*Journal of Microscopy, Vol. 240, Pt 1 2010, pp. 38–49*
*Received 12 October 2009, accepted 27 January 2010*

doi: 10.1111/j.1365-2818.2010.03383.x

# Evaluating sharpness functions for automated scanning electron microscopy

M.E. RUDNAYA, R.M.M. MATTHEIJ & J.M.L. MAUBACH
*Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, The Netherlands*

**Key words.** Autofocus, automatic focusing, focusing, scanning electron microscopy, SEM.

## Summary

Fast and reliable autofocus techniques are an important topic for automated scanning electron microscopy. In this paper, different autofocus techniques are discussed and applied to a variety of experimental through-focus series of scanning electron microscopy images with different geometries. The procedure of quality evaluation is described, and for a variety of scanning electron microscope samples it is demonstrated that techniques based on image derivatives and Fourier transforms are in general better than statistical, intensity and histogram-based techniques. Further, it is shown that varying of an extra parameter can dramatically increase quality of an autofocus technique.

## Introduction

Nowadays, scanning electron microscopy (SEM) still requires an expert operator to obtain in-focus and astigmatism-free images. To obtain sharp images, the operator uses his eyes (to estimate defocus and astigmatism) and his hands to adjust the controls. For the next SEM generations, the manual operation has to be automated, at least to an extent that a non-expert can efficiently obtain sharp images. One of the reasons is that the work of a human operator has a low repeatability level. After focusing of hundreds of images manually, a human gets tired and loses concentration, which can influence output image quality; and a lot of automated SEM application nowadays requires hundreds and thousands of high-quality images. Therefore, a robust and reliable autofocus algorithm is a necessary tool for the automation of SEM operation.

The existing autofocus techniques can be divided into five groups, viz. based on the image derivatives, statistical information (image variance or autocorrelation), histogram, image intensity or Fourier transform. In the literature, a number of autofocus techniques have been compared for different microscopy forms. For fluorescence microscopy, the Vollath autocorrelation algorithm (Vollath 1987, 1997) was found to be the best (Santos *et al.* 1997). For non-fluorescence microscopy, it has been shown by Sun *et al.* (2004) and Liu *et al.* (2007) that the normalized variance algorithm leads to optimal results. However, the Fourier transform-based autofocus techniques were not taken into account (Santos *et al.* 1997; Sun *et al.* 2004; Liu *et al.* 2007). These techniques were mentioned (Santos *et al.* 1997) but not evaluated due to the slow computations of the image Fourier transform. For the scanning transmission electron microscopy, statistical and Fourier transform-based techniques have been discussed (Van den Broek 2007). The application of derivative, variance, autocorrelation and Fourier transform-based techniques to SEM were examined (Batten 2000), but not subjected to rigorous statistical testing.

The goal of this paper is to find an appropriate autofocus technique for SEM that will work for a large variety of samples. To this end, we exploit the full range of autofocus techniques that were earlier evaluated for fluorescence microscopy in classical paper on autofocus (Santos *et al.* 1997). In addition, we extend derivative-based and autocorrelation-based techniques with extra parameters. We show that varying these parameters leads to improvements. Moreover, we include a new group of autofocus techniques based on Fourier transform. With the desktop CPU's available nowadays, the calculation of a discrete image Fourier transform with the Fast Fourier transform algorithm only takes milliseconds. In particular, in electron microscopy the Fourier transform computation of an image in a through-focus series takes less time than the acquisition of the next image. The image Fourier transform can easily be computed without slowing down the autofocus procedure.

This paper applies explained autofocus techniques to 14 experimental SEM through-focus series of samples with

Correspondence to: M.E. Rudnaya, Department of Mathematics and Computer Science, Eindhoven University of Technology, HG 8.87, PO Box 513, 5600 MB, Eindhoven, The Netherlands. Tel: +31-40-247-3162; fax:+31-40-244-2489; e-mail: m.rudnaya@tue.nl

different geometries. Some of the series contain typical SEM astigmatism effects that usually do not occur in light microscopy and make the autofocusing more complicated. A modified procedure for SEM evaluation based on Santos *et al.* (1997) and Liu *et al.* (2007) is discussed and applied to estimate the quality of each of the listed autofocus techniques. This paper shows that for a wide variety of SEM samples the best results are obtained by derivative-based autofocus techniques, varying the pixel difference parameter. The Fourier transform techniques score almost equally good as derivative-based techniques. This result coincides with result of our preliminary research (Rudnaya *et al.* 2009).

Section 2 of this paper describes the SEM image formation and related problems that are encountered by autofocus techniques. The next section gives an overview of existing autofocus techniques, including the Fourier transform-based techniques. Derivative-based and autocorrelation-based techniques are adorned with an extra parameters. The Sections 4 and 5 describe the experimental SEM data sets and the autofocus techniques evaluation procedure. Finally, Section 6 discusses the experiments and results.

## Scanning Electron Microscope

The electron beam in SEM scans a sample spot by spot in horizontal direction. The electrons reflected from the sample are captured by a detector. The amount of electrons reflected from each spot indicates the image intensity in the current pixel.

The image post-processing in SEM is more complicated and challenging than in light microscopy. Ideally, SEM image formation can be considered as a linear model (Erasmus & Smith 1982), the same way as in light microscopy (Nayar & Nakagawa 1994)

$$f_{\mathbf{p}}(\mathbf{x}) = \psi_0(\mathbf{x}) * h(\mathbf{x}, \mathbf{p}) = \iint_\Omega \psi_0(\mathbf{x}')h(\mathbf{x} - \mathbf{x}', \mathbf{p})d\mathbf{x}'. \quad (1)$$

In Eq. (1), $*$ denotes the convolution, $\mathbf{x} = [x, y]^{\mathrm{T}} \in \Omega$ the spatial coordinates, $\mathbf{p} \in \mathbb{R}^n$ the SEM parameters, $\psi_0(\mathbf{x})$ the object function that describes the sample geometry and $f_{\mathbf{p}}(\mathbf{x})$ the SEM image. The point spread function $h(\mathbf{x}, \mathbf{p})$ that describes electron beam satisfies

$$\iint_\Omega h(\mathbf{x}, \mathbf{p})d\mathbf{x} = 1. \quad (2)$$

The SEM point spread function can be approximated with a Gaussian function (Erasmus & Smith 1982)

$$h(\mathbf{x}, \mathbf{p}) = \frac{1}{2\pi\sigma_x(\mathbf{p})\sigma_y(\mathbf{p})} e^{-\frac{x^2}{2\sigma_x^2(\mathbf{p})} - \frac{y^2}{2\sigma_y^2(\mathbf{p})}}, \quad (3)$$

where astigmatism aberration of the magnetic lens leads to $\sigma_x(\mathbf{p}) \neq \sigma_y(\mathbf{p})$. In light microscopy where the lenses are astigmatism-free it is assumed that (Nayar & Nakagawa 1994)

$$\sigma_x(\mathbf{p}) = \sigma_y(\mathbf{p}). \quad (4)$$

Next to astigmatism the other aberrations, such as spherical and chromatical aberrations, affect the SEM point spread function. The more aberrations are present, the worse is the Gaussian approximation (3) of the point spread function. In particular, it is important to find an astigmatism-stable autofocus technique, because existing astigmatism correction algorithms rely on the fact that the in-focus image parameter can be found in a robust way (Erasmus & Smith 1982; Ong *et al.* 1997).

The SEM's signal-to-noise ration is worse than in light microscopy due to the limited electron dose that can be applied to the specimen. Also to obtain a sufficient signal-to-noise ratio, the microscope exposure time has to be increased. Normally, this is not acceptable for automated applications that require fast image acquisition. The low signal-to-noise ratio causes difficulties for the proper functioning of autofocus techniques. We consider two noise factors $n_1(\mathbf{x}, \mathbf{p})$ and $n_2(\mathbf{x}, \mathbf{p})$ in Eq. (1)

$$f_{\mathbf{p}}(\mathbf{x}) = (\psi_0(\mathbf{x}) + n_1(\mathbf{x}, \mathbf{p})) * h(\mathbf{x}, \mathbf{p}) + n_2(\mathbf{x}, \mathbf{p}). \quad (5)$$

During the image formation process, the SEM sample can be damaged, contaminated or charged. This is another limiting factor for the number of images that we can take. The other problem is that due to the microscope stage drift the geometry of the images in through-focus series can slightly change. Thus, the object function $\psi_0(\mathbf{x})$ is a function of time $t$. Due to instability of the electron beam, the condition (2) is not always true and the function $h(\mathbf{x}, \mathbf{p})$ is a function of time

$$f_{\mathbf{p},t}(\mathbf{x}) = (\psi_0(\mathbf{x}, t) + n_1(\mathbf{x}, \mathbf{p}, t)) * h(\mathbf{x}, \mathbf{p}, t) + n_2(\mathbf{x}, \mathbf{p}, t). \quad (6)$$

The mean value of the image

$$\bar{f}_{\mathbf{p},t} = \frac{\iint_\Omega f_{\mathbf{p},t}(\mathbf{x})d\mathbf{x}}{\iint_\Omega d\mathbf{x}} \quad (7)$$

also changes in time.

The existing autofocus techniques are usually based on the assumptions (1)–(4). As we have explained, these assumptions do not always hold for electron microscopy, thus, the autofocus techniques can fail. Further, we investigate which of the existing autofocus techniques can be successfully adopted for electron microscopy.

## Sharpness functions

The microscopy images are discrete images that can be represented by the matrix

$\mathbf{F} = \left((f_{i,j})_{i=1}^{N}\right)_{j=1}^{M}$ with the discrete image mean value

$$\bar{F} = \frac{\sum f_{i,j}}{NM}. \quad (8)$$

We focus on iterative autofocus techniques with the use of a sharpness function (SF), a real-valued estimation of the discrete image's sharpness. For a through-focus series, the ideal SF should reach its maximum for the in-focus image and have no other maximum. The ideal SF should change monotonically according to the change of defocus absolute value. The dotted line in Fig. 4 shows the ideal SF for the Gold-on-Carbon SEM through-focus series plotted versus defocus.

The autofocus procedure can be established in two different ways, which are as follows:

- *The static approach*: An amount of images is taken within a wide defocus range and the SF maximum is found (course focusing). The same procedure is done within the smaller defocus range around the maxima, found on the previous step (fine focusing).
- *The dynamic approach*: After taking two images with different defocus values, we use a search algorithm to reach the maximum of the SF (Liu *et al.* 2007, use the Fibbonachi algorithm to this end).

The goal of the second approach is to minimize the amount of images necessary to perform the autofocus procedure. The disadvantage of this approach is that it requests an almost perfect SF shape, which is often not the case in SEM. As it will be shown further, SF in SEM can obtain local optima due to the presence of astigmatism. Also, due to the low signal-to-noise ratio, the SF can be very noisy (has a lot of local minima and maxima). In this case, search algorithms often end up in one of the local maxima, which can be far away from the actual in-focus value. In this paper we focus only on the first approach (the static approach). It will still work in the case of noisy SF. To get a more precise output in this case, SF can be fitted with an analytical function (e.g. Gaussian). Both aproaches will fail, if the SF does reach its global maximum far away from the in-focus image. Thus, it is important to find an SF that will reach its maximum at the in-focus image for a wide variety of SEM samples.

The later sections present different existing SFs. Most of them are based on the assumptions (1)–(4). The goal of the described experiment is to investigate which of them is more stable for the real SEM situations (6)–(7). To stabilize the SFs, they are applied to scaled images

$$\hat{\mathbf{F}} = \left((\hat{f}_{i,j})_{i=1}^{N}\right)_{j=1}^{M}, \quad \text{where } \hat{f}_{i,j} = \frac{f_{i,j}}{\bar{F}}, \tag{9}$$

as a consequence, the mean value of a scaled image

$$\bar{\hat{F}} = 1$$

is a constant, which does not change in time. This brings the situation closer to Eqs (1) and (2), when a mean value of a final image does not depend on time and microscope parameters.

Further we use the notation $s_{\mathbf{par}}^{(\mathbf{name})}$, where $s$ denotes the SF value, **par** denotes the SF parameters and **name** denotes the SF name.

### *Derivative-based sharpness functions*

Derivative SFs are based on the fact that due to the image blur the intensity difference between neighbouring pixels in the defocused image decreases. Derivative-based SFs, described by Brenner *et al.* (1976) and Santos *et al.* (1997) are of the form

$$s_{p,k,\theta}^{(d_{\text{hor}})} = \sum_{i,j} |f_{i,j} - f_{i,j+k}|^p, \quad \text{where } |f_{i,j} - f_{i,j+k}|^p > \theta,$$
$$p \in \{1, 2\}, \quad k \in \mathcal{N}, \quad \theta \in \mathbb{R}^+, \tag{10}$$

where $k$ (pixel difference) and $\theta$ (threshold) adjust the sensitivity of the SF to the noisy images. For the particular parameter values $k = 1$, $p = 1$ in Eq. (10), we obtain SF, known in literature as an *absolute gradient*; for $k = 1$, $p = 2$—*squared gradient*; for $k = 2$, $p = 2$—*Brenner function*. The SF (10) for $\theta > 0$ is also known as a *threshold gradient*. In Eq. (10), only the difference between the pixels in horizonal direction is taken into account, because the SEM scanning is performed in horizontal direction and therefore the noise is correlated there. This SF can fail for certain image geometries (e.g. a number or uniform horizontal stripes). Let $s_{l,k,\theta}^{(d_{\text{ver}})}$ be the function that computes the norm of the pixel difference in vertical direction. Then the form that generalizes derivative-based SFs is

$$s_{p,k,\theta,v}^{(d)} = \frac{1}{2}\left(s_{p,k,\theta}^{(d_{\text{hor}})} + v s_{p,k,\theta}^{(d_{\text{ver}})}\right), \quad v \in \{0, 1\}. \tag{11}$$

Brenner *et al.* (1976) and Santos *et al.* (1997) used only pixel difference parameters $k = 1, 2$. Batten (2000) considered $k = 1, \ldots, 10$. Later we experimentally show that the larger values of $k$ often give more satisfactory results.

The other derivative-based SF used by Santos *et al.* (1997) was proposed and discussed by Tenenbaum (1970), makes use of Sobel operators

$$\mathbf{S}_1 = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}, \quad \mathbf{S}_2 = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix},$$

$$s_{\theta}^{(dt)} = \sum_{i,j} \left((\mathbf{F} * \mathbf{S}_1)_{i,j}^2 + (\mathbf{F} * \mathbf{S}_2)_{i,j}^2\right),$$
$$\text{where } \left((\mathbf{F} * \mathbf{S}_1)_{i,j}^2 + (\mathbf{F} * \mathbf{S}_2)_{i,j}^2\right) > \theta. \tag{12}$$

In Eq. (12), $*$ denotes discrete convolution.

### *Statistical sharpness functions*

Statistical SFs are variance-based and autocorrelation-based SFs (Sun *et al.* 2004). The autocorrelation $\mathbf{C}$ of the image $\mathbf{F}$ is given by $\mathbf{C} = \mathbf{F} * (-\mathbf{F}^*)$, where $^*$ is a complex conjugate. We deal with a real-valued image, as a consequence $\mathbf{F}^* = \mathbf{F}$, and discrete autocorrelation $\mathbf{C} = ((c_{k,l})_{k=1}^{N})_{l=1}^{M}$ is defined (Vollath 1987)

$$c_{k,l} = \sum_{i,j} f_{i,j} f_{i+k-1,j+l-1}. \tag{13}$$

The autocorrelation difference SF is defined (Vollath 1987; Vollath 1997)

$$s_{k,l}^{(acr_h)} = c_{1,k} - c_{1,k+l}. \tag{14}$$

The use of these SFs is based on the fact that the autocorrelation peak of the in-focus image is sharper than autocorrelation peak of the defocused one. If in Eq. (14), $k = 2, l = 1$ then the second term can be approximated (Vollath 1987)

$$s^{(sd_h)} = c_{1,2} - NM\bar{F}. \tag{15}$$

Similarly to derivative-based SF (11) deviations in both vertical and horizontal directions can be considered

$$s_v^{(sd)} = \frac{1}{2}(s^{(sd_h)} + vs^{(sd_v)}), \quad v \in \{0, 1\}, \tag{16}$$

$$s_{k,l,v}^{(acr)} = \frac{1}{2}(s^{(acr_h)} + vs^{(acr_v)}), \quad v \in \{0, 1\}. \tag{17}$$

The image variance SFs are based on the fact that the in-focus image has higher contrast than the defocused one. Image variance and normalized variance SFs are defined by Vollath (1987), Santos *et al.* (1997) and Liu *et al.* (2007) as

$$s^{(v)} = \sum_{i,j}(f_{i,j} - \bar{F})^2, \tag{18}$$

$$s^{(vn)} = \frac{1}{\bar{F}}s^{(v)}. \tag{19}$$

For the scaled discrete image (9), SFs (18) and (19) are equivalent.

*Histogram-based sharpness functions*

In most applications, the unscaled image **F** is a matrix of natural intensity values. Let

$$\tilde{\mathbf{f}} = (\tilde{f}_k)_{k=1}^L, \quad \tilde{f}_{k-1} < \tilde{f}_k,$$

be a set of all the pixel values in the image **F**, that is $f_{i,j} \in \mathbf{F} \Leftrightarrow \exists k$ such that $f_{i,j} = \tilde{f}_k \in \tilde{\mathbf{F}}$. The vector $\mathbf{h} = (h_k)_{k=1}^L$, where $h_k$ is the amount of pixels with the value $\tilde{f}_k$ in the image **F**, is called the histogram of the image **F**. Then the probability of the pixel with the value $\tilde{f}_k$ is equal to $\frac{h_k}{NM}$.

Histogram-based SFs are the range and the entropy (Santos *et al.* 1997)

$$s^{(hr)} = \max_{k, h_k \neq 0} h_k - \min_{k, h_k \neq 0} h_k. \tag{20}$$

$$s^{(he)} = -\sum_{k, h_k \neq 0} \frac{h_k}{NM} \log_2 \frac{h_k}{NM}. \tag{21}$$

Another histogram-based function is threshold image count (Santos *et al.* 1997)

$$s_\theta^{(ht)} = \sum_{k=1}^{n} h_k, \quad \text{where } n \text{ is such that } \tilde{f}_n \le \theta \text{ and } \tilde{f}_{n+1} > \theta. \tag{22}$$

*Intensity-based sharpness functions*

Erasmus & Smith (1982) used the modified version of variance (18) with neglected term $\bar{F}$ for SEM simultaneous autofocus and astigmatism correction. Santos *et al.* (1997) extended this SF with a power $p$ and threshold parameter

$$s_{p,\theta}^{(vm)} = \sum_{i,j} |f_{i,j}|^p, \quad \text{where } f_{i,j} > \theta, \quad p \in \{1, 2\}. \tag{23}$$

*Fourier transform-based sharpness functions*

Fourier transform sharpness functions are based on the fact that the magnitudes of the high frequencies in the in-focus image are higher than in a defocused one. For our evaluation, we choose an SF based on work of Ong *et al.* (1997) and Vladar *et al.* (1998)

$$s_{l,h}^{(ft)} = \sum_{i,j} |g_{i,j}|, \quad \text{where } i \in [h, n_0 - l] \cap [n_0 + l, n - h],$$
$$j \in [h, m_0 - l] \cap [m_0 + l, m - h], \tag{24}$$

where $l$ and $h$ are the low- and the high-band frequencies, $n_0 = \mathrm{mod}(n/2) + 1$, $m_0 = \mathrm{mod}(m/2) + 1$. $\mathbf{G} = ((g_{i,j})_{i=1}^N)_{j=1}^M$ is the discrete Fourier transform of the image **f**, multiplied by the window function to avoid the Gibbs phenomenon due to the discontinuity and non-periodicity.

**Experimental images**

Figures 1 and 2 show the images from experimental through-focus series of typical SEM samples. They were obtained with an FEI Strata SEM (Eindhoven, The Netherlands) at magnifications from 15.000 to 25.000. The size of each image is $442 \times 442$ pixels. The number of images in a through-focus series varies from 11 to 53. The images with a pixel depth of both 8- and 16-bit were tested. The images shown in Fig. 1 are from through-focus series which are free of astigmatism, Those in Fig. 2 are from stigmatic through-focus series.

Figures 1(c–e) and 2(c) and (d) show the images of tin balls samples; Figs 1(a) and (b) and 2(a) and (b)—the images of cross sections; Figs 1(f) and (g)—the images of integrated circuits; Figs 1(h) and (i)—the images of the hard disk heads; Fig. 2(e)—an image of Gold-on-Carbon sample.

The images contain different amounts of details. Series of images, such as in Figs 1(b) and (b), contain fine details, that is the difference between neighbouring objects in an image is several pixels. Other series, such as in Figs 2(b) and 1(h), contain only course details (the difference between two objects is more than 100 pixels). There are also samples with periodic structure, such as Fig. 1(a).

**Evaluation procedure**

In the recent work with evaluation of SFs (Liu *et al.* 2007) several evaluation criteria, based on the previous works (Santos *et al.* 1997; Sun *et al.* 2004) are explained. For

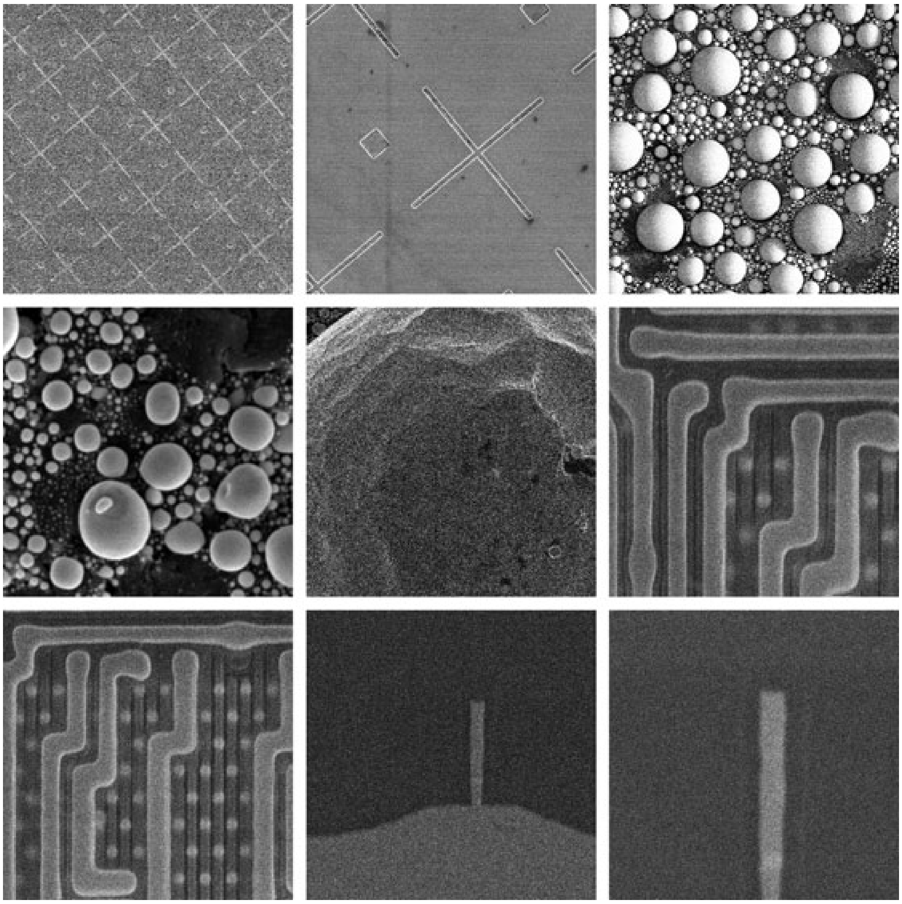

**Fig. 1.** Images from experimental SEM through-focus series without astigmatism.

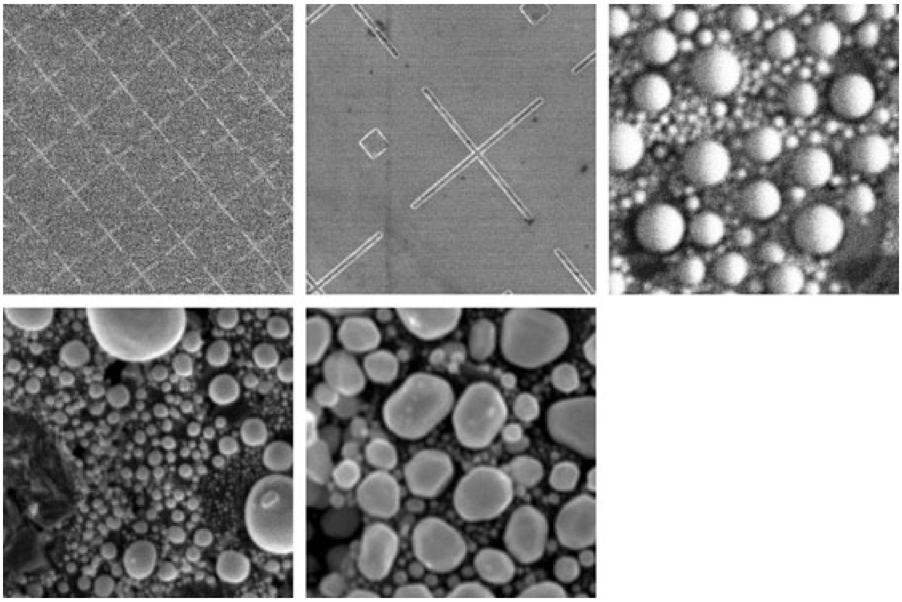

**Fig. 2.** Images from experimental SEM through-focus series with astigmatism.

evaluation of earlier defined SFs, we estimate four evaluation criteria described in the recent work (Liu *et al.* 2007):

- *Accuracy* estimates the distance between the best focus position, determined by a professional human operator and the maximum of an SF.
- *Local maxima* estimates the number of local (false) maxima in an SF.
- *Range* estimates the monotonicity interval range on both sides of an SF global maximum.
- *Noise* estimates the noise amplitude in an SF.

We do not consider dynamic evaluation criteria (Liu *et al.* 2007), because in this paper we discuss only static autofocus approach, for the reason that it is more robust to the local optima in SF, which is often the case in SEM due to the noise and astigmatism. Instead dynamic accuracy (Liu *et al.* 2007) we use a static accuracy described in (Santos *et al.* 1997; Sun *et al.* 2004). We also neglect the evaluation criteria, which estimate the width of an SF peak (Sun *et al.* 2004; Liu *et al.* 2007). The narrow SF peak is not always beneficial for SEM autofocusing. If defocus step in through-focus series is large (like in the course focusing) and the SF is noisy due to the noise in the images, then the SF peak can be skipped and autofocus procedure will fail. Earlier the computational time required by an SF was also considered as an evaluation criterion (Santos *et al.* 1997). Nowadays, the computational time of any of the SFs described earlier is lower, then an SEM image acquisition time. Thus, the SF value for a current image in a focus series can be computed in parallel with the acquisition of the next image in the focus series, which costs overall no time. Also, for this reason of available computational power, the time evaluation criteria was neglected in later works (Sun *et al.* 2004; Liu *et al.* 2007).

Figure 3 shows five different simulated SFs. Before evaluation the SF values are scaled between 0 and 1

$$\bar{\mathbf{s}} = \frac{\mathbf{s} - \min \mathbf{s}}{\|\mathbf{s} - \min \mathbf{s}\|_\infty}. \tag{25}$$

We imagine that these SFs are computes for one focus series of seven images, with the fourth image determined to be in-focus by an experienced human operator. Based on these simulated SFs, we are going to illustrate how our evaluation procedure works. In the evaluation, procedure criteria described earlier are represented by the real numbers $z_{\mathrm{acc}}$, $z_{\mathrm{ran}}$, $z_{\mathrm{lm}}$, $z_{\mathrm{noise}} \in [0, 1]$. The numbers are equal to 1 in the ideal situation and equal to 0, if the situation is as far from ideal as possible. All $z$-values between 0 and 1 are achieved in respect to the two boundary cases. The evaluation criteria values for SFs from Fig. 3 are shown in Table 1.

**Table 1.** Evaluation scores for simulated SFs.

| SF | Accuracy $z_{\mathrm{acc}}$ | False maxima $z_{\mathrm{lm}}$ | Range $z_{\mathrm{ran}}$ | Noise $z_{\mathrm{noise}}$ | Overall score $z_{\mathrm{tot}}$ |
|---|---|---|---|---|---|
| SF1 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| SF2 | 1.00 | 0.00 | 0.00 | 0.45 | 0.58 |
| SF3 | 0.67 | 0.506 | 0.25 | 0.68 | 0.57 |
| SF4 | 1.00 | 0.00 | 0.00 | 0.00 | 0.50 |
| SF5 | 0.00 | 1.00 | 1.00 | 1.00 | 0.50 |

*Accuracy*

The accuracy of SF1, SF2 and SF4 is equal 1, because the maxima of SF1, SF2 and SF4 coincides with the position, identified by the human operator. The accuracy of SF5 is equal 0, because the maximum of SF5 is as far from the maximum, determined by the human operator, as possible: $\delta_{\mathrm{acc}}^{\max} = 3$ defocus steps. For SF3, the amount of defocus steps between SF maximum and the in-focus image is $\delta_{\mathrm{acc}} = 1$, thus

$$z_{\mathrm{acc}} = 1 - \frac{\delta_{\mathrm{acc}}}{\delta_{\mathrm{acc}}^{\max}} = \frac{2}{3}.$$

*Local maxima*

*Analogically for the local maxima criterion.* Maximum amount of false maxima in SF for seven images is $\delta_{\mathrm{lm}}^{\max} = 3$, because we do not consider a global maximum as a false maximum and

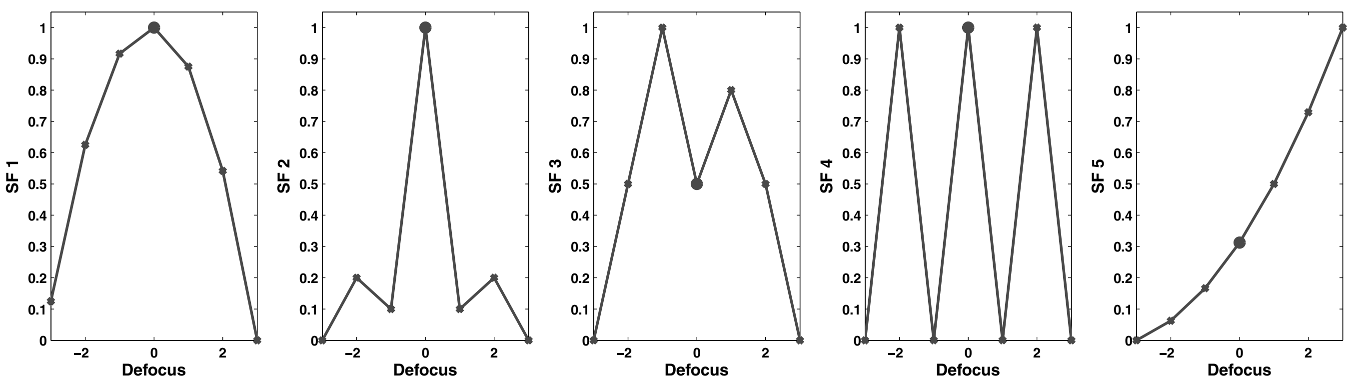


**Fig. 3.** Simulated SFs for the focus series of seven images with the fourth image considered to be in-focus. The functions are ordered according to the computed evaluation score, that is SF1 is a function of the best quality and SF5 is the worst from the five represented functions.

we do not consider boundary values. Thus, for SF1 and SF5 $z_{\mathrm{lm}} = 1$, for SF2 and SF4 $z_{\mathrm{lm}} = 0$ and for SF3 with the amount of false maxima $\delta_{\mathrm{lm}} = 1$

$$z_{\mathrm{lm}} = 1 - \frac{\delta_{\mathrm{lm}}}{\delta_{\mathrm{lm}}^{\mathrm{max}}} = \frac{1}{2}.$$

For the general case of SF for a focus series of $K$ images

$$\delta_{\mathrm{lm}}^{\mathrm{max}} = \left\lfloor \frac{K-1}{2} \right\rfloor - 1.$$

If the global maximum coincides with a boundary point, $\delta_{\mathrm{lm}}^{\mathrm{max}} = \lfloor \frac{K-1}{2} \rfloor - 1$.

*Range*

SF1 and SF5 have $z_{\mathrm{ran}} = 1$, because they are monotone around the global maximum. SF2 and SF4 have $z_{\mathrm{ran}} = 0$, because because their monotonicity interval around the global maximum is minimal (2 defocus steps). By subtraction 2 from the amount of steps

$$\delta_{\mathrm{ran}}^{\mathrm{max}} = (K-1) - 2$$

we get the area of the largest possible domain size outside the monotonicity range around the global maximum. If the global maximum is located in the boundary point (like in SF5), we consider $\delta_{\mathrm{ran}}^{\mathrm{max}} = K - 2$. For SF3, this area is $\delta_{\mathrm{ran}} = 3$ defocus steps, then for SF3

$$z_{\mathrm{ran}} = 1 - \frac{\delta_{\mathrm{ran}}}{\delta_{\mathrm{ran}}^{\mathrm{max}}} = \frac{1}{4}.$$

*Noise level*

The noise evaluation criteria estimates amplitude of local maxima and minima. SF1 and SF5 have $z_{\mathrm{noise}} = 1$, because they do not have local maxima and minima. SF4 have $z_{\mathrm{noise}} = 0$, because it has as much maxima and minima as possible, with the highest sum amplitude possible

$$\delta_{\mathrm{noise}}^{\mathrm{max}} = K - 3,$$

because the SF values are scaled between 0 and 1 (25), and we do not consider the global maximum or boundary points as local maxima or local minima. If the global maximum is located in the boundary point (like in SF5), $\delta_{\mathrm{noise}}^{\mathrm{max}} = K - 2$. Then for SF3, the total amplitude of global minima and maxima is $\delta_{\mathrm{noise}} = 0.8 + (1 - 0.5) = 1.3$

$$z_{\mathrm{noise}} = 1 - \frac{\delta_{\mathrm{noise}}}{\delta_{\mathrm{noise}}^{\mathrm{max}}} = 1 - \frac{1.3}{4} = 0.675.$$

*Overall score*

For the total overall score of SF values, we use the weighted sum of criteria described earlier

$$z_{\mathrm{tot}} = \frac{1}{2}\left(z_{\mathrm{acc}} + \frac{1}{3}(z_{\mathrm{ran}} + z_{\mathrm{lm}} + z_{\mathrm{noise}})\right) \in [0, 1].$$

We choose the weight in such a way that we give equal importance to the criteria $z_{\mathrm{acc}}$ and the sum of criteria $z_{\mathrm{ran}}$, $z_{\mathrm{lm}}$ and $z_{\mathrm{noise}}$. The reasoning for this choice is the importance of accuracy for static autofocus algorithms and the fact, the criteria $z_{\mathrm{ran}}$, $z_{\mathrm{lm}}$ and $z_{\mathrm{noise}}$ are related. If SF does not have false maxima then $z_{\mathrm{ran}} = z_{\mathrm{lm}} = z_{\mathrm{noise}} = 1$, and the overall score could be high. However, if the maximum of such SF is far away from in-focus position (see SF5, Fig. 3), such SF is meaningless for the autofocus, and it should not get high overall score. For ideal SF, like SF1, $z_{\mathrm{tot}} = 1$.

Evaluation systems described previously are comparative (Santos *et al.* 1997; Sun *et al.* 2004; Liu *et al.* 2007; Rudnaya *et al.* 2009), that is each criteria is set to a value with one of the boundaries that cannot be estimated in advance, and further results are compared between each other, to find the best quality SF. For example, the accuracy evaluation criteria would be equal to the difference between SF maximum and in-focus image in steps. The minimal value would be equal to zero and considered to be the best. The maximal value here would be limited by amount of images in focus series, which could be different in different cases. In our evaluation procedure we know in advance that every criteria lies in the range [0, 1] independently of amount of images in the focus series.

## Results and discussion

SFs defined in Section 3 are applied to each of the SEM through-focus series described in Section 4. The results are evaluated according to the procedure described in the previous section, and averaged for all given through-focus series. Parameterized SF are applied with different parameter values. For example, the gradient-based SF (11) is applied to the Gold-on-Carbon through-focus series (Fig. 2e). We fix function parameters $v = 0$, $p = 2$, $\theta = 0$ and vary the pixel difference parameter $\mathbf{k} = 1, \ldots, 441$. For each $k$, we get an SF, and each SF is evaluated. As we have discussed before, the higher overall evaluation criteria $z_{\mathrm{tot}} \in [0, 1]$, the higher we estimate SF. The results with the highest overall score $z_{\mathrm{tot}}$ is chosen. The same procedure is repeated for each through-focus series. Evaluation scores obtained for each focus series are averaged, and the final result can be seen in the first row ($\mathcal{N} = 1$) of Table 2. Table 2 shows the performance of derivative-based SFs with varying of different parameters separately (for the details, see columns *SF parameters*). The results for each of the five SF families are shown in five different tables Tables 2–6. Table 7 summarizes the results according to the highest overall scores $z_{\mathrm{tot}}$ from Tables 2 to 6. The derivative-based and Fourier transform-based SFs have shown the best overall performance.

**Table 2.** Derivative-based SFs (11) and (12) average evaluation score for all through-focus series combined.

| | SF parameters | | | | Evaluation scores | | | | |
|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{N}$ | Vertical direction $v$ | Power $p$ | Pixel difference $k$ | Threshhold $\theta$ (%) | Accuracy $z_{acc}$ | False max. $z_{lm}$ | Range $z_{ran}$ | Noise $z_{noise}$ | Overall $z_{tot}$ |
| 1 | 0 | 2 | 1, . . . , 441 | 0 | 0.9767 | 0.9627 | 0.9033 | 0.9591 | 0.9592 |
| 2 | 1 | 1 | 1, . . . , 441 | 0 | 0.9940 | 0.9449 | 0.8766 | 0.9485 | 0.9587 |
| 3 | 1 | 2 | 1, . . . , 441 | 0 | 0.9583 | 0.9565 | 0.9033 | 0.9569 | 0.9486 |
| 4 | 0 | 1 | 1, . . . , 441 | 0 | 0.9767 | 0.9140 | 0.8376 | 0.9291 | 0.9351 |
| 5 | 1 | 1 | 1 | 0, . . . , 100 | 0.8773 | 0.8081 | 0.5574 | 0.8297 | 0.8045 |
| 6 | 0 | 2 | 1 | 0, . . . , 100 | 0.8267 | 0.8449 | 0.5589 | 0.8471 | 0.7885 |
| 7 | 0 | 1 | 1 | 0, . . . , 100 | 0.8273 | 0.8068 | 0.5856 | 0.8331 | 0.7846 |
| 8 | 1 | 2 | 1 | 0, . . . , 100 | 0.8178 | 0.8506 | 0.5186 | 0.8427 | 0.7776 |
| 9 | – | – | – | 0, . . . , 100 | 0.7809 | 0.8277 | 0.5744 | 0.8554 | 0.7667 |

Note: The squared gradient without threshold that takes into account the difference between pixels in vertical and horizontal directions has the highes overall score. Tenenbaum gradient has the lowest overall score.

**Table 3.** Fourier transform-based SFs (24) average evaluation score for all through-focus series combined.

| | SF parameters | | Evaluation scores | | | | |
|---|---|---|---|---|---|---|---|
| $\mathcal{N}$ | Low-frequency band l | High-frequency band h | Accuracy $z_{acc}$ | False max. $z_{lm}$ | Range $z_{ran}$ | Noise $z_{noise}$ | Overall $z_{tot}$ |
| 1 | 0 | 1,. . . ,219 | 0.9708 | 0.9472 | 0.9070 | 0.9604 | 0.9545 |
| 2 | 2 | 1,. . . ,219 | 0.9734 | 0.9416 | 0.8850 | 0.9518 | 0.9498 |
| 3 | 1 | 1,. . . ,219 | 0.9734 | 0.9241 | 0.8725 | 0.9421 | 0.9432 |

Note: Fourier transform-based SF without low-frequency band has the highest overall score.

**Table 4.** Statistical SFs (17) and (18) average evaluation score for all through-focus series combined.

| | SF parameters | | | Evaluation scores | | | | |
|---|---|---|---|---|---|---|---|---|
| $\mathcal{N}$ | Vertical direction $v$ | Autocorrelation coefficient $k$ | Autocorrelation coefficient $l$ | Accuracy $z_{acc}$ | False max. $z_{lm}$ | Range $z_{ran}$ | Noise $z_{noise}$ | Overall $z_{tot}$ |
| 1 | 0 | 1 | 1,. . ., 50 | 0.9311 | 0.9208 | 0.8587 | 0.9450 | 0.9197 |
| 2 | 1 | 1 | 1,. . ., 50 | 0.9726 | 0.8839 | 0.7801 | 0.8868 | 0.9115 |
| 3 | 0 | 1,. . ., 50 | 1 | 0.9264 | 0.8840 | 0.8226 | 0.9097 | 0.8992 |
| 4 | 1 | 1,. . ., 50 | 1 | 0.9482 | 0.8579 | 0.7131 | 0.8839 | 0.8832 |
| 5 | – | – | – | 0.8713 | 0.9038 | 0.7383 | 0.9191 | 0.8625 |

Note: Autocorrelation-based SF has the highest overall score. Variance-based SF has the lowest overall score.

**Table 5.** Histogram-based SFs (20) and (22) average evaluation score for all through-focus series combined.

| | SF parameters | | Evaluation scores | | | | |
|---|---|---|---|---|---|---|---|
| $\mathcal{N}$ | SF name | Threshhold $\theta$ (%) | Accuracy $z_{acc}$ | False max. $z_{lm}$ | Range $z_{ran}$ | Noise $z_{noise}$ | Overall $z_{tot}$ |
| 1 | Threshold count | 0, . . . , 100 | 0.8787 | 0.8855 | 0.5671 | 0.8811 | 0.8283 |
| 2 | Entropy | – | 0.7359 | 0.8273 | 0.4769 | 0.8662 | 0.7297 |
| 3 | Range | – | 0.2856 | 0.6944 | 0.1954 | 0.7367 | 0.4139 |

Note: Histogram-based SF has the highest overall score. Range SF has the lowest overall score.

**Table 6.** Intensity-based SFs (23) average evaluation score for all through-focus series combined.

| | SF parameters | | Evaluation scores | | | | |
|---|---|---|---|---|---|---|---|
| $\mathcal{N}$ | Power $p$ | Threshhold $\theta$ (%) | Accuracy $z_{\text{acc}}$ | False max. $z_{\text{lm}}$ | Range $z_{\text{ran}}$ | Noise $z_{\text{noise}}$ | Overall $z_{\text{tot}}$ |
| 1 | 2 | $0, \ldots, 100$ | 0.8207 | 0.8376 | 0.6294 | 0.8679 | 0.7995 |
| 2 | 1 | $0, \ldots, 100$ | 0.1304 | 0.8110 | 0.3574 | 0.8045 | 0.3940 |

Note: Squared intensity has the highest overall score. Absolute intensity has the lowest overall score.

For derivative-based SFs it is clear that varying the pixel difference parameter gives much better results, then varying the threshold. Two SFs with pixels difference parameters $k = 1$ and 10, corresponding to the $\mathcal{N} = 1$ in Table 2 are plotted for the Gold-on-Carbon through-focus series with astigmatism (Fig. 2e). We can see that for $k = 1$ the SF has a double-peak effect. The double-peak effect in an SF was shown analytically earlier (Erasmus & Smith 1982) for certain types of samples and SFs. In Fig. 4 for $k = 1$ $z_{\text{tot}} = 0.7431$ due to the error in accuracy, one local maxima, noise amplitude and low monotonicity range. For $k = 10$, the double-peak effect disappears and $z_{\text{tot}} = 1$. Figures 5 and 6 show the SF surface for $k \in [1, 441]$ for the Gold-on-Carbon series with astigmatism present (Fig. 2e). In Fig. 5, the SFs are not scaled. In Figs 6 and 4, the SF values are scaled between 0 and 1 (25) for each parameter $k$. Fig. 7 shows SF values for several defocus values plotted versus pixel difference $k$.

The size of each image in the through-focus series is $442 \times 442$ pixels. Evaluation shows that for $k \in [6, 421]$ $z_{\text{tot}} = 1$. This effect was observed for most of the other experimental through-focus series as well. The derivative-based SF does not fail even if we take the differences between the pixels located far from each other. For low values of $k$, it can be explained as noise robustness growth. It was also shown by Batten (2000) that the derivative-based SF performs better with $k = 10$ than for smaller $k < 10$, for the Gold-on-Carbon SEM through-focus series. However, the possibilities of using larger $k$ were not explored. To give an idea about the reasoning behind this phenomena, we will look at derivative based SFs (11) closer. Assuming that the threshold parameter $\theta = 0$, the power

**Table 7.** Average evaluation score for all through-focus series combined.

| $\mathcal{N}$ | SF family | Accuracy $z_{\text{acc}}$ | False maxima $z_{\text{lm}}$ | Range $z_{\text{ran}}$ | Noise $z_{\text{noise}}$ | Overall score $z_{\text{tot}}$ |
|---|---|---|---|---|---|---|
| 1 | Derivative | 0.9767 | 0.9627 | 0.9033 | 0.9591 | 0.9592 |
| 2 | Fourier transform | 0.9708 | 0.9472 | 0.9070 | 0.9604 | 0.9545 |
| 3 | Statistical | 0.9311 | 0.9208 | 0.8587 | 0.9450 | 0.9197 |
| 4 | Histogram | 0.8787 | 0.8855 | 0.5671 | 0.8811 | 0.8283 |
| 5 | Intensity | 0.8207 | 0.8376 | 0.6294 | 0.8679 | 0.7995 |

Note: For each SF family the results with the highest overall scores are chosen from Tables 2–6.

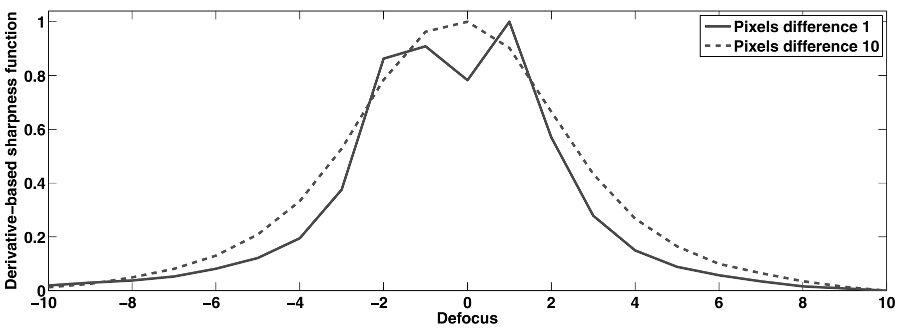


**Fig. 4.** Derivative-based SF (11) for experimental Gold-on-Carbon through-focus series with astigmatism. Then horizontal axis represents the defocus in arbitrary units. The solid line is the function related to parameters $k = 1$, $p = 2$, $\theta = 0$ and the dotted line is a the function related to $k = 10$, $p = 2$, $\theta = 0$. The SF values are scaled between 0 and 1.

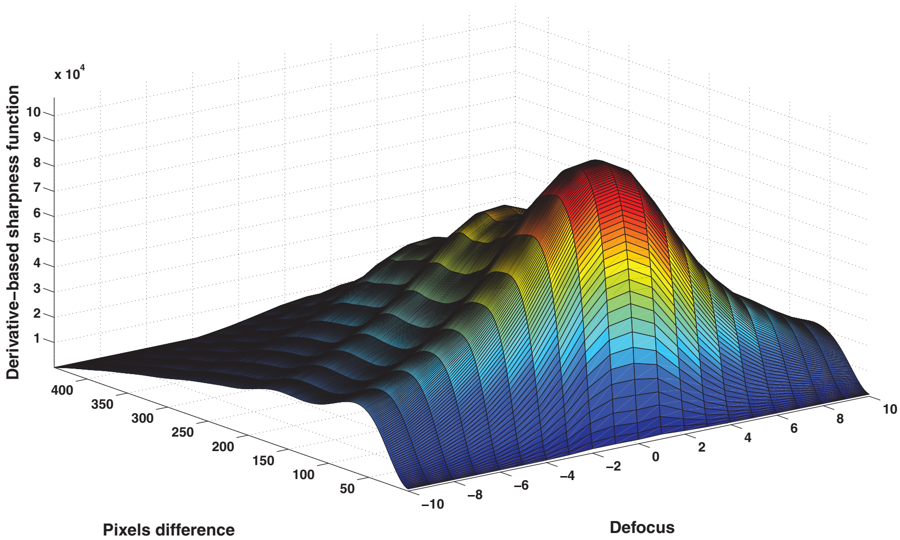


**Fig. 5.** Scaled derivative-based SF (11) for Gold-on-Carbon through-focus series with astigmatism. The axes represent the defocus in arbitrary units, the pixel difference parameter $k$ and the SF values for parameters $p = 2, \theta = 0$.

parameter $p = 2$ and $v = 0$, we get an SF that takes into account the difference between pixels only in horizontal direction with the highest overall score from table Table 2 ($\mathcal{N} = 2$)

$$s_{2,k,0}^{(d_{\mathrm{hor}})} = \sum_{i=1}^{N}\sum_{j=1}^{M-k}(f_{i,j} - f_{i,j+k})^2 = \sum_{i=1}^{N}\sum_{j=1}^{M-k} f_{i,j}^2 + \sum_{i=1}^{N}\sum_{j=k+1}^{M} f_{i,j}^2 - 2\sum_{i=1}^{N}\sum_{j=1}^{M-k} f_{i,j} f_{i,j+k}. \quad (26)$$

The first two terms of Eq. (26) are simply SF (26), applied to the parts of the image. The third term is an autocorrelation coefficient. When $k$ is increasing the autocorrelation coefficient tends to zero. Thus, derivative SF (26) tends to the sum of SF (26) applied to small parts of discrete image. We can clearly see it in Figs 5 and 7, which with increasing of the pixel difference parameter $k$ the width of the SF peak first increases and then decreases. This could be also explained by Eq. (26). It was shown by Erasmus & Smith (1982) that the SF (26) has a local minimum in the in-focus position in the case of astigmatism for amorphous samples. Due to non-amorphous nature of most of the samples, this minimum could be shifted to left or right. Because the first and the second terms of (26) are the SF (26) applied to

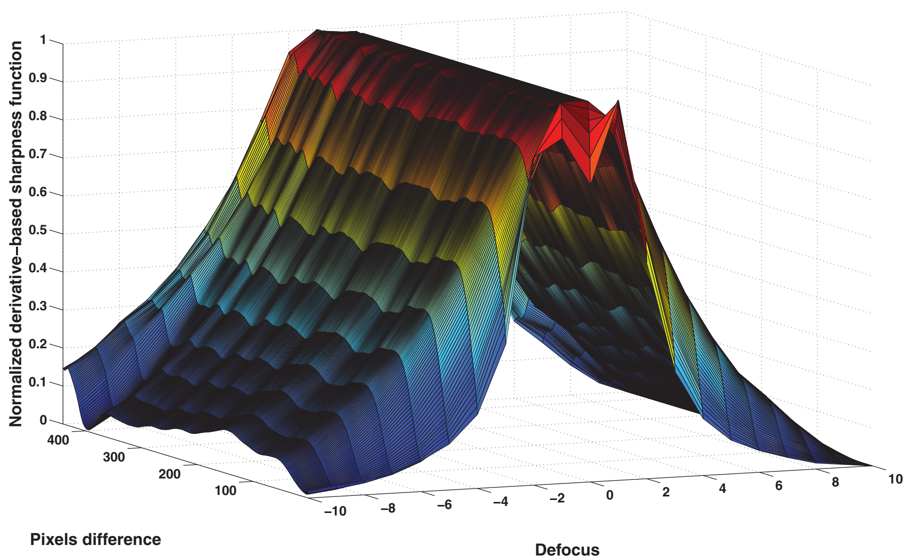


**Fig. 6.** The same results as in Fig. 5, but SF values are scaled between 0 and 1 for every $k$ value.

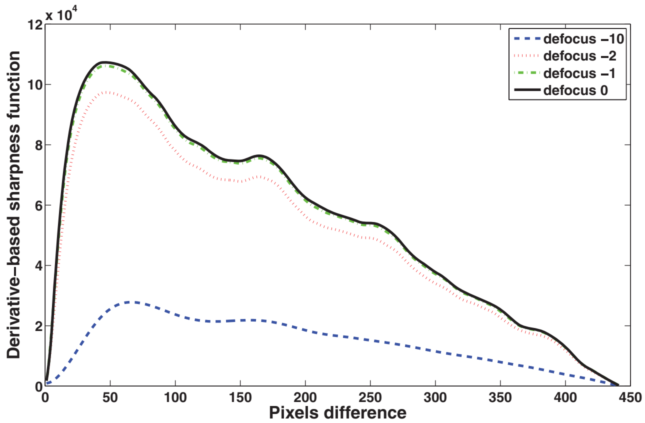


**Fig. 7.** Derivative-based SF (11) for experimental Gold-on-Carbon through-focus series with astigmatism. The horizontal axis shows pixel difference $k$. The vertical axis shows SF values. The several plots correspond to the different defocus values.

different parts of an image, there is a high chance that the local minimum of both functions will be obtained in different positions. Thus, the composition of the two will average the local minimum effect or even will help to get reed of it.

Figure 8 shows Fourier transform-based SFs for a low-frequency band equal to 2 with different high-frequency bands for Gold-on-Carbon series with astigmatism. The high-frequency band is changing from 1 to 219. For certain high frequencies, the similar double peak effect, as for the derivative-based SF is observed. It clearly shows, how the SF parameter varying can influence the SF quality.

According to Table 4 variance-based SF ($\mathcal{N} = 5$) has much lower overall evaluation score than autocorrelation based SFs ($\mathcal{N} = 1, \ldots, 4$). It is important to note that the variance-based SF is applied without varying any parameter. Probably, a smart parameter choice for this function could lead to higher evaluation results. Through Table 5 we can see the histogram-based SF with threshold count (22) has much higher evaluation score then entropy (21) and range (20). It is remarkable, that enthropy and range functions are as well, as variance, are applied without an extra parameter. Table 6 shows that the intensity-based SF has higher quality with the power parameter $p = 2$.

A number of autofocus techniques has been applied to variety of the SEM through-focus series, including through-focus series with astigmatism. The modified evaluation score procedure has been introduced. We have seen that Fourier transform-based autofocus techniques with varying frequency band parameter and derivative-based autofocus techniques with varying pixel difference parameter show better performance, compared to most of the other autofocus techniques.

It has been shown earlier that the Fourier transform-based SFs perform as well as the variance-based SFs, and the derivative-based SFs fail due to the noise in SEM. However, our research has shown that by extra variation of SF parameter the sensitivity to noise can dramatically increase. Not only Fourier transform and derivative, but also the autocorrelation-based SFs show better performance for certain parameters than the variance-based SFs. These techniques with a wide parameter variation were not evaluated before, neither for the variety of SEM samples, nor for light microscopy.

The varying parameters influence the SF quality (e.g. in the case of astigmatism) and the SF peak width. However, the quantification of this influence has still to be studied in order to develop robust and independent SEM autofocus algorithm. Surprisingly, satisfactory results were achieved for large values of the pixel difference parameter with derivative-based autofocus techniques. This can be explained through the

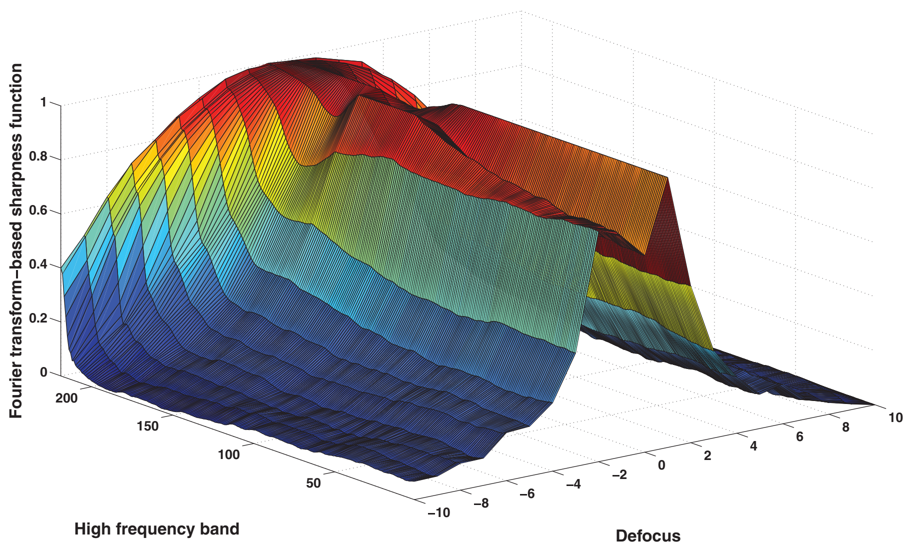


**Fig. 8.** Scaled Fourier transform-based SF for Gold-on-Carbon through-focus series with astigmatism.

fact, that derivative-based SF can be seen as a combination of intensity and autocorrelation-based SFs.

## Acknowledgements

This work has been carried out as a part of the Condor project at FEI Company under the responsibilities of the Embedded Systems Institute (ESI). This project is partially supported by the Dutch Ministry of Economic Affairs under the BSIK program.

We kindly acknowledge R. van Vucht and M. Janus (FEI, The Netherland) for assistance with obtaining the experimental data.

## References

Batten, C.F. (2000) *Autofocusing and Astigmatism Correction in the Scanning Electron Microscope*. Master's thesis, University of Cambridge.

Brenner, J.F., Dew, B.S., Horton, J.B., King, T., Neurath, P.W. & Selles, W.D. (1976) An automated microscope for cytologic research a preliminary evaluation. *J. Histochem. Cytochem.* **24**, 100–111.

Erasmus, S.J. & Smith, K.C.A. (1982) An automatic focusing and astigmatism correction system for the SEM and CTEM. *J. Microsc.* **127**, 185–199.

Liu, X.Y., Wang, W.H. & Sun, Y. (2007) Dynamic evaluation of autofocusing for automated microscopic analysis of blood smear and pap smear. *J. Microsc.* **227**, 15–23.

Nayar, S.K. & Nakagawa, Y. (1994) Shape from focus. *IEEE Trans. Pattern Anal. Mach. Intell.* **16**, 824–831.

Ong, K.H., Phang, J.C.H. & Thong, J.T.L. (1997) A robust focusing and astigmatism correction method for the scanning electron microscope. *Scanning* **19**, 553–563.

Rudnaya, M.E., Mattheij, R.M.M. & Maubach, J.M.L. (2009) Iterative autofocus algorithms for scanning electron microscopy. *Microsc. Microanal.* **15**, 1108–1109.

Santos, A., DeSolórzano, C.O., Vaquero, J.J., Peña, J.M., Malpica, N. & Del Pozo, F. (1997) Evaluation of autofocus functions in molecular cytogenetic analysis. *J. Microsc.* **188**, 264–272.

Sun, Y.U., Duthaler, S. & Nelson, B.J. (2004) Autofocusing in computer microscopy: selecting the optimal focus algorithm. *Microsc. Res. Tech.* **65**, 139–149.

Tenenbaum, J.M. (1970) *Accomodation in computer vision*. Ph.D. thesis, Standford University.

Van den Broek, W. (2007) *Advanced focus methods in electron microscopy: tomographic reconstruction of the EELS data cube autofocus of HAADF-STEM images*. Ph.D. thesis, University of Antwerp.

Vladar, A.E., Postek, M.T. & Davidson, M.P. (1998) Image sharpness measurement in scanning electron microscopy–Part II. *Scanning* **20**, 24–34.

Vollath, D. (1987) Automatic focusing by correlative methods. *J. Microsc.* **147**, 279–288.

Vollath, D. (1997) The influence of the scene parameters and of noise on the behaviour of automatic focusing algorithms. *J. Microsc.* **151**, 133–146.