Bioimaging 2 (1994) 193–203. Printed in the UK

Autofocusing in microscopy based on the OTF and sampling

Frank R Boddeke†, Lucas J van Vliet, Hans Netten and Ian T Young

Pattern Recognition Group of the Faculty of Applied Physics, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands

Submitted 11 August 1994, accepted 22 November 1994

Abstract. In the literature many autofocus algorithms have been proposed and compared (Groen *et al* 1985, Firestone *et al* 1991, Yeo *et al* 1993, Price and Gough 1994) for use in optical microscopy (bright field and fluorescence microscopy). Most of the focus criteria measure the high frequency contents of a recorded image as a measure of focus. In this paper we show that a focus criteria should measure the signal power of the middle frequency, since defocusing mainly reduces the frequencies around half the cut-off frequency of the *optical system*. The filter that provides the required *band-pass* filtering depends strongly on the sampling density of the camera. There are two practical combinations of sampling density and one-dimensional digital band-pass filter:

- Sampling at the Nyquist frequency and the {1, 0, −1} filter;
- Sampling at half the Nyquist frequency and the {1, −1} filter.

The latter is to be preferred due to noise considerations and the fact that it uses four times fewer sample points. Calculation speed can also be increased by further reducing the sampling density perpendicular to the filter (on chip or in software) down to 1/8 of the Nyquist frequency. We have designed a three-phase autofocus algorithm that works well in fluorescence and bright field microscopy. The phases are:

- *Coarse*: find the region near focus (step size of typically a few microns);
- *Fine*: find a quadratic region around focus (step size around one micron);
- *Refine*: use a quadratic fit on samples around the peak to find the in-focus position.

We found that the final focus error is smaller than the mechanical reproducibility of our z-axis (≈50 nm) for light levels down to 400 photo-electrons per pixel (sampling at the Nyquist frequency using a cooled CCD camera with pixels of 6.8 × 6.8 μm).

Keywords: autofocusing, focus criteria, focus algorithm, CCD camera, microscopy, spatial frequency, optical transfer function

1. Introduction

Autofocusing is essential in automated microscope systems where a large number of slides needs to be scanned. It also provides the objective and reproducible focusing that is required in quantitative microscopy in order to perform accurate measurements on an imaged object (e.g. the analysis of metaphase chromosomes, Mendelsohn and Mayall 1971).

Autofocus algorithms find the maximum of a focus function. A focus function is a measure of focus as a function of the axial (z) position and is sampled at different positions along the z-axis. The value of the focus function is calculated from an image acquired at that z position. The shape of the focus function is determined by the focus criterion, the microscope and camera parameters, and the imaged object. The properties of a useful focus function are (Groen *et al* 1985):

- Unimodality: only one maximum;
- Accuracy: maximum of the focus function at the in-focus position;
- Reproducibility: a sharp top of the focus function;
- Range: must provide focus information over the desired range;
- Implementation: fast calculation of the focus value.

The goal of the focus algorithm is to find the in-focus position using a small number of samples of the focus

† e-mail: boddeke@ph.tn.tudelft.nl

0966-9051/94/040193+11$19.50

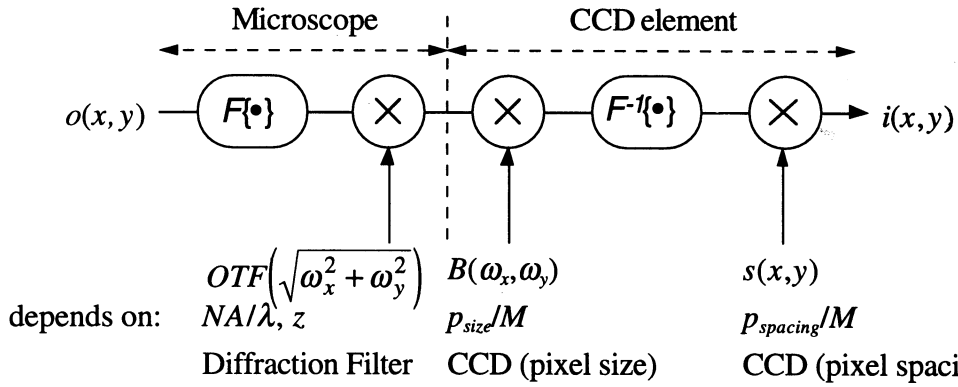

Figure 1. Model representing image acquisition of microscopic images with a CCD camera. The input image $o(x, y)$ (assumed to be a flat—two dimensional—illuminated object) is filtered by the microscope (OTF) and sampled (s) with square pixels (size $p \times p$) by the CCD element into an output image $i(x, y)$. The magnification of the microscope M is incorporated in the pixel size of the CCD element. In the model, $\mathcal{F}\{\bullet\}$ denotes the Fourier transformation. More precise explanation of the parameters of the model will be given below.

function. A small number of images results in minimal focusing time.

2. Image formation and image acquisition

To determine the focus criterion, we need to model the image formation through a microscope and the image acquisition with a CCD camera.

2.1. The microscope as linear shift invariant system

To use linear system theory (Oppenheim *et al* 1983) we have to show that a high-quality optical system behaves (to a good approximation) as a linear, shift-invariant (LSI) system (Young 1989). Linear superposition requires that the combined effect of multiple sources at one time is equal to the sum of the individual responses. In fluorescence microscopy this is guaranteed due to the random nature of the photon emission of the fluorescence process. In bright field microscopy, Köhler illumination (Köhler 1893, Inoué 1986) provides an almost perfect linear response. Within an isoplanatic patch, the point-spread-function of a microscope lens system is also shift invariant. For now we assume that the object is 'thin' with respect to the depth-of-focus of the optical system and can therefore be treated as two dimensional. Image acquisition with a microscope system can then be modeled as a linear, shift-invariant system as depicted in figure 1. This model is the same for bright field and fluorescence microscopy, since the image formation is equal for both types of microscopy.

2.2. Transfer function and sampling of the CCD element

The microscope projects the object onto the CCD element. Apart from magnification of the object, a diffraction-limited

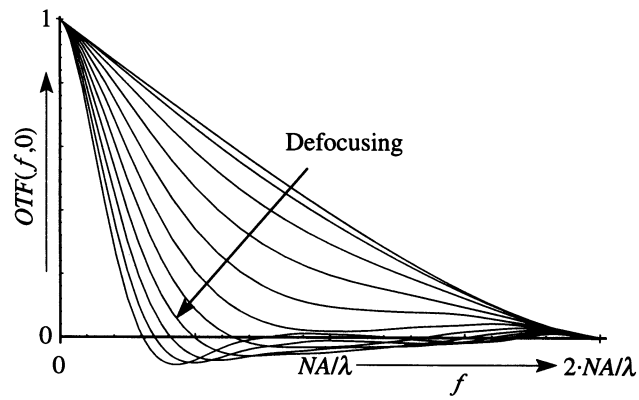

Figure 2. The optical transfer function (OTF) of a microscope for different degrees of defocusing:
$\Delta z = 0, 0.2, 0.4, \ldots, 2\ \mu$m (with $NA = 0.75$, $\lambda = 540$ nm).

microscope acts as a low-pass filter. The projected image is then sampled by the square pixels of the CCD element.

In our model, $o(x, y)$ represents the illuminated object. The filtering effect of the microscope is described by the optical transfer function OTF (Hopkins 1955, Born and Wolf 1959, Williams and Becklund 1989). The back focal plane of the microscope objective lens contains the Fourier transform of the image plane. The finite aperture, therefore, ensures bandwidth limitation (low-pass filtering). Considering intensity images, the input image is multiplied in the Fourier domain by the OTF. The OTF (which is circularly symmetric in the ω_x, ω_y-plane, ω indicating continuous spatial frequencies in radians per second) depends on the numerical aperture of the objective lens divided by the wavelength of light (NA/λ) and on the degree of defocusing (Δz).

The OTF for different degrees of defocusing (with $NA = 0.75$ and $\lambda = 540$ nm; NA relates the defocusing Δz to the aberration from a spherical wave front relative

to λ) as a function of the spatial frequency f is shown in figure 2. It was evaluated according to the formulas given in Hopkins. The image filtered by the OTF is then sampled by the pixels of the CCD element. This can be described by first a convolution with a block function representing the size ($p_{size} \times p_{size}$ μm) and shape of the pixels followed by a multiplication with a square grid of unit impulse functions (spaced by $p_{spacing}$) representing the position of the pixels. The 'fill factor' of the element is defined as the squared ratio between p_{size} and $p_{spacing}$. Without loss of generality, the magnification M of the microscope (objective lens and possible relay optics) is incorporated in the CCD element by reducing the size and spacing of the pixels by a factor M.

The convolution with a block function in the spatial domain is equal to a multiplication with $B(\omega_x, \omega_y)$ (see figure 3) in the Fourier domain where $B(\omega_x, \omega_y)$ is the Fourier transform of $b(x, y)$. The result is then sampled by multiplication with $s(x, y)$, a square grid of unit impulse functions $\delta(\bullet)$, which is the same as a convolution with $S(\omega_x, \omega_y)$ in the Fourier domain. The block and sample functions in both spatial and Fourier representation are listed in table 1.

2.3. Image acquisition model

The image acquisition model of the microscope and the CCD element is given by:

$$I(\bar{\omega}) = \frac{1}{2\pi}\{O(\bar{\omega}) \cdot OTF(\bar{\omega}) \cdot B(\bar{\omega})\} \otimes S(\bar{\omega}) \qquad (1)$$

in which $\bar{\omega} = (\omega_x, \omega_y)$, and $\otimes$ denotes a convolution.

Finally, the continuous image $i(x, y)$ is read out as a discrete output image $i_d[n, m]$. This is represented by a rescaling in both domains (equations (2) and (3)).

$$i_d[n, m] = i\big(n \cdot p_{spacing}/M,\ m \cdot p_{spacing}/M\big) \qquad (2)$$

$$I_d\big(\Omega_x + 2\pi \cdot n,\ \Omega_y + 2\pi \cdot m\big) = I\big(\Omega_x \cdot M/p_{spacing},\ \Omega_y \cdot M/p_{spacing}\big) \qquad (3)$$

in which $|\Omega_x| \leqslant \pi \wedge |\Omega_y| \leqslant \pi$ and $n, m \in \mathbb{N}$. The Ω_x, Ω_y-plane represents the periodic discrete spatial frequency domain (Ω in radians per sample). Sampling at the Nyquist frequency (Oppenheim *et al* 1983) is retrieved when the folding frequency ($\Omega = \pi$) equals the cut-off frequency of the optical system ($\Omega = (p_{spacing}/M) \cdot (4\pi NA/\lambda)$). The sampling density is then twice the highest possible spatial frequency.

Figure 3 shows what happens in the Fourier domain when an object is imaged by a microscope and sampled with a CCD element. For simplicity, an imaged point object $\delta(x, y)$ is used as input image $o(x, y)$. This image has a flat Fourier spectrum. The shape of the spectrum after filtering with the OTF ($I_{OTF}(\omega_x, \omega_y)$ in figure 3(a)) equals that of the OTF. The transfer function as a results of pixel shape $B(\omega_x, \omega_y)$ is shown in figure 3(b). The pulses in figure 3(c) denote the periodicity as a result of sampling $S(\omega_x, \omega_y)$. Figure 3(d) shows the resulting spectrum $I(\omega_x, \omega_y)$. All graphs in figure 3 have the same horizontal scale, which can be expressed in the continuous space (ω-axis) and the discrete space (Ω-axis).

3. A focus criterion near focus

A robust autofocus algorithm has a focus criterion which shows a sharp peak in the focus function at the in-focus position. In the previous section we showed that the OTF only depends on the degree of defocusing and not the signal spectrum $O(\omega_x, \omega_y)$. Therefore the focus criterion should only depend on the OTF. Figure 2 shows the OTF for various degrees of defocusing. It is clear that the dependency of the OTF on Δz is small for the low frequencies and the high frequencies (around and above the cut-off frequency of the optical system). At low frequencies the OTF will always be close to 1.0 and at high frequencies the OTF will always be close to 0.0. The largest dependency on the focus error is to be found at the middle frequencies, just below $f = \omega/2\pi = NA/\lambda$. Another argument in favour of the middle frequencies is the SNR, which is far better for the middle frequencies then in the region near the cut-off frequency, where virtually no signal energy is passed.

3.1. Band-pass filters in focus criteria

Considering the above, the best focus criterion near focus is the signal power after band-pass filtering. The band-pass filter should select only that part of the spectrum which depends most on Δz. In the spectrum of the discrete image the central frequency of the band is given by: $\Omega_{bp} = (2\pi NA/\lambda) \bullet (p_{spacing}/M)$. Therefore, the focus criterion does not only depend on the microscope parameters (NA and λ) but also on the sampling density of the CCD element (M and $p_{spacing}$). The resulting focus criterion is given by:

$$F(z) = \sum_x \sum_y \left| i_z[x, y] \otimes h_{\Omega_{bp}}[x, y] \right|^2 \qquad (4)$$

in which $i_z[x, y]$ is the discrete image acquired at z-position z and $h_{\Omega_{bp}}[x, y]$ is the band-pass filter with central frequency $\Omega_{bp} = (2\pi NA/\lambda) \bullet (p_{spacing}/M)$. The width of the pass-band should be chosen small enough but is not critical (see also 4.2.). Using Parseval's relation this equation is equal to the following expression in the Fourier domain:

$$F(z) = \frac{1}{2\pi} \int_2 \int_\pi \left| I_z(\Omega_x, \Omega_y) \cdot H_{\Omega_{bp}}(\Omega_x, \Omega_y) \right|^2 d\Omega_x d\Omega_y. \qquad (5)$$

It is remarkable that none of the focus criteria found in the literature have incorporated the effects of sampling. We show that this is essential.

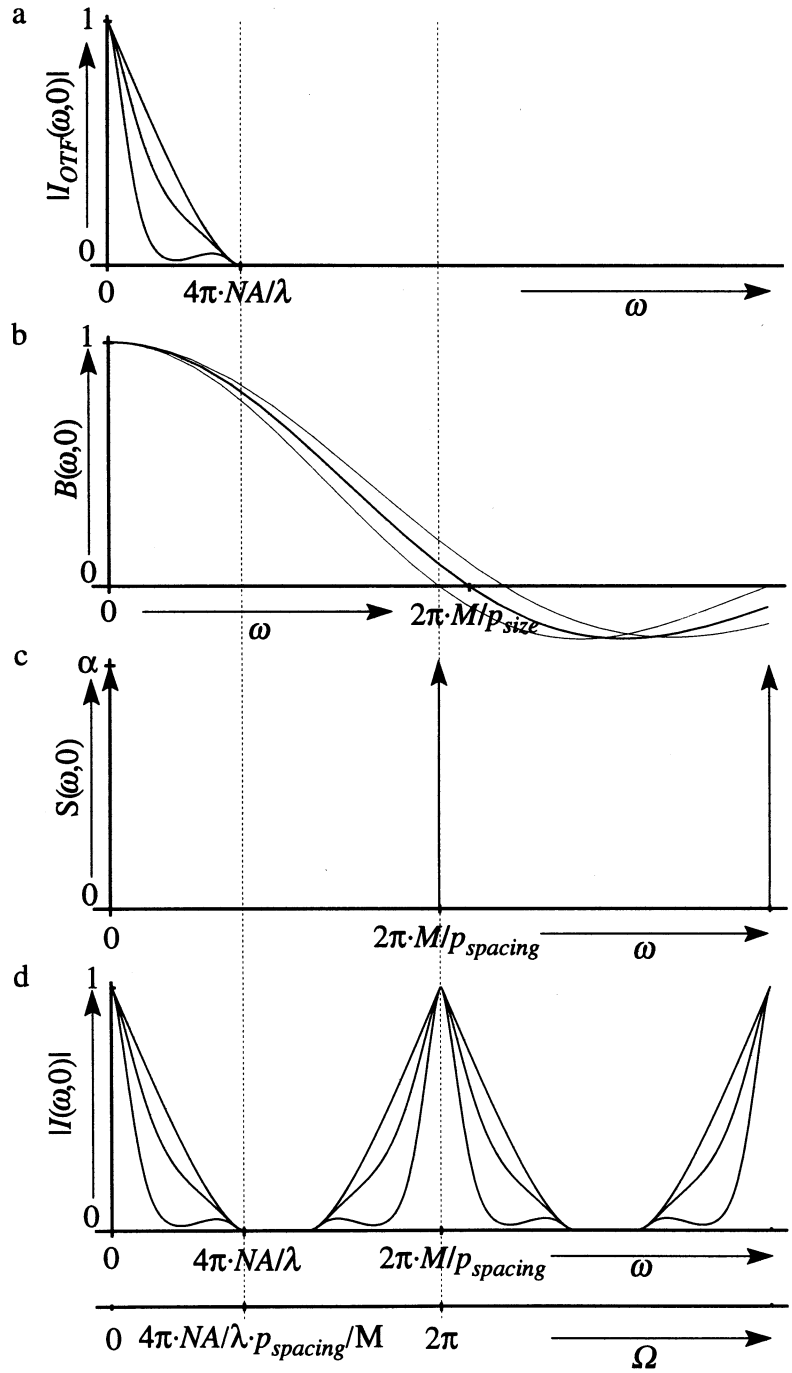

Figure 3. These four graphs show what happens in the Fourier domain when a point object ($O(\omega_x, \omega_y)$ with a flat spectrum) is imaged by a microscope and sampled with a CCD element. The projected image is sampled at 1.25 times the Nyquist frequency, with a filling factor of 83%. (a) $I_{OTF}(\omega_x, \omega_y)$ is the spectrum after filtering with the OTF ($\Delta z = 0$, 0.5 and 1 μm, $NA = 0.75$ and $\lambda = 540$ nm). (b) The transfer function $B(\omega_x, \omega_y)$ of the square aperture of the pixels with a filling factor of 83% (thin line right-side: 70%, left-side 100%). (c) The sampling function $S(\omega_x, \omega_y)$ denotes the periodicity as a result of sampling ($\alpha = (2\pi \cdot M/p_{spacing})^2$). (d) The resulting spectrum, $I(\omega_x, \omega_y)$. All graphs have the same horizontal scale, which can be expressed in the continuous space (ω-axis) and discrete space (Ω-axis).

4. Combinations of filter and sampling frequency

The power, after band-pass filtering, proposed as the focus criterion is not practical to use. The spatial representation of an arbitrary circular band-pass filter will involve many coefficients (a large filter) which makes application rather slow. A first simplification is to use only a one-dimensional filter. This filter applied to an image will select a part of all

Table 1. The transfer function of the square aperture of the pixels (b and B), and the sampling function (s and S). Both in spatial and Fourier representation.

Spatial domain	Fourier domain
$b(x, y) = \begin{cases} \frac{4 \cdot \pi^2}{p_M^2} & \lvert x\rvert < \pi \cdot p_M \wedge \lvert y\rvert < \pi \cdot p_M \\ 0 & \text{elsewhere} \end{cases}$	$B(\omega_x, \omega_y) = \operatorname{sinc}(\omega_x \cdot p_M) \cdot \operatorname{sinc}(\omega_y \cdot p_M)$
$s(x, y) = \sum_{m,n} \delta(x - m \cdot 2\pi \cdot p'_M) \cdot \delta(y - n \cdot 2\pi \cdot p'_M)$	$S(\omega_x, \omega_y) = \frac{1}{p'^2_M} \sum_{m,n} \delta(\omega_x - \frac{m}{p'_m}) \cdot \delta(\omega_y - \frac{n}{p'_M})$

$p_M = \frac{p_{size}}{2\pi \cdot M}$, $p'_M = \frac{p_{spacing}}{2\pi \cdot M}$, $n, m \in \mathbb{N}$

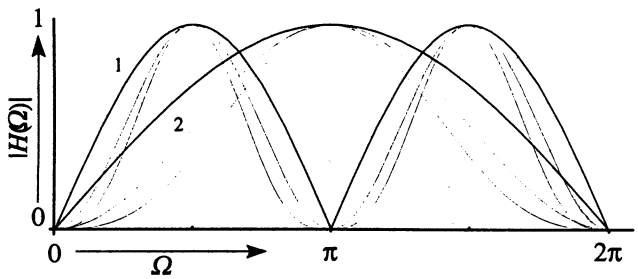

Figure 4. The normalized transfer function of the {1, 0, −1} filter (1) and of the {1, −1} filter (2). The normalized transfer function of double and triple application of the single filter are drawn in thin lines. Note that $\Omega = \pi$ is the folding frequency.

middle frequencies. But even a one-dimensional filter with an arbitrary central frequency has too many coefficients to be practical. In the literature on focus criteria two simple one-dimensional filters are used as discrete approximations of the derivative filter (Brenner *et al* 1976): the {1, −1} high-pass filter and the {1, 0, −1} band-pass filter.

4.1. Simple filters

The {1, 0, −1} filter is a band-pass filter with $\omega_{bp} = \pi/2$. This satisfies our requirements for band-pass filters in one dimension when applied to images sampled at the Nyquist frequency. For images sampled at half the Nyquist frequency the spectrum is folded around half the cutoff frequency of the optical system, $f = NA/\lambda$. The desired central frequency for the band-pass filter becomes the folding frequency. The {1, −1} filter applied to the folded spectrum sampled at half the Nyquist frequency achieves the same effect as the {1, 0, −1} filter applied to images sampled at the Nyquist frequency. The aliasing that occurs in the discrete image does not harm the focus criterion since it does not affect the signal power at the selected frequencies in the analog image. Table 2 lists the spatial representation, the magnitude of the transfer function and the corresponding focus criterion of the two filters. The normalized magnitudes of the corresponding transfer functions are drawn in figure 4.

The two combinations are depicted in figure 5. The spectra of an imaged point object (as in figure 3) are drawn for different degrees of defocusing together with the transfer functions of the filters.

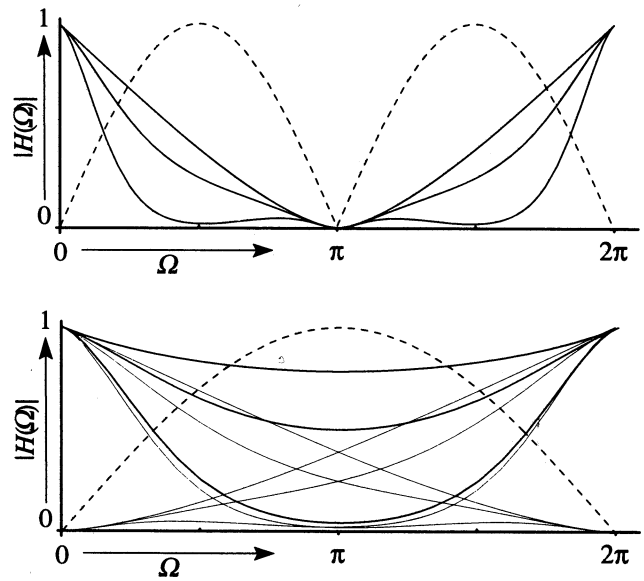

Figure 5. Combinations of filter and sampling density. Upper graph: sampling at the Nyquist frequency and the {1, 0, −1} filter. Lower graph: sampling at half the Nyquist frequency and the {1, −1} filter. Spectra are drawn for different amounts of defocusing of a point object (plain lines) together with the transfer function of the filter (dashed line). In the lower graph, the non aliased contributions of the spectrum are drawn in thin lines.

4.2. Multiple application of simple filters

It is also possible to use focus criteria which are based on the multiple application of one of the two filters. As a single application of the {1, −1} and {1, 0, −1} filter is used to approximate differentiation, multiple application of these filters can be used to approximate higher derivatives (Groen *et al* 1985, Linge *et al* 1982). Groen *et al* do not make a distinction between continuous and discrete images nor between the different discrete approximations for differentiation. They reject all focus criteria based on third and higher derivatives (discrete approximations), because Erteza (1976) argued that they were too noise sensitive for continuous images (real differentiation). Higher derivatives of continuous images (no approximation) give more weight to high frequencies as the transfer function associated with the n^{th} derivative is $(j\omega)^n$. Since the signal frequencies are limited by the OTF, noise effects will indeed dominate.

Multiple application of the {1, 0, −1} and {1, −1} filter

Table 2. Spatial representation, magnitude of the transfer function $|H_{bp}(\Omega_x)|$ and the corresponding focus criterion for the {1, −1} filter and {1, 0, −1} filter.

Filter	Spatial representation	$\|H_{bp}(\Omega_x)\|$	Corresponding focus criterion
{1, −1}	$\delta[x, y] - \delta[x-1, y]$	$2\|\sin(\Omega_x/2)\|$	$F(z) = \sum_x \sum_y (i[x, y] - i[x-1, y])^2$
{1, 0, −1}	$\delta[x+1, y] - \delta[x-1, y]$	$2\|\sin(\Omega_x)\|$	$F(z) = \sum_x \sum_y (i[x+1, y] - i[x-1, y])^2$

causes narrower band-pass filtering (see figure 4). The effect of these filters on sampled images cannot be predicted without considering the OTF and the sampling density. For example, (multiple) application of the {1, −1} filter and sampling at the Nyquist frequency yields the same problems as the (higher) derivatives applied to continuous images that are described above. Multiple application of the {1, 0, −1} filter and sampling at the Nyquist frequency, however, only narrows the width of the band-pass filter and makes the focus criterion more selective for the frequencies around $\Omega = \pi/2$. Since these frequencies are most sensitive for focus errors the resulting focus function will have an even sharper peak; (see also the following results). The same hold for the {1, −1} filter and sampling at half the Nyquist frequency.

4.3. Results

An optimal focus criterion is reached when the middle spatial frequencies passed by the optical system (analog) matches the central frequency of the band-pass filter (discrete). Considering the two simple filters, two combinations emerge:

- Sampling at the Nyquist frequency and the {1, 0, −1} filter based focus criterion.
- Sampling at half the Nyquist frequency and the {1, −1} filter based focus criterion.

For these combinations, multiple application of the filter gives a sharper focus function (i.e. more sensitive to focus errors).

These results have been verified with a set of experiments. Two sets of images of a pinhole were acquired at successive z-positions. One set was sampled at the Nyquist frequency and one set at half the Nyquist frequency. Both focus criteria (based upon the {1, −1} and {1, 0, −1} filters) were applied to both sets of images. The resulting focus functions are shown in figure 6. The graphs in the figure clearly show better focus functions for the combinations mentioned above. It is also clear that multiple application of the filters in the correct combination yields better results. They produce sharper focus functions without suffering from noise.

The lower right graph in figure 6 shows, although not as sharp, also fairly good focus functions. In this combination, the filter selects a part of the spectrum which is passed well by the OTF but does not depend as much on defocusing as around half the cut-off frequency of the optical system.

5. Three-dimensional objects

Until now, all objects were assumed to be thin so that they could be treated as two-dimensional. A three-dimensional object can be considered as a stack of infinite thin two dimensional layers. When three-dimensional objects are imaged through a microscope, the image projected onto the CCD element consists of contributions from all layers within the object.

In three dimensions, the spectrum of an image formed from an object by the OTF is given by:

$$I_{3D}(\omega_x, \omega_y, \omega_z) = O_{3D}(\omega_x, \omega_y, \omega_z) \cdot OTF_{3D}(\omega_x, \omega_y, \omega_z) \tag{6}$$

in which the subscript 3D denotes a three dimensional spectrum. O_{3D} is the spectrum of the illuminated object. The inverse Fourier transform along the z-coordinate equation (6) gives:

$$I_{stacked\ 2D}(\omega_x, \omega_y, z) = O_{stacked\ 2D}(\omega_x, \omega_y, z) \otimes_z OTF_{stacked\ 2D}(\omega_x, \omega_y, z) \tag{7}$$

in which the subscript '*stacked* 2D' denotes a two-dimensional spectrum in the spatial frequency coordinates ω_x and ω_y as a function of the spatial z-coordinate. (Note that the separability of the complex exponential in the Fourier transform, $e^{j(\omega_x x+\omega_y y+\omega_z z)} = e^{j(\omega_x x+\omega_y y)} \cdot e^{j\omega_z z}$, permits considering the inverse transform only along the z axis.) The convolution is only taken over the z variable and can be rewritten as:

$$I_{stacked\ 2D}(\omega_x, \omega_y, z) = \int O_{stacked\ 2D}(\omega_x, \omega_y, z-h) \cdot OTF_{stacked\ 2D}(\omega_x, \omega_y, h)\mathrm{d}h. \tag{8}$$

The effects of decreasing light intensity through the object (absorption) is not taken into account. The total focus value at a certain z-position can thus be seen as the sum of the focus values of all infinite thin layers of the object:

$$F_{3D\ object}(z) = \int_{object} F_{2D\ layer}(z+h)\mathrm{d}h. \tag{9}$$

The maximum in the focus function is reached at the z-position where the sum of the focus value of all the layers of the object is maximal. The detected in-focus position lies within the object, but the position depends on the three dimensional structure of the object. It is not assured that the focus function still contains only one maximum. This can easily be determined considering the focus function of objects above each other: there are two in-focus z-positions. This is true for all focus criteria and will in general not cause any problems in practice.

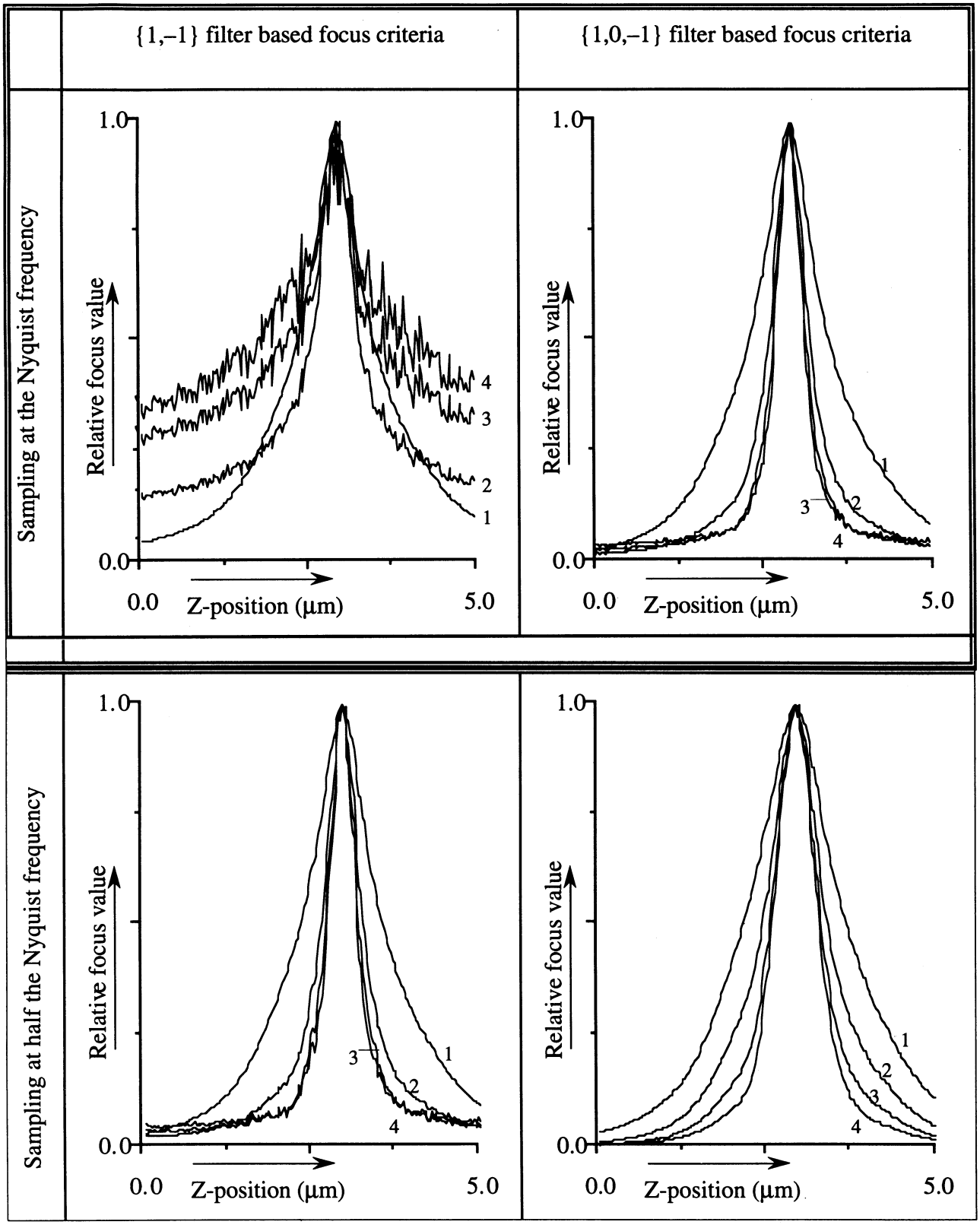

Figure 6. The effect of sampling frequency on focus functions. The two left graphs are obtained by applying the {1, −1} filter based focus criteria, the two right graphs by applying the {1, 0, −1} filter based focus criteria. The images, used to calculate the focus functions in the upper two graphs, were sampled at the Nyquist rate, the images in the lower two graphs were sampled at half the Nyquist rate. The numbers besides the curves indicate how often the filter was applied.

6. Focus criteria in practice

In the discussion on focus criteria we did not include noise. Furthermore is it not always possible to acquire images sampled at exactly (half) the Nyquist frequency. Since we use one-dimensional filters, another parameter is the sampling density perpendicular to the filter.

6.1. Noise considerations

The two combinations from section 4.3 will give identical results in the absence of noise. Due to the quantum nature of light, all images acquired contain Poisson-distributed photon noise. The SNR of the acquired image increases with an increasing number of photons per pixel in the image. For a single pixel this means that the SNR is lower for smaller pixel values. Whereas in bright field microscopy enough light can be accumulated to achieve high SNRs, fluorescence microscopy usually suffers from low light levels producing low SNR images. The Poisson noise in the image results in noise in the spectrum of the image. The noise is assumed to be uniformly distributed (white) over the spectrum of the image (Priestley 1981).

The noise contribution to the focus function (using the band-pass based focus criteria) is related to the SNR of the selected part of the spectrum. An increased SNR can be achieved in three ways:

- by increasing the intensity of the light source (not practical in fluorescence microscopy);
- by using longer exposure times, which is also not practical in fluorescence microscopy, because long exposure times slow down the process considerably and decrease the fluorescence intensity due to photo bleaching of the fluorescent molecules;
- by combining pixels—binning—to form super pixels.

Binning is a feature of the CCD elements used in this study. The charge of a selected number of adjacent pixels (in both directions) can be combined before digitization. A binning factor of two by two combines the charge of four pixels. Binning increases the number of photons per super pixel and therefore increases the SNR. It also reduces the number of pixels which speeds up the image transfer to the host computer and the calculation of the focus value.

Recalling the two combinations of filter and sampling frequency and considering the noise suppression and the data-reducing effect of binning, the combination of the $\{1, -1\}$ filter and sampling at half the Nyquist frequency (achieved by binning) is to be preferred. An experiment was done to compare the two combinations of sample densities and filters for low light levels using a 20× Nikon objective and a Photometrics camera with the KAF1400 CCD element. Using a built-in relay lens, the sampling density without binning and Nyquist frequencies are $f_{sampling} = 7.6$ pixels/μm and $f_{\text{Nyquist}} = 5.6$ pixels/μm. This shows that it is not possible to obtain a sampling density at exactly (half) the Nyquist rate by changing the binning factor. Two sets of 100 images were acquired of a pinhole at successive z-positions ($\Delta z = 25$ nm). The contrast in the images was about 50 ADU (approximately 400 photo-electrons result in 50 AD-converter units) between object and background pixels in an image without binning. One set was sampled without binning (i.e. as close as possible to the Nyquist rate), the other set was sampled with a binning factor of two (i.e. as close as possible to half the Nyquist rate). Both sets were acquired with exactly the same optical setup. The images acquired with binning contain four times fewer pixels. Figure 7 shows the resulting focus functions using the corresponding focus criteria. The figure shows that the combination of binning and $\{-1, 1\}$ filter performs better.

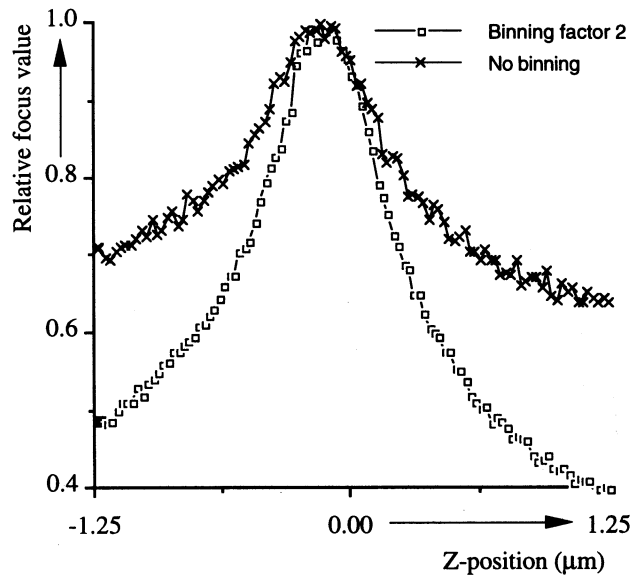

Figure 7. The effect of binning on focus functions. The figure shows two graphs: A focus function derived by applying the $\{1, 0, -1\}$ filter to images sampled at 1.36 times the Nyquist frequency and a focus function derived by applying the $\{1, -1\}$ filter to images sampled at 0.68 times the Nyquist frequency caused by a binning factor of two.

6.2. Sampling density perpendicular to the 1-D filter

In our previous examination of focus criteria we did not consider the sampling density perpendicular to the 1-D filter. Electronic binning on the CCD-element or combining pixels in the computer memory will decrease the time needed for the calculation of a focus value. Electronic binning also reduces the readout time of the image. Combining all pixels in a row perpendicular to the filter will result in a cross section of the Fourier spectrum at $\Omega_y = 0$. This one-dimensional spectrum and the corresponding filter exactly matches the criterion but only those spatial frequencies that are exactly along the filter contribute to the focus value. The result is a sharp peaked focus function that is dominated by noise. Several

Table 3. The three phases of the proposed autofocus algorithm with step sizes and average number of images needed. z is the unknown initial distance to the in-focus z-position.

Phase:	Step size:	Average number of images:
Coarse	Δz_{coarse}	$\lfloor \frac{z}{\Delta z_{coarse}} \rfloor + 2$
Fine	$\Delta z_{fine} = D_{quadratic}$	$\lfloor \frac{\Delta z_{coarse}}{4 \cdot \Delta z_{fine}} \rfloor + 1$
Refine	$\Delta z_{refine} = \frac{D_{quadratic}}{N_{samples}}$	$N_{samples}$

$[x]$ denotes the largest integer equal or less than x (floor function).

experiments with different objects have shown that the sampling density perpendicular to the filter can be reduced to a 1/8 of the Nyquist frequency. The band-pass filter then selects a part of the middle frequencies. The maximum electronic binning factor is given by the dynamic range of the analog-to-digital converter.

7. Autofocus algorithm

Now that a focus criterion has been found, an algorithm is needed to actually focus the image of the object onto the camera. Focusing has to be done as fast as possible, that is, with a minimum number of images.

In figure 8, a typical focus function is shown for a wide range of z-positions (100 μm). Three different regions can be distinguished: a near-flat region, a sloped region and a quadratic region. The flat region is characterized by a relatively low gradient towards focus. The sloped region has a distinct gradient towards focus. The quadratic region lies just around the peak of the focus function. The observation that the top matches a quadratic function can be used to find the in-focus z-position by interpolation. This has been described by (Mendelsohn and Mayall 1971) and (Mason and Green 1975). Interpolation has two advantages:

- An accurate estimation with only a few images (high speed);
- A robust estimation for noisy images by minimizing the mean square-error.

The proposed autofocus algorithm has three phases, adapted to each region of the focus function.

7.1. Three-phase autofocus algorithm

The *coarse phase* steps along the z-axis with steps of size Δz_{coarse}, sampling the focus function by acquiring an image and calculating the focus value. The first two samples are used to determine the direction towards focus. The algorithm steps in that direction until the last focus value is lower than the previous one. The in-focus z-position is then located between the z-position with the highest and second highest focus value.

The *fine phase* repeats the same procedure with a smaller steps. The starting point and initial direction towards focus are already known from the previous phase. The step size is chosen so, that the distance between two samples (Δz_{fine}) is equal to the size of the quadratic region. The *refine phase* samples the focus function with $N_{samples}$ equidistant samples around the in-focus z-position (all samples taken in the quadratic region). The final estimation of the in-focus z-position is calculated through a quadratic fit of the samples in the quadratic region. Table 3 lists the three phases with step size and average number of images needed.

Suitable values have to be found for the step size in the coarse phase, the size of the quadratic region, and the number of samples in the quadratic region. The step size in the coarse phase should be chosen as follows:

- The total number of images should be minimized;
- A step size that is too small may cause unreliable information about the direction towards focus;
- A larger step size causes more samples in the fine phase.

Furthermore, if the initial distance to focus is known to be small, it is faster to use only the last two phases. This situation exists for example. in scanning applications (see Netten *et al* 1994).

7.2. The quadratic region: size and number of samples

The choice of the size of the quadratic region is somewhat arbitrary. Tests with different objects to determine a suitable quadratic region size for a 20× objective lens ($NA = 0.77$) showed that the quadratic shape extends to 1.4 μm around focus. Assuming that the maximum wave front aberration (Hopkins 1955) in the quadratic region is the same for all objective lenses, the quadratic region sizes for other objective lenses can be derived from:

$$w_{20} \leqslant \frac{NA^2}{\eta} \cdot \frac{D_{quadratic}}{8} = 0.1\ \mu\text{m} \qquad (10)$$

in which w_{20} is the wavefront aberration due to defocusing and η is the refractive index. This gives for example with a 60× oil immersion ($\eta = 1.52$) objective lens with $NA = 1.4$, a quadratic region of 0.6 μm.

The number of samples in the quadratic region depends on the amount of noise in the focus values. For a hypothetical noise-free focus function, three samples in the quadratic region are enough to give an accurate estimate of the in-focus z-position. When the noise contributions to the

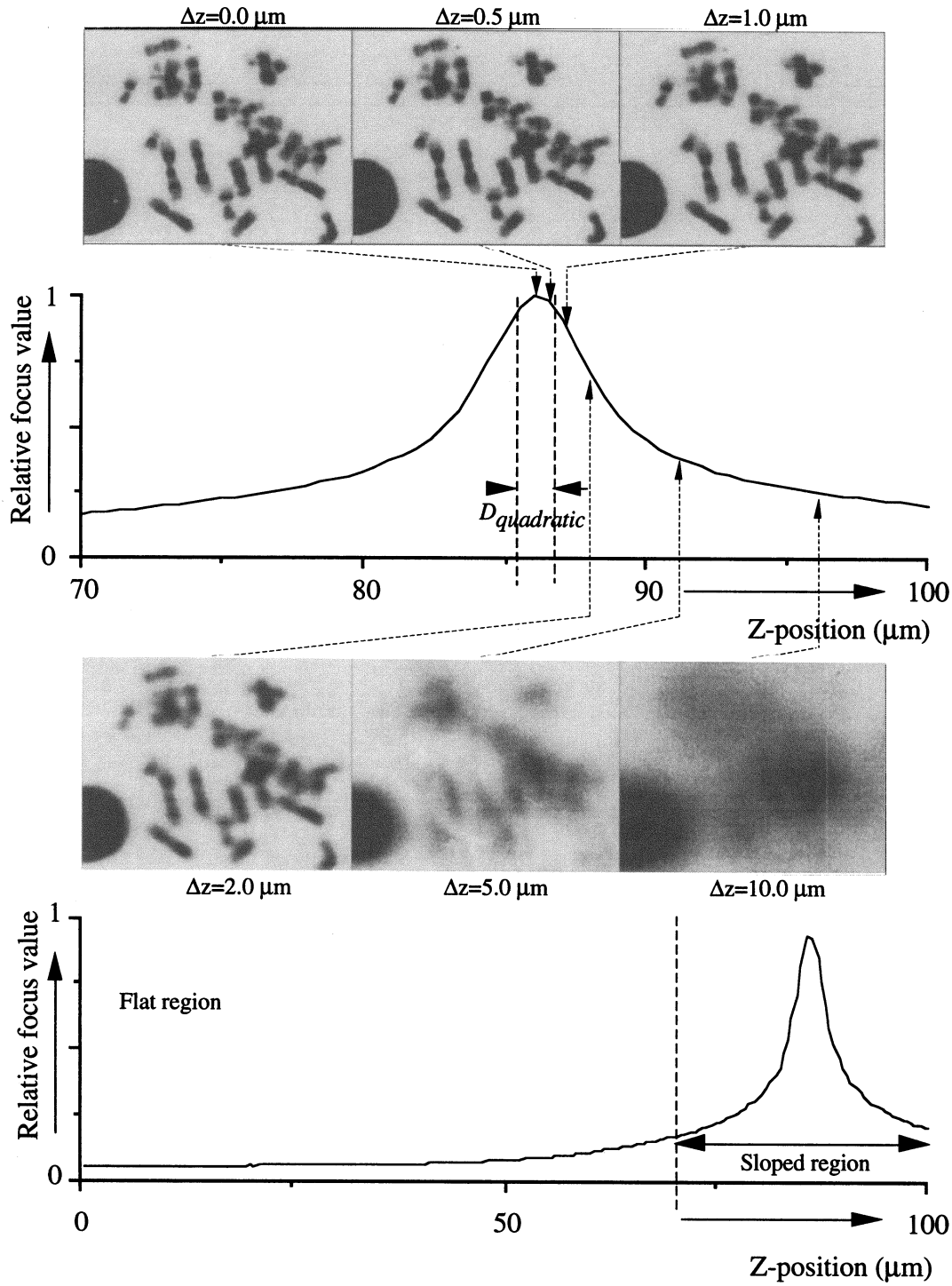

Figure 8. Different regions which can be distinguished in a focus function, near-flat, sloped, and quadratic. The focus function consists of 200 focus values calculated from images acquired at successive z-positions ($\Delta z = 0.5$ μm). The lower graph is an enlarged part of the upper graph. The accompanying images show Giemsa stained human metaphase chromosomes acquired at the indicated z-position (image size 315 × 250 pixels; setup as in section 6.1), to visualize typical defocusing.

Table 4. Typical focus error using a quadratic fit on $N_{samples}$ samples in the quadratic region of the focus function of a pinhole for different light levels.

Contrast	$N_{samples}$	Focus error
50 ADU	9	≈ 50 nm
100 ADU	5	≈ 25 nm
> 200 ADU	3	< 25 nm

focus function increase (lower light levels), more samples in the quadratic region should be taken to obtain a good estimate of the position.

Some insight into the accuracy of the final estimated in-focus position is gained by the following experiment: The quadratic region of a focus function is sampled with the smallest step size (25 nm). These samples are then used to get all possible estimates of the focus position, using the refine algorithm with $N_{samples}$ samples. The experiment was repeated for different illumination levels of the object (pinhole). Typical focus errors are listed in table 4.

8. Conclusions and discussion

Evaluation of image formation through a microscope shows that the best focus criterion without *a priori* knowledge of the imaged object, is the signal power after applying a band-pass filter that selects the frequencies most sensitive for focus errors, that is, a band-pass filter which has a ring-shaped band in the Fourier domain, passing the frequencies around $|\bar{\Omega}| = (p_{spacing}/M) \bullet (4\pi NA/\lambda)$.

Since this filter is not practical to use, we investigated the use of simple (i.e. fast) one-dimensional filters resembling this criterion. Two optimal combinations of focus criterion and sampling density have been found:

- Sampling at the Nyquist frequency and applying the $\{1, 0, -1\}$ filter based criterion
- Sampling at half the Nyquist frequency and applying the $\{1, -1\}$ filter based criterion.

Due to noise and speed considerations, the latter combination is to be preferred.

Using that criterion, we propose a focus algorithm designed in three phases, each stepping along the focal axis but with a decreasing step size. The last phase uses a quadratic fit to estimate the final in-focus position by interpolation.

Perhaps the most important conclusion which can be drawn is the necessity to include the sampling density in the evaluation of focus criteria. The microscope parameters determine the spatial frequency at which the OTF is most dependent on focus errors (just below $f = NA/\lambda$) in continuous domain. The sampling density gives the relation between the spatial frequencies in the continuous and discrete domain. Resulting in the spatial frequency in the discrete domain which is most sensitive for focus errors (just below $\Omega = (p_{spacing}/M) \bullet (4\pi NA/\lambda)$). Comparisons between filter based focus criteria cannot be made without considering the sampling density.

Acknowledgments

We would like to thank the Project SPIN-3D (NWO-MW) grant 900-538-016 and the STW Project 2987 in The Netherlands, NATO grant RG.85/0324, and Imaginetics Corporation (Naperville, Illinois, USA) for their partial support in this work.

References

Born M and Wolf E 1959 *Principles of Optics* (London: Pergamon)

Brenner J F, Drew B S, Horton J B *et al* 1976 An automated microscope for cytologic research *J. Histochem. Cytochem.* **24** 100–11

Erteza A 1976 Sharpness index and its application to focus control *Appl. Opt.* **15** 877–81

Firestone L F, Cook K *et al* 1991 Comparison of autofocus methods for automated microscopy *Cytometry* **12** 195–206

Groen F C A, Young I T and Ligthart G A 1985 A comparison of different focus functions for use in autofocus algorithms *Cytometry* **6** 81–91

Hopkins H H 1955 The frequency response of a defocused optical system *Proc. R. Soc.* A **23** 91–103

Inoué S 1986 *Video Microscopy* (New York: Plenum)

Linge H, Zimmerman H G and Neuhoff V 1982 Focus adjustments in linear systems *Proc. ISMIII* Berlin

Köhler A 1893 A new system for photomicrographic purposes *Zeitschrift für Wissenschaftl. Mikroskopie* **10** 430–40

Mason D C and Green D K 1975 Automatic focusing of a computer-controlled microscope *IEEE Trans. on Biomed. Eng.* **22** 312–7

Mendelsohn M L and Mayall B H 1971 Computer-oriented analysis of human chromosomes-III: Focus *Comput. Biol. Med.* **2** 137–50

Netten H, v Vliet L J, Boddeke F R and Young I T 1994 A fast 2D CCD scanner for fluorescence microscopy using time delayed integration *Bioimaging* accepted

Oppenheim A V, Willsky A S and Young I T 1983 *Signals and Systems* (Englewood Cliffs, NJ: Prentice-Hall)

Price J H and Gough D A 1994 Comparison of phase-contrast and fluorescence digital autofocus for scanning microscopy *Cytometry* **16** 283–97

Priestley M B 1981 *Spectral Analysis and Time Series* (London: Academic)

Williams C S and Becklund O A 1989 *Introduction to the Optical Transfer Function* (New York: Wiley)

Yeo T, Ong S H, Jayasooriah and Sinniah R 1993 Autofocusing in tissue microscopy *Image and Vision Computing* **11** 629–39

Young I T 1989 Image Fidelity: Characterizing the Imaging Transfer Function *Methods in Cell Biology* (New York: Academic) pp 1–45