

Comparison of Autofocus Methods for Automated Microscopy¹

Lawrence Firestone, Kitty Cook, Kevin Culp, Neil Talsania, and Kendall Preston, Jr²

Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213

Received for publication December 14, 1988; accepted November 24, 1990

Traditional autofocus methods were designed for microscopes driven by single processor computers. As computers are developed that exploit massive parallelism when acquiring and analyzing images, parallel cellular logic techniques became available to focus automatically. This paper introduces the reader to both cellular logic techniques for autofocus and a new spectral moment autofocus measure. It then compares these methods with more traditional autofocus methods. It is shown that traditional methods

based on measurements of image power give the best results when tested on one set of real images and two sets of synthetic images. The next best methods are the cellular logic and spectral moment techniques, while the worst are those based on the image probability density function or histogram.

Key terms: Autofocus, cellular logic, spectrum estimation, histograms, automated microscopy

When applying image processing and pattern recognition techniques to images of cells and tissue, researchers need to start with precisely focused images. Automated autofocus systems save time and focus better and more consistently than do humans. Researchers and product developers are constantly searching for ways to improve existing autofocus systems. New parallel systems employing cellular logic have already been shown to offer great potential in military systems (13), and they potentially could dramatically speed up automated focus in both research and clinical environments. This paper describes autofocus routines that exploit cellular logic techniques and then compares these techniques with an array of representative autofocus techniques described in the literature. A comparison is also made with a new spectral moment technique developed during this study.

Ingram and Preston (7) implemented autofocus capability as early as 1967. Mendelsohn and Mayall (12), Mason and Green (11), Jarvis (8), Brenner et al. (1), Erteza (2,3), and Johnson and Goforth (9) conducted pioneering work on autofocus methods in the 1970s. Harms and Aus (6), and Groen et al. (5) performed comprehensive comparative studies of autofocus methods in the 1980s.

The independent variable for all techniques considered herein is the distance (or number of focus steps) from the optimal focal plane. Techniques that involve hardware modification of cameras (such as introducing

multiple sensors), the acquisition of three-dimensional images, or those applicable only to multispectral images were not considered. All of the techniques studied could be implemented on any automated microscope with stage height control (and user access to the control software).

MATERIALS AND METHODS

To compare and contrast various focusing algorithms with regard to near-focus and far-focus performance, we applied each autofocus algorithm (see Autofocus Methods, below) to a set of tissue section images, a set of two-dimensional sinusoidal images, and a set of random noise images. The tissue images were taken of a lymph node section, furnished by the Los Angeles County Hospital. They were digitized using a Coulter diff3/50 40× oil-immersion robot microscope. For this image set, 64 × 64 frames were digitized at a resolution corresponding to a 10× objective using six bits per picturepoint, green illumination, and a field-of-view of approximately 170 × 170 μm.

After establishing an initial focal plane with the internal diff3/50 Golay transform computer, the micro-

¹This research was supported in part by grant CA45047 from the National Cancer Institute.

²Address reprint requests to Kendall Preston Jr, Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213.

scope was programmed to defocus the image automatically by moving the microscope slide 31 focus steps (1 step = 0.8 μm) below the initial focal plane. At this point, the specimen was clearly out of focus. Under computer control, the sample was raised, one focus step at a time, to the initial focal plane and past it until the image was again completely defocused. At each focal position a 64×64 image was digitized and incorporated into a mosaic of 64 images on disk. Figure 1 shows six images in the through-focus series generated in this manner.

The sinusoidal and random noise images were produced numerically. We generated a sinusoidal image using a spatial frequency characteristic of the spatial frequency corresponding to the size of cells in the lymph node section images. The function used to compute the gray-level value at row i , column j , was

$$G_{ij} = 160 - 30(1 - \cos 2\pi fi)(1 - \cos 2\pi fj), \quad (1)$$

where f was set equal to one-sixth cycle per picture-point. The constants were chosen so that the brightness and contrast corresponded to those of the tissue section images.

We also synthesized a random noise image to permit us to determine how the autofocus methods performed on a completely unstructured input. The random noise image was generated using a random number generator that produces numbers uniformly distributed between 0 and 1. These values were then scaled to the greatest range possible for eight bit images: 0–255.

The random noise image contains random changes in brightness, while the sinusoidal image features periodic changes. A robust autofocus algorithm should work well for the random noise images, the sinusoidal images, and the tissue images. To compare the performance of the autofocus algorithms on the synthetic sinusoidal and random noise images, we needed to defocus them in the same way that they would appear to be defocused by the optical system of the microscope. For an axially symmetric diffraction limited optical system with no distortion, blurring can be modeled as the convolution of the focused image with blur function B_{ij} , where B_{ij} is a two-dimensional Gaussian defined as

$$B_{ij} = K_A e^{-A(i^2 + j^2)}, \quad (2)$$

where A is inversely proportional to the radius and K_A is a normalizing constant that ensures that the volume of the blur function is 1. Small values of A correspond to the most defocused versions of the image. When $1/A$ reaches 0, we are convolving with an impulse. In this case the image is unchanged. We synthesized 32 defocused images corresponding to linearly increasing values of $1/A$ for both the sinusoidal image and the random noise image. Figure 2 shows three examples in the series of in-focus to defocused sinusoidal images (Fig. 2B,D,F) and random noise images (Fig. 2A,C,E). The 22nd sinusoidal image was defocused to the point that all points in the image had the same intensity throughout.

Autofocus Methods

This section details the nine autofocus methods that we analyzed.

Cellular logic measures ($F_{\text{Triakis7d}}$, $F_{\text{Triakis11s}}$). The mathematical morphology methods of Serra (15) and the cellular logic filters pioneered by Golay (4) opened the potential for very-high-speed image analysis through the use of parallel processing. Various applications of these techniques in microscopy and other areas are reviewed by Preston and Duff (14). One of the most promising methods of cellular logic views gray-level images as binary surfaces in three dimensions. The two three-dimensional cellular logic autofocus techniques evaluated here were written using the Triakis software system written for the Concurrent 3230, host computer for the Coulter diff3/50.

Triakis operates on a gray-level image as follows. Initially any gray-level image is stored as a two-dimensional matrix of integers, where the row and column coordinates of the matrix correspond to the row and column coordinates of the picturepoints. In our study the integer stored is proportional to image opacity. In Triakis, the image opacity at each picturepoint is encoded as a column of 1-voxels in a three-dimensional array of bits (binary voxels). The height of the column is proportional to opacity. Each column of 1-voxels is topped by a column of 0-voxels to fill the $64 \times 64 \times 64$ Triakis workspace. Thus the two-dimensional gray-level image is transformed by Triakis into a three-dimensional array of binary voxels, where two of the three dimensions correspond to row and column coordinates, and the third dimension corresponds to the opacity. This encoding process is described as column encoding of the image. It is important to note that Triakis works with three-dimensional binary representations of two-dimensional matrices of integers. The third dimension does not represent an attempt to capture three-dimensional image information. Some column encodings of in-focus and defocused tissue images are shown in Figure 3 as explained below.

Triakis uses neighborhood operators to process the column-encoded images. A neighborhood operator takes as its input the value of an individual binary voxel and the values of its neighboring voxels. While the simplest neighborhood to visualize is a Cartesian neighborhood, such a neighborhood requires that each voxel has 26 neighbors—therefore resulting in 2^{26} possible neighborhood configurations. This would preclude the possibility of building lookup tables (LUTs) listing all 2^{26} binary output values for a particular transform. The best alternative is to decrease the neighborhood size. Triakis employs a tetradecahedral neighborhood, so that each voxel has 12 neighbors. The LUTs can be implemented in hardware since there are only 2^{12} neighborhood configurations. The tetradecahedral neighborhood is derived from a face centered cubic packing of voxels.

The filters employed in Triakis use two fundamental

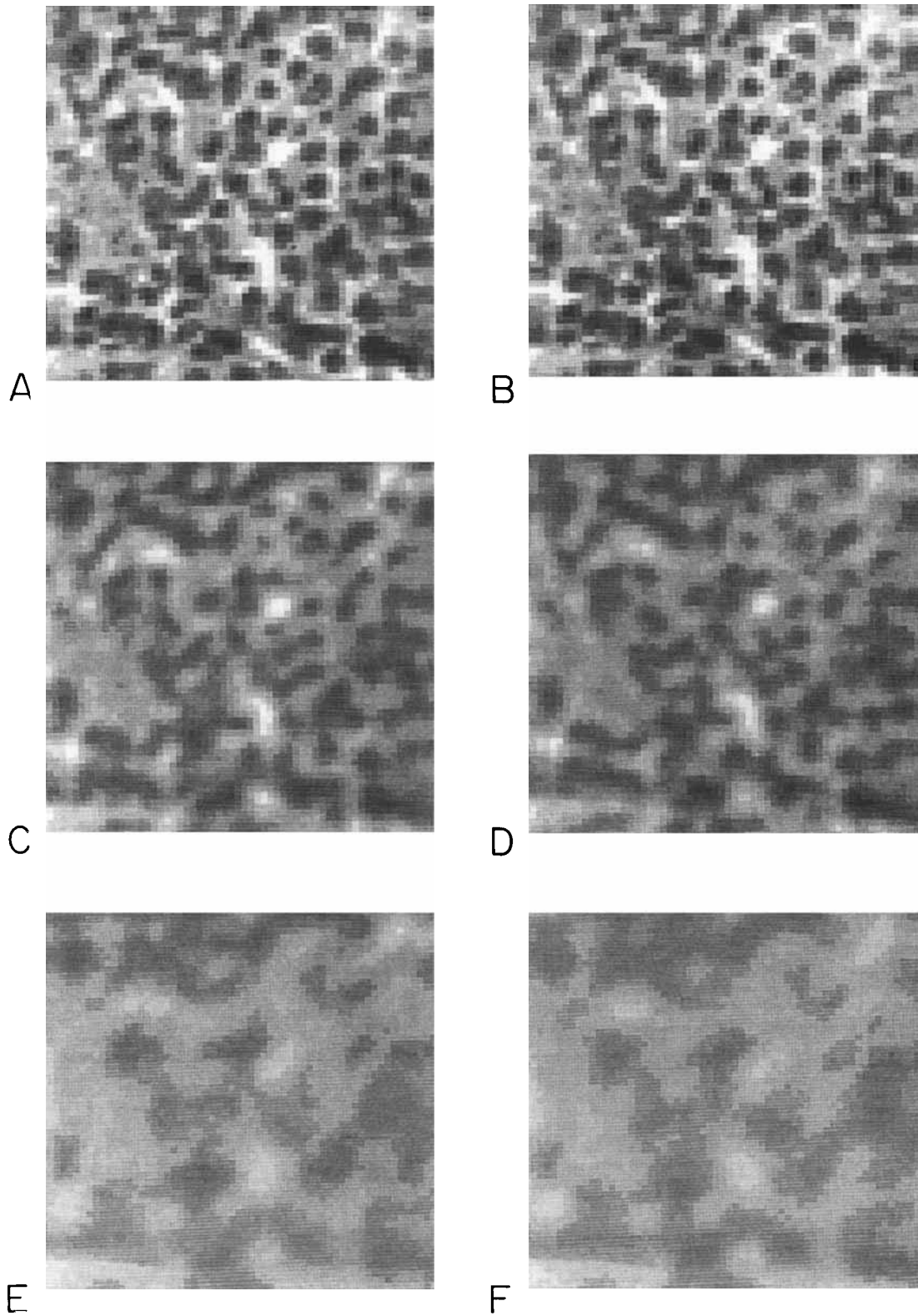


Fig. 1. Six images from the through focus set of 64 images produced using the Coulter diff3/50 robot microscope imaging a $170 \times 170 \mu\text{m}$ portion of a lymph node tissue section. In-focus (A), nearly in-focus

(B), intermediate focus (C, D), and defocused (E, F) images are shown, illustrating how the image of a single scene changes with different degrees of focus.

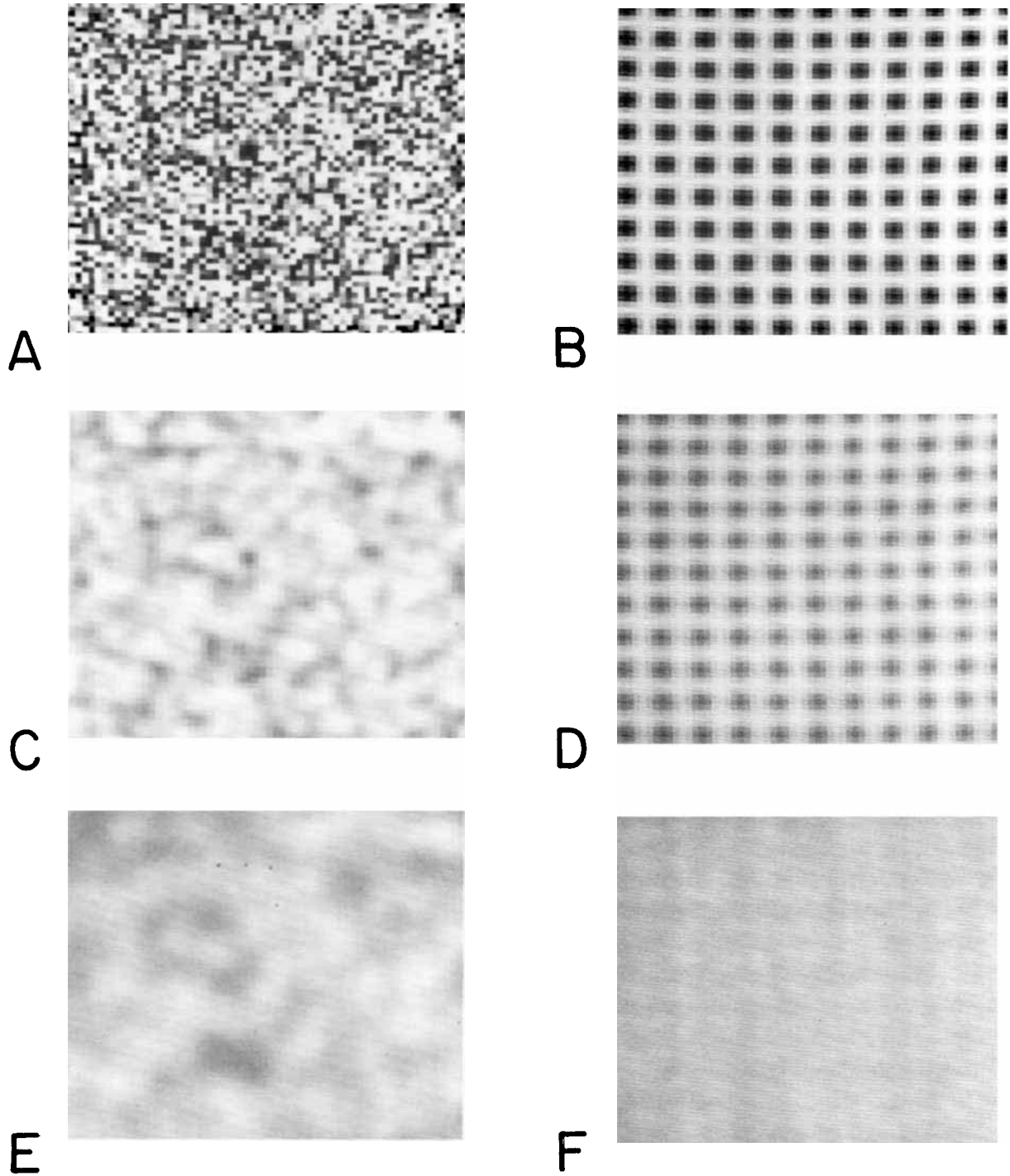


Fig. 2. In-focus, intermediate focus, and defocused random noise images (A, C, E) and in-focus, intermediate focus, and defocused sinusoidal images (B, D, F).

neighborhood parameters called *factor* and *crossing number*. The parameter *factor* equals the number of voxels in a neighborhood that are of value one. There are two *crossing numbers*. They describe connectivity relationships of the 0-voxels and 1-voxels in the neigh-

borhood (14). Filters can be built that pass voxels with particular *factor* and *crossing number* combinations to select surface voxels, ends of columns, interior voxels, residues (a single 1-voxel surrounded by zeros), and many other voxel combinations.

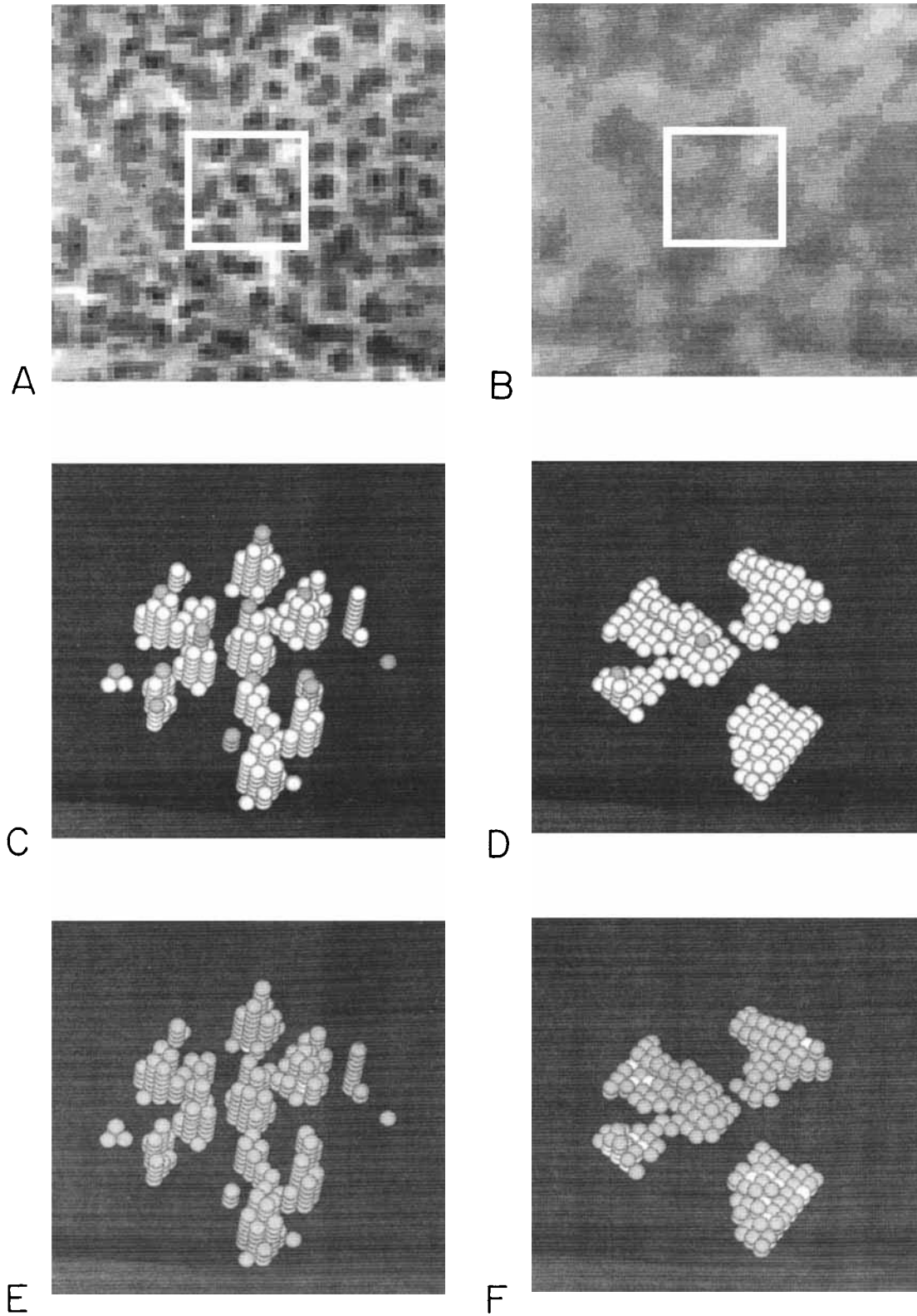


Fig. 3. In-focus (A) and partially defocused (B) tissue images with their corresponding column encodings of their opacity shown for those pixels contained within the central 19×19 pixel square (C–F). A single-sided Triakis filter using a *factor* value of 2 removes those

voxels marked as gray in C and D. The corresponding results for a *factor* value of 9 is shown in E and F. Clearly more voxels are removed for the in-focus image than for the defocused image. Thus the measure of voxels removed is a measure of relative focus.

The filters that we tested for autofocus are based on 1) executing a single iteration of a filter using a given *factor* value, EXORing with the original column-encoded image, and counting the 1-voxels that result; or 2) executing the filter, inverting, reexecuting the filter, inverting again, and then EXORing and counting. Neither filter employs the *crossing number* parameter. The first is called a *single-sided filter* and the second a *double-sided* (because of inversion) *filter*. Both filters are high pass filters and use only the *factor* value as their parameter. For example, we found that *factor* value 2 worked poorly in comparison to *factor* value 9 for both the single-sided and double-sided filters.

Figure 3 shows the voxels (gray) passed by *factor* 2 (Fig. 3C,D) and *factor* 9 (Fig. 3E,F) for single-sided filters for a small segment of an in-focus and defocused tissue image. In the study reported here filters using all possible *factor* values were tested; the filters that produced the best results were the single-sided filter using a *factor* value of 11, denoted $F_{\text{triakis11s}}$, and the double-sided filter using the *factor* value 7, denoted $F_{\text{triakis7d}}$.

Spectral analysis (F_{spectral}). The system described by Ingram and Preston over 20 years ago (7) measured high-frequency energy in the video signal for autofocus. At that time high-frequency maximization was the only method used. Ingram and Preston also used the phase of the high frequency as an error signal to determine the direction in which to move to come into focus. The utility of such spectral analysis is apparent when we consider how images change as they are focused. As an image comes into focus, edges sharpen and fine details become more visible. The gray level changes rapidly from picturepoint to picturepoint. In contrast, gray levels change very slowly as one traverses a defocused image. Sharp edges and fine details correspond to high spatial frequencies, and large objects with slowly changing gray levels correspond to low spatial frequencies. Therefore, we can expect focused images to exhibit more power at high frequencies than at low frequencies.

Spectral focus measures characterize the degree of focus of an image by first estimating the power spectrum and then characterizing the relative amounts of high- and low-frequency power. We used standard spectral estimation techniques (10) to estimate the power spectra of each of the images. All spectra were then normalized so that the sum of the components were equal to 100. Figure 4 shows the relationship between power spectrum and focal position for the sinusoidal images. Note that, close to the optimal focal position, the high-frequency power is maximized. The spectral peak corresponding to the spatial frequency of the sinusoid is very pronounced for in-focus position 0 and decreases steadily with defocus. The tissue and random noise images showed the same relationship between focus and high-frequency power.

Initially, we considered the power in various spectral bands. While we were able to find spectral bands that

worked well for one set of images, the same spectral bands were not necessarily optimal for the other sets of images, since their intrinsic spectral contents were different. Therefore, we rejected measures that used specific bands as too context specific. Instead, we developed a new method using spectral moments. The log moment given in equation 3 below was found to be a better focus measure than the linear moment:

$$F_{\text{spectral}} = \sum_i P_i \log_{10}(i), \quad (3)$$

where P_i = percent power in the i th spectral component.

Power measures. If $v(t)$ is the voltage across a 1Ω resistor, the instantaneous power of the current produced is defined to be $v^2(t)/1\Omega$. The average power for the current is defined to be the time average of the instantaneous power:

$$P = (1/T) \int_T |v(t)|^2 dt, \quad (4)$$

Power can be divided into its AC and DC components, where DC power is the mean squared value of $v(t)$, and the AC power is the power in the time varying components of $v(t)$.

Power can be defined for images in a similar fashion:

$$P = (1/N) \sum_i \sum_j (G_{ij})^2 = E\{G^2\}, \quad (5)$$

where G_{ij} is the gray level at row i and column j of the image, N is the number of picturepoints in the image, and E is the expected value operator. DC power is the mean squared $(E\{G\})^2$ and AC power the space varying $E\{G^2\} - (E\{G\})^2$ power components. As an image comes into focus, the DC power component does not change significantly; however, the AC power component increases as the image sharpens. The two autofocus measures described in this section exploit this fact.

Variance (F_{var}). The AC power of an image is directly measured by computing the variance of the picturepoint values, which is expressed by the equation:

$$F_{\text{var}} = E\{G^2\} - (E\{G\})^2. \quad (6)$$

In addition to variance, Groen et al. (5) investigated measures based on other simple statistics such as standard deviation (the square root of the variance). In addition, they investigated measures normalized by average intensity. We found this to be unnecessary in that the means of our images did not vary significantly during defocus.

Brenner's methods (F_{brenner}). Another autofocus that has the dimensions of power was developed by Brenner (1). Brenner noted that as an image comes into focus, differences between a picturepoint and its neighbor two points away increase. He computed:

$$F_{\text{brenner}} = \sum_i \sum_j (G_{ij} - G_{i+2,j})^2, i+2 < \text{image size}. \quad (7)$$

Histogram measures. Completely defocused images are a single shade of gray. Focused images contain

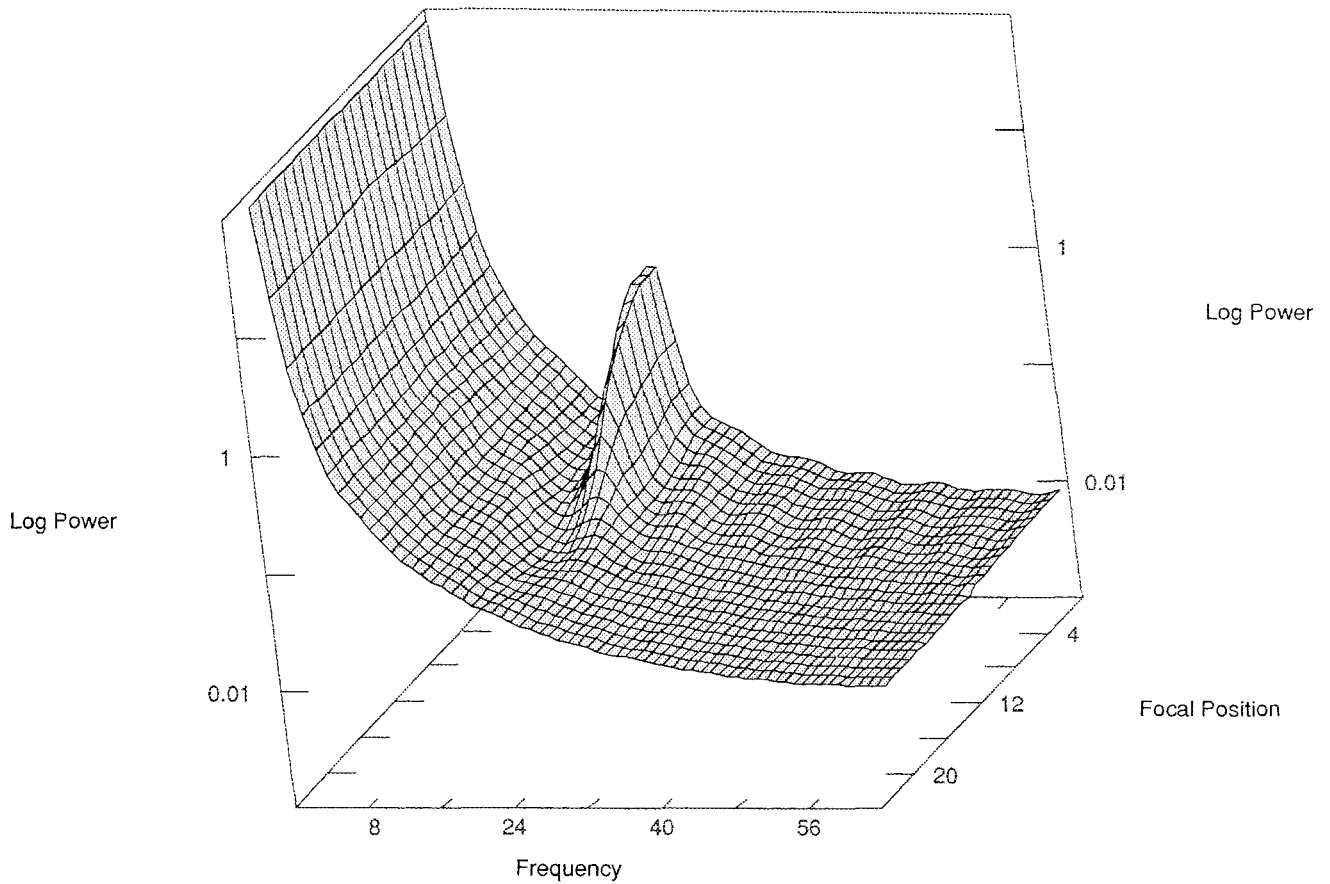


Fig. 4. Three-dimensional plot of log spectral power, frequency component, and focal position for the sinusoidal images. The curve at each focal position is normalized so that the height at each frequency is the percent power in that frequency band. Note that the peak cor-

responding to the defining spectral frequency of the sinusoid is well defined for the in-focus focal position and decreases with defocus. The log spectral moment exploits this relationship when estimating relative focus.

much discernible detail because of the many different gray levels exhibited. The histogram (unnormalized probability density function estimate) of picturepoint values shows the number of occurrences of individual gray levels. For most in-focus images, the histogram contains occurrences in many bins; as the images defocus, the number of bins in the histogram that contain occurrences usually decreases (and the number of occurrences in each bin increases).

However, some focused images start with occurrences in just a few bins. Purely black and white images would just contain occurrences in bins 0 and 255 (completely black and white, respectively). Sinusoidal images whose period is an integer number of samples will only fill those bins corresponding to intensities at the periodic sample points. In this case, the number of bins with occurrences first increases as blurring adds new gray levels to the images and then decreases with further defocusing. When images are totally out of focus, all of the occurrences are at one gray level.

The remaining autofocus measures are derived from the histogram.

Range (F_{range}). Range is the difference between the maximum gray level and the minimum gray level: Let H_k = image histogram (so that H_k = number of picturepoints with intensity k); then

$$F_{range} = \max(k|H_k > 0) - \min(k|H_k > 0). \quad (8)$$

As an image comes into focus, the histogram range increases.

Mendelsohn and Mayall's histogram method (F_{menmay}). The Mendelsohn and Mayall focus function is found by computing the weighted sum of picturepoints in histogram bins that are above a given threshold (12). The value of the threshold is chosen to be near the mean gray level value. Let T = the threshold, then

$$F_{menmay} = \sum_{k>T} k H_k. \quad (9)$$

We set the threshold equal to the first integer greater than or equal to the mean gray level value of the first image in each series. Note that since the mean did not vary significantly during defocus, we could have computed the mean for an infocus image (as we did for the

sinusoidal and random noise image) or for a defocused image (as we did for the tissue image).

Mason and Green's histogram method (F_{masgrn}). Mason and Green's histogram method (11) differs from the Mendelsohn and Mayall histogram method in the way the threshold is selected. They weighed the importance of picturepoints by estimates of the gradient at that point. We implemented their method of threshold calculation as follows:

$$\Delta_{ij} = 2(G_{i,j-1} - G_{i,j+1})^2 + 2(G_{i-1,j} - G_{i+1,j})^2 + (G_{i-1,j-1} - G_{i+1,j+1})^2 + (G_{i-1,j+1} - G_{i+1,j-1})^2. \quad (10)$$

$$T = \left(\sum_i \sum_j \Delta_{ij} G_{ij} \right) / \left(\sum_i \sum_j \Delta_{ij} \right). \quad (11)$$

Using the threshold T calculated in equation 11, the focus measure can be computed:

$$F_{masgrn} = \sum_{k \geq T} H_k (k - T). \quad (12)$$

Histogram measure entropy ($F_{entropy}$). For a focused image, the diversity of picturepoint values is usually greater than for a defocused one. The traditional measure of diversity is information content or entropy (16). For an in-focus image, the information content is usually higher because the probability of occurrence of each gray level is low. As the image is defocused, the picturepoint values eventually tend toward a single gray level, so the information content decreases. The entropy function is a measure of information content: Let N be the number of points in the image, and let $p_i = H_i/N$ be the probability that gray level i occurs in the image, then

$$F_{entropy} = - \sum_k p_k \log_b p_k, \quad (13)$$

where $b = 2$ is the base of the log function to express entropy in bits.

The usefulness of the entropy measure for autofocus is predicated on the assumption that more histogram bins will contain occurrences as an image comes into focus; thus entropy has fundamental problems when used to measure the relative focus of those images that fill few bins when in focus (such as sinusoidal images or black and white images without shades of gray).

Performance Criteria

Automated microscopes focus slides by increasing or decreasing the distance between slide and lens. When this distance is controlled by a stepper motor (as it often is), a finite number of focal positions is available to the microscope. The distance between steps should be very small (i.e., $<1 \mu\text{m}$) so that fine focusing is possible. To determine in which direction to move the slide to sharpen focus, images are captured at two sequential positions. The slide is then moved in the direction of ascent to the focus maximum. This continues until the maximum of the focus function is found. In practical applications, this procedure can be refined by adaptively varying both the number of focal heights exam-

ined and the number of steps between them, and also by recording focus data to remove redundancy. In addition, care must be taken to ensure convergence.

For an autofocus method to work well in an automated microscope, the values generated should come to a maximum at the correct focal position, should be monotonically decreasing as the distance from the focal plane increases, should not have any phase reversals or false maxima, and should come to a sharp peak at the position of best focus for maximal sensitivity. These four characteristics are defined as 1) *accuracy*, expressed here as the number of steps by which the maximum of a particular focus function departs from the correct focal position; 2) *range*, the number of steps from the true focal position for which a particular focus function is monotonically decreasing; 3) *number of false maxima*, or the number of spurious focus function maxima; and 4) *width at 50% maximum*, which is inversely related to the sharpness of the peak of the curve of the focus function. Each of these performance criteria was applied to focus function curves obtained by each of the nine methods for the tissue, sinusoidal, and random noise images. The methods were then ranked with regard to each of the criteria for each of the three sets of images. The best method was assigned a score of 1, the next best 2, and so on. In case of ties, identical ranks were assigned. A total rank score was found by summing the individual ranks for the four criteria.

RESULTS

Figures 5–7 depict all focus function curves for the tissue, sinusoidal, and random noise images, respectively. Note that the data for each curve were scaled so that all maxima were unity. Table 1 summarizes the focus function characteristics; the rows correspond to the nine focus methods listed in descending order of merit, with the best autofocus method first. The “accuracy column” gives the position of the focus function maximum for the particular method. For all images, position 0 was the position of best focus. The “range” column gives the monotonic range for each method. Note that full range for the tissue section images is -31 to 32 ; for the sinusoidal images 0 to 21 , and for the random noise images 0 to 31 . The “No. of false maxima” column specifies the number of positions that would incorrectly be deemed to be in focus because the function reaches a local maximum at these positions even though they are not the true focus positions. The “50% width” column specifies width of the particular focus function curve at 50% of its maximum height. The ranks for each criterion are given in parentheses. The ranks were then summed to determine overall relative performance. Table 2 summarizes the ranks for all methods and then gives an overall ranking.

Tissue Section Images

All of the methods found the correct focal plane within one step for the tissue section images. The $F_{brenner}$, $F_{triakis11s}$, and $F_{entropy}$ methods were right on target.

Table 1
Autofocus Function Characteristics and Relative Ranks Based on Accuracy, Range, Number of False Maxima, and Width at 50% of Maximum for the Tissue Section Images, Sinusoidal Images, and Random Noise Images^a

Method	Accuracy	Range	No. of false maxima	50% Width	Score
Tissue section images					
F _{brenner}	0 (1)	-31 to 32 (1)	0 (1)	20.98 (1)	4
F _{triakis11s}	0 (1)	-31 to 32 (1)	0 (1)	33.53 (3)	6
F _{var}	1 (4)	-31 to 32 (1)	0 (1)	36.36 (4)	7
F _{triakis7d}	-1 (4)	-31 to 32 (1)	0 (1)	24.95 (2)	8
F _{spectral}	-1 (4)	-31 to 32 (1)	0 (1)	>64 (9)	12
F _{entropy}	0 (1)	-1 to 32 (6)	1 (6)	63.78 (8)	21
F _{menmay}	1 (4)	-1 to 25 (7)	7 (7)	50.91 (5)	23
F _{range}	-1 (4)	-14 to 0 (8)	11 (8)	61.19 (6)	26
F _{masgrn}	1 (4)	-1 to -3 (9)	15 (9)	63.61 (7)	29
Sinusoidal images					
F _{brenner}	0 (1)	0 to 21 (1)	0 (1)	4.69 (1)	4
F _{var}	0 (1)	0 to 21 (1)	0 (1)	5.89 (2)	5
F _{menmay}	0 (1)	0 to 18 (3)	0 (1)	6.68 (3)	8
F _{triakis11s}	0 (1)	0 to 18 (3)	0 (1)	7.06 (4)	9
F _{range}	0 (1)	0 to 18 (3)	0 (1)	7.54 (6)	11
F _{masgrn}	0 (1)	0 to 12 (8)	0 (1)	8.15 (7)	17
F _{spectral}	0 (1)	0 to 17 (6)	0 (1)	>22 (9)	17
F _{triakis7d}	0 (1)	0 to 13 (7)	1 (8)	7.45 (5)	21
F _{entropy}	11 (9)	10 to 12 (9)	4 (9)	17.11 (8)	35
Random noise images					
F _{brenner}	0 (1)	0 to 31 (1)	0 (1)	3.15 (1)	4
F _{var}	0 (1)	0 to 31 (1)	0 (1)	3.16 (2)	5
F _{spectral}	0 (1)	0 to 31 (1)	0 (1)	3.54 (3)	6
F _{triakis7d}	0 (1)	0 to 31 (1)	0 (1)	3.55 (4)	7
F _{triakis11s}	0 (1)	0 to 31 (1)	0 (1)	3.79 (5)	8
F _{menmay}	0 (1)	0 to 31 (1)	0 (1)	3.81 (6)	9
F _{range}	0 (1)	0 to 31 (1)	0 (1)	6.96 (7)	10
F _{masgrn}	0 (1)	0 to 31 (1)	0 (1)	>32 (8)	11
F _{entropy}	0 (1)	0 to 31 (1)	0 (1)	>32 (8)	11

^aThe ranks for each criterion are given in parentheses.

Table 2
Rank Summary for Autofocus Methods on All Test Images^a

Method	Tissue	Random noise	Sinusoidal	Overall
F _{brenner}	4	4	4	12
F _{var}	7	5	5	17
F _{triakis11s}	6	8	9	23
F _{spectral}	12	6	17	35
F _{triakis7d}	8	7	21	36
F _{menmay}	23	9	8	40
F _{range}	26	10	11	47
F _{masgrn}	29	11	17	57
F _{entropy}	21	11	35	67

^aNote that low ranks correspond to better performance.

F_{brenner}, F_{spectral}, F_{triakis11s}, F_{triakis7d}, and F_{var} were unimodal, showing no false maxima over the full range. F_{brenner} and F_{triakis7d} had the smallest widths at half maximum.

Sinusoidal Images

F_{entropy} was the only method that did not find the correct focal plane for the sinusoidal images; in fact, it reached a maximum at step 11. This occurred because

only a few bins in the histogram were filled for the in-focus image because of the periodicity of the sinusoid.

F_{brenner} and F_{var} operated over the full range of focal positions. All of the methods except for F_{entropy} and F_{triakis7d} produced no false maxima. F_{brenner} and F_{triakis7d} had the smallest widths at half maximum.

Random Noise Images

The random noise images provided the least challenge. All of the methods found the correct focal plane, operated over the full range, and had no false maxima. The only discrimination among the methods is predicated on width at 50% maximum; F_{brenner}, F_{var}, F_{spectral}, F_{triakis7d}, and F_{triakis11s} were best in this regard.

DISCUSSION

The purpose of this study was to compare cellular logic, spectral, power, and image probability density function (histogram) measures with regard to accuracy, range, number of false maxima, and width at 50% maximum so that researchers and systems developers can make a more informed decision when selecting an autofocus method. While more images might be needed

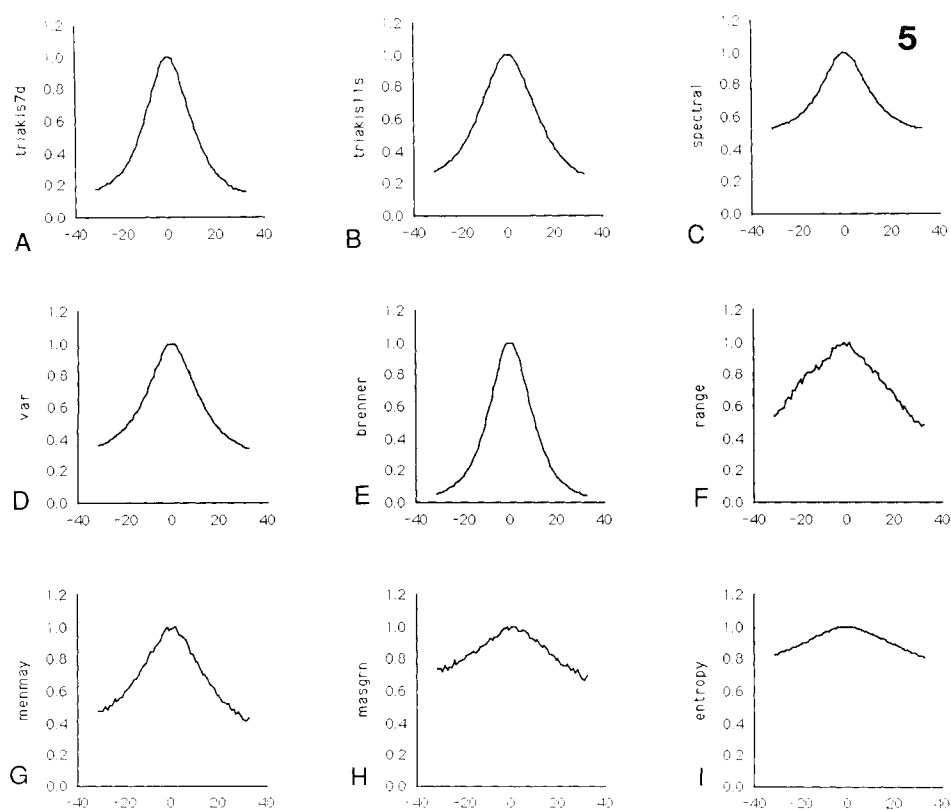


Fig. 5.

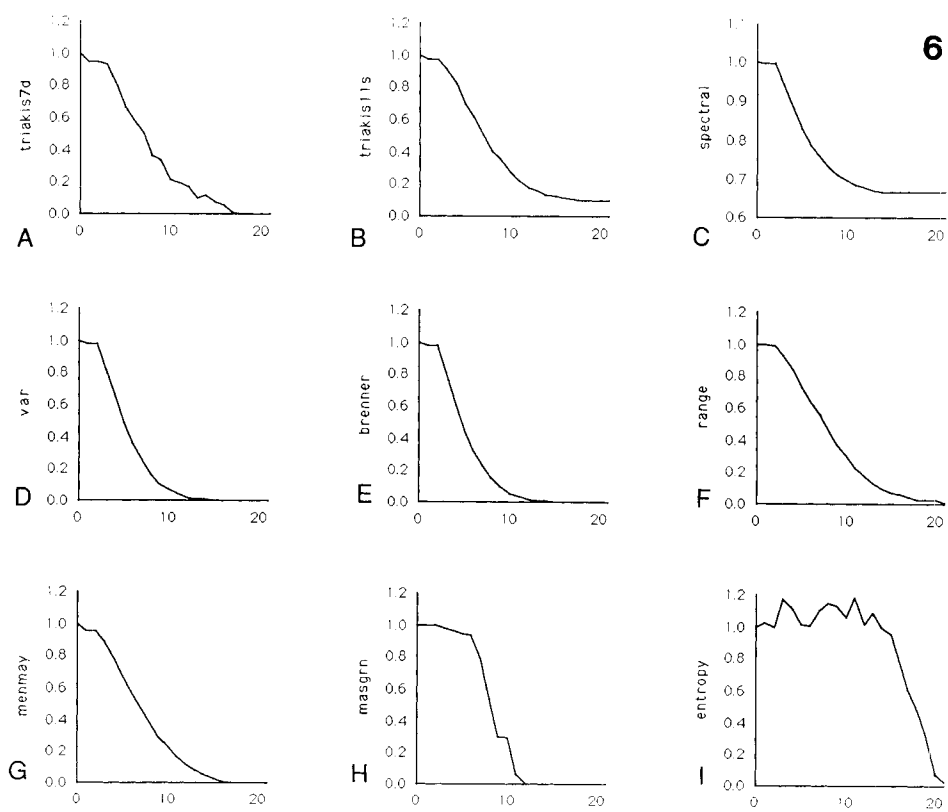


Fig. 6.

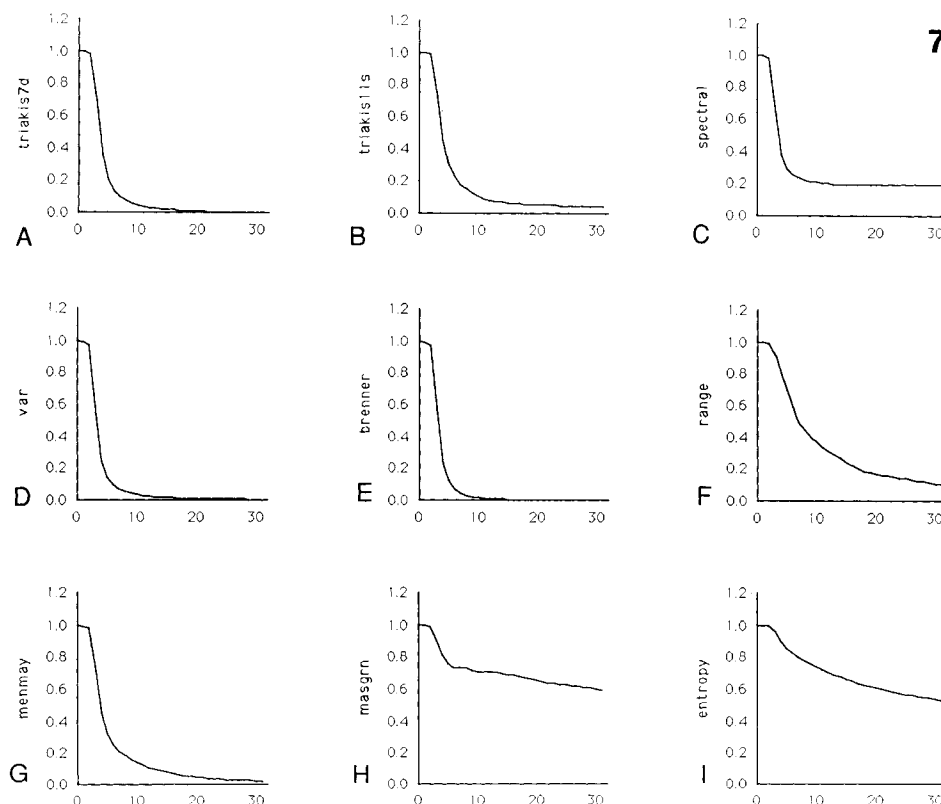


Fig. 5. Tissue image focus curves, showing the focus measurements computed at focal positions -31 to 32 for autofocus methods $F_{\text{triakis7d}}$ (A), $F_{\text{triakis11s}}$ (B), F_{spectral} (C), F_{var} (D), F_{brenner} (E), F_{range} (F), F_{menmay} (G), F_{masgrn} (H), and F_{entropy} (I). Note that an ideal autofocus curve should reach a single sharp maximum at position 0 and decay monotonically both above and below the optimal focal position.

Fig. 6. Sinusoidal image focus curves, showing the focus measurements computed at focal positions 0 to 21 for autofocus methods $F_{\text{triakis7d}}$ (A), $F_{\text{triakis11s}}$ (B), F_{spectral} (C), F_{var} (D), F_{brenner} (E), F_{range} (F),

F_{menmay} (G), F_{masgrn} (H), and F_{entropy} (I). Note that an ideal autofocus curve should reach a single sharp maximum at position 0 and decay monotonically with increasing focal position.

Fig. 7. Random noise image focus curves, showing the focus measurements computed at focal positions 0 to 31 for autofocus methods $F_{\text{triakis7d}}$ (A), $F_{\text{triakis11s}}$ (B), F_{spectral} (C), F_{var} (D), F_{brenner} (E), F_{range} (F), F_{menmay} (G), F_{masgrn} (H), and F_{entropy} (I). Note that an ideal autofocus curve should reach a single sharp maximum at position 0 and decay monotonically with increasing focal position.

to test fully and to draw definitive conclusions, the best three measures, F_{brenner} , F_{var} , and $F_{\text{triakis11s}}$, performed well for all three sets of images.

Interestingly, the two best measures, F_{brenner} and F_{var} , are measures of signal power; essentially, all we need is an AC watt meter on the output of our TV camera to determine focus. The next best methods all produced measures of high spatial frequency content: $F_{\text{triakis11s}}$, F_{spectral} , and $F_{\text{triakis7d}}$. The remaining methods, based on data in the histogram, did not perform as well.

In addition to the objective performance criteria, researchers and systems designers need to consider the speed of the various methods in their implementations. Focus time depends on many factors, including the distance between focal steps; the number of steps from the actual focus plane; the size of the image (or the section of image used to determine focus); whether the method is implemented in hardware, assembly code, or a high-level language; and the inherent computing speed of

the host computer system. The best three methods, F_{brenner} , F_{var} , and $F_{\text{triakis11s}}$, should now be compared with respect to speed. The fourth best method, F_{spectral} , would be substantially slower than both the power and cellular logic methods because of the complicated spectral estimation calculations that need to be performed.

Assuming that the focus method is applied to a 64×64 window of the image, the host computer microprocessor is a Motorola 68030 running at 16.7 MHz, and the algorithms are implemented in assembly, the single frame time for both F_{brenner} and F_{var} is 15–20 ms. If we assume that 10–20 frames need to be visited in the focus process, then it would take 150–400 ms to compute the focus measures for a microscope slide. In contrast, the $F_{\text{triakis11s}}$ method runs in 3–5 ms using a Triakis chip designed for a Motorola 68030-based system (such as Macintosh II); therefore, it would take 30–100 ms to visit 10–20 focus frames. The cellular logic approach using Triakis is therefore approximately five times faster than the power methods.

In conclusion, F_{brenner} and F_{var} are very straightforward methods that are recommended for conventional automated microscopes for which focusing times of up to half a second are acceptable. To improve speed, the researcher or systems developer would need to consider implementing the method in hardware or using a smaller image segment for focusing. Care should be taken when decreasing the size of the image segment, since estimate uncertainty increases with a decreasing number of picturepoints. Very fast focusing can be accomplished with $F_{\text{triakis11s}}$ on systems that have access to high-speed cellular logic transforms.

ACKNOWLEDGMENTS

The authors acknowledge the help received from the staff of the Biomedical Imaging Processing Unit, Division of Radiological Imaging, University of Pittsburgh, which houses the diff3/50 microscope and the supercomputer used for executing Triakis procedures. Gary Thomas and his staff, particularly Mary Ann Waleski and Mark Martin of the photography and drafting group, Mellon Institute, Carnegie Mellon University, are acknowledged for preparing the illustrations used in this paper.

LITERATURE CITED

1. Brenner JF, Dew BS, Horton JB, et al.: An automated microscope for cytologic research. *J Histochem Cytochem* 24:100-111, 1976.
2. Erteza A: Sharpness index and its application to focus control. *Appl Opt* 15:877-882, 1976.
3. Erteza A: Depth of convergence of a sharpness index autofocus system. *Appl Opt* 16:2273-2277, 1977.
4. Golay MJE: Hexagonal parallel pattern transformations. *IEEE Trans Comput* C18:733-740, 1969.
5. Groen FCA, Young IT, Lighthart G: A comparison of different focus functions for use in autofocus algorithms. *Cytometry* 6:81-91, 1985.
6. Harms H, Aus HM: Comparison of digital focus criteria for a TV microscope system. *Cytometry* 5:236-243, 1984.
7. Ingram M, Preston K Jr: Automatic analysis of blood cells. *Sci Am* 223(5):72-82, 1970.
8. Jarvis RA: Focus optimization criteria for computer image processing. *Microscope* 24:163-171, 1976.
9. Johnson ET, Goforth LJ: Metaphase spread detection and focus using closed circuit television. *J Histochem Cytochem* 22:563-572, 1974.
10. Kay S: Spectrum estimation. In: *Advanced Topics in Signal Processing*, Lim JS and Oppenheim AV (eds). Prentice Hall, Englewood Cliffs, NJ, 1990, pp 58-122.
11. Mason DC, Green DK: Automatic focusing of a computer controlled microscope. *IEEE Trans Biomed Eng BME* 22:312-317, 1975.
12. Mendelsohn ML, Mayall BH: Computer-oriented analysis of human chromosomes. *Comput Biol Med* 2:137-150, 1972.
13. Preston K Jr: Detection of Weak, Subpixel Targets Using Mesh-Connected Cellular Automata. Naval Ocean Systems Center, San Diego, CA, Final Report, Contract/PR N66001-85-D-0203, Delivery No. 0055, 1988.
14. Preston K Jr, Duff MJB: *Modern Cellular Automata: Theory and Applications*. Plenum Press, New York, 1984.
15. Serra J: *Image Analysis and Mathematical Morphology*. Academic Press, London, 1982.
16. Shannon CE: A mathematical theory of communications. *Bell Sys Tech J* 27:379-423, 623-656, 1948.