On the extended depth of focus algorithms for bright field microscopy

A.G. Valdecasas[a,*], D. Marshall[b], J.M. Becerra[a], J.J. Terrero[a]

[a]*Museo Nacional de Ciencias Naturales, c/Jose Gutierrez Abascal, 2. 28006 Madrid, Spain*
[b]*Department of Computer Science, Cardiff University, P.O. Box 916, Cardiff CF24 3XF, UK*

Abstract

Microscopes offer a limited depth of focus which precludes the observation of a complete image of a three-dimensional (3D) object in a single view. Investigations, by a variety of researchers, have led to the development of extended depth of focus algorithms for serial optical slices of microscopic 3D objects in recent years. However, to date, no quantitative comparison of the different algorithms has been performed, generally leaving the evaluation to the subjective qualitative appreciation of the observer. In this paper we use three different tests for extended depth of focus algorithm evaluation and test 10 different algorithms, some of them have been adapted (by us) for a series of optical slices. However, the main contribution of the paper is a new improved algorithm for computing the extended depth of focus. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Bright field microscopy; Extended depth of focus; Image processing; Wavelets

1. Introduction

Algorithms for extended depth of focus images through digital processing have appeared regularly in the past 20 years, either applied to reflected or transmitted light systems (Ito et al., 1989; Pieper and Korpel, 1983; Sugimoto and Ichioka, 1985; Tympel, 1996; Tympel, 1997; Willis et al., 1993). Conceptually, there is no reason why any of the above extended depth of focus algorithms cannot be applied to images obtained from either reflected or transmitted light systems. It is remarkable, however, that all the evaluations of performance of the algorithms have been generally based on the subjective impression of the human observer.

Digital extended depth of focus processing methods operate on a set of optical slices that fully span the object under study. Only a small proportion of each image slice will be in focus due to the limited depth of focus of the microscope. The purpose of the extended depth of focus algorithms is to recover from each slice those pixels that are in focus and build a final single composite image from the selected in-focus pixels. There are many objects and situations where this approach could provide an efficient way of portraying such information. For instance, many taxonomic type slides deposited in museums could be made available to specialist worldwide in this new "virtual" form, whilst securing and preserving the original "material" slide (Garcia-Valdecasas et al., 1997) at the host repository. Furthermore, if one records not only the in-focus pixel (intensity) values but also the slice index from which the in-focus pixels were selected then the three-dimensional (3D) structure information of the underlying organism may be preserved, provided that certain microscope characteristics (e.g. slice step size, optical parameters) are also recorded. Thus extended depth of focus algorithms allow the incorporation of 3D information into a single two-dimensional (2D) composite image.

The main contributions of this paper are as follows:

1. Two new wavelet-based algorithms for extended depth of focus are presented which, we show in this paper, offer improved performance over existing algorithms.
2. The performance of 10 different algorithms, including our new ones, is evaluated by using three tests to evaluate the performance of extended depth of focus algorithms.

2. Materials and methods

We now describe in further detail the methods we use to compute extended depth of focus, the methodology we adopt in comparing various methods and the metrics used to perform the comparison.

All images were acquired with a Xillix Microimager

---

* Corresponding author.
E-mail addresses: antonio@albia.museo.csic.es (A.G. Valdecasas), dave@cs.cf.ac.uk (D. Marshall), josem.becerra@retemail.es (J.M. Becerra), jjose.terrero@amena.es (J.J. Terrero).
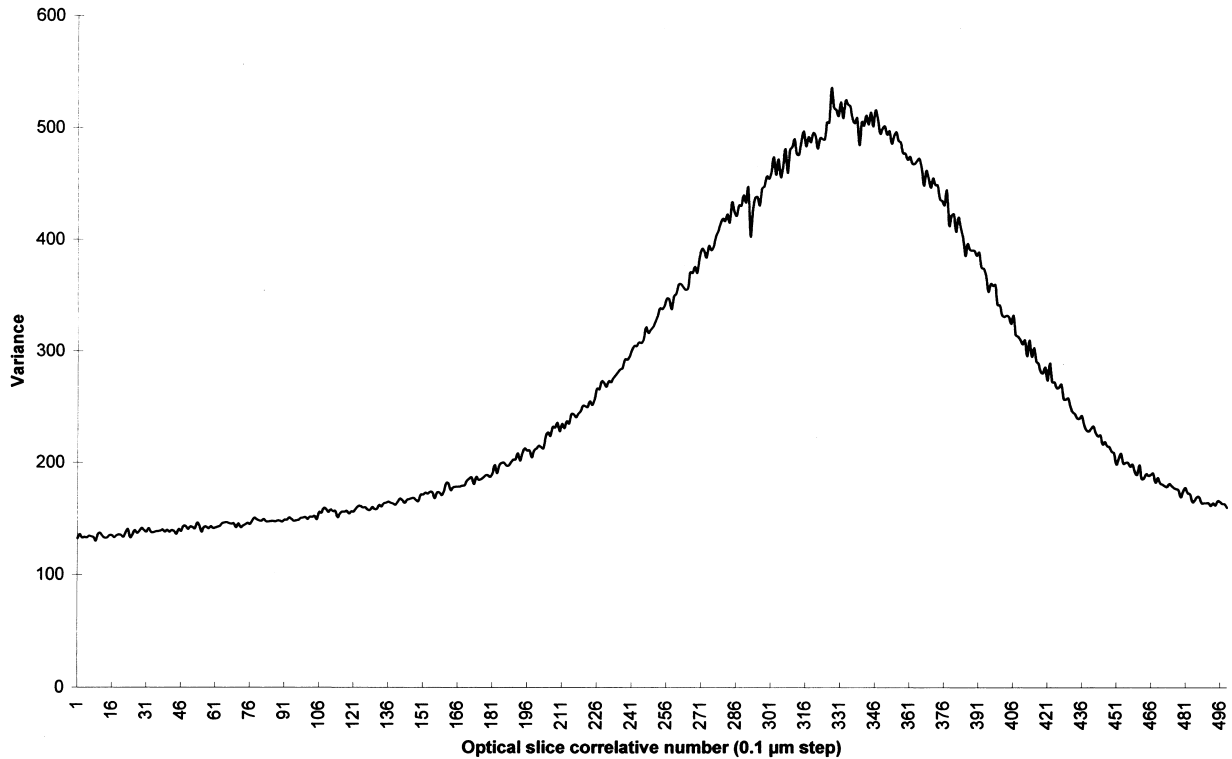
Fig. 1. Plot of the variance for a 500 image set. Peak focus is at image slice number 327.

1400-10X camera on a Zeiss Axiolab microscope interfaced to a Sun Sparc workstation. An optical axis focus controller, MFC-1 model from Applied Scientific Instruments Inc., with a declared repeatability of $\pm 0.2\ \mu m$ was used to acquire image slices at some small (e.g. $0.1\ \mu m$) step of the controller. The microscope objective was a Zeiss CP-Achromat $10 \times /0.25$ giving a final magnification of approximately $100 \times$ with a depth of field of $4.5\ \mu m$ above and below the plane of focus.

Three different tests are used to evaluate the performance of extended depth of focus algorithms. Two of the tests involve comparison to a reference object, a Zeiss stage micrometer. Stage micrometers are a relatively common subject for microscopic imaging testing (Jaggi et al., 1993; Oliva et al., 1998). In our first test we are able to determine an exact *reference plane* with which to test our methods. The stage micrometer is 1 mm long and is divided in 100 parts. The second test employs the micrometer tilted at an angle so that we can empirically examine the 3D effectiveness of our algorithms. Finally, we test our algorithms on a real 3D biological specimen.

In the first test, the stage micrometer is used as an almost ideal 2D object. There is no *above* or *below* material that would contribute blurred areas to the line's plane. Therefore, there is only one focal plane "where the light rays form a minimum spot" (Inoué and Spring, 1997). Images taken above and below this reference focal plane will show an amount of blur that increases with the distance to reference plane.

Initially five hundred 1024×1024 gray-scale images of a Zeiss stage micrometer that occupy the center of the image were taken at $0.1\ \mu m$ steps. The set of the 500 images was chosen including the focal plane of objective initially by visual evaluation. The main aim here was to establish the *reference plane* to be approximately in the middle of the set of image slices. Acquisition extended well beyond the upper and lower limits of depth of field as calculated using the methods proposed by Pluta (1988) or Willis et al. (1993). Images were further reduced by taking a 256×256 or 512×512 window from the center of the image. There are two reasons to do this. Firstly, to avoid interference due to microscope lens aberrations which are more prominent in the periphery of an image than in the center. Secondly, this improves the practical computational aspects: it is not possible to process several ultra-resolution image slices with standard computer memory limitations; and the selection of a smaller image portion also speeds computation.

Throughout the paper, the following notation is used: $I_p(i, j)$ represents the pth image slice, where $p = 1$ to P — P is the number of image slices; $i = 1$ to N — N is the horizontal array size in pixels; $j = 1$ to M — M is the vertical array size in pixels. For simplification, it is assumed that $M = N$, that is, images are squared in all our tests.

The basic approach for the test is as follows:

- Firstly, an autofocus function is used to select an image that is nearest to the focal plane of the stage micrometer. We call this the *reference image*. Two functions were

used: the variance and Tenengrad function (Yeo et al., 1993). These are applied to the whole set of 500 images in order to select the image in the set that was closest to the focal plane of the objective. These two functions have been shown to perform adequately with bright field microscopy (Yeo et al., 1993) although they may produce local maxima or side peaks with phase contrast microscopy (Oliva et al., 1998; Price and Gough, 1994).The variance function is:

$$V(I) = \sum_i \sum_j [I(i,j) - \mu]^2$$

where μ is the mean gray level given by:

$$\mu = \frac{1}{N^2} \sum_i \sum_j I(i,j)$$

the slice with the highest variance was selected. The Tenengrad function is:

$$T(I) = \sum_i \sum_j [S(i,j)]^2$$

and S is a Sobel operator:

$$S(i,j) = \sqrt{G_i^2(i,j) + G_j^2(i,j)}$$

with the i kernel

$$\begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

and the j kernel

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

Two windows were used with the variance measure: the whole (256×256), image, and a centered window of 40×40. The Tenengrad function was applied with a centered window of 40×40 pixels due to limitations of the hardware to process larger windows. In both cases the reduced 40×40 window was set in the center of the image, and was applied to the complete set of images. Both functions selected the same image, which was the optical slice number 327. Fig. 1 shows the plot of variance values for the 500 images set where the maximum corresponds to the 327 image. Similar plots were described by Yeo et al. (1993) for their image data set. This image was the one used as reference image against which all the resultant depth of focus images were compared.

• Secondly, in order to effectively test our algorithms a selection of images above and below the reference image (and including it) are used to test the performance of extended depth of focus algorithms for their ability to select the pixels of the reference image.

• Thirdly, the reference image and the composited extended depth of focus image are compared using the mean square error of the difference between the reference image and the depth of focus image:

$$\rho = \sqrt{\frac{\sum_{i=1}^{N} \sum_{j=1}^{N} [I_{\text{ref}}(i,j) - I_{\text{ext}}(i,j)]^2}{N^2}}$$

where I_{ref} is the reference image and I_{ext} is the extended depth of focus image. I_{ref} and I_{ext} represent the (pixel) array of intensity of brightness values in each image, the index $I_{\text{ref}}(i,j)$, for example indicates a reference to the pixel at location i,j in the image. N is the dimension of the image, in our cases $N = 256$ or 512. This is a common measure and Li et al. (1995) used it for the evaluation of their wavelet algorithms. As our new algorithm suggests improvements over those of Li et al. it is useful to directly compare our results with those of Li et al.

The second test used a second set of 20 images made using the same stage micrometer tilted approximately 25° over the microscope stage plane. Twenty images were taken sequentially at 1 μm steps. All the extended depth of focus algorithms were run over the set, and results were obtained by noting the selected pixel's slice location and also compared visually.

A third and final evaluation test consisted of a visual comparison among the resultant images after processing a set of images of the palp of a water mite (*Piona* sp.). This set is comprised of 20 images at 3 μm steps.

2.1. Depth of focus algorithms

Algorithms for extended depth of focus can be classified as those which work on a point process basis, those which work on an area process basis and those which work in frequency space of a transformed image (Fourier or wavelet transforms) (Gonzalez and Woods, 1992). These form the three major classes of image processing approaches to the solution of such problems. The algorithms selected for implementation represent a large proportion of contemporary methods available from the appropriate scientific literature. We implemented, modified for our purposes and tested eight existing methods. Based on our early experiences we have developed two further methods, which bring together the best features of the studied methods. Ultimately we conclude that one of these methods performs the best. However, to ease explanation and further discussion we still include our other method as it forms part of our final solution.

The basic approach of all the algorithms is as follows:

1. Initially a series of p images slices, I_p, are obtained by serially stepping the microscope through a fixed small step using the optical axis focus controller.
2. The aim is to assemble a composite image by returning a selected pixel at every pixel location or small image

region depending on the algorithm (the size of the composite image is the same in all the image slices) corresponding the best in-focus material across all slices at that pixel location. For the sake of clarity in the following discussion we refer simply to selecting appropriate pixels, the selection of small image areas should be interchanged for the respective algorithms.

3. For all image slices $I_k : k = 1...p$, every pixel is scanned and for every pixel $(I_p(i,j)\ i = 1...N$ and $j = 1...M$ for an image of dimension $N \times M$. In our case $M = N = 256$ or 512) location a measure is made for every slice at the corresponding location on how in-focus the slice region appears at the pixel location.

4. The selected pixel is chosen by the evaluation of function that returns some measure of likelihood of the pixel being in focus. The main difference in the algorithms below is in how they formulate this measure. Some use a point process, others work on an area and others work in the frequency domain.

Thus essentially every pixel is scanned and for every pixel location a measure is made for every slice at the corresponding location. The candidate pixel with the best measure is then selected, according to criteria laid out for each algorithm below. For further details on justification of the algorithms tested, refer to the original references.

The following algorithms were tested.

1. A function evaluating the distance from the mean of maximum and minimum values in a series of slices. This is referred to as index Q in Pieper and Korpel (1983) which uses the $|\text{maximum} - \text{average}| - |\text{minimum} - \text{average}|$ as the decision function:

$$Q(i,j) = |I_{\max}(i,j) - \overline{I_p(i,j)}| - |I_{\min}(i,j) - \overline{I_p(i,j)}|$$

where

$$I_{\max}(i,j) = \text{maximum}[I_1(i,j), I_2(i,j), ...I_p(i,j)],$$

$$I_{\min}(i,j) = \text{minimum}[I_1(i,j), I_2(i,j), ...I_p(i,j)]$$

and $\overline{I_p(i,j)}$ is the intensity average taken over the set of image slices. If $Q(i,j) \geq 0$ then the value chosen for the composite image will be $I_{\max}$ otherwise if $Q(i,j) < 0$ then $I_{\min}$. The composite image is built with the $I_{\max}$ or $I_{\min}$ selected pixels from the set of slices for each image pixel.

2. A nondirectional difference operator, index D in Pieper and Korpel (1983). For every corresponding pixel location across all k slices, $k = 1...p$, a measure D_k is computed:

$$D_k(i,j) = |I_k(i-1,j+1) - I_k(i+1,j-1)|$$

$$+ |I_k(i+1,j+1) - I_k(i-1,j-1)|$$

$$+ |I_k(i,j+1) - I_k(i,j-1)|$$

$$+ |I_k(i-1,j) - I_k(i+1,j)|$$

The image closest to focus at pixel (i,j) will satisfy:

$$D_{\max}(i,j) = \text{maximum}[D_1(i,j), D_2(i,j), ...D_k(i,j), ...D_p(i,j)]$$

The composite image is built in a manner similar to algorithm 1, where the pixel corresponding to the selected $D_{\max}$ value is selected.

This algorithm is similar to an algorithm presented by Wall et al. (1981). In both cases the procedure implies taking the differences between the gray levels of pairs on nearby pixels in each optical slice. Wall et al. divide the image in square subsections and calculate a measure of gray level change using the difference of two pairs of pixels at a time. Pieper and Korpel do not subdivide the optical slice in sections but calculate the gray level difference based on four pair of pixels at a time. Differences are compared for corresponding coordinates in all the slices in the stack and the set that present the higher difference is selected for the composite image. Pieper and Korpel's algorithm does not require the arbitrary division of images in subregions and was therefore selected for this testing.

3. The Sobel operator (e.g. Castleman, 1996; Gonzalez and Woods, 1992), $S(i,j)$, as explained earlier in relation to the Tenengrad function is employed to select the best in-focus slice. Essentially the Sobel operator is an edge detector with some noise smoothing incorporated. The Sobel operator returns a measure of the strength of an edge being present at a given pixel. Clearly in-focus regions will have strong edges present. The composite image is composed by selecting the pixel with the corresponding strongest (maximum) edge value through the set of slices for the pixel coordinates. The process is repeated for every pixel in the composite image.

4. The Tympel (1996, 1997) selection algorithm. This is an area algorithm based on a 5×5 kernel which selects pixels for the composite image in terms of a maximum function of local area focus.

$$V_{\text{ext}}(i,j,p) = \frac{\sum_{l=-2}^{+2} \sum_{m=-2}^{+2} (V_h(i,j,p) - V(i+l, j+m, p))^2}{V_h^2(i,j,p)}$$

$$V_h(i,j,p) = \frac{1}{25} \sum_{l=-2}^{+2} \sum_{m=-2}^{+2} V(i+l, j+m, p)$$

The maximum of $V_{\text{ext}}(i,j,p)$ fixes the position of the focused area (Tympel, 1997) and contributes to the composite image by selecting the image pixel value corresponding to the position of the maximum $V_{\text{ext}}(i,j,p)$ for every pixel in the image.

5. The focusing algorithm of Wu et al. (1996) is a two-step procedure. Images are partitioned into square blocks of 8×8 pixel size. Each image is high pass filtered by $(\delta(i,j) - h_{g_{\text{tr}}}(i,j))$, where δ, the 2D discrete delta function $\delta(i,j)$ is 1 if $(i,j) = (0,0)$ and 0 otherwise, and $h_{g_{\text{tr}}}(i,j)$ is a truncated

Gaussian function defined as:

$$h_{g_{tr}}(i,j) = \begin{cases} C \exp\left(-\dfrac{i^2 + j^2}{\sigma_g^2}\right) & \text{for } |i| \leq l_1 \text{ and } |j| \leq l_2 \\ 0 & \text{otherwise} \end{cases}$$

where C is a normalization constant, σ_g^2 is the variance of the Gaussian distribution and l_1 and l_2 are the integers to restrict the nonzero region of $h_g(i,j)$. Wu et al. suggested $\sigma_g^2 = 20.0$ and $l_1 = l_2 = 7$. We have found that a value of $\sigma_g^2 = 10.0$ and $l_1 = l_2 = 7$ gives better results for our test images. The energy index in the (i,j)th block is given by:

$$e_{p,(x,y)} = \sum_{i=xM}^{(x+1)M-1} \sum_{j=yM}^{(y+1)M-1} \left(C(i,j;l_1,l_2;S) \right.$$

$$\times \sum_{m_1=-(l_1-1)/2}^{(l_1-1)/2} \sum_{m_2=-(l_2-1)/2}^{(l_2-1)/2} h_{g_{tr}}(m_1,m_2)$$

$$\left. \times I_p(i-m_1, j-m_2) \right)^2$$

where:

$$C(i,j;l_1,l_2;S) = \left(\sum_{m_1=-(l_1-1)/2}^{(l_1-1)/2} \sum_{m_2=-(l_2-1)/2}^{(l_2-1)/2} h_{g_{tr}}(m_1,m_2) \right)^{-1}$$

where S is defined in the image domain as:

$$S = \{(i,j)|0 \leq i \langle M; 0 \leq j \langle M\}$$

and $m_1 = m_2 = 8$

In-focus areas in an image clearly give rise to sharp edges in an image. When transformed into frequency space sharp edges give rise to high frequency components. Therefore the effect of high pass filtering is to attempt to remove (low frequency) out of focus data. The block with the maximum energy in each subset of corresponding blocks in the stack of images is selected for the composite image. The Wu algorithm derives its methods from Fourier theory even though the actual implementation operates in the spatial domain.

Alternative frequency analysis methods have utilized wavelets. The next five algorithms, 6–10, are variations based upon the wavelet transform methods initially proposed by Burt and Lolczynski (1993) and Li et al. (1995), the final 2 being new improved versions. The wavelet algorithms of Burt and Lolczynski and Li et al. were originally described for a single pair of images, and we have, therefore, needed to modify the basic algorithms to deal with a set of several images (typically 30–60 slices). First, the input image slices are decomposed with a discrete wavelet transform and extended depth of focus follows by the application of an area-based maximum selection rule and a consistency verification step. Different versions of the two last steps essentially distinguish the five different algorithms. Li et al. (1995) should be consulted for a detailed justification of the fusion wavelet-based algorithm.

Readers wishing to find more background information to the wavelet transforms applied to image processing can find good introductions in Castleman (1996), Parker (1997), Prasad and Iyengar (1997) and Rioul and Veterli (1991). Computer code is also widely published and routines used in our implementations may be found in Press et al. (1986) and Parker (1997). The term wavelet means *small wave*. Therefore any function that is a wavelet must be a wave — it has some periodicity and it must be small — the amplitude of the function decreases as function of distance from its center. These properties mean that a wavelet is local — it only has a significant value in a small region of space — and it has frequency characteristics similar to the Fourier transform. A distinguishing feature of the wavelet transform over the Fourier transform is that it also has a scale property. The above properties are both exploited in the extended depth of focus processing. Like the Fourier-based approach of Wu (algorithm 5) we exploit the frequency properties of the wavelet transform to perform high pass filtering. Recall that in-focus material will give rise to high-frequency data in the transform. The scale property means that we can perform our measures of focus (algorithms 6–10) in a localized area of the image (*a window*). We now explain the basic wavelet transform and the extended depth of focus algorithms with a little more detail, for complete implementation details of the wavelet transform readers should consult Press et al. (1986) or Parker (1997). In order to perform a wavelet transform of an input signal, a family of wavelets is created from dilations and translations of a prototype or mother wavelet ψ that for a one-dimensional function in the discrete version can be defined by:

$$\psi_{m,n}(i) = 2^{-m/2}\psi(2^{-m}i - n)$$

at resolutions separated by a factor of 2 and m and n being integers. The wavelet decomposition is then:

$$f(i) = \sum_{m,n} c_{m,n}\psi_{m,n}(i)$$

and ψ chosen such that $\psi_{m,n}(i)$ constitute an orthonormal basis.

Similarly, a scaling function is introduced as:

$$\phi_{m,n}(i) = 2^{-m/2}\phi(2^{-m}i - n)$$

Again, the $\phi_{m,n}(i)$ are orthonormal for fixed m.

In the case of a 2D function (an image) the scaling function is:

$$\phi(i,j) = \phi(i)\phi(j)$$

The decomposition at each resolution is given by three wavelets:

$$\psi^1(i,j) = \phi(i)\psi(j)$$

$$\psi^2(i,j) = \psi(i)\phi(j)$$
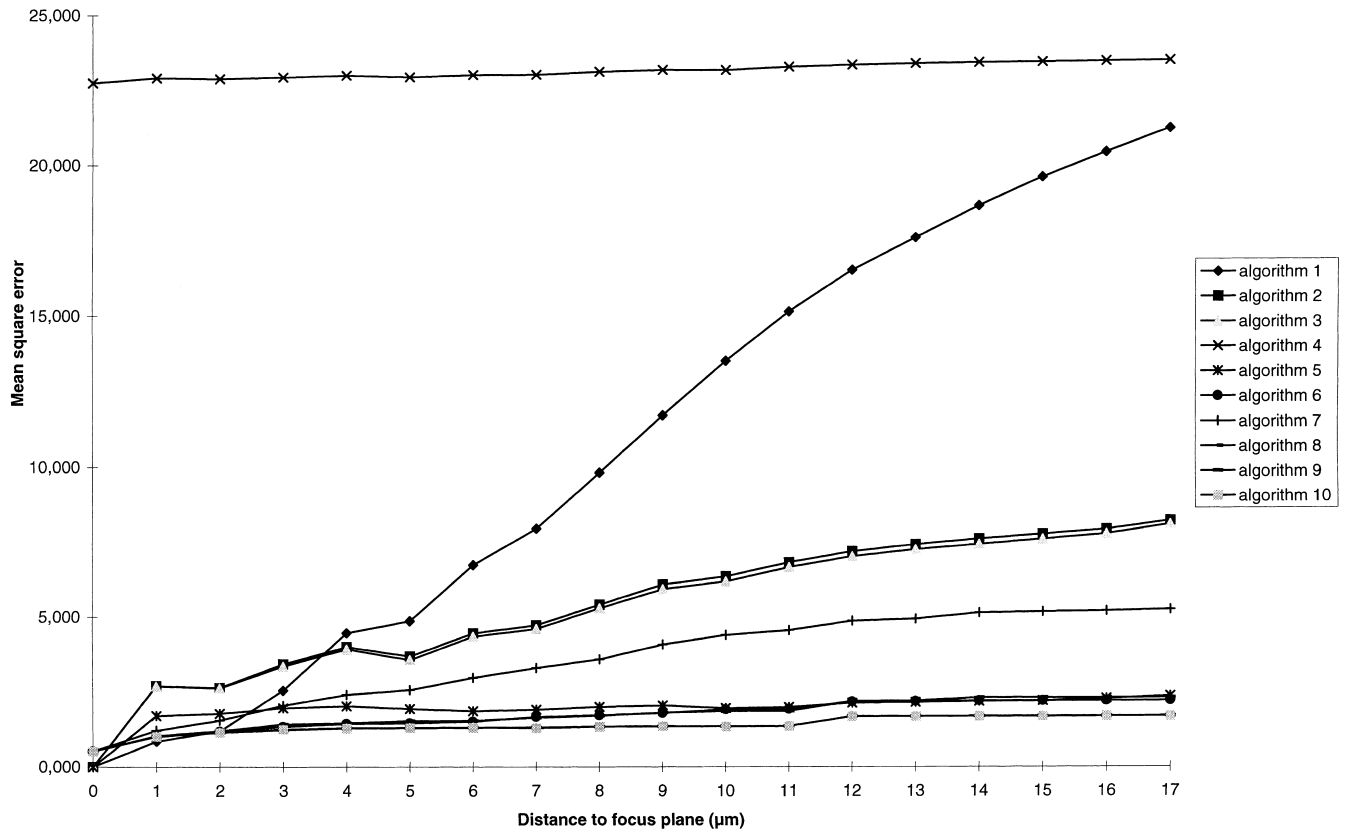
$$\psi^3(i,j) = \psi(i)\psi(j)$$

Fig. 2. Mean square error for 10 extended depth of focus algorithms on a set of images at different distances of the peak focus image slice.

The set of wavelet coefficients typically used are Harr, Daubechies 4, 12 and 20, and coefficients and code for such basis functions are readily published (Press et al., 1986). Harr wavelet basis functions are frequently used in image compression algorithms. However, the Daubechies functions offer a smooth (orthonormal and compact) basis for the computation of wavelet transforms. Efficient implementations of this format using scaling function sequences or quadrature mirror filters exist. In our examples we use a Daubechies 4 basis function which has been popular in many image processing applications and also fits into the classes of filters used by Li et al.

Individual steps in the production of the extended depth of focus algorithm are now given with these wavelet methods and are included under each specific algorithm which we now briefly outline.

6. In the Burt and Lolczynski (1993) method, the maximum activity measure is made by computing the variance over a 5×5 pixel window. If the measure at the corresponding center pixels of neighboring slices is close (has very small variance; we chose a variance of a value within the range $-0.5...0.5$) then an average value is selected, otherwise the maximum value is chosen. The use of the wavelet transform and the computation of a variance can be considered as a nonlinear high pass filter. Thus, the wavelet transform algorithms provide effective methods for determining extended depth of focus features where in-focus features correspond to high-frequency components of the transform.

7. In the Li et al. (1995) method, a wavelet maximal function which computes the maximum absolute value of (wavelet) image values within a 5×5 pixel window for corresponding pixel location at different slices is used as the selection criteria.

8. In the modified wavelet function from Li et al. (1995), a binary decision map was introduced to help select cohesive regions of in-focus from a pair of slices. This approach has been modified to allow for p image slices in the following manner: The maximum absolute value is computed as in algorithm 7 above. The results of the selection are then stored in a decision map that records the selected slice number. If the majority of the neighboring pixels (decided by a simple voting scheme over a small dimension window) come from a different slice than that of the center pixel then the center pixel (image) value is changed to the value of the corresponding pixel value from the majority slice, determined from the voting scheme.

9. We have developed a new method based on a modified version Burt and Lolczynski (1993) method. The maximum activity measure is still determined by computing the variance over a small dimension window. However, ultimate selection of the representative value (center pixel) is made using an adaptation of the modified voting scheme of Li et al. described above in algorithm 8.
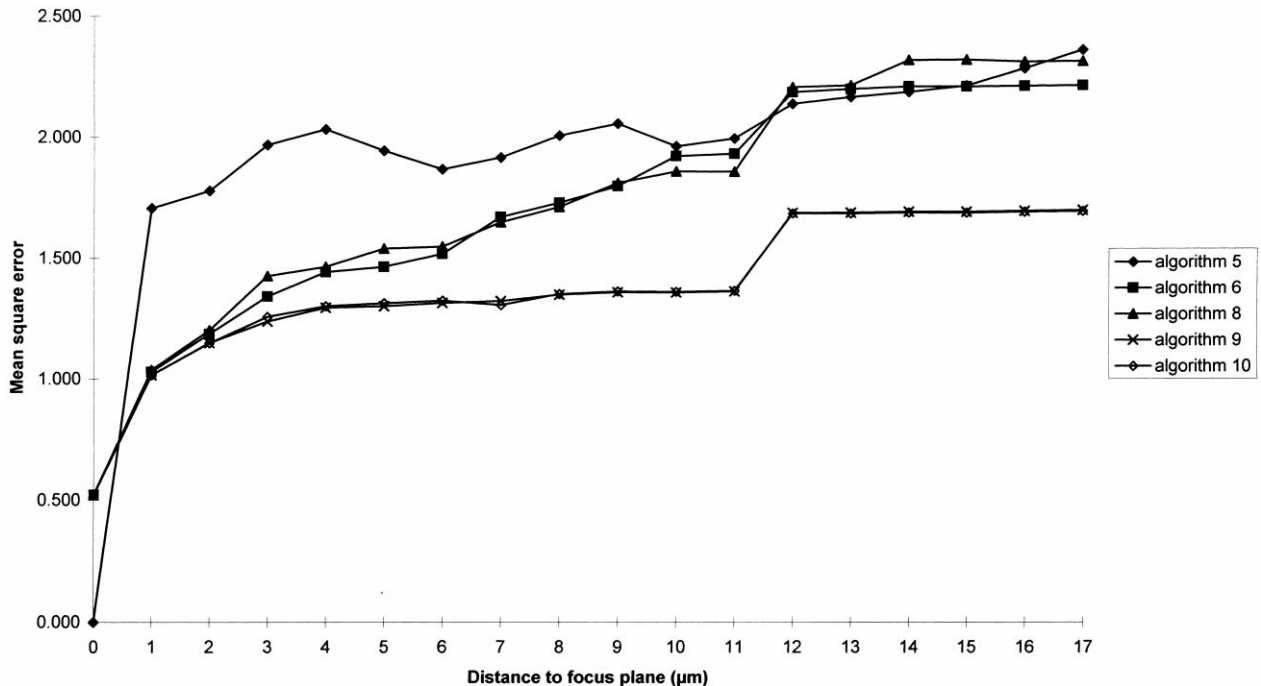
Fig. 3. Mean square error for the five best extended depth of focus algorithms on a set of images at different distances of the peak focus image slice and on an expanded *y* axis.

10. Finally we present a final new method that is essentially a hybrid of the best ideas of the versions of Burt and Lolczynski (1993) and Li et al. (1995). In our tests (see Section 3 below) this hybrid method has the best overall performance. As mentioned above the wavelet transform and appropriate selection function can be regarded as a high pass filter. Li et al. argue that their measure is more effective in selecting dominant in-focus features, and the voting scheme they developed is efficient in selecting appropriate features. We have shown this by applying such a scheme to the basic Burt and Lolczynski (algorithm 9). However, we have found that the Li et al. algorithm is occasionally biased toward other high-frequency components (e.g. noise and other image artifacts) and may, in the presence of a not too dominant feature, make an inappropriate selection. We, therefore, propose a hybrid approach fusing together the approaches of Burt and Lolczynski and Li et al. We compute the respective activity measures as described above (algorithms 6 and 7) and apply an independent vote count for each method as described in algorithms 8 and 9. The representative final (center pixel) value is then selected as the pixel that comes from the slice which has the most votes from both counts, considered independently. Thus, a dominant feature in both counts will always be selected (as is the case with detailed organism structure in sharp focus) where an anomaly between the counts arise (as is the case in areas of less well-defined organism structure in focus) the best representative feature is selected according to the above criteria. This is a selection criterion that works well in improving the effectiveness of the methods in our application domain.

The extra overheads in allowing for such modified voting scheme are minimal in terms of time and memory considerations.

3. Results

First, the results of the quantitative evaluation of the algorithms will be presented and then their performance with biological objects will be assessed.

3.1. Test 1

In this test we control the amount of out of focus material that is presented to the algorithms. We have previously established the exact position of the *reference image* slice (slice index 327). Our experiments summarized in Figs. 2 and 3 show the mean square error at increasing distance from the plane of focus for the 10 algorithms and for the five best, respectively. The "Distance to the focus plane" on the *x* axis of the figures represents adding simultaneously out of focus material above and below the reference plane: each increment step in the *x* axis includes image slices up to the specified distance in micrometers. Thus at a 1 μm step we include the plane of focus (image 327) plus one above and one below at 1 μm, respectively (images 317 and 337, in this instance). For a 2 μm step we include those images at 1 μm from the reference plane plus those at 2 μm above and below the plane of focus (the set is formed by the following images: 307, 317, 327, 337 and 347). And so on. As we move further in distance from the reference image we obviously increase the number of images to be processed.
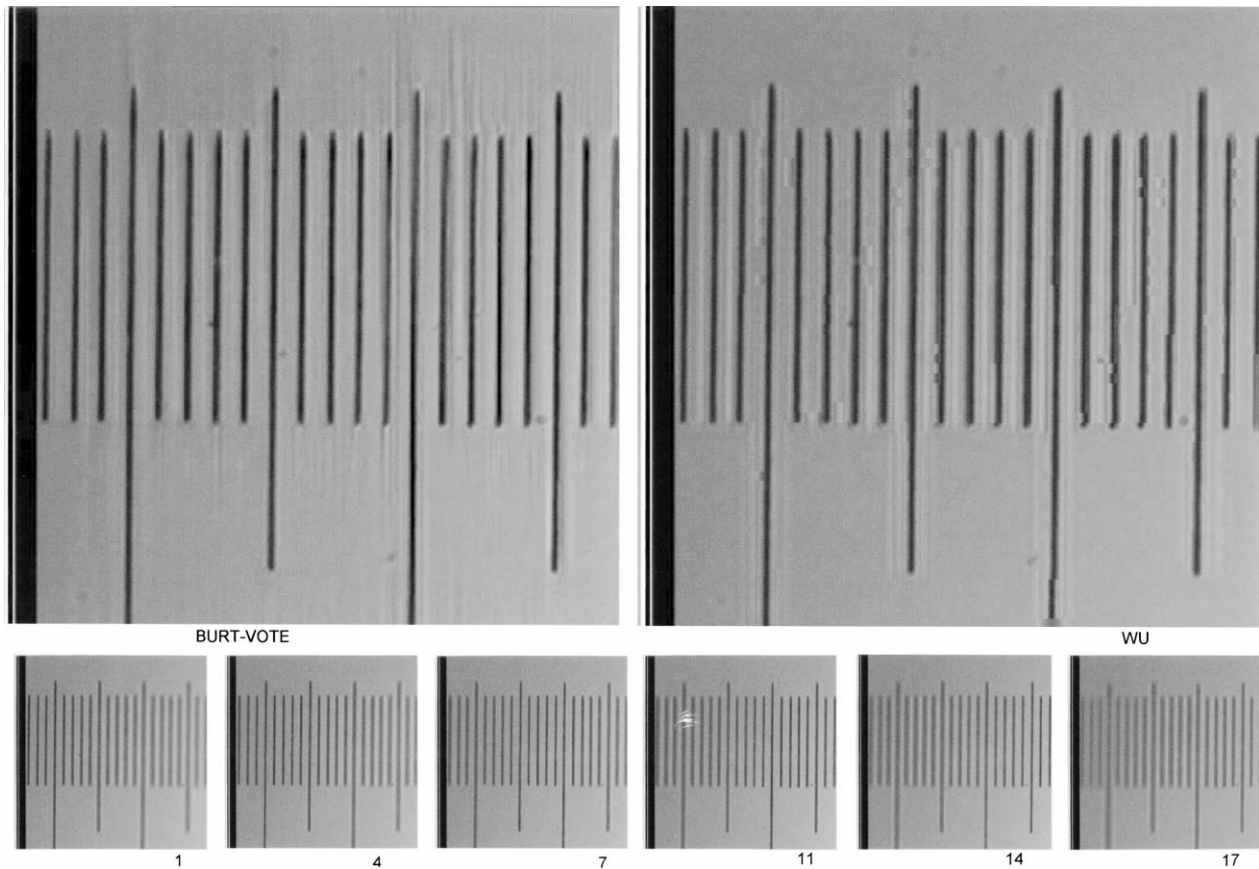
Fig. 4. Results on tilted stage micrometer of algorithm 10 (Burt-Vote) and algorithm 5 (Wu). The lower row is a sample of the testing image set.

This imposes another test on the algorithms in terms of computational speed and memory requirements.

One minor difference in image selection is for the distance "null" where we chose a set of three optical slices that were all of the same reference image slice. This was used to calculate the "residual" error of the algorithms. Only the wavelet functions and Tympel algorithm gave values different from zero with this set. This means that these algorithms introduce some image artifacts into the final image. For the Tympel algorithm the residual error introduced is initially poor and remains relatively uniform. For wavelet methods one would expect some artifacts to be introduced by their nature of processing, however we show that wavelet methods are generally the best-performing methods in the presence of out of focus material. Our main emphasis is naturally on the algorithm's ability to accurately choose correct in-focus portions of the image.

The best-performing algorithms, i.e. the ones that produce the lowest mean square error, are the modified Burt and Lolczynski (algorithm 9) and the decision sorting (algorithm 10) wavelet fusion methods. Both perform almost equally well, algorithm 10 having a slight advantage when more slices are used above and below the slice nearer the plane of focus. This clearly shows that the addition of the voting scheme, which accounts for area coherence, has a marked improvement on the original Burt and Lolczynski

method (algorithm 6). The best performance of algorithm 10 with images that extend beyond the limit of the best focus (sets 12 and onward) can be further explained as the additional (with respect to algorithm 9) dominant feature selection of Li et al. contributes a slight but noteworthy improvement. The modified algorithms 9 and 10 show a significant improvement on the original methods (algorithms 6–8). Li et al. report values for the mean square error of 3.279 for their best fusion algorithm, the one that included an area-selection rule and consistency verification. It should be noted that this value was obtained for a two-image fusion over a manually fused reference. Our test has been done on a more complex set, a higher number of images to filter and fuse and still our values for the mean square error are lower. It is notable that this wavelet fusion algorithm procedure keeps its performance to a very high level without doubling its initial error with a large slice selection. These wavelet algorithms do not deteriorate heavily with increasing distance from the plane of focus. The Wu et al. algorithm performs close to the modified wavelet algorithms. The worst-performing algorithm is that described by Tympel, which begins at a poor performance level and, although it does not degrade very much with distance, it is not very useful due to its poor initial starting point. All the other algorithms performed correctly (as expected) at short distances from the target but deteriorate quickly with
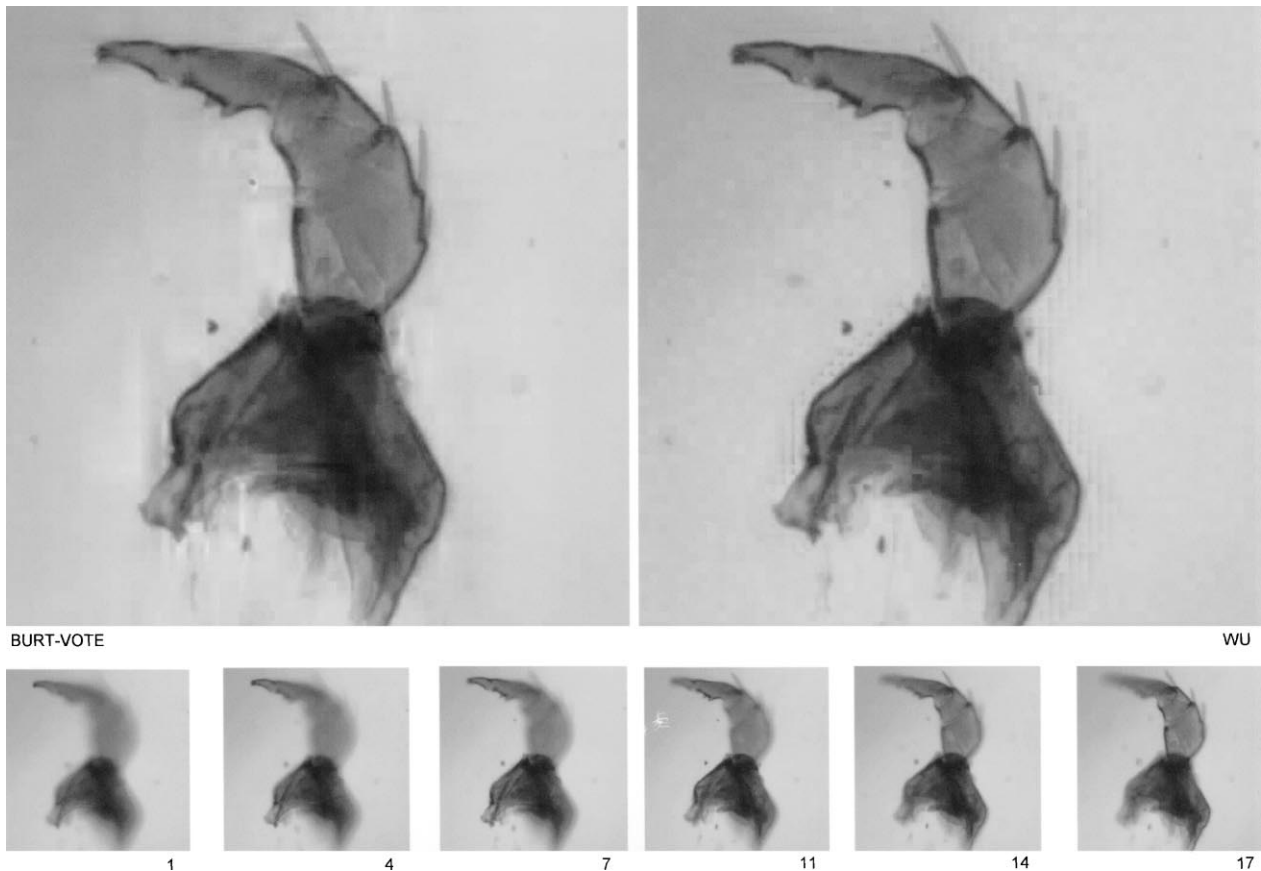
BURT-VOTE

WU

1 4 7 11 14 17

Fig. 5. Results on a palp of water mite *Piona* sp of algorithm 10 (Burt-Vote) and algorithm 5 (Wu). The lower row is a sample of the testing image set.

increasing distance. The best-performing algorithm of those based in pixel and area processing in the spatial domain is the Sobel operator (algorithm 3) but this does not compare very favorably with the frequency-based methods.

3.2. Test 2

The processing of the set of images of a tilted slide reinforce the previous results. Behavior of all the algorithms was similar as with the 2D test target. Fig. 4 shows a detailed view of the result using our algorithm 10 (labeled Burt-Vote in the figure), which is representative of the wavelet-based algorithms and that of Wu et al. (algorithm 5, labeled Wu) which is the best-performing (nonwavelet) algorithm (Test 1) in the other subgroup of algorithms. Below there is a row of images from the set used as test. It is visually clear that algorithm 10 performs much better that algorithm 5, which yields some clearly poor image artifacts. This can be further verified by recording and analyzing the pixel slices indexes selected by the algorithms. Whilst it is ultimately difficult to assess the exact empirical nature of the algorithms as in regions of no detail (focus) there is no definite notion of a perfect selection, and in many cases the same feature is in focus, in adjacent slices some notion of the algorithms' performance can be gained. By looking at pixels on the

actual micrometer rule we note a definite steady linear selection of pixels across the slices.

3.3. Test 3

A test with a 3D object is presented in Fig. 5 which shows the results of algorithms 10 and 5. As in Fig. 4 the image row below is a subset of the images used for the test. The resulting images are shown in their original size because slight differences between both can only be discerned at this resolution. Again, algorithm 10 produces a better resulting image with fewer artifacts than algorithm 5.

The last two tests confirm the numerical results obtained with the stage micrometer as a reference object.

4. Discussion

Extended depth of focus algorithms are a promising tool for biological microscopy. Simple pixel by pixel or area algorithms have had some success in certain areas (Tieman et al., 1986), but the true potential is still waiting for a more generalized usage. Computer limitations (processor speed, memory, storage and bandwidth overheads, in particular) have been the main reason in the past for their restricted applications. These limitations have almost disappeared nowadays.

The analysis of depth of focus algorithms with a simple target allows a clear comparison of robustness of the depth of focus algorithms. The main objective of this test is that the function used to select the reference image is robust enough to point to that which is nearest to the focal plane of the objective. It is expected that the robust extended depth of focus algorithms will select the majority of pixels from the reference image, and the higher this number the better. Only the algorithm that is based on a point process procedure (algorithm 1) performs worst. This is an expected behavior as the point spread function (PSF) of the microscope is an area-based function that depends on the distance to the focus plane.

Area- and frequency-based transforms perform much better. This is especially illustrated in the case of algorithms 6–10. Algorithm 6 is a method that operates on a pixel by pixel basis, while the other algorithms are based on area (3×3 and 5×5 windows). The introduction of an area coherence measure (to algorithm 6) in algorithm 9 shows a significant improvement. The best-performing method, algorithm 10, uses two area-based measures (both) in the frequency domain to perform the analysis.

The Wu et al. method is similar in philosophy (they operate in the frequency domain) to the Burt and Lolczynski and Li et. al methods but different in implementations. Wu et al. use a Gaussian function to filter the image. Their method is also based in the Fourier domain. Firstly, they perform a high pass filter to extract the high-frequency in-focus portions of the image and then perform a low pass filter to obtain an energy measure for selection of maximally in-focus areas of slices. The high pass, low pass energy approach is therefore similar in principle to the Burt and Lolczynski and Li et al. methods. However, the last two methods employ wavelet theory to perform the high pass filtering and energy measures.

It is clear that, for all the algorithms tested here, increasing the distances between slices has a detrimental effect (a kind of mosaicing effect) beyond a certain distance.

The Burt and Lolczynski maximal variance of a wavelet transform approach is clearly good because it is a nonlinear adaptive high pass filter and one would expect it to behave well. The modified Li et al. methods might be expected to outperform the basic Li method as it attempts to account for image coherence in the voting scheme. Li et al. claim that using maximum absolute value within a window accounts for the presence of dominant features. However, occasionally their measure can give a biased selection. This we believe is particularly true in regions where there is a lack of structure in the underlying image, as is the case in many biological images. In this case, a simple nonlinear high pass filter (Burt and Lolczynski, 1993) behaves more robustly. Note, however, that the addition of the voting scheme (algorithm 9), which accounts for area coherence whilst high pass filtering, has a marked improvement on the original Burt and Lolczynski method. The best performance of algorithm 10 can be further explained as the additional (with respect to

algorithm 9) dominant feature selection procedure of Li et al. contributes to a better selection of higher definition in-focus features. It should be noted that the modified algorithms 9 and 10 show a significant improvement over the original methods (algorithms 6–8).

Although wavelet transforms do not necessarily require image sizes to be of the order of a power of 2, the actual implementation of the wavelet algorithms is based on quadrature mirror banks (Press et al., 1986) that implicity requires width and height be powers of 2. However, image width and height can always be "padded" to the next power of 2 with 0 in order to allow for processing. This minor implementation detail should not detract from the fact that our study (also Li et al., 1995) clearly shows that the wavelet-based approach is the most promising technique for extended depth of focus processing of images.

Acknowledgements

References

Burt, P.J., Lolczynski, R.J., 1993. Enhanced image capture through fusion. Proceedings of the Fourth International Conference on Computer vision, Berlin, Germany, May 1993, pp. 173–182.

Castleman, K.R., 1996. Digital Image Processing. Prentice Hall, Englewood Cliffs, NJ.

Garcia-Valdecasas, A., Becerra, J.M., Marshall, D., 1997. Extending the availability of microscopic type material for taxonomy and research. Trends Ecol. Evol. 12, 211–212.

Gonzalez, R.C., Woods, R.E., 1992. Digital Image Processing. Addison-Wesley, Reading, MA.

Inoué, S., Spring, K.R., 1997. Video Microscopy. The Fundamentals. Plenum Press, New York.

Ito, K., Hayashi, A., Ichioka, Y., 1989. Digitized optical microscopy with extended depth of field. Appl. Opt. 28, 3487–3493.

Jaggi, B., Pontifex, B., Swanson, J., Poon, S.S.S., 1993. Performance evaluation of a 12-Bit, 8Mpel/s digital camera. Proceedings of SPIE, Cameras, scanners and image acquisition systems, vol. 1901. pp. 99–108.

Li, H., Manjunath, B.S., Mitra, S.K., 1995. Multisensor image fusion using the wavelet transform. Graph. Models Image Process. 57, 235–245.

Oliva, M.A., Bravo-Zanoguera, M., Price, J.H. 1998. Autofocus for phase-contrast microscopy: Investigation of causes of non-unimodality. In: Farkas, D.L., Leif, R.C., Tromberg, B.J. (Eds.), Proceedings of SPIE, vol. 3260, Optical Investigations of Cells In Vitro and In Vivo. pp. 174–179.

Parker, J.R., 1997. Algorithms for Image Processing and Computer Vision. Wiley, New York.

Pieper, R.J., Korpel, A., 1983. Image processing for extended depth of field. Appl. Opt. 22, 1449–1453.

Pluta, M., 1988. Principles and Basic Properties. Advanced Light Microscopy, vol. 1. Elsevier, Amsterdam.

Prasad, L., Iyengar, S.S., 1997. Wavelet Analysis with Applications to Image Processing. CRC Press, Boca Raton, FL.

Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T., 1986. The Art of Scientific Computing. Numerical Recipes. Cambridge University Press, Cambridge, MA.

Price, J.H., Gough, D.A., 1994. Comparison of phase-contrast and fluorescence digital autofocus for scanning microscopy. Cytometry 16, 283–297.

Rioul, O., Veterli, M., 1991. Wavelets and signal processing. IEEE Signal Process. Mag., 14–38.

Sugimoto, S.A., Ichioka, Y., 1985. Digital composition of images with increased depth of focus considering depth information. Appl. Opt. 24, 2076–2080.

Tieman, D.G., Murphey, R.K., Schmidt, J.T., Tieman, S.B., 1986. A computer-assisted video technique for preparing high resolution pictures and stereograms from thick specimens. J. Neurosci. Methods 17, 231–245.

Tympel, V., 1996. New high-level image capture system for conventional microscopy. In: Kim, Y. (Ed.), Medical Imaging 1996: Image Display, Proceedings of SPIE, vol. 2707. pp. 529–536.

Tympel, V., 1997. Three dimensional animation with a conventional light microscopy. In: Cogswell, C.J., Conchello, J.-A., Wilson, T. (Eds.), Three-Dimensional Microscopy: Image Acquisition and Processing IV, Proceedings of SPIE, vol. 2984. pp. 190–198.21.

Wall, R.J., Sobin, S.S., Karspeck, M., Lindal, R.G., Tremer, H.M., Fung, Y.C.B., 1981. Computer-derived image compositing. J. Appl. Physiol. 51, 84–89.

Willis, B., Turner, J.N., Collins, D.N., Roysam, B., Holmes, T.J., 1993. Developments in three-dimensional stereo brightfield microscopy. Microsc. Res. Tech. 24, 437–451.

Wu, H-S., Barba, J., Gil, J., 1996. A focusing algorithm for high magnification cell imaging. J. Microsc. 184, 133–142.

Yeo, T.T.E., Ong, S.H., Jayasooriah, S.R., 1993. Autofocusing for tissue microscopy. Image Vis. Comput. 11, 629–639.