Use of a general imaging model to achieve predictive autofocus in the scanning electron microscope

F.C. Nicolls [a] G. de Jager [a] B.T. Sewell [b]

[a] *Department of Electrical Engineering, University of Cape Town, South Africa*

[b] *Electron Microscope Unit, University of Cape Town, South Africa*

This work outlines the development of a general imaging model for use in autofocus, astigmatism correction, and resolution analysis. The model is based on the modulation transfer function of the system in the presence of aberrations, in particular defocus.

The signals used are related to the ratios of the Fourier transforms of images captured under different operating conditions. Methods are developed for working with these signals in a consistent manner.

The model described is then applied to the problem of autofocus. A fairly general autofocus algorithm is presented and results given which reflect the predictive properties of this model.

The imaging system used for the generation of results was a scanning electron microscope (SEM), although the conclusions should be valid across a far wider range of instruments. It is however the specific requirements of the SEM that make the generalisation presented here particularly useful. We have therefore confined our investigation to SEM.

1 Introduction

Traditional autofocus methods involve the use of a focus measure which exhibits an extremum for the condition of best focus [4,14,1]. In general a search is then required over the possible range of focal lengths. This requires that a large number of images be taken, which places limits on the maximum speed that the algorithm can achieve.

This search through focal lengths is unnecessary, however, in that the effects on the image vary smoothly and predictably as a function of defocus [11]. In this paper a very general model is developed for the dependence, and it is shown how it may be used to efficiently effect autofocus.

The section which follows characterises certain aspects of linearity in the image formation process. It is demonstrated how a point-spread function can be defined for the system. Following that is an overview of the basic idea that is used in the development of an autofocus algorithm in the subsequent sections. Noise reduction is an important aspect which needs to be contended with, and section 4 discusses some of the associated problems and ways in which they can be circumvented. Section 5 shows how a general model of the image formation process can be constructed, and section 6 extends the use of this model to autofocus. Following that is a discussion of some of the implementation details involved in the algorithm. Section 8 then assesses the performance of the proposed method by providing results generated from a through-focus sequence of images.

The imaging system for which results are presented is a scanning electron microscope (SEM), operating in the secondary electron detection mode. The instrument used was the Leica S440. The in-focus image for the sample set used to generate the results in this paper is shown in Figure 1.
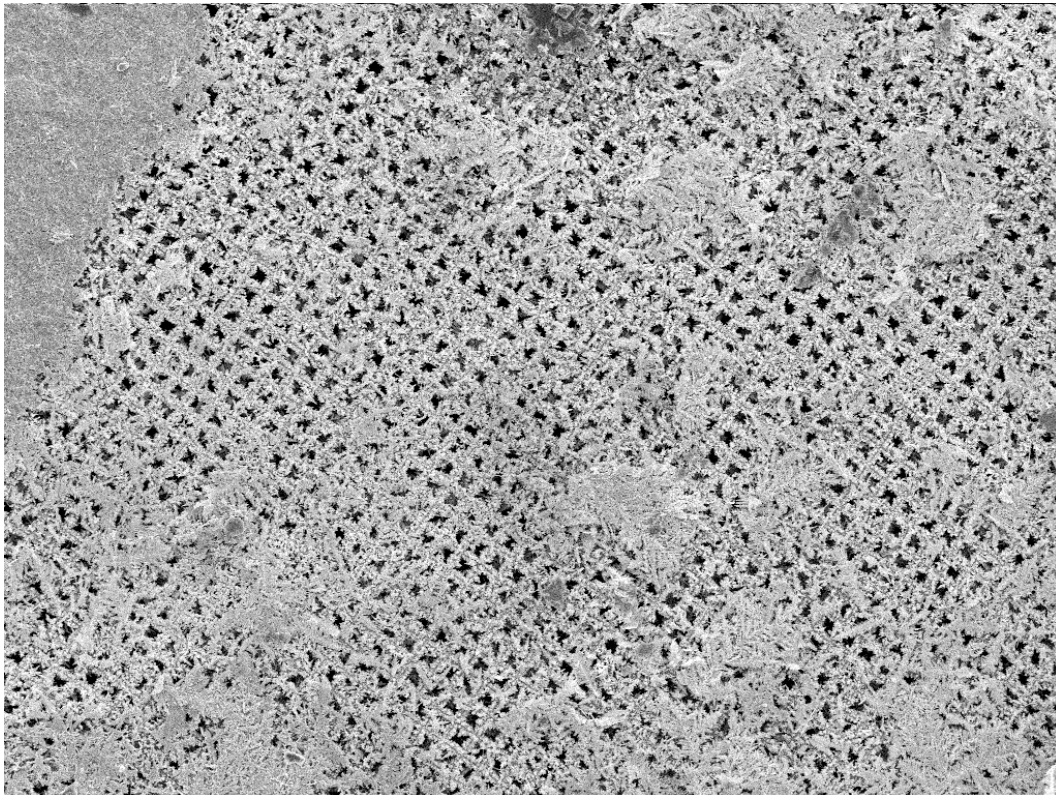

Fig. 1. In-focus image for sample set used in experiments.

2

2 Image formation in the SEM

For the purposes of this paper, the restriction will be made to image formation in terms of the detection of secondary electrons (SEs) with an Everhart-Thornley detector.

For a single primary electron incident on a position (x, y) of a specimen, a yield coefficient $\delta(x, y)$ can be assigned such that δ is the average number of resultant secondary electrons emitted. This means that for n_B electrons incident at (x, y), the number of secondary electrons liberated will be

$$n_{\mathrm{SE}} = \delta(x, y) n_B \tag{1}$$

If an infinitesimal area element $dxdy$ centred on (x, y) is considered, and time dependence is introduced, it can be said that the SE current emitted from this area is

$$di_{\mathrm{SE}} = \delta(x, y) \mathcal{J}(x, y) dxdy \tag{2}$$

where $\mathcal{J}(x, y)$ is the incident current density at the point (x, y). The total SE current emitted from the specimen is then

$$i_{\mathrm{SE}} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x, y) \mathcal{J}(x, y) dxdy \tag{3}$$

If the system is assumed isoplanatic [6] then the current density distribution of the beam can be written

$$\mathcal{J}(x, y) = \mathcal{J}_o(x - x_o, y - y_o) \tag{4}$$

where (x_o, y_o) is the centre of the incident beam current distribution. Now, for every point (x_o, y_o) of the centre of the beam, the total resulting secondary electron current is

$$i_{\mathrm{SE}}(x_o, y_o) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x, y) \mathcal{J}_o(x - x_o, y - y_o) dxdy \tag{5}$$

which is just the convolution product

$$i_{\mathrm{SE}}(x_o, y_o) = \delta(x_o, y_o) \otimes \mathcal{J}'_o(x_o, y_o) \tag{6}$$

where $\otimes$ represents the convolution operator, and $\mathcal{J}'_o(x_o, y_o) = \mathcal{J}_o(-x_o, -y_o)$.

To a first approximation, the output of the detector is linear with respect to i_{SE} [12, p.178]. Thus, if $f(x_o, y_o)$ is the resulting signal for the beam centred on (x_o, y_o), then

$$f(x_o, y_o) = G_D \ \delta(x_o, y_o) \otimes \mathcal{J}'_o(x_o, y_o) \tag{7}$$

where G_D is the assumed constant gain of the detector. With $h(x_o, y_o) = G_D \ \mathcal{J}'_o(x_o, y_o)$,

$$f(x_o, y_o) = \delta(x_o, y_o) \otimes h(x_o, y_o) \tag{8}$$

It has thus been shown that, under the conditions described here, image formation in the SEM can be considered to be a linear convolution of two quantities:

– A specimen dependent component, namely a two-dimensional field of secondary electron yield coefficients, and
– A system dependent point-spread function $h(x_o, y_o)$ which, in conjunction with the image field defined here, is effectively the current density distribution of the electron beam scaled and reflected through the origin of the coordinate system.

This result is an extremely important one because it characterises the linearity of the image formation process in terms of observables.

The primary assumption made in the derivation is that of isoplanatism, which implies the existence of a shift invariant PSF. Even if this assumption should be invalid, the notion of an average or effective PSF is still useful, however. The PSF would then no longer correspond to the current density distribution of the beam as derived earlier.

3 Autofocus approach

Before continuing any further, an outline of the general approach to developing the autofocus algorithm will be discussed.

In the Fourier transform domain, the convolution product in the previous section can be represented by a simple point-by-point multiplication

$$F(\omega_x, \omega_y) = \Delta(\omega_x, \omega_y) H(\omega_x, \omega_y) \tag{9}$$

4

where the upper case functions are the Fourier transforms of the respective lower case functions. ω_x and ω_y are the spatial frequency coordinates in the Fourier transform domain. The quantity $|H(\omega_x, \omega_y)|$ is the modulation transfer function (MTF) corresponding to the point spread function $h(x, y)$. If two images are now taken of exactly the same area of the specimen, but using different imaging conditions, then

$$
\begin{aligned}
F_1(\omega_x, \omega_y) &= \Delta(\omega_x, \omega_y) H_1(\omega_x, \omega_y) \\
F_2(\omega_x, \omega_y) &= \Delta(\omega_x, \omega_y) H_2(\omega_x, \omega_y)
\end{aligned}
\tag{10}
$$

where the difference between H_1 and H_2 comes about because of this change in settings. For purposes of this paper this change is considered to come about due to a change in focal length, although in general other factors such as changes in aperture can also be utilised [3]. The corresponding spatial domain PSFs h_1 and h_2 can be said to relate to the current density distribution of the beam at those positions where it intercepted the specimen. The ratio $|F_1/F_2|$ is now

$$
\left| \frac{F_1(\omega_x, \omega_y)}{F_2(\omega_x, \omega_y)} \right| = \frac{|H_1(\omega_x, \omega_y)|}{|H_2(\omega_x, \omega_y)|}
\tag{11}
$$

which is an image independent quantity that varies according to the two PSFs used in the formation of the images.

Note that the expressions derived above are for the ideal case of infinite-extent signals. In practice some form of windowing is required to implement calculation of the ratio $F_1(\omega_x, \omega_y)/F_2(\omega_x, \omega_y)$ [5]. The effect of this windowing is that the dependence of the specimen $\Delta(\omega_x, \omega_y)$ is never entirely removed from the ratio. However, this effect is minimised if the number of pixels in the transformed image is large with respect to the width of the corresponding MTF.

The approach described here has been used for light-optical systems in depth-from-defocus work [13,3,15,7], which has direct application to autofocus. Two common methods in which this ratio is used for autofocus are:

– Lookup-tables are used which relate the ratio signal obtained to the distance of the object from the aperture plane, or
– An expression for the defocus dependence on the MTF is derived, usually using the thin lens approximation in conjunction with the aperture function.

A further common assumption is that of the MTF being roughly Gaussian in shape. This condition facilitates storage of lookup-tables and derivation of closed-form solutions for the specimen distance. Simulations for the S440 [9],

however, demonstrated that this assumption is not very accurate, except for the in-focus situation.

For the case of the SEM, the first of these approaches is unfavourable due to the large number of variable parameters in the image formation process. Storage and calibration present a problem for parameter spaces of such high dimensionality. The second approach, although better, proves to be restrictive and sometimes inaccurate in its assumptions. A generalisation of this model will therefore be developed, which appeals to fundamental aspects of the imaging process. In section 5 characteristics are isolated which prove to be of use in an autofocus context. Firstly, however, noise reduction in the ratio estimate will be discussed.

4 Noise reduction techniques

Noise in the two-dimensional ratio signal presents a serious problem. It can be shown that the probability distribution of the ratio of two Gaussianly-distributed random variables is such that the mean does not exist. Thus the average of a number of estimates of this ratio is not guaranteed to exhibit any noise reduction. Because of this, alternative noise-reduction techniques have to be developed and their validity justified. Some aspects of these techniques are covered in this section:

– A model of the noise in the transformed images is given, and and the problem with simple averaging presented,
– Noise-reduction methods which circumvent the problems are described and their use justified.

An analysis of a microscope image in the frequency domain suggests that the desired signal is contaminated by an additive noise component as well as an overall offset. If $F_{\text{image}}(\omega_x, \omega_y)$ is the transform and $F(\omega_x, \omega_y)$ is the desired signal, then

$$|F_{\text{image}}(\omega_x, \omega_y)| = |F(\omega_x, \omega_y)| + n(\omega_x, \omega_y) + C \tag{12}$$

where $n(\omega_x, \omega_y)$ is the random noise contribution (which is approximately Gaussianly distributed), and C is a constant offset which is independent of defocus. It is assumed that this value C can be estimated from the transformed data.

Given the accuracy of this model, an estimate of the required ratio can be formed as

$$\frac{|F_1(\omega_x, \omega_y)| + n_1(\omega_x, \omega_y)}{|F_2(\omega_x, \omega_y)| + n_2(\omega_x, \omega_y)} = \frac{|F_{\text{image1}}(\omega_x, \omega_y)| - C}{|F_{\text{image2}}(\omega_x, \omega_y)| - C} \qquad (13)$$

It can be seen that this ratio is contaminated by noise fields n_1 and n_2, which because of the inclusion of C are guaranteed to be zero-mean. Thus, for any (ω_x, ω_y), the values of the numerator and denominator can be considered to be Gaussian random variables with means $|F_1(\omega_x, \omega_y)|$ and $|F_2(\omega_x, \omega_y)|$ respectively. This estimate of the required ratio becomes more accurate as the noise standard deviation decreases.

Consider now the random variable $Z = X/Y$, where X and Y are Gaussian random variables with mean values x_0 and y_0, and the same standard deviation σ. The probability distribution of Z can be calculated to be [10]

$$p_Z(z) = 2K \left\{ \frac{\nu(z)}{\mu(z)} \sqrt{\frac{\pi}{\mu(z)}} e^{\frac{\nu(z)^2}{\mu(z)}} \Phi\left(\frac{\nu(z)}{\sqrt{\mu(z)}}\right) + \frac{1}{2\mu(z)} \right\} \qquad (14)$$

where $\mu(z)$ and $\nu(z)$ are

$$\mu(z) = \frac{1}{2}z^2 + \frac{1}{2} \qquad (15)$$
$$\nu(z) = -\frac{zx_0}{2\sigma} - \frac{y_0}{2\sigma}$$

and $\Phi(x)$ is the usual error function defined as

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \qquad (16)$$

Plots of this distribution for a fixed value of $x_0/y_0 = 1/2$ and changing σ are shown in Figure 2. This value of x_0/y_0 was chosen because it is fairly representative of usual values in the MTF ratio.

The expected value $E\{Z\} = \int_{-\infty}^{\infty} z p_Z(z) dz$ of this distribution does not exist since the integral does not converge [10]. This outlines the basic difficulty of noise reduction by means of averaging of the ratio signal.

A possible solution to this problem lies in performing averaging before the formation of the ratio [8, p.559]. Suppose we have p independent estimates of
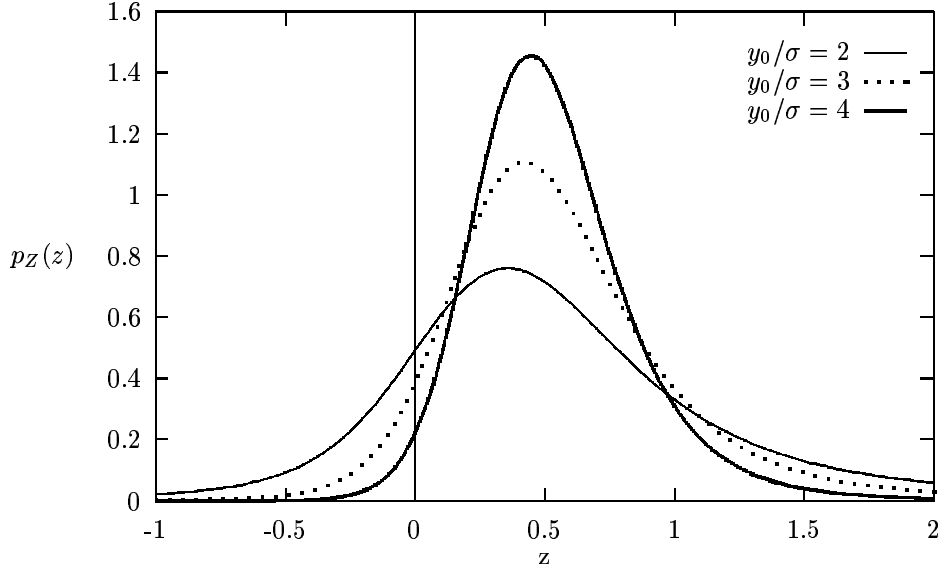
7

Fig. 2. Probability distribution of Z for $x_0/y_0 = 1/2$ and changing noise

the numerator and denominator images, designated by $N_{\text{image k}}(\omega_x, \omega_y)$ and $D_{\text{image k}}(\omega_x, \omega_y)$ for $k = 1, \ldots, p$. The quantity

$$\frac{(1/p)\sum_{k=1}^{p}|N_k(\omega_x, \omega_y)| - C}{(1/p)\sum_{k=1}^{p}|D_k(\omega_x, \omega_y)| - C} \tag{17}$$

can then be shown to be a valid (and noise reduced) estimate of the required MTF ratio. This comes about essentially because the mode of the distribution $p_Z(z)$ becomes a better estimate of the value x_0/y_0 in equation 14 for lower values of σ. This statement is justified by Figure 3, which demonstrates that the mode is a more effective estimator of the ratio value for smaller values of σ. Since the averaging reduces σ, the estimate derived from equation 17 represents an improvement on the unaveraged estimate.

Rather than working with the two-dimensional MTF ratio, one-dimensional representations can be obtained for the MTF in any plane through the optical axis at a specific angle. This has advantages in that it is easier to work with the 1-D signals. In the absence of coma, astigmatism, and distortion, the MTF ratio is circularly symmetric [2], and the redundancy can be used to reduce noise in the mapping from 2-D to 1-D. In this case if the median of all samples at each distance from the origin is found, the result forms a noise-reduced estimate of a cut through the MTF ratio at any angle. Here this occurs because the median of $p_Z(z)$ becomes a better estimate of x_0/y_0 for smaller σ. This is demonstrated by Figure 4. In the presence of astigmatism, the same conclusion applies in a limited sense if a small sector surrounding a radial line through the MTF ratio is considered.

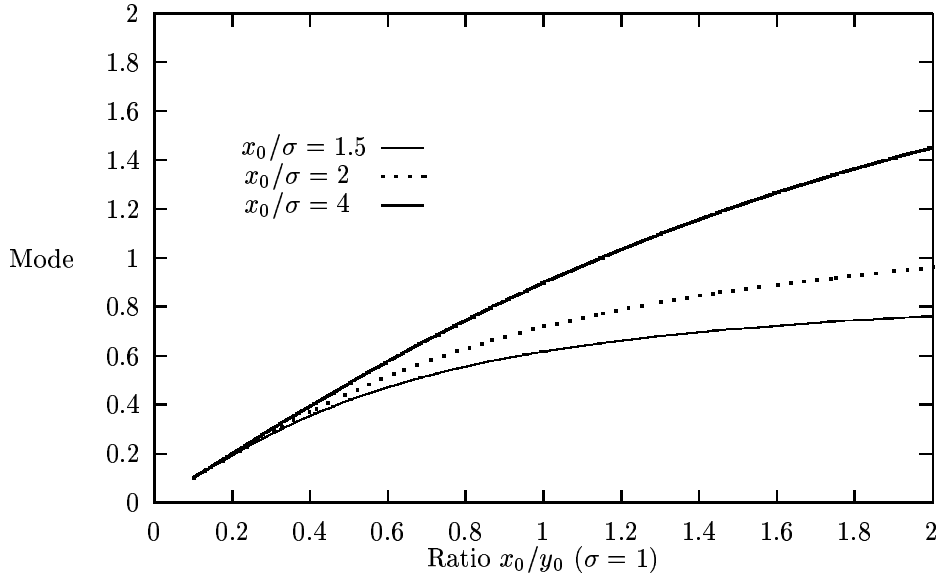In summary, noise reduction is achieved as follows:

8

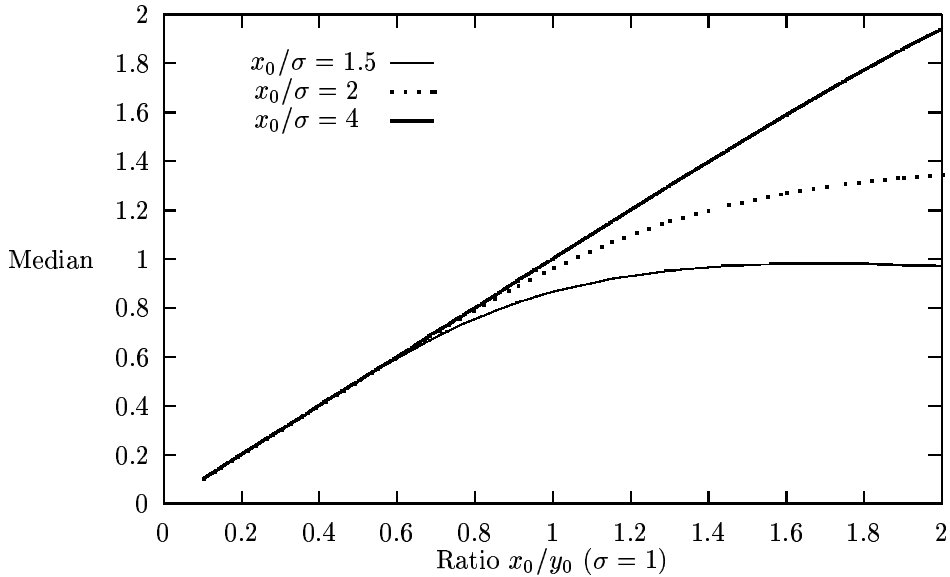Fig. 3. Mode vs. x_0/y_0 ratio for changing noise

Fig. 4. Median vs. x_0/y_0 ratio for changing noise

- A number of estimates of each of the numerator and denominator images are averaged and then the ratio formed of these averages. This results in a noise-reduced representation of the 2-D ratio.
- If the profile of a radial cut through the centre of the MTF is required, this is obtained by considering the portion of the 2-D ratio in a sector surrounding the line of interest. The median of the samples in this sector at each distance from the origin forms a noise-reduced estimate of the 1-D ratio at that angle.

For the remainder of this paper it is assumed that wherever possible these noise reduction techniques have been applied.

9

5 Model of image formation

If the denominator image F_2 in equation 11 is chosen to be an in-focus image, then $|H_2(\omega_x, \omega_y)|$ is approximately unity and the ratio reduces to the MTF corresponding to the first image. In this way it is possible to form estimates of the MTF of the microscope for varying degrees of defocus.

The method was applied to a through-focus series of images, and profiles of three of the resulting MTFs is given in Figure 5. The MTFs correspond
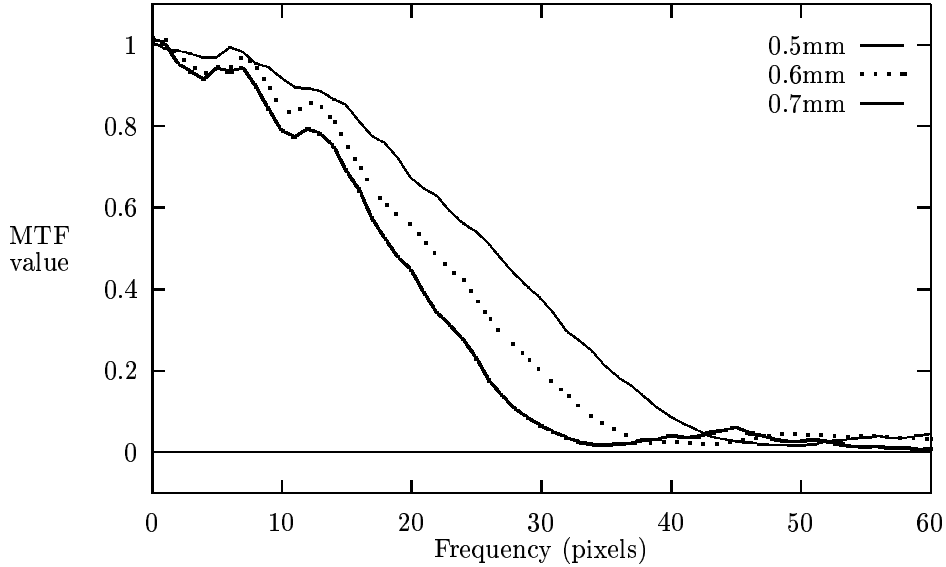

Fig. 5. Radial cuts through MTFs corresponding to images taken of a specimen at three different distances from best focus.

to distances of $0.5mm$, $0.6mm$, and $0.7mm$ from best focus, with a working distance of around $15mm$ and a magnification of $500\times$.

Of note is the apparent similarity between the MTFs for the different defocus levels. Each MTF is to a good approximation just a dilation or contraction in the horizontal direction of any other MTF. This property will be referred to as the self-similarity in the MTF at different positions along the beam.

In order to get an idea of the width of the MTFs, Gaussian functions were fitted to each. These Gaussians were best-fit in the sense that they minimised the distance between the functions in the l^2-norm. Converting these standard deviations back to the spatial domain, the dependence of PSF width with respect to defocus can be determined. Such a plot is shown in Figure 6. The width can be seen to vary linearly with the amount of defocus.

Furthermore, the self-similarity of MTFs ensures self-similarity of the corre-
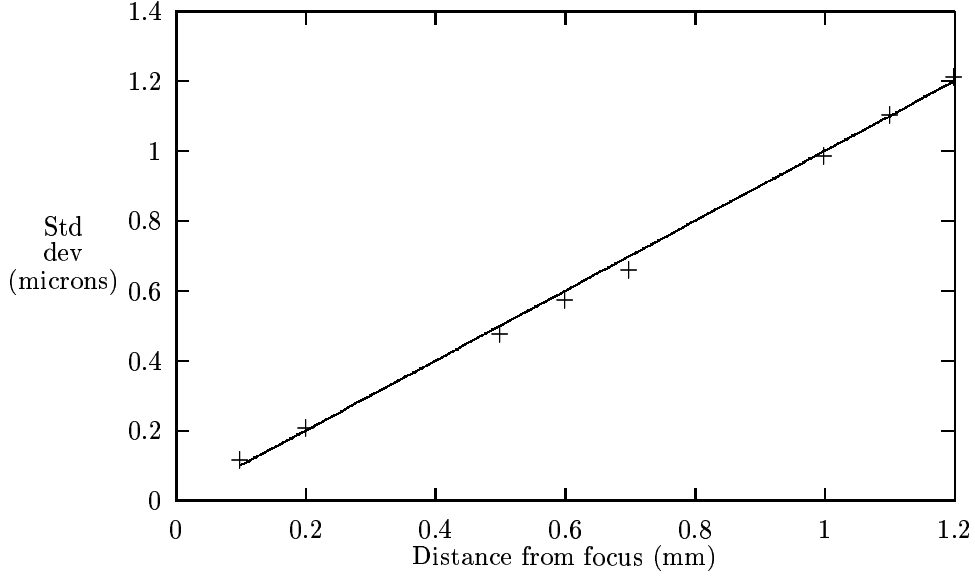
10

Fig. 6. Standard deviation of best-fit Gaussians referred to the spatial domain plotted as a function of the distance from focus. Also indicated is the best-fit straight line to the data points

sponding PSFs by the Fourier relation

$$\mathcal{F}\{f(ax, by)\} = F(\frac{\omega_x}{a}, \frac{\omega_y}{b}) \tag{18}$$

where $\mathcal{F}$ represents the Fourier transform operator. Thus, it can be said that if the distance from focus of an image is doubled, the resulting PSF will change only by a stretching in the x-direction by a factor of 2.

These findings represent a very general model of the MTF for changing levels of defocus. It can be shown that the properties outlined conform to the notion of an idealised beam, where the motion of the electrons is strictly rectilinear and the crossover infinitely small. The cross-section along the optical axis through such a beam is depicted in Figure 7. The width of the beam at the specimen is then given by

$$s = a\frac{|f_x - d|}{f_x} \tag{19}$$

Note that even if the beam does not precisely conform to this requirement, the expression for the width can still apply differentially for a modified a, which can be said to represent the effective aperture for the region of the beam in the vicinity of the specimen. For this reason it may not be possible in general to assign a fixed value to a for a given microscope under all viewing conditions.

The model presented here for the spatial distribution of the beam may seem
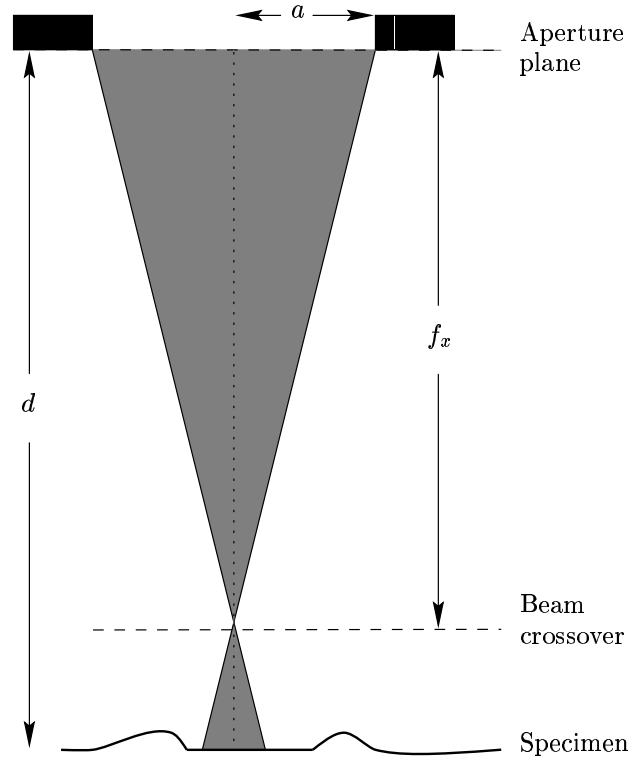
11

Fig. 7. Cross-section through an ideal beam

to be based on rather dubious assumptions. Simulations show that spatial self-similarity of the beam only becomes valid about $0.1mm$ from the crossover [9]. Furthermore, the idea of a zero-width crossover cannot be reconciled with reality. With regard to the first point, it must be stressed that the model was derived from frequency-domain representations, and will only be applied to information gathered from this domain. In practice the space domain beam profiles can exhibit a pronounced lack of self-similarity while the MTFs largely retain this property. The reason for this can in general be attributed to the fact that the spatial detail that introduces the discrepancies resides outside the main frequency lobe representing the bulk of the signal energy. A second source of possible error lies in the approximation that the signal from the detector is a linear function of the scattered electrons. Experiments suggest that the transfer function of the detector might be significantly more complex. With regard to the nonzero beam crossover, this only becomes significant if the sampling distance on the specimen for neighbouring pixels becomes small enough that the beam distributions overlap. This is more likely to occur close to focus with high magnifications, and the model will fail under these circumstances. However, in practice the failure is not catastrophic, again because it is the frequency domain representation that is ultimately used.

6 Autofocus

The presentation of the autofocus method using the described model is greatly facilitated my making use of a concrete example. Let us assume the situation of a two-dimensional radially symmetric Gaussian beam,

$$h(x, y) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}} \tag{20}$$

where the standard deviation is related to the focal length and specimen distance by a function

$$\sigma(a, d, f_x) = a\frac{|f_x - d|}{f_x} \tag{21}$$

d is the aperture-specimen distance, which, if it can be found, yields the optimum focal length. As has been mentioned, the assumption of a beam with a Gaussian PSF is widely made in the literature. Note that this restriction is only made here for demonstration purposes; except for theoretical evaluation it is dropped when the general autofocus algorithm is introduced.

The corresponding MTF is

$$H(\omega_x, \omega_y) = e^{-\frac{\sigma(d,f)^2}{2}(\omega_x^2 + \omega_y^2)} \tag{22}$$

Considering the MTF ratio for two focal lengths f_1 and f_2, we have

$$\frac{|H_1(\omega_x, \omega_y)|}{|H_2(\omega_x, \omega_y)|} = e^{-\frac{(\sigma^2(d,f_1)^2 - \sigma^2(d,f_2))}{2}(\omega_x^2 + \omega_y^2)} \tag{23}$$

which has a resulting variance of

$$\sigma_{\text{res}}^2(d, f_1, f_2) = \frac{1}{\sigma^2(a, d, f_1) - \sigma^2(a, d, f_2)} \tag{24}$$

This relation contains two unknowns, a and d, so the system is under-constrained. Note that the variance may take on negative values in the expression. This occurs when the numerator MTF is wider than the denominator, and the ratio grows without bound for large values of the independent variable.

Since, for reasons described at the end of the previous section, it may be impossible to fix a for all conditions, more information is required to solve for d. This information can be provided by means of a second MTF ratio which

will yield an independent relation in the same unknowns. In order to form this second ratio, a minimum requirement is that one further image be taken at focal length f_3, and the ratio formed with either of the existing two images. a can then be eliminated from the two resulting relations, and the desired value for the focal length found. This value for d can be shown to be unique. Thus under the Gaussian assumption a closed-form solution for d exists using three images taken at different focal lengths.

The way in which the Gaussian assumption is overcome is by making use of a template which is representative of the system MTF. It is therefore assumed that the MTF at every defocus level can be accurately described by a scalar value which gives the factor by which this template must be compressed to yield the actual MTF. For the general case where the shape of the MTF is specified by a template function rather than restricted to Gaussian, similar requirements for autofocus may be expected to exist since the number of free parameters is unchanged. Three images taken at different focal lengths are necessary and sufficient to completely specify d.

Suppose that the template function for the MTF is $H_t(x)$. It is assumed that the MTF of the second image is given by

$$\text{MTF}_2 = H_t(kx) \tag{25}$$

for some value of k. The models that the two MTF ratios conform to are then

$$\frac{\text{MTF}_1}{\text{MTF}_2} = \frac{H_t(kk_1x)}{H_t(kx)}$$
$$\frac{\text{MTF}_2}{\text{MTF}_3} = \frac{H_t(kx)}{H_t(kk_2x)} \tag{26}$$

Here k_1 and k_2 represent the factors by which MTF_2 is wider than MTF_1 and MTF_3 respectively. Equation 19 can be used to derive relations between these two factors and the imaging configuration, namely

$$k_1(d, f_1, f_2) = \frac{s_1(a, d, f_1)}{s_2(a, d, f_2)} = \frac{|f_1 - d|}{|f_2 - d|}\frac{f_2}{f_1}$$
$$k_2(d, f_2, f_3) = \frac{s_3(a, d, f_3)}{s_2(a, d, f_2)} = \frac{|f_3 - d|}{|f_2 - d|}\frac{f_2}{f_3} \tag{27}$$

with s the beam width at the specimen. Note that since the ratios are independent of a, no precise definition of what is meant by the width of the beam is required if this development is followed.

14

If the shape of the template is Gaussian, the example at the start of the section shows that a unique solution exists for d. It can be shown that the same is true for k (to within an arbitrary sign change). Thus, given the variance of the two MTF ratios, there is only one combination of (k, d) that will result in the model being consistent. The method can now be inverted: if the MTF ratios are calculated according to the model for each combination of (k, d), there will be a unique case where the calculated ratios coincide with the actual ratios derived from the data. This lays the foundations for a search method to find the required aperture-specimen distance for an arbitrary template.

Note that for a non-Gaussian template function, the guarantee of uniqueness is compromised. However, experience shows that actual MTFs generally adopt the form of a Gaussian, if not the precise shape. Thus it is not expected that this will constitute a problem.

A further point of interest is that the value obtained for k is representative of the width of MTF_2. The procedure demonstrated here is applied to one-dimensional signals which are derived from a cross-section of the beam through the optical axis. By reapplying the procedure to cross-sections at different angles through this axis, a complete representation of the intermediate MTF can be constructed. The same can be said for MTF_1 and MTF_3. This can be used to detect and correct for astigmatism, and can give an indication of the resolution of the instrument at any defocus level.

7   Algorithm implementation

The way in which the ideas discussed are implemented is to define a distance measure for comparing the sets of actual ratios and calculated ratios. These measures prove to be something of a difficulty due to noise in the images, which tends to obscure the ratio data in regions where the numerator and denominator used in forming the ratios become small. The only fundamental restriction on the distance measure is that it be zero if and only if the sets of ratios are identical. This is a necessary and sufficient condition to ensure a unique solution under the Gaussian assumption. Finding the optimal focal length then reduces to finding the minimum of this distance measure over all space of (k, d).

A number of distance measures were tried in the implementation. It was found that a modification by a weighting function of the squared difference metric was most successful. Although this modification violated the condition required for uniqueness, this proved not to be a problem as long as a good starting point for the search could be estimated. The Gaussian example described previously provides the grounding for making such an estimate. The

measure that was ultimately used was

$$d_{s_{(\mathrm{ratio1})}}(R_{\mathrm{act1}}, R_{\mathrm{gen1}}) =$$

$$\min\left\{ |\int_{-\infty}^{\infty} (R_{\mathrm{act1}}(x) - R_{\mathrm{gen1}}(x))^2 H_t(kk_1x)dx|, \right.$$

$$\left. |\int_{-\infty}^{\infty} (R_{\mathrm{act1}}^{(\mathrm{inv})}(x) - (R_{\mathrm{gen1}}(x))^{-1})^2 H_t(kx)dx| \right\} \tag{28}$$

where $d_{s_{(\mathrm{ratio1})}}$ is the difference contribution for the first set of ratios of which R_{act1} is the actual ratio signal obtained from the data and R_{gen1} the ratio calculated for and specified combination of (k, d). $R_{\mathrm{act1}}^{(\mathrm{inv})}(x)$ is the reciprocal ratio, which may not be equal to $R_{\mathrm{act1}}^{-1}(x)$ due to the noise reduction procedure. The combined distance is then

$$d_s = d_{s_{(\mathrm{ratio1})}} + d_{s_{(\mathrm{ratio2})}} \tag{29}$$

where $d_{s_{(ratio2)}}$ is similarly defined. Note that the operation of taking the minimum of the two options ensures that only ratios that go to zero for large values of x are compared. This is necessary because if ratios are compared which grow without bound the distance would become infinite for compared ratios with very small discrepancies. The weighting function H_t in the measure emphasises those portions of the signal which are less prone to influence by noise.

Figure 8 shows a contour plot of the theoretical search space under the Gaussian assumption for a particular imaging configuration. The situation corresponds to images taken at defocus distances of $0.2mm$, $0.5mm$, and $0.7mm$, again for a working distance of around $15mm$. Shown in Figure 9 is a comparative plot obtained experimentally from actual image data and a representative MTF. The change in overall value is simply a result of a different normalisation for the two plots.

8   Results

A series of results were generated by applying the algorithm to all combinations of three images of a through-focus image series. The images spanned a range of defocus distances from $0mm$ to $1.2mm$, with in-focus working distance of $15mm$ and magnification $500\times$. Results are given in Figure 10. Each plot (a)-(f) presents the results for a fixed distance from best focus of the intermediate image used in the prediction. The solid horizontal line in each
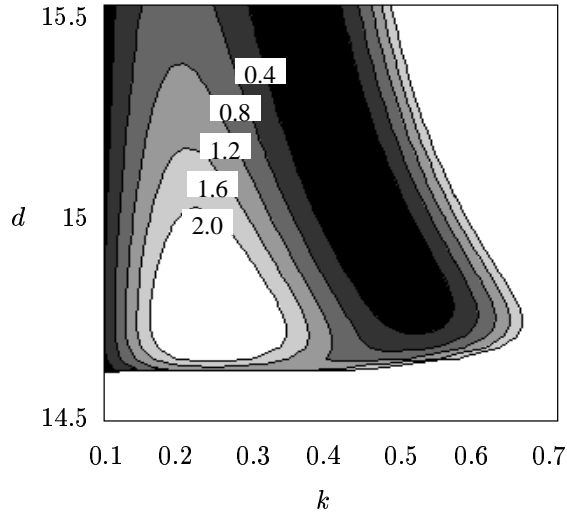
16

Fig. 8. Contour plot of theoretical search space to demonstrate sensitivity for specific case (working distance $15mm$) under weighted squared difference measure. The plot shows how the value of the difference measure changes when varying values of k (the MTF stretch factor) and d (the aperture-specimen distance) are assumed.

Fig. 9. Contour plot of the actual search space for specific case (working distance $15mm$ under the weighted squared difference measure. The plot shows how the value of the difference measure changes when varying values of k (the MTF stretch factor) and d (the aperture-specimen distance) are assumed.
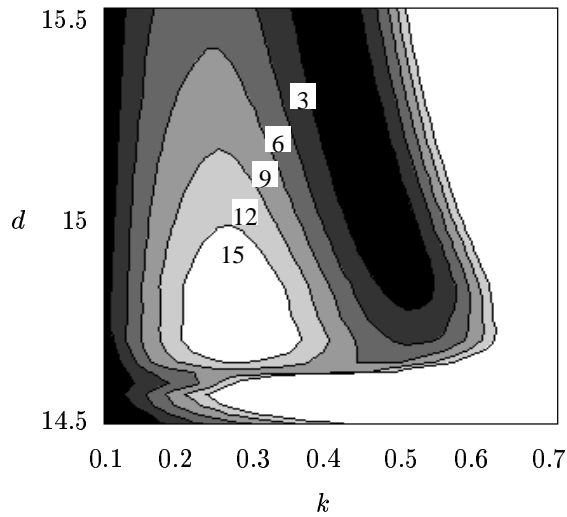
case indicates this distance. The vertical dotted lines partition the plot into separate prediction sets, each of which correspond to an experiment using different input data. Since three images are required for the predictions, the two crosses in each set indicate the distances from focus of the additional images. For example, consider plot (a) in Figure 10: the leftmost interval corresponds to input images at the three defocus distances $0.1mm$, $0.2mm$, and $0.5mm$, with $0.2mm$ being the distance for the intermediate input image. For each case the prediction process is applied, and the distance as calculated for this

intermediate image plotted using a star. The situation for correct prediction thus corresponds to all the stars lying on the horizontal line. Note that in some cases where the prediction procedure fails, the location of the star may not appear in the plotted range, and is hence absent.

The plot demonstrates two basic trends:

– The prediction is fairly accurate if the intermediate MTF is close to focus. For points where the actual distance is large, the reliability tends to break down.
– If the prediction fails for images which are close to focus, then generally it seems that it occurs when at least one of the pairs of images used in the formation of the ratios are close together.

The cause of the first trend can be explained in reference to the width of the MTF as a function of the distance from the crossover. In a previous section it was shown that this relation is hyperbolic. Thus, near to focus the MTF is very wide, falling off rapidly as the distance is increased. Now the first zero of the ratio will be the same as the first zero of the numerator MTF. Thus the basic shape of the ratio is dictated by the numerator. The effect of the denominator MTF is then to alter the shape of this ratio between the origin and the first zero. Clearly this is a far smaller effect to detect than the change in the zero position. If the numerator MTF now becomes narrow, the deviation caused by the denominator MTF profile becomes less significant. Huge changes in this denominator will then cause only marginal changes in the ratio signal. In the presence of noise this becomes catastrophic to the prediction process because the distance measure no longer exhibits a suitable minimum.

The observation that the prediction fails more readily if the images forming the ratio are close together can also be simply explained. In the limit as the distance between the images goes to zero, no prediction can be made because data has been irrecoverably lost. It follows then that if the images are close together then the effectiveness of the prediction will be hindered. The discussion of the first trend also makes it apparent that this closeness has to be considered in relation to the actual distance of the images from the crossover. MTFs separated by a fixed distance far from focus will differ from one another far less than MTFs separated by the same distance but nearer to the beam crossover.

Essentially, the prediction fails more readily for image sets far from the crossover, which correspond to an overall narrower system MTF. In the case of the SEM, this failure is not catastrophic. This is because of the ability to change the magnification of the instrument, and hence affect the overall MTF width. For example, if it is found that the MTF is too narrow for effective prediction to be made, a reduction in magnification can effectively shift the process into a

region where the MTFs are wider and the algorithm has more chance of success. This suggests an iterative application of the algorithm at progressively higher magnifications, until the desired accuracy is obtained. An investigation into the detailed effects of change in magnification on the autofocus algorithm has not been performed, so results relating to this will not be presented.

It should be emphasised that at this stage there is no existing real-time implementation of the algorithm presented here. All the calculations have been done off-line on an image sequence which was captured under known conditions. The speed of an implementation would naturally be hardware dependent.

9   Conclusions

It has been demonstrated in this paper that aspects of the image formation process in the electron microscope constitute a very general model for use in autofocus and astigmatism correction. This model is not restricted to electron microscopy; many imaging systems exhibit the property of a self-similar PSF which varies linearly in size with the distance from focus.

An autofocus algorithm was developed using this model. The algorithm is very general: all that is required is a template function representative of the system MTF, and three images taken at different focal lengths. Using this information it is possible to calculate the out-of-focus distance for each of these images, and consequently the optimal focal length can be found. Except for finding the template function, no calibration of the system is required.

References

[1] F.R. Boddeke, L.J. van Vliet, H. Netten, and I.T. Young. Autofocusing in microscopy based on the OTF and sampling. *Bioimaging*, 2:193–203, 1994.

[2] M. Born and E. Wolf. *Principles of Optics*. Pergamon Press, 1959.

[3] J. Ens and P. Lawrence. An investigation of methods for determining depth from focus. *IEEE Transactions on pattern analysis and machine intelligence*, 15(2):97–108, February 1993.

[4] S.J. Erasmus and K.C.A. Smith. An autofocusing and astigmatism correction system for the SEM and STEM. *Journal of Microscopy*, 127:185–200, 1982.

[5] F.J. Harris. On the use of windows for harmonic analysis with the discrete fourier transform. *Proceedings of the IEEE*, 66(1):51–83, January 1978.

[6] H.H. Hopkins. The frequency response of a defocused optical system. *Proceedings of the Royal Society of London: A*, 231:91 – 103, 1955.

[7] Shang-Hong Lai, Chang-Wu Fu, and Shyang Chang. A generalised depth estimation algorithm with a single image. *IEEE transactions on pattern analysis and machine intelligence*, 14(4):405–411, April 1992.

[8] J.S. Lim. *Two-dimensional signal and image processing*. Prentice-Hall Signal Processing Series. Prentice-Hall, 1990.

[9] Eric Munro. Simulations carried out: MEBS Ltd, 14 Cornwall Gdns, London, SW7 4AN, 1995.

[10] Frederick Nicolls. The development of a predictive autofocus algorithm using a general image formation model. Master's thesis, University of Cape Town, Rondebosch 7700, South Africa (http://www.dip.ee.uct.ac.za/~nicolls), February 1996.

[11] A.P. Pentland. A new sense for depth of field. *IEEE transactions on pattern analysis and machine intelligence*, PAMI-9(4):523–531, July 1987.

[12] Ludwig Reimer. *Scanning Electron Microscopy*. Springer Series in Optical Sciences Vol. 45. Springer-Verlag, Berlin, Heidelberg, 1985.

[13] M. Subbarao. Parallel depth recovery by changing camera parameters. In *Second international conference on computer vision*, pages 149–155, Innisbrook Resort, Tampa, Florida, USA, December 1988. Computer society press.

[14] M. Subbarao, T. Choi, and A. Nikzad. Focusing techniques. *Optical Engineering*, 32(11):2824–2836, November 1993.

[15] Tse-Chung Wei and Murali Subbarao. Continuous focusing of moving objects using DFD1F. In *Machine Vision Applications in Industrial Inspection II*, volume 2183 of *SPIE Proceedings Series*, pages 290–300, San Jose, California, February 1994. SPIE.

(a) Case: $0.2mm$

(b) Case: $0.5mm$

(c) Case: $0.6mm$

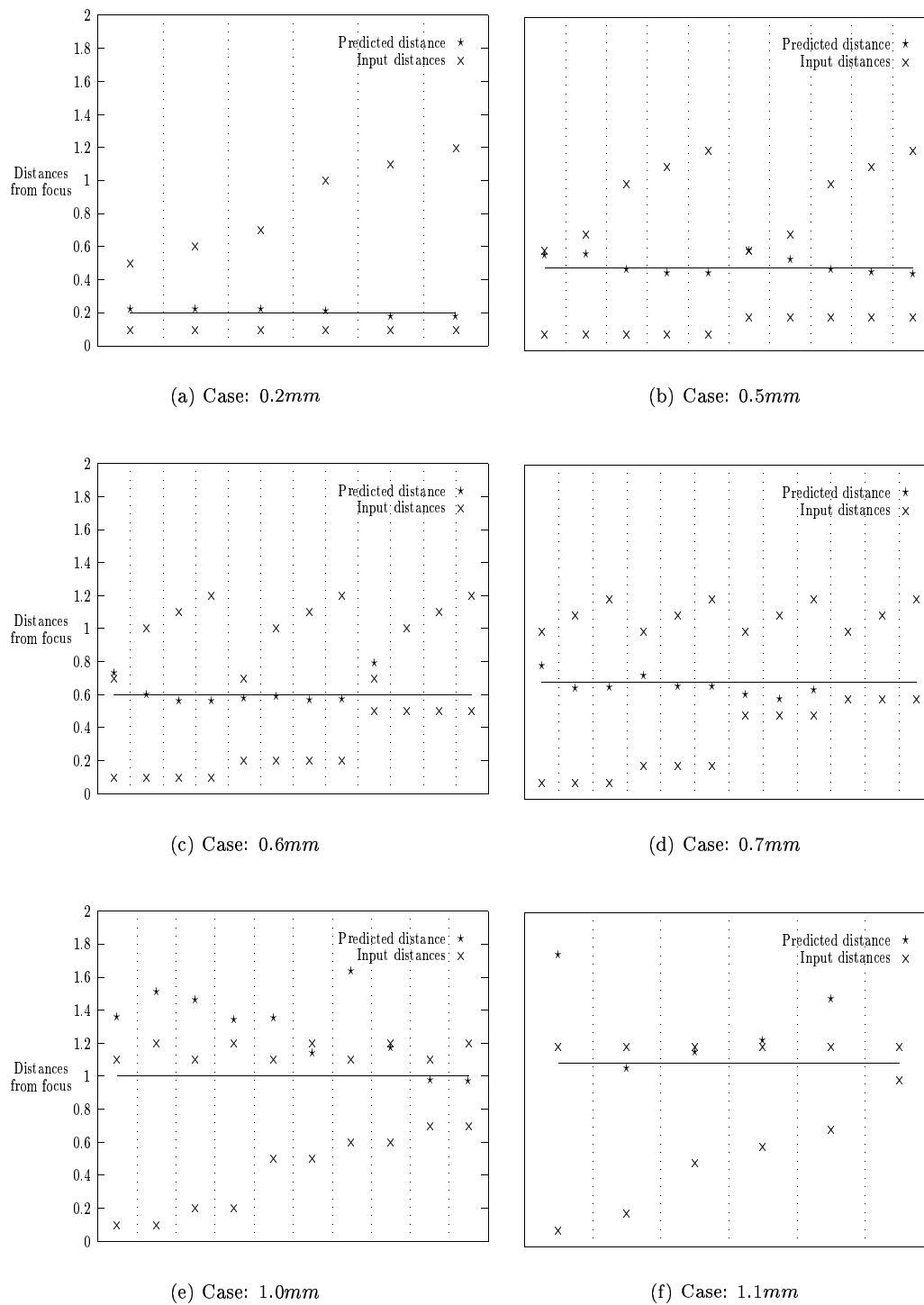(d) Case: $0.7mm$

(e) Case: $1.0mm$

(f) Case: $1.1mm$

Fig. 10. Prediction results for through-focus image series under the weighted squared difference distance measure.