Autofocus and two-fold astigmatism correction in HAADF-STEM

EINDHOVEN UNIVERSITY OF TECHNOLOGY
Department of Mathematics and Computer Science

CASA-Report 10-09
February 2010

Autofocus and two-fold astigmatism correction in HAADF-STEM

by

M.E. Rudnaya, W. van den Broek, R. Doornbos,
R.M.M. Mattheij, J.M.L. Maubach

Centre for Analysis, Scientific computing and Applications
Department of Mathematics and Computer Science
Eindhoven University of Technology
P.O. Box 513
5600 MB Eindhoven, The Netherlands
ISSN: 0926-4507

Autofocus and two-fold astigmatism correction in HAADF-STEM

M.E. Rudnaya[a], W. van den Broek[b], R. Doornbos[c], R.M.M. Mattheij[a] , J.M.L. Maubach[a]

[a] CASA, Dept. of Mathematics and Computer Science, Eindhoven University of Technology, PO Box 513, 5600 MB, Eindhoven, The Netherlands

[b] EMAT, University of Antwerp, Groenenborgerlaan 171, B-2020 Antwerp, Belgium

[c] Embedded Systems Institute, PO Box 513, Building LG 0.10 NL - 5600 MB Eindhoven, The Netherlands

Key words. Automatic focusing, astigmatism, Scanning transmission electron microscopy, HAADF-STEM, Nelder-Mead simplex method, Variance.

February 2, 2010

Abstract

A new simultaneous autofocus and two-fold astigmatism correction method is proposed for High Angle Annular Dark Field Scanning Transmission Electron Microscopy (HAADF-STEM). The method makes use of a modification of an image variance, which has already been used before as an image quality measure for different types of microscopy. In this paper we describe numerical simulations based on a classical HAADF-STEM linear image formation model showing that the modified variance reaches it's maximum for Scherzer focus and zero astigmatism. In order to find this maximum in a three-parameter space we employ the well-known Nelder-Mead simplex optimization algorithm. The method is implemented and tested on a FEI Tecnai F20. It successfully finds the optimal defocus and zero astigmatism with the time and accuracy, compared with the human operator. The method is iterative, and finding the optimal defocus and zero astigmatism requires obtaining typically 20-50 images.

1 Introduction

High Angle Annular Dark Field Scanning Transmission Electron Microscopy (HAADF-STEM) is a powerful material science research tool. Nowadays HAADF-STEM still requires an expert operator in order to obtain in-focus and astigmatism-free images manually. Both the defocus and the two-fold astigmatism have to be adjusted regularly during the image obtaining process. The possible reasons are for instance the instabilities of the electron microscope and the magnetic nature of some samples. For the next HAADF-STEM instrument generations the manual operation has to be automated. One of the reasons is that for some applications the high level of repetition severely strains the required concentration. Therefore, a robust and reliable simultaneous autofocus and two-fold astigmatism correction algorithm is a necessary tool for HAADF-STEM automation.

The problem of autofocus and two-fold astigmatism correction has been addressed before for different types of microscopy. The autofocus techniques were investigated for fluorescent light microscopy [11] and non-fluorescence light microscopy [8]. For fluorescent light microscopy

it has been shown experimentally [11] that out of a large number of image quality measures, image variance was the better. However, the astigmatism is not considered in [11, 8], because it is not important for light microscopy nowadays. Few methods for simultaneous autofocus and astigmatism correction for scanning electron microscopy were proposed [6, 9]. One of them is based on a Gaussian model for the point spread function [6], which is not accurate enough for HAADF-STEM image formation description: Gaussian point spread function can only describe defocus and astigmatism effects, while HAADF-STEM point spread function is also influenced by spherical abberation. Variance-based, autocorrelation-based and Fourier transform-based iterative autofocus techniques were implemented, tested and compared for electron tomography [3], but astigmatism aberration was not considered. Defocus and two-fold astigmatism in HAADF-STEM, as well as higher order aberrations, can be corrected using Ronchigrams [12]. However, this method is not applicable for correction of defocus and astigmatism for an arbitrary non-amorphous sample. Non-iterative methods based on the image's Fourier transform, where defocus and astigmatism are estimated from only two HAADF-STEM images have been proposed [14, 1]. However, they as well as the method described in [9], can hardly be used for situations, when the image Fourier Transform has a limited frequency range.

For HAADF-STEM applications this paper introduces a new method for simultaneous autofocus and two-fold astigmatism correction. The method makes use of an image variance, which has already been used before as an image quality measure for different types of microscopy [6, 11, 8, 3]. We modify it in order to make it more robust against machine or sample instabilities. Further, we describe numerical simulations based on a classical HAADF-STEM linear image formation model [7, 5] extended with an extra astigmatism parameter. Numerical simulations show that the modified variance reaches it's maximum for Scherzer focus and zero astigmatism for amorphous sample and ellipsoid particles for noise-free and noisy images. Also, the behavior of a variance function further away from the ideal microscope parameters (like, local minima and maxima in one-parameter space) coincides with the analytical observations made before for scanning electron microscopy [6]. From the observations about the variance maximum position and its behavior we draw the conclusions that it is important to look for the maximum in three-parameter space at the same time. For this purpose we employ the well-known Nelder-Mead Simplex optimization algorithm. Based thereon, we show numerically and experimentally that the algorithm converges to the optimal image of a general sample in typically 20-50 function evaluations (depending on a first given image and the algorithm's input parameters), where a function evaluation means obtaining a single image with the HAADF-STEM.

Section 2 of this paper gives an overview of the classical linear model of image formation in HAADF-STEM. This model is extended in a way that the problem of simultaneous autofocus and astigmatism correction becomes a three parameter problem (one parameter is defocus and the two other parameters are x-astigmatism and y-astigmatism), which corresponds to the real-world situation. Section 3 describes the image variance and its modification. Section 4 describes the numerical experiment based on the linear image formation model from Section 2 and realistic physical values. We show that the modified variance reaches it's optimum for the simulated images of amorphous sample and elliptic particles for ideal HAADF-STEM parameters (Scherzer focus and zero astigmatism). Section 5 gives a brief overview of the Nelder-Mead simplex method for the three-parameter optimization. Section 6 describes the implemented automated application. The chosen experimental samples are difficult for analyzing with Fourier-based methods. The variances of the sample's images are optimized with the

Nelder-Mead simplex method. It successfully finds the optimal defocus and zero astigmatism with the time and accuracy, compared with the human operator.

2 HAADF-STEM linear image formation model

A detailed explanation of the HAADF-STEM image formation can be found in [7]. Here we provide only a short overview. Let $f(\mathbf{x}, \mathbf{p})$ be a HAADF-STEM image, where $\mathbf{x} = (x, y) \in \mathbf{X}$ are the spatial coordinates and

$$\mathbf{p}^T = (d, C_b, C_c) \tag{1}$$

is a vector of defocus and two-fold astigmatism HAADF-STEM parameters. We consider a linear image formation model

$$f(\mathbf{x}, \mathbf{p}) = \psi_0(\mathbf{x}) * h(\mathbf{x}, \mathbf{p}) + \epsilon(\mathbf{x}) \tag{2}$$

where $*$ denotes convolution, $\psi_0(\mathbf{x})$ is the object function, $\epsilon(\mathbf{x})$ is noise, and $h(\mathbf{x}, \mathbf{p})$ is the intensity of the probe function, such that

$$\iint_{\mathbf{X}} h(\mathbf{x}, \mathbf{p}) d\mathbf{x} = 1. \tag{3}$$

In the Fourier space (2) becomes

$$F(\mathbf{u}, \mathbf{p}) = \Psi_0(\mathbf{u}) H(\mathbf{u}, \mathbf{p}) + \varepsilon(\mathbf{u}), \tag{4}$$

where $\mathbf{u} = (u, v)$ are the frequency coordinates and $F, \Psi_0, H, \varepsilon$ are the Fourier transforms of f, ψ_0, h, ϵ.

The wave function that enters the specimen is given in frequency space by assuming a fully coherent point source of electrons in the far field

$$G(\mathbf{u}, \mathbf{p}) = A(\mathbf{u}) e^{-i\chi(\mathbf{u}, \mathbf{p})}. \tag{5}$$

Here the aperture function $A(\mathbf{u})$ is

$$A(\mathbf{u}) = \begin{cases} 1, \text{ if } |\mathbf{u}| \leq q_0 \\ 0, \text{ elswise,} \end{cases} \tag{6}$$

and the wave aberration function $\chi(\mathbf{u}, \mathbf{p})$ is defined as in [7]

$$\chi(\mathbf{u}, \mathbf{p}) = \pi\lambda|\mathbf{u}|^2 (d + \frac{1}{2}\lambda^2 |\mathbf{u}|^2 C_s + C_a \cos(2(\phi - \phi_a))), \text{ where } \phi = \arctan(\frac{v}{u}), \tag{7}$$

$\lambda, d, C_s, C_a, \phi_a$ represent the wavelength, the defocus, the spherical aberration, the two-fold astigmatism amplitude and the two-fold astigmatism rotation angle respectively. The aperture radius q_0 in (6) controls the convergence semi angle α_0 of the beam by

$$q_0 = \frac{\alpha_0}{\lambda}. \tag{8}$$

The point spread function is the intensity of the scanning probe, that is the inverse Fourier transform of the wave function (5) [7]:

$$h(\mathbf{x}, \mathbf{p}) = C \left| \mathfrak{F}^{-1}[G] \right|^2, \tag{9}$$

where C is a normalization constant, such that (9) satisfies (3).

In (7) the problem of automated defocus and two-fold astigmatism adjustment is a two-dimensional problem. We are adjusting two variables d and C_a in order to get the best image quality. The other two-fold astigmatism variable in (7) ϕ_a represents the point spread function rotation. It does not influence the image quality: When C_a is equal to 0, the ϕ_a value are irrelevant. However, in real HAADF-STEM ϕ_a can not be adjusted. It is a constant value $\phi = \phi_{a_0}$ that depends on machine calibration, magnification and other parameters. The two-fold astigmatism in HAADF-STEM is adjusted by changing the values of the x-stigmator and the y-stigmator, based on the quadrupole lenses [13]. We modify (7) according to the real world HAADF-STEM situation

$$\chi(\mathbf{u},\mathbf{p}) = \pi\lambda|\mathbf{u}|^2(d + \frac{1}{2}\lambda^2|\mathbf{u}|^2 C_s + C_b\cos(2(\phi-\phi_{a_0})) + C_c\sin(2(\phi-\phi_{a_0}))). \tag{10}$$

The equivalence of (7) and (10) is shown in Appendix.

The Scherzer conditions for incoherent image formation are defined in [7] as

$$q_0 := \frac{1}{\lambda}(\frac{6\lambda}{C_s})^{\frac{1}{4}}, \tag{11}$$

$$d_0 := -(1.5C_s\lambda)^{1/2}. \tag{12}$$

Then the ideal machine parameters are

$$\mathbf{p}^T = \mathbf{p}_0^T := (d_0, C_{b_0}, C_{c_0}) = (d_0, 0, 0). \tag{13}$$

3 Variance as an image quality measure

We define the mean value of the image $f(\mathbf{x},\mathbf{p})$ as

$$\bar{f}(\mathbf{p}) := \frac{\iint_{\mathbf{X}} f(\mathbf{x},\mathbf{p})d\mathbf{x}}{\iint_{\mathbf{X}} d\mathbf{x}}. \tag{14}$$

The image variance is defined as

$$\tilde{s}[f(\mathbf{x},\mathbf{p})] := \iint_{\mathbf{X}} \left|f(\mathbf{x},\mathbf{p}) - \bar{f}(\mathbf{p})\right|^2 d\mathbf{x}. \tag{15}$$

The image variance has been used before as an image quality measure for autofocus for different types of microscopy [11, 8, 3]. The use of variance is heuristic in nature. It is based on the assumption that an in-focus image has higher contrast then a defocused one. For fluorescent light microscopy it has been shown experimentally [11] that out of a large number of image quality measures, image variance was the better.

It has been proved that the variance of an image reaches its maximum for the ideal defocus and astigmatism parameter in scanning electron microscopy under the following assumptions [6]: 1) Noise-free image formation, i.e. $\epsilon(\mathbf{x}) = 0$ in (2); 2) a normalized image, such that $\bar{f}(\mathbf{p}) = 0$; 3) $|\Psi_0(\mathbf{u})|^2 = \text{const}$; 4) a gaussian shape of the point spread function. It was shown that under the listed assumptions the variance (15) is inversely proportional to the Gaussian widths a and b

$$\tilde{s}[f(\mathbf{x},\mathbf{p})] = \frac{|\Psi_0|^2}{4\pi ab}, \tag{16}$$

therefore a minimum beam width corresponds to the maximal variance. The above assumptions can not be accurately applicable to the HAADF-STEM images: real world images are always noisy, the power spectrum of the object function is never a constant, and the Gaussian function is inaccurate description of a HAADF-STEM point spread function, for instance, due to the presence of spherical abberation. In the next section we will show numerically that the variance reaches its maximum in three-parameter space (1) for ideal parameters (13) for simulated HAADF-STEM images.

Due to the linear image formation model (2)-(3), the total image intensity of a particular sample is expected to be constant independent on machine parameters $\mathbf{p}$. However, due to instabilities, such as sample charge due to the collected electrons, this is not always true, i.e.

$$\bar{f}(\mathbf{p}) \neq \text{const} \quad \forall \mathbf{p}. \tag{17}$$

In order to avoid the instabilities to influence the variance as an image quality measure, we introduce the modified image variance with the normalized image intensity:

$$s[f(\mathbf{x},\mathbf{p})] := \frac{1}{(\bar{f}_{\mathbf{p}})^2} \iint_{\mathbf{X}} \left| f(\mathbf{x},\mathbf{p}) - \bar{f}(\mathbf{p}) \right|^2 d\mathbf{x}. \tag{18}$$

For the remainder of this paper we refer to (20) as variance. In HAADF-STEM we deal with discrete images that can be represented by matrices

$$\mathbf{F}(\mathbf{p}) := ((f_{i,j}(\mathbf{p}))_{i=1}^{N})_{j=1}^{M}. \tag{19}$$

After the variance discretization, which is explained in Appendix, we obtain:

$$s[\mathbf{F}] = \frac{1}{\bar{f}^2 NM} \sum_{i,j} (f_{i,j} - \bar{f})^2. \tag{20}$$

The definition (20) coincides with [11, 3] with the coefficient $\frac{1}{\bar{f}^2}$ in front as the only difference, which is our contribution.

4 Numerical experiments

4.1 Description

The goal of these numerical experiments is to show that for a given object function the image variance (20) reaches it's maximum in a three parameter space for Scherzer focus and zero astigmatism (13). Figure 1 shows the modeled object functions of the amorphous sample and ellipsoid particles samples and their Fourier transforms. We model the amorphous object function as a random variable drawn from a Poisson distribution with expected value γ. Thus, the probability, that the object pixel $f_{i,j} = k$ is

$$p(k,\gamma) := \frac{\gamma^k e^{-\gamma}}{k!}. \tag{21}$$

We set $\gamma = 10$, in order to obtain on average 10 counts per pixel, which is usually the case in the real world situation. The amorphous object function Fourier transform has a circle shape, while the ellipsoid particles object function Fourier transform has elliptic shape. This shape

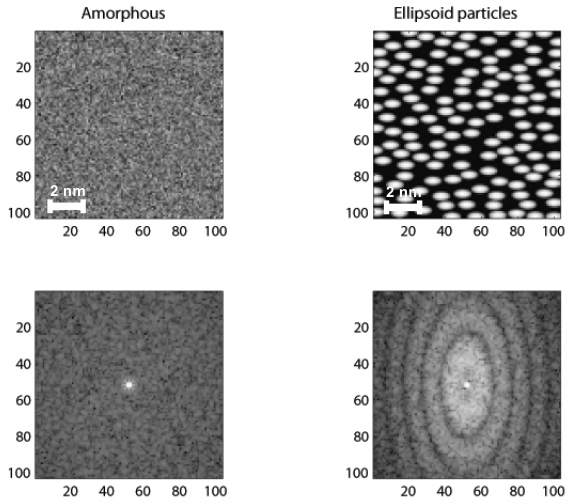

Figure 1: On top: Object functions modeled as amorphous sample and ellipsoid particles samples. On bottom: Their Fourier transforms.

Table 1: Parameters used for numerical calculations.

C_s	E	λ	q_0	d_0	α_0	Δd	ΔC_b	ΔC_c
1.07 mm	300 keV	1.9×10^{-2} nm	5.3 nm	-55.2 nm	10.2 mrad	9.2 nm	9.2 nm	9.2 nm

is a result of the oval particles that are oriented along the x-axis. This type of sample could cause difficulties for the Fourier transform-based correction algorithms, because in general we expect elliptic shape of Fourier transform in the case of astigmatism. The size of both images is 100×100 pixels. We assume the pixel width and height for both samples

$$\delta_1 = 0.1 \text{ nm}. \tag{22}$$

We use the physical model described in Section 2 in order to get a simulated HAADF-STEM experimental image. The convolution is carried out in Fourier space, see (4). The noise $\epsilon(\mathbf{x})$ in (2) is simulated as a poisson distribution (21).

Table 1 describes the parameter values used in the computations. Realistic values for the spherical aberration C_s and the electron voltage E are taken. The wavelength λ is computed through the electron voltage E [5]. Scherzer values α_0, q_0, d_0 are computed from (8),(11) and (12). The tolerable defocus error is defined as [3]

$$d_e = \sqrt{(\frac{w}{2})^2 + (\frac{t}{2})^2},$$

where t is the specimen thickness and w is the depth of field defined in [2] as

$$w := \frac{\lambda}{\alpha_0^2}.$$

Table 2: Parameter range used for numerical calculations.

N_d	N_{C_a}	N_{C_b}	d_{min}	d_{max}	$C_{b_{min}}$	$C_{b_{max}}$	$C_{c_{min}}$	$C_{c_{max}}$
40 steps	20 steps	20 steps	-239.2 nm	+128.8 nm	-92 nm	+92 nm	-92 nm	92 nm

We take the defocus step Δd as the lower bound set by the depth of field

$$\Delta d = \frac{\lambda}{4\alpha_0^2}. \tag{23}$$

Parameters C_b and C_c in (10) indicate the distance between the two focal points in the case of astigmatism. Thus, we can set the astigmatism tolerance equal to the defocus tolerance:

$$\Delta C_b = \Delta C_c = \Delta d. \tag{24}$$

We vary the parameters d, C_b and C_c in order to show numerically that the variance reaches its maxima for ideal parameters $\mathbf{p}_0$. The numerical computations are performed for N_d defocus d steps of around the Scherzer focus d_0, N_{C_b} steps of C_b around C_{b_0} and N_{C_c} steps of C_c around C_{c_0}, i.e. the parameter domain Ω is

$$\Omega = [d_{min}, d_{max}] \times [C_{b_{min}}, C_{b_{max}}] \times [C_{c_{min}}, C_{c_{max}}], \text{ where}$$

$$d_{min} := d_0 - \frac{N_d}{2}\Delta d, \quad d_{max} := d_0 + \frac{N_d}{2}\Delta d,$$

$$C_{b_{min}} := C_{b_0} - \frac{N_{C_b}}{2}\Delta C_b, \quad C_{b_{max}} := C_{b_0} + \frac{N_{C_b}}{2}\Delta C_b,$$

$$C_{c_{min}} := C_{c_0} - \frac{N_{C_c}}{2}\Delta C_c, \quad C_{c_{max}} := C_{c_0} + \frac{N_{C_c}}{2}\Delta C_c.$$

The actual used values are shown in Table 2.

4.2 Results of the numerical experiment

In total four numerical experiments, based on amorphous sample without noise, amorphous sample with noise, ellipsoid particles without noise, ellipsoid particles with noise, were carried out. We have shown numerically that the variance function reaches it's maximum in 3-parameter space (1) for ideal microscope parameters (13) in all the four cases.

Figure 2 shows the results of numerical computations for the pixel width (22) for simulated ellipsoid particles sample without noise and with high level of noise (signal-to-noise ratio is equal to 3). The figure shows the variance functions in the space of parameters d and C_b for a fixed $C_c = C_{c_0}$. We can clearly see that when noise is added to the images, the variance function becomes noisy as well: It has a lot of local optima, especially further away from the main peak. A similar shape of the variance function was obtained for the amorphous sample.

It is important to note that the variance has a local minimum at Scherzer focus for constant astigmatism value of $C_b \gg C_{b_0}$. We can see this phenomenon in Figure 3. For $C_b = 0$ nm, variance has its maximum at Scherzer focus. However, for $C_b = -92$ nm variance has a local minimum close to Scherzer focus and two maxima on the left and the right. This local minimum is not a minimum any more, when we look at the problem in two-parameter space, like in Figure 2, or three-parameter space. This also corresponds to the analytical observations made before in [6]. The positions of the two local maxima are not exactly symmetric. Their nature has to deal with the nature of the object function geometry. Similar observations were made by examining the shapes of the variance in two dimensions far away from the ideal parameters: the function has local optima, which are not optima any more in three-parameter space. From the facts that the variance reaches its maximum at Scherzer focus and zero astigmatism, and

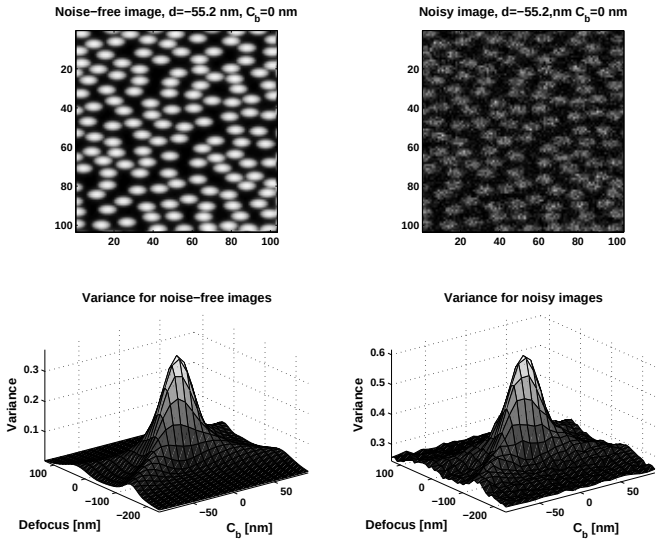

Figure 2: From left to right, from top to bottom: The simulated image of ellipsoid particles sample for Scherzer focus and zero astigmatism, the same image influenced by noise, variance of simulated images of ellipsoid particles sample without noise, variance of simulated images of ellipsoid particles sample with noise.

from the fact of existence of the local maxima with unpredictable behavior in lower dimensions space we draw the conclusion that it is important to optimize over the three parameters (1) at the same time.

It is also important for practical reasons to optimize in a reasonable amount of function evaluations, e.g. the sample can be destroyed by the electron beam. Also, obtaining the necessary number of images is the main time limiting factor for the algorithm. It is important that we do not have any analytical derivative information is available. We could compute the derivatives from finite differences, but this would dramatically increase the amount of function evaluations. Thus, we will use a derivative free optimization method described in the next section.

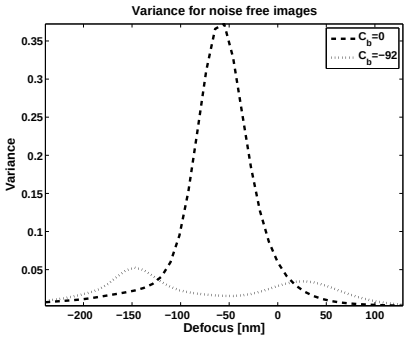

Figure 3: Ellipsoid particles sample variance for different astigmatism values. For $C_b = 0$ nm, variance has its maximum at Scherzer focus. However, for $C_b = -92$ nm variance has a local minimum for Scherzer focus and two maxima on the left and the right.

Table 3: HAADF-STEM parameters during the experiment.

	Carbon cross grating specimen	Gold-particle
Electron voltage	200 keV	200 keV
Pixel size	[5 ... 20] nm	[0.5 ... 2] nm
Dwell time	[1 ... 71.7] μs	[5 ... 50] μs
Camera length	200 mm	200 mm
Spot size	7	7
Image size	512x512 px	512x512 px
Image pixel depth	16 bit	16 bit

5 The Nelder-Mead simplex method

The Nelder-Mead simplex method is designed to find a local optimum of a function. It makes no assumptions about the shape of the function and does not use derivative information. A detailed description of the Nelder-Mead simplex method for n dimensions can be found in [10, 4]. In this paper we just give a short overview.

The simplex is a geometrical object in n dimensions that consists of $n+1$ points. In the case of defocus and two-fold astigmatism correction $n = 3$, and the simplex is a tetrahedron. The Nelder-Mead simplex method must be started with $n + 1$ points, defining the initial simplex. Let the initial value of the HAADF-STEM defocus and astigmatism parameters be (d^0, C_b^0, C_c^0). Then the other three points of the initial simplex are $(d^0 + \Delta d^0, C_b^0, C_c^0)$, $(d^0, C_b^0 + \Delta C_b^0, C_c^0)$, $(d^0, C_b^0, C_c^0 + \Delta C_c^0)$, where

$$\Delta \mathbf{p}^T := (\Delta d^0, \Delta C_b^0, \Delta C_c^0) \tag{25}$$

is the input parameter for the optimization.

On every iteration the Nelder-Mead simplex method evaluates the function for a finite number of points and replaces the point corresponding to the lowest variance in the simplex, with a new point, corresponding to a higher variance. The algorithm stops, according to the input tolerance parameters

$$s_{tol}, p_{tol} \in \mathbb{R}, \text{ when} \tag{26}$$
$$\|s(\mathbf{p}^{(1)}) - s(\mathbf{p}^{(2)})\| < s_{tol}, \text{ and } \|\mathbf{p}^{(1)} - \mathbf{p}^{(2)}\| < p_{tol},$$

where $\mathbf{p}^{(1)}$ and $\mathbf{p}^{(2)}$ are the two points of the simplex, corresponding to the highest variance. We have to keep in mind that the defocus and stigmator values in HAADF-STEM have different scalings. Thus, the parameter tolerance should be separated for each of the parameters and extended to $\mathbf{p}_{tol} \in \mathbb{R}^3$.

The Nelder-Mead simplex method was successfully applied to the simulated images, described in the previous section. The parameters $\mathbf{p}_{tol} \in \mathbb{R}^3$ and $\Delta\mathbf{p}$ were choosen according to the defocus step (23) and the stigmatic steps (24). In the next section we describe the application of the Nelder-Mead simplex method to the variance of real images.

6 Application

The physical experiment was performed on a FEI Tecnai F20 STEM electron microscope ($C_s = 1.3mm$). Our Java-based experimental platform (called EXPLA) consists of a core that connects to the TEMScripting interface for FEI microscope control, and an application

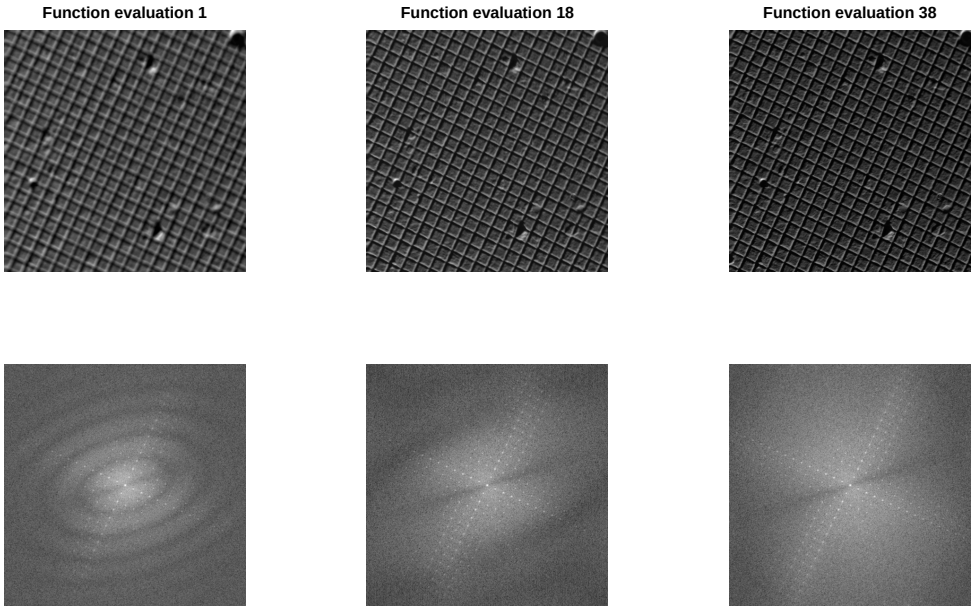

Figure 4: Experimental carbon cross grating images and the absolute values of their Fourier transforms before, during and after correction.

Table 4: Experimental values for carbon cross grating images.

Function evaluation number	1	10	18	38
Iteration number	1	5	9	20
Variance	-0.0131	-0.0412	-0.0410	-0.0465
Defocus [nm]	2190	-32.2222	-133.0453	138.461
X-stigmator [a.u.]	-0.00308	0.00047556	0.00045909	-0.010065
Y-Stigmator [a.u.]	-0.00499	0.0021211	-0.00087477	0.0039236

control framework. EXPLA also allows efficient and flexible computation and control from a Matlab environment. For our experiment the variance and the Nelder-Mead Simplex method was implemented in Matlab V7.5 (R2007b). The Nelder-Mead Simplex implementation is based on the code available in the Matlab Optimization Toolbox V3.1.2.

Further we describe experimental results for two samples: carbon cross grating sample and gold-particle sample.

6.1 Carbon cross grating specimen

The first experiment was performed on the carbon cross grating sample. The first column of Table 3 shows the HAADF-STEM parameters during the experiment. Figure 4 shows images of this sample on different optimization iterations and the logarithmically scaled absolute values of their Fourier transforms. The image's Fourier transforms are influenced by the sample's geometry, which could cause difficulty for non-iterative autofocus and 2-fold astigmatism correction procedures based on the image Fourier transform analysis. We can see, that the frequency values in the Fourier transform on 38th evaluation have increased, comparing to 1st and 18th

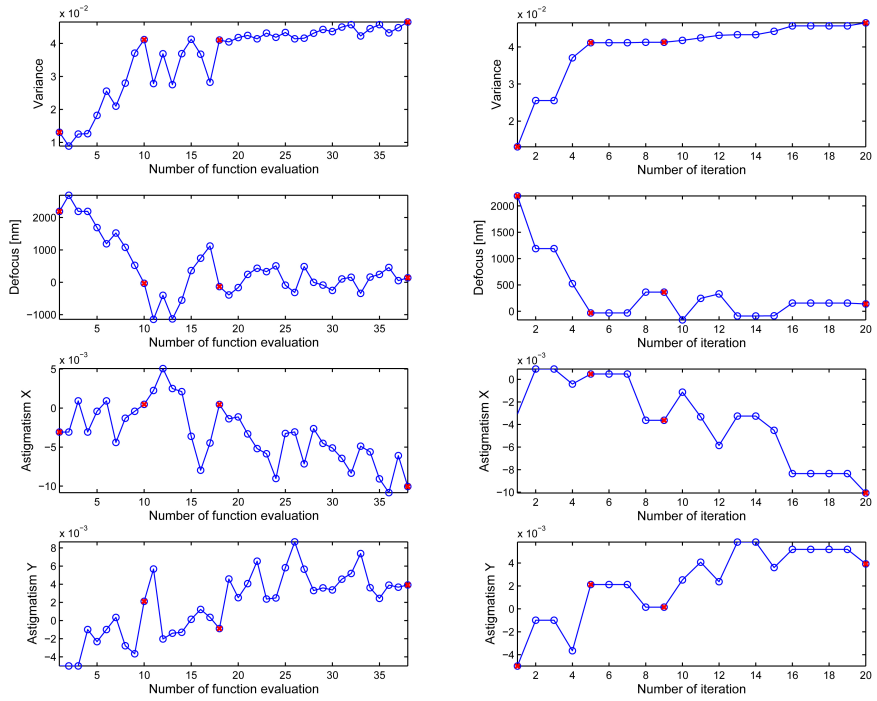

Figure 5: Modified variance, defocus, x-stigmator and y-stigmator values plotted versus function evaluation and optimization iteration numbers.

evaluations, which also indicates the progress of optimization.

Further an example of one optimization is given in details. The initial microscope values are shown in Table 4. The initial step to create the initial simplex $\Delta\mathbf{p}$ (25) was chosen 500 nm for defocus and 0.004 for both x-stigmator and y-stigmator. The total optimization procedure took 20 iterations with 38 function evaluations in total (from 1 to 4 function evaluations for each iteration). Every function evaluation corresponds to one image recording.

Figure 4 shows the three images that correspond to the 1st, 18th and 38th function evaluation. The image corresponding to the 18th function evaluation shows sufficient quality improvement compared to the 1st function evaluation. Also, we do not observe a large difference between the 18th and 38th function evaluation. Only if we zoom in and observe the details (Figure 6), we will see a difference in contrast and quality.

Figure 5 plots the variance, defocus, x-stigmator and y-stigmator values versus function evaluation and iteration number. The 1st, 10th and 18th function evaluations are indicated with a cross in the plots of Figure 5. The experimental values for these evaluations are shown in Table 4. For every iteration the highest variance found thus far is chosen to plot. Although, the maximum is achieved at 38th function evaluation, we can see from the plots in Figure 5 that the variance is stable and close to maximum already from the 18th function evaluation in the 9th iteration. Also, the first sufficient quality image was obtained already at the 10th function evaluation (5th iteration).

The sum processor time required for all 38 images for variance computations and estimation of the next step of the Nelder-Mead is less then one second on a standard PC 2.2 GHz. Thus, the main time limitation factor of the algorithm is the time necessary for image recording and

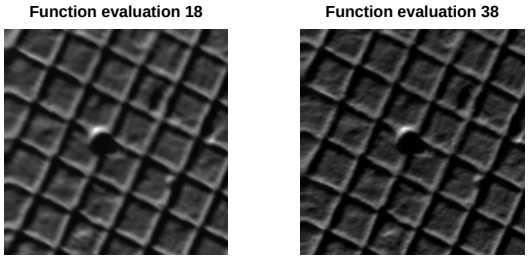

Figure 6: Only fine details of zoomed-in images show the quality difference between the 18th and the 38th function evaluations.

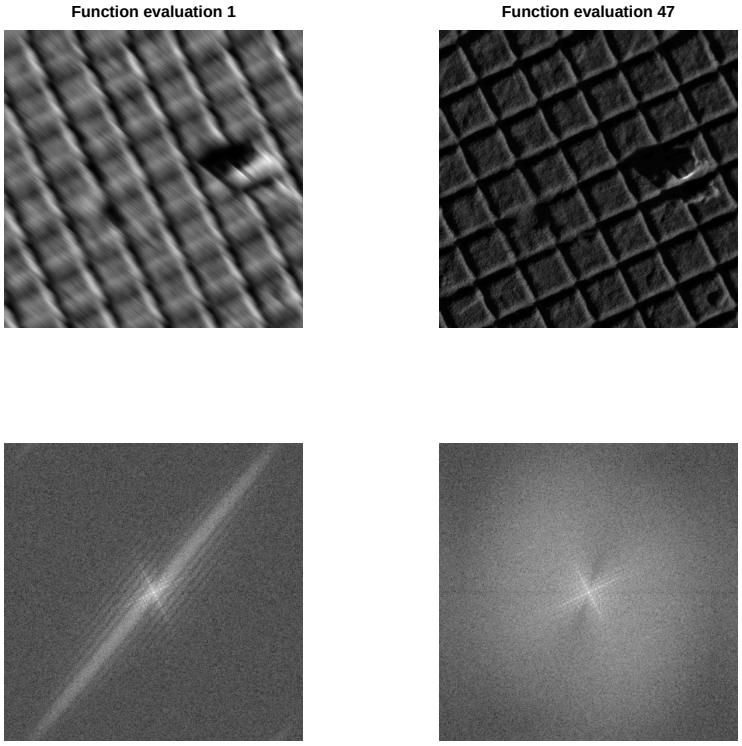

Figure 7: Experimental carbon cross grating HAADF-STEM images and the absolute values of their Fourier transforms before and after correction. The first function evaluation image has a high level of astigmatism.

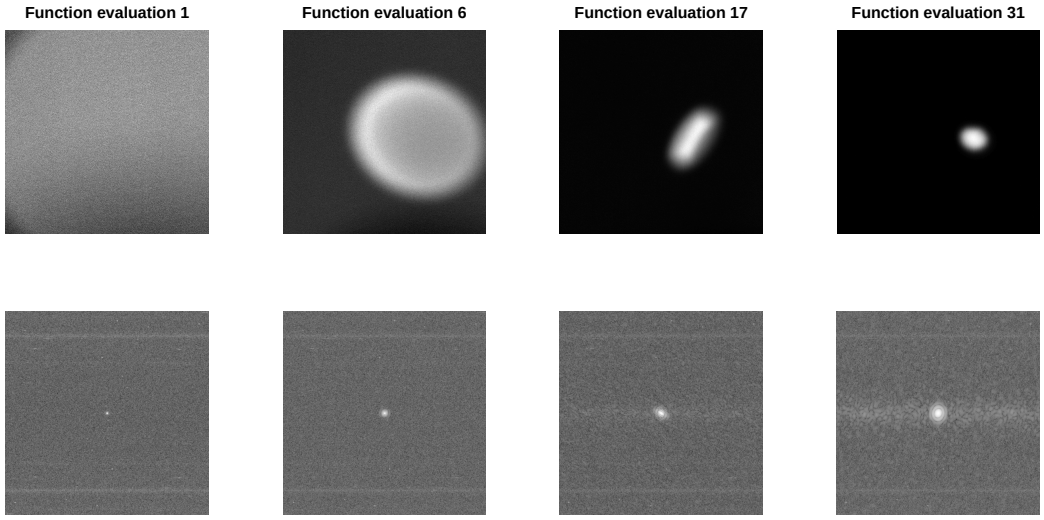

Figure 8: Experimental gold particle HAADF-STEM images and the absolute values of their Fourier transforms before, during and after correction.

image transfer to the remote computer. The total time for all 38 function evaluation is 30 seconds. This time could be decreased down to approximately 15 seconds if the algorithm will be implemented directly on the microscope. We have to take into account that the image recording time depends on the microscope dwell time and the image dimensions in pixels. By changing these parameters, the total time could be decreased even further.

A few more experiments for different magnifications, dwell time, initial microscopic parameters and initial optimization parameters were performed on the carbon cross gratingsample. In each of the cases the algorithm converged in 30-50 function evaluations. However, images of sufficient quality were usually obtained earlier, in 20-30 function evaluations. In order to decrease the amount of function evaluations the accuracy of the algorithm could be adjusted with the tolerance parameters (26). High levels of astigmatism in the beginning require more iterations. Figure 7 shows the first image from a run with high initial astigmatism, which is clearly seen from the stretched Fourier transform. It took 47 function evaluations until the image was completely corrected.

6.2 Gold-particle sample

The algorithm was tested on another sample as well: A gold particle with a radius of about 20 nanometers. The second column of Table 3 shows the parameters during the experiment. Figure 8 shows the images of four Nelder-Mead function evaluations from one of optimization runs. The Fourier transforms of the images have a low amount of frequencies. This type of sample would be difficult to analyze with a Fourier transform-based technique, because it is difficult to gain information from such a Fourier transform. The algorithm converged in 31 function evaluations. The image quality on the last function evaluation is typically equal to the image quality obtained by a human operator.

The dwell time is equal to 5 μs. We can clearly see the noise in the images, especially during the first evaluation. The signal-to-noise ration here is estimated to be 10.

7 Discussion

A new simultaneous defocus and two-fold astigmatism correction method for HAADF-STEM is proposed. Though the method is automated, it requires input parameters, such as the tolerance, used as a stopping criterion and the size of the initial simplex. These parameters can be the same for similar samples and microscope settings. It means that the method can already be used for automated applications from different fields, where the input parameters can be usually easily predicted. But in order to automate the method completely the choice of the input parameters has to be as well automated.

The working time of the method is only limited by the time required for acquiring of a single image. The computations of variance and the Nelder-Mead do not cost time. Because the method is iterative, and requires about 30-50 images, in general, it will take more time then the alternative Fourier transform-based methods proposed before. However, it has one important advantage: it works for the cases, when the images of a sample do not have a sufficient amount of frequencies to be analyzed by the Fourier transform-based methods.

At the moment the proposed method takes time comparable to the time required by a human operator. Further, the working time of the method can be decreased by decreasing the input image size, subsampling the input image, decreasing the dwell time or choosing for a faster optimization derivative-free technique, alternative to the Nelder-Mead method. However, we have to take into account that decreasing of the dwell time will increase the level of noise in the image.

The real-world experiments have shown that the method is robust to a reasonable amount of noise (signal-to-noise ratio equal 10). The numerical computations have shown that the variance function can have local optima further away from the ideal microscope parameters due to the lower levels of noise (signal-to-noise ratio equal 2). This means that if the initial microscopic defocus and astigmatism parameters are far away from ideal, and the size of the initial simplex is not chosen properly (it is too small compared to the broadness of a variance peak), and in the same time the images are too noisy, the method may fail. In general case, when the signal-to-noise ratio is too low, one can use the noise reduction techniques, such as median filtering, pixel averaging or others.

Our numerical computations and real-world experiments have shown that the method is working well for the use on HAADF-STEM. Furthermore, it is based on assumptions, which are general enough that it could be used for other types of microscopy (such as scanning electron microscopy, etc.).

8 Conclusion

A new method for automated defocus and astigmatism correction in HAADF-STEM was discussed. The method uses a modification of image variance, which was already used before as an image quality measure for different types of microscopy. Numerical computations based on a HAAF-STEM classical image formation model extended with an extra astigmatism parameter were performed. The numerical computations have shown that the variance reaches it's maximum in a three-parameter space (defocus and two-fold astigmatism) for Scherzer focus and zero astigmatism for amorphous sample and ellipsoid particles sample. Also, the computations have shown that the variance has local optima in one-parameter and two-parameter spaces, which are not the optima any more in three-parameter space. From the observations about position

of the variance maximum and the presence of local optima in lower then three dimensions, we have drawn the conclusion that it is important to optimize variance in the three-parameter space at the same time. For this purpose we are using the well-known Nelder-Mead simplex optimization algorithm is used. It proved to work both for simulated images and in a real-world application. The proposed defocus and astigmatism correction method also works for the cases, when alternative Fourier transform-based methods could fail. The method can be used for the real-world HAADF-STEM applications already now. However, it has been discussed that the method's performance and robustness to noise could be still improved.

Appendix

Modified abberation function

We want to show that the aberration function (7) and the modified aberration function (10) are equivalent. We will use the trigonometric sum formulas

$$\cos(x+y) = \cos x \cos y - \sin x \sin y, \tag{27}$$

$$\sin(x+y) = \sin x \cos y + \cos x \sin y. \tag{28}$$

We will examine only the parts of the functions (7) and (10) that are different, i.e. the last terms that deal with two-fold astigmatism. We introduce two new functions

$$g_1(\phi) = C_a \cos(2(\phi - \phi_a)), \tag{29}$$

$$g_2(\phi) = C_b \cos(2(\phi - \phi_{a_0})) + C_c \sin(2(\phi - \phi_{a_0})). \tag{30}$$

Equations (29) and (30) have to be equal for every ϕ. Then the functions $g_1(\phi)$ and $g_2(\phi)$ are equivalent if and only if

1. $\forall C_a, \phi_a, \phi_{a_0} \in \mathbb{R} \quad \exists C_b, C_c$, such that $g_2(\phi) \equiv g_1(\phi)$,
2. $\forall C_b, C_c, \phi_{a_0} \in \mathbb{R} \quad \exists C_a, \phi_a$, such that $g_1(\phi) \equiv g_2(\phi)$.

We rewrite (29) and (30) using the trigonometric sum (27-28)

$$g_1(\phi) = (C_a \cos 2\phi_a) \cos 2\phi + (C_a \sin 2\phi_a) \sin 2\phi, \tag{31}$$

$$g_2(\phi) = (C_b \cos 2\phi_{a_0} - C_c \sin 2\phi_{a_0}) \cos 2\phi + (C_b \sin 2\phi_{a_0} + C_c \cos 2\phi_{a_0}) \sin 2\phi. \tag{32}$$

The functions (31) and (32) are equivalent if and only if

$$\begin{cases} C_a \cos 2\phi_a = C_b \cos 2\phi_{a_0} - C_c \sin 2\phi_{a_0} \\ C_a \sin 2\phi_a = C_b \sin 2\phi_{a_0} + C_c \cos 2\phi_{a_0} \end{cases}, \tag{33}$$

If we sum the squares of the both equations (33), we get

$$C_a^2 = (C_b \cos 2\phi_{a_0} - C_c \sin 2\phi_{a_0})^2 + (C_b \sin 2\phi_{a_0} + C_c \cos 2\phi_{a_0})^2, \text{ or} \tag{34}$$

$$C_a^2 = C_b^2 + C_c^2. \tag{35}$$

The angle ϕ_a can then be found according to the second equation in (33)

$$\phi_a = \begin{cases} \operatorname{sign}(C_a)\arcsin(C_b \sin 2\phi_{a_0} + C_c \cos 2\phi_{a_0})/C_a, & \text{if } C_a \neq 0 \\ \mathbb{R}, & \text{if } C_a = 0 \end{cases}. \tag{36}$$

If we multiply the first equation in (33) with $\cos y$, the second with $\sin y$, sum the results, and apply further the trigonometric sum (27), we achieve

$$C_b = C_a \cos(\phi_a - \phi_{a_0}). \tag{37}$$

If we multiply the first equation in (33) with $\sin y$, the second equation in (33) we multiply with $\cos y$, substitute the results, and apply further the trigonometric sum 28, we achieve

$$C_c = C_a \sin(\phi_a - \phi_{a_0}). \tag{38}$$

We have shown, that for all C_a, ϕ_a, ϕ_{a_0}, such C_b, C_c can be found with the formulas (37), (38), that $g_2(\phi) \equiv g_1(\phi)$. Also, for all C_b, C_c, ϕ_{a_0}, such C_a, ϕ_a can be found with the formulas (35), (36), that $g_2(\phi) \equiv g_1(\phi)$. It means that the functions (29) and (30) are equivalent, and as a consequence the aberration function (7) and the modified aberration function (10) are equivalent.

Variance discretization

Image domain is given as

$$\mathbf{X} = [x_{min}, x_{max}] \times [y_{min}, y_{max}]$$

We define the mean value of the image $f_{\mathbf{p}}(\mathbf{x})$ as

$$\bar{f}(\mathbf{p}) := \frac{\iint_{\mathbf{X}} f(\mathbf{x}, \mathbf{p}) d\mathbf{x}}{\iint_{\mathbf{X}} d\mathbf{x}}. \tag{39}$$

We introduce the modified image variance with the normalized image intensity

$$s[f(\mathbf{x}, \mathbf{p})] := \frac{1}{(\bar{f}_{\mathbf{p}})^2} \iint_{\mathbf{X}} \left| f(\mathbf{x}, \mathbf{p}) - \bar{f}(\mathbf{p}) \right|^2 d\mathbf{x}. \tag{40}$$

We consider pixel discretization of the given domain

$$x_0 = x_{min} < x_1 < \ldots < x_n = x_{max}, \quad \delta_x := x_i - x_{i-1} = \frac{x_{max} - x_{min}}{N}, \tag{41}$$

$$y_0 = y_{min} < y_1 < \ldots < y_m = y_{max}, \quad \delta_y := y_j - y_{j-1} = \frac{y_{max} - y_{min}}{M}, \text{ and} \tag{42}$$

In (41) and (42) δ_x and δ_y are the sizes of a single pixel in the image, or the pixel widths. In the real world applications the pixel widths are usually equal

$$\delta = \delta_x = \delta_y.$$

We define

$$f_{i,j}(\mathbf{p}) := f((x_i, y_j), \mathbf{p}), \quad i \in \{0, \ldots, N\}, \quad j \in \{0, \ldots, M\}. \tag{43}$$

In HAADF-STEM we deal with discrete images that can be represented by matrices

$$\mathbf{F}(\mathbf{p}) := ((f_{i,j}(\mathbf{p}))_{i=1}^{N})_{j=1}^{M}. \tag{44}$$

We approximate (39) with the Riemann sum

$$\bar{f} \doteq \frac{\sum_{i,j} f_{i,j}\delta_x\delta_y}{\sum_{i,j}\delta_x\delta_y} = \frac{\delta_x\delta_y\sum_{i,j} f_{i,j}}{\delta_x\delta_y NM} \text{ or}$$

$$\bar{f} := \frac{\sum_{i,j} f_{i,j}}{NM}. \tag{45}$$

We approximate (40) with the Riemann sum

$$s[f(\mathbf{x},\mathbf{p})] \doteq \frac{1}{\bar{f}}\sum_{i,j}\delta_x\delta_y\left|f_{i,j}-\bar{f}\right|^2 = \frac{\delta_x\delta_y}{\bar{f}}\sum_{i,j}\left|f_{i,j}-\bar{f}\right|^2 =$$

$$\frac{(x_{max}-x_{min})(y_{max}-y_{min})}{\bar{f}NM}\sum_{i,j}\left|f_{i,j}-\bar{f}\right|^2. \tag{46}$$

If we set $(x_{max}-x_{min})(y_{max}-y_{min})=1$ in (46), we obtain (20)

$$s[\mathbf{F}] = \frac{1}{\bar{f}^2 NM}\sum_{i,j}(f_{i,j}-\bar{f})^2.$$

Acknowledgment

We kindly acknowledge Dries Langsweirdt (Technical University of Leuven) and Nol Venema (Technolution, The Netherlands) for the assistance with software development and the technical support. We are also thankful to S. Sluyterman and M. Otten (FEI, The Netherlands) for their useful comments and suggestions.

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.

References

[1] N. Baba, K. Terayama, T. Yoshimizu, N. Ichise, and N. Tanaka. An auto-tuning method for focusing and astigmatism correction in HAADF-STEM, based on the image contrast transfer function. *Ultramicroscopy*, 50:163–176, 2001.

[2] M. Born and E. Wolf. *Principles of optics : Electromagnetic theory of propagation, interference and diffraction of light*. Cambridge University Press, 7th expanded edition edition, 1999.

[3] W. Van den Broek. *Advanced Focus Methods in Electron Microscopy: Tomographic Reconstruction of the EELS Data Cube Autofocus of HAADF-STEM Images*. PhD thesis, University of Antwerp, 2007.

[4] A. R. Conn, K. Scheinberg, and L. N. Vicente. *Introduction to derivative-free optimization*. MPS-SIAM series on optimization, 2009.

[5] M. de Graef. *Introduction to Conventional Transmission Electron Microscopy*. Cambridge University Press, 2003.

[6] S.J. Erasmus and K.C.A. Smith. An automatic focusing and astigmatism correction system for the sem and ctem. *Jour. of Micr.*, 127(2):185–199, 1982.

[7] E. J. Kirkland. *Advanced Computing in Electron Microscopy*. Plenum Press, 1998.

[8] X.Y. Liu, W.H. Wang, and Y. Sun. Dynamic evaluation of autofocusing for automated microscopic analysis of blood smear and pap smear. *Jour. of Micr.*, 227:15–23, 2007.

[9] K.H. Ong, J.C.H. Phang, and J.T.L. Thong. A robust focusing and astigmatism correction method for the scanning electron microscope. *Scanning*, 19:553–563, june 1997.

[10] W.H. Press, S.A. Teukolsky, and W.T. Vetterling. *Numerical recipes : the art of scientific computing*. Cambridge University Press, 3rd edition edition.

[11] A. Santos, C. Ortiz De Solrzano, J. J. Vaquero, J. M. Pea, N. Malpica, and F. Del Pozo. Evaluation of autofocus functions in molecular cytogenetic analysis. *Jour. of Micr.*, 188:264–272, 1997.

[12] H. Sawada, T. F. Sannomiya, F. Hosokawa, T. Nakamichi, T. Kaneyama, T. Tomita, Y. Kondo, T. Tanaka, Y. Oshima, Tanishiro Y., and Takayanagi K. Measurement method of aberration from ronchigram by autocorrelation function. *Ultramicroscopy*, 108:1467–1475, 2008.

[13] S. Sluijterman. *Innovative use of Magnetic Quadrupoles in Cathode-Ray Tubes*. PhD thesis, Eindhoven University of Technology, 2002.

[14] N. Tanaka, J.J. Hu, and N. Baba. An on-line correction method of defocus and astigmatism in haadf-stem. *Ultramicroscopy*, 78:103–110, 1999.

PREVIOUS PUBLICATIONS IN THIS SERIES:

Number	Author(s)	Title	Month
10-05	J.C. van der Meer	Singularities of Poisson structures and Hamiltonian bifurcations	Jan. '10
10-06	V.N. Kornilov R. Rook J.H.M. ten Thije Boonkkamp L.P.H. de Goey	Experimental and numerical investigation of the acoustic response of multi-slit Bunsen burners	Jan. '10
10-07	J.H.M. ten Thije Boonkkamp M.J.H. Anthonissen	The finite volume-complete flux scheme for advection-diffusion-reaction equations	Jan. '10
10-08	A.Muntean O. Lakkis	Rate of convergence for a Galerkin scheme approximating a two-scale reaction-diffusion system with nonlinear transmission condition	Febr. '10
10-09	M.E. Rudnaya W. van den Broek R. Doornbos R.M.M. Mattheij J.M.L. Maubach	Autofocus and two-fold astigmatism correction in HAADF-STEM	Febr. '10

Ontwerp: de Tantes,
Tobias Baanders, CWI