_3 and M_5. He concluded his experiments by recommendation to use the energy of image Laplacian because of its tolerance to additive noise. However, the differences among individual measures were not significant.

The only existing focus measure which is beyond the framework of high-pass filters was proposed by Zhang and Wen (2000). Their focus measure is based on image moments and employs the theoretical results achieved by Flusser and Suk (1998). Flusser and Suk proved that odd-order moments exhibit certain blur-invariant properties while even-order moments change under blur. Using their results, Zhang and Wen (2000) proposed to take second and fourth order moments as a focus measure and proved the monotonicity of this measure. However, moment-based focus measure is very sensitive to boundary effect and therefore it performs well only on images with zero background.

3. Wavelet-based focus measure

The new focus measure is based on discrete wavelet transform (DWT). It decreases with blurring, i.e., the more focused the image the larger is the measure.

Given an image f of the size $N_r \times N_c$ pixels, a DWT using decomposition to depth d with a wavelet w of multiplicity (dilation factor) m produces a low-pass band $l_w(f)$ of size $(N_r/dm) \times (N_c/dm)$ and several high-pass bands which we denote collectively by $h_w(f)$; the total number of coefficients in these bands is $N_rN_c(1 - 1/(d^2m^2))$. We propose an absolute wavelet-based measure

$$W = \frac{\|h_w(f)\|}{\|l_w(f)\|},$$

where $\|\cdot\|$ denotes the discrete Euclidean norm, and its relative modification

$$\widetilde{W} = \frac{W}{\sqrt{d^2m^2 - 1}}.$$

The measure $\widetilde{W}$ reflects average square values per pixel in the high-pass and low-pass bands. It is useful when comparing the performance of two DWT's with different values of dm.

We will consider only orthogonal DWT which preserves the image energy (i.e. $\|f\|^2 = \|h_w(f)\|^2 + \|l_w(f)\|^2$) so that we have

$$W^2 = \frac{\|h_w(f)\|^2}{\|f\|^2 - \|h_w(f)\|^2},$$

which increases monotonously with $\|h_w(f)\|$. High-pass bands of DWT are m-decimated convolutions of the image with high-pass filters for which we have monotonicity with respect to blurring (or defocusing) by the same argument as given in Subbarao et al. (1993).

It is easy to understand intuitively why the DWT should have this property. The short filters in DWT identify local areas of smoothness in the image by expressing them almost exactly in the low-pass band. The high-pass bands contain large coefficients only where the image is not smooth; so smoothing will decrease energy in these bands. Longer filters, based on higher approximation order wavelets, would do this better but the extend of the local area of smoothness must be appropriately larger. [1]

The form of W makes it invariant to the scaling (or choice of units) of the image. That is not the only advantage. Because of the norm preservation, blurring (or defocusing) the image decreases the energy in the high-pass bands and simultaneously increases the energy in the low-pass band; this increases the discrimination power of the measure (if f and g are two images of the same scene which differ only slightly in focus, then $W(f) - W(g)$ is still noticeable which may not be true in case of other focus measures).

4. Robustness of the wavelet-based measure

In this section we present a deterministic robustness analysis of our wavelet-based focus measure W. Following the idea of stability margin, which is well-known from control theory, we define an analogous concept.

Let M be a focus measure and g_1 and g_2 be two images such that $M(g_1) < M(g_2)$. By *margin of robustness* $\varrho_M(g_1, g_2)$ for this measure and given images we understand the maximal relative error such that for any two images $\tilde{g}_1$ and $\tilde{g}_2$ "close" to g_1 and g_2, respectively,

$$\frac{\|\tilde{g}_i - g_i\|}{\|g_i\|} < \varrho, \quad i = 1, 2,$$

there is $M(\tilde{g}_1) < M(\tilde{g}_2)$.

If g_1 and g_2 are differently focused versions of the same original image then we can relate the margin of robustness directly to the levels of blurring (and the original image, of course).

If the DWT of the difference between the close images has the same mean-square value in its low-pass and high-pass bands, i.e. if

$$\|h_w(\tilde{g}_i - g_i)\| \approx (m^2d^2 - 1)\|l_w(\tilde{g}_i - g_i)\|, \quad i = 1, 2,$$

holds (true, e.g., for a "typical" realization of white noise) then we have a simple estimate

$$\varrho_W \approx \frac{1}{C}(W(g_2) - W(g_1)),$$

where

$$C = \left(\sqrt{1 - \frac{1}{m^2d^2}} + \frac{W(g_1)}{md}\right)\sqrt{1 + W^2(g_2)} + \left(\sqrt{1 - \frac{1}{m^2d^2}} + \frac{W(g_2)}{md}\right)\sqrt{1 + W^2(g_1)} \approx 2.$$

The estimate follows from triangular inequalities for the norms

$$\frac{\|h_w(g_i)\| - \|h_w(\tilde{g}_i - g_i)\|}{\|l_w(g_i)\| + \|l_w(\tilde{g}_i - g_i)\|} \leqslant \frac{\|h_w(\tilde{g}_i)\|}{\|l_w(\tilde{g}_i)\|} \leqslant \frac{\|h_w(g_i)\| + \|h_w(\tilde{g}_i - g_i)\|}{\|l_w(g_i)\| - \|l_w(\tilde{g}_i - g_i)\|}$$

by some straightforward algebra. This is a conservative estimate in the sense that if

$$\max\left(\frac{\|\tilde{g}_1 - g_1\|}{\|g_1\|}, \frac{\|\tilde{g}_2 - g_2\|}{\|g_2\|}\right) < \frac{1}{C}(W(g_2) - W(g_1))$$

then $W(\tilde{g}_1) < W(\tilde{g}_2)$ is guaranteed. The true value of ϱ_W may be even larger for the type of changes we assume but it would be difficult to establish.

Although the exact value of the factor $1/C$ in our estimate depends on the measures of the images and on the type of DWT, approximating it by $1/2$ is quite reasonable for practical purposes. Generally, $W(g_i) \ll 1$, particularly for images without many sharp edges, so $C \to 2$ as $W(g_i)$ become small and md large. On the other hand, $C = 2$ exactly also for a variety of non-limiting cases, e.g., for $W(g_i) = \sqrt{2} - 1$ and $md = 2$.

Regardless of the exact value of C, we may conclude that the measure W is as robust as we may expect. The information about the relative blurring of the images g_1 and g_2 is preserved if the

[1] Although one can, for a given wavelet, construct an artificial example with $W = 0$ on which blurring would increase the measure, on real images the properties of DWT are very close to the undecimated transform for which we have monotonicity. Even with the mentioned anomaly, let us comment, that shifting such artificial image by $j = 1, \ldots, m-1$ pixels and taking an average of the obtained norms would restore the desired monotonicity of the measure (recall that DWT is only m-shift invariant). To use such a theoretically guaranteed to be monotonous measure would not justify an extra cost (m^2 times increase) to evaluate it.

magnitude of the change in the images is comparable or less than $W(g_2) - W(g_1)$.

5. Numerical experiments

5.1. Artificial data

First, we tested the new focus measure on artificial data where the level of blurring as well as the amount of additive noise can be fully controlled. We used the well-known Lena image and we blurred it gradually by the convolution with square averaging masks of sizes 3×3, 5×5, 7×7 and 9×9. Moreover, each of the images was corrupted by Gaussian white noise with standard deviations (STD) 5, 15, 30, and 45, respectively. In this way we got 25 test images (four of them can be seen in Fig. 1) that were examined by the following focus measures: gray-level variance M_1, energy of Laplacian M_5, and the proposed measure W. M_1 and M_5 were chosen because M_1 is the simplest and most cited measure and M_5 was the best measure in the comparative study (Subbarao and Tyan, 1998). To compare the influence of various wavelets and decomposition depth, we evaluated measure W using Daubechies wavelets with 2, 4, 6, and 10 taps (the number of taps denotes the number of coefficients in the wavelet filter, e.g., Haar wavelet has 2 taps) and the left-tree decomposition scheme of depth 1 and 2 (see Daubechies, 1992 for detailed description of these wavelets). The major results of this experiment can be summarized as follows.

- All investigated measures were proved to be monotonic in the noise-free case; we consider a measure robust if this property prevails in the presence of noise.
- Gray-level variance M_1 is robust to noise but has limited discrimination power.

Fig. 1. Four examples of our test images: Lena blurred by 3×3 averaging mask with slight additive noise of STD = 5 (top left), the same with heavy noise (STD = 30) (top right), blurring by 9×9 mask with slight noise (bottom left), both heavy blurring and noise (bottom right).

- Energy of Laplacian M_5 performs well in a noise-free case but it is very sensitive even to small noise. This is understandable as the second-order derivatives cannot distinguish between sharpness of the image and noise.
- Wavelet-based measure W exhibits both good robustness (as can be expected from the theoretical consideration in Section 4) and discriminability. Robustness increases with the depth of the wavelet decomposition. In our experiments, depth 2 yielded sufficient robustness even if heavy noise was present. It follows from the fact that the high-pass bands of level 1 do not change when going deeper to depth 2 (traditional left-tree decomposition scheme was applied in all cases) while the low-pass band is further smoothed and decomposed, which leads to noise suppression. It should be noted that the decomposition from level 1 to level 2 requires only 25% of extra cost.
- Wavelets with more taps are appropriate for images without sharp edges and many details, shorter wavelets perform better on images with prominent high-frequency information. This is because longer wavelets, i.e. with higher approximation order, result in noise-like high-pass bands when applied to images with many edges, which thus diminishes their focus discrimination power. On the other hand, shorter filters are able to capture more details and can, consequently, register their fading due to image blurring. Based on the experiments we carried out, we recommend 6-taps wavelets as a good compromise for common images.

Representative results are visualized in detail in Figs. 2–4. In each graph, one can see the behavior of the particular focus measure under various image blurring and noise. An "ideal" focus measure should decrease with an increasing blurring and this property should not be affected by noise. (However, the monotonicity requirement cannot be fulfilled for extremely heavy noise. In that case, the high-frequency information introduced by noise beats that of the image itself and W becomes "noise measure" rather than "focus measure". The same is true for the other focus measures.) The higher the difference between the consecutive values, the better discrimination power of the measure. Thus, for such blurring and noise level when the graph becomes constant or even increasing, the focus measure is useless. Considering these criteria we can clearly see that W of the 10-taps wavelet and decomposition depth 2 (Fig. 4) is the best measure among those three shown here. On the other hand, the same measure calculated on the decomposition depth 1 yields poor results (Fig. 3). The performance of the energy of Laplacian M_5 (Fig. 2) is in between them.

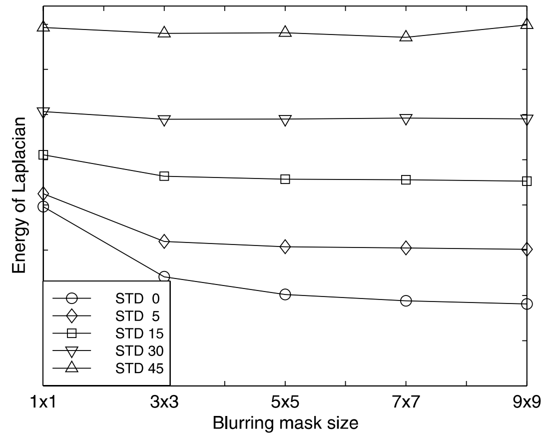

Fig. 2. Energy of Laplacian M_5 calculated on Lena image for various amount of blurring and noise.

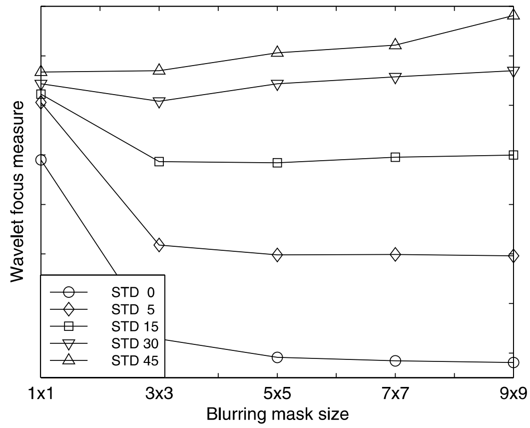

Fig. 3. Focus measure W calculated on Lena image for various amount of blurring and noise. Daubechies wavelet of 10 taps and the decomposition depth 1 was used.

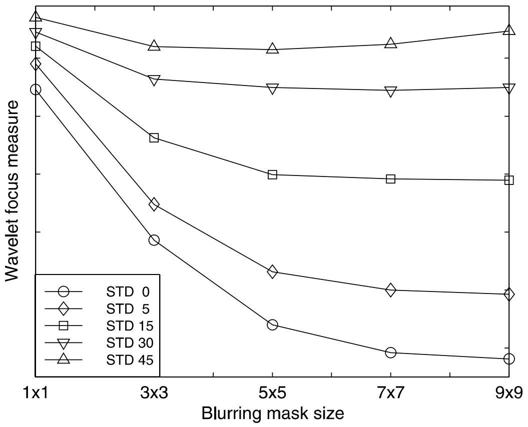

Fig. 4. Focus measure W calculated on Lena image for various amount of blurring and noise. Daubechies wavelet of 10 taps and the decomposition depth 2 was used.

5.2. Indoor scene

In this experiment, a static indoor scene was captured four times by digital camera Nikon Coolpix 950 (see Fig. 5). The first image was taken

Table 1
Various focus measures calculated for the indoor images depicted in Fig. 5. Proportional values are used, 1 stands for the best focused image in the sequence

Image	Focused	Slight defocus	Medium defocus	Heavy defocus
M_1	1	0.92	0.88	0.82
M_5	1	0.22	0.17	0.14
W (4-taps depth 1)	1	0.34	0.30	0.28
W (4-taps depth 2)	1	0.48	0.39	0.32
W (10-taps depth 2)	1	0.41	0.32	0.29

Fig. 5. Indoor scene: Focused image (top left), slight out-of-focus (top right), medium out-of-focus (bottom left), heavy out-of-focus (bottom right).

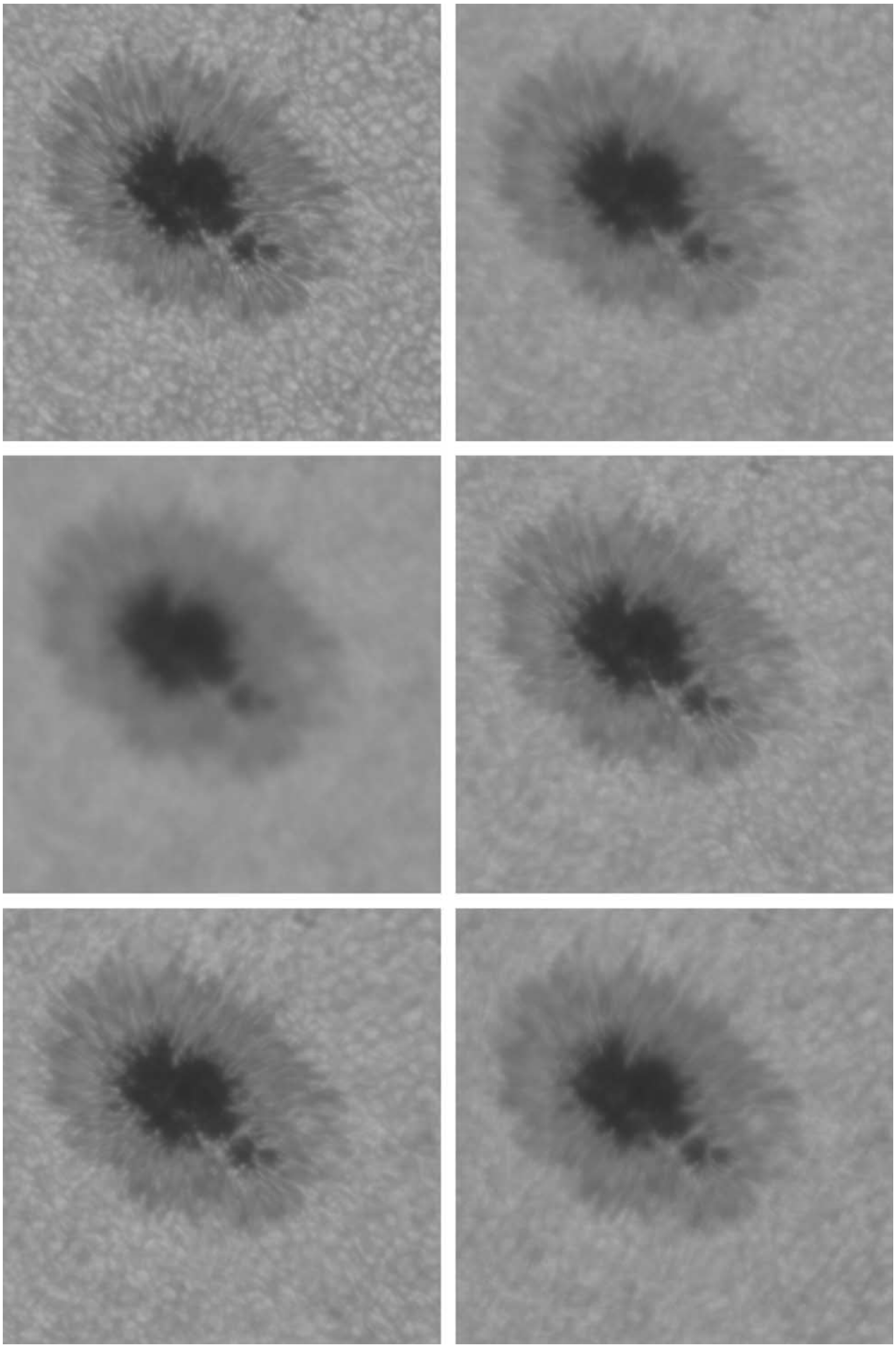

Fig. 6. Six consecutive sunspot images corrupted by variable blurring caused by atmospheric turbulence.

Table 2
Various focus measures calculated for the sunspot images in Fig. 6. Proportional values are used, 1 stands for the best image in the sequence

Image	Top left	Top right	Middle left	Middle right	Bottom left	Bottom right
Human	1st	5th	6th	3rd	2nd	4th
M_1	1	0.906	0.864	0.942	0.964	0.909
M_5	1	0.970	0.967	0.987	0.991	0.969
W (4-taps depth 1)	1	0.921	0.905	0.937	0.948	0.923
W (4-taps depth 2)	1	0.702	0.619	0.764	0.812	0.718
W (10-taps depth 2)	1	0.807	0.767	0.845	0.870	0.821

using camera autofocus, the others were manually defocused in such a way that the degree of out-of-focus was changing gradually from slight to heavy. Thus, we know exactly the ordering of the images with respect to the defocus level.

For each image we calculated focus measures M_1 (gray-level variance), M_5 (energy of image Laplacian) and our measure W using three different Daubechies wavelets: 4-taps depth 1, 4-taps depth 2, and 10-taps depth 2. The results are summarized in Table 1. For better insight we present proportional rather than absolute values of the focus measures (the measure of the focused image is taken as 1). This form of presentation allows better to observe the discrimination power of each measure.

One can see from the monotonicity of the rows of Table 1 that each measure ordered the images correctly. There are, however, some differences in the discrimination power. All three wavelets provide very good discrimination because the most defocused image has low-proportional measures — 28%, 32%, and 29%, respectively. The same is true for Laplacian that discriminates even better in this case. On the other hand, the most defocused image measured by gray-level variance yields 82%, which does not provide enough space to distinguish among the test images reliably.

5.3. Solar images

In this experiment, we illustrate the performance of the new focus measure when solving a real task in astronomical imaging. In Fig. 6, one can see six images of the spot in the solar photosphere taken by a telescope with a CCD camera in visible spectral band (venue: Observatory Ondřejov, Czech Republic; wavelength: $\lambda = 590$ nm). Since the time interval between each two consecutive acquisitions in this sequence was very short (few seconds), the scene can be considered still. Nevertheless, atmospheric conditions were changing during the acquisition period. Thus, some of the images are blurred considerably due to the atmospheric turbulence. The task is to order the images according to their suitability for visual interpretation, in this case inversely to the level of blurring. (It should be noted that additive noise also affects the visual perception of the images. Fortunately, its impact is not significant here. Thanks to high quality of the imaging device, the signal-noise-ratio is higher than 50 dB.)

To order the acquired sequence we applied the same five focus measures as in the indoor experiment. Moreover, an expert in astronomical imaging was asked to order these images visually. His opinion can be considered as the "ground truth". As can be seen from Table 2 (as in the indoor experiment, the proportional measure values are used instead of the absolute ones), the wavelet-based measures ordered the images correctly while the depth 2 decomposition provided a better discrimination than the decomposition to the depth 1. Gray-level variance gave good results too but at low discrimination (the proportional measure of the most defocused image was 86%). The energy of Laplacian failed and two images were misordered. The reason of this failure is the poor discrimination power of this measure on the given image set. Even slight noise or other disturbances can cause wrong ordering.

6. Conclusion

In this paper, we proposed a new measure of the image focus. It is based on the wavelet transform of the image and is defined as a ratio of high-pass band and low-pass band norms. We showed that this measure fulfils the basic requirements usually imposed on a focus measure: it is monotonic with respect to the degree of defocusation and sufficiently robust. We experimentally illustrated its good performance on simulated as well as on real data and compared it with existing focus measures (gray-level variance and the energy of the Laplacian). We also discussed the influence of using various wavelets and decomposition depths. Since our primary motivation came from the area of astronomical imaging, we tested this measure on real example of a sequence of sunspot images, which were differently blurred by atmospheric turbulence. The ordering of this sequence achieved by means of the new focus measure corresponded to the ordering provided by a human expert, which is a satisfactory result.

Acknowledgements

This work was done when J. Kautsky was visiting the Institute of Information Theory and Automation in Prague. Financial support of this research was provided by the Grant Agency of the Czech Republic under the projects no. 102/00/1711 and no. 102/01/P065.

References

Daubechies, I., 1992. Ten Lectures on Wavelets. SIAM.

Flusser, J., Suk, T., 1998. Degraded image analysis: An invariant approach. IEEE Trans. Pattern Analysis and Machine Intelligence 20, 590–603.

Krotkov, E., 1987. Focusing. Int. J. Comput. Vision 1, 223–237.

Nayar, S.K., Nakagawa, Y., 1994. Shape from focus. IEEE Trans. Pattern Analysis and Machine Intelligence 16, 824–831.

Subbarao, M., Choi, T., 1995. Accurate recovery of three-dimensional shape from image focus. IEEE Trans. Pattern Analysis and Machine Intelligence 17, 266–274.

Subbarao, M., Choi, T., Nikzad, A., 1993. Focusing techniques. J. Opt. Eng. 32, 2824–2836.

Subbarao, M., Tyan, J.K., 1998. Selecting the optimal focus measure for autofocusing and depth-from-focus. IEEE Trans. Pattern Analysis and Machine Intelligence 20, 864–870.

Zhang, Y., Zhang, Y., Wen, C., 2000. A new focus measure method using moments. Image and Vision Computing 18, 959–965.