Feature Extraction of Chromosomes From 3-D Confocal Microscope Images

Matthew J. Kyan, Ling Guan*, *Senior Member, IEEE*, Matthew R. Arnison, and Carol J. Cogswell

***Abstract*—An investigation of local energy surface detection integrated with neural network techniques for image segmentation is presented, as applied in the feature extraction of chromosomes from image datasets obtained using an experimental confocal microscope. Use of the confocal microscope enables biologists to observe dividing cells (living or preserved) within a three-dimensional (3-D) volume, that can be visualised from multiple aspects, allowing for increased structural insight. The Nomarski differential interference contrast mode used for imaging translucent specimens, such as chromosomes, produces images not suitable for volume rendering. Segmentation of the chromosomes from this data is, thus, necessary.**

A neural network based on competitive learning, known as Kohonen's self-organizing feature map (SOFM) was used to perform segmentation, using a collection of statistics or features defining the image. Our past investigation showed that standard features such as the *localized mean* and *variance* of pixel intensities provided reasonable extraction of objects such as mitotic chromosomes, but surface detail was only moderately resolved. In this current work, a biologically inspired feature known as *local energy* is investigated as an alternative image statistic based on phase congruency in the image. This, along with different combinations of other image statistics, is applied in a SOFM, producing 3-D images exhibiting vast improvement in the level of detail and clearly isolating the chromosomes from the background.

***Index Terms*—DIC, differential interference contrast, feature extraction, feature space, image segmentation, local energy, Morlet wavelet, phase congruency, self organizing feature map, SOFM.**

I. Introduction

CONFOCAL microscopy is widely becoming a popular method for imaging specimens across three-dimensional (3-D) volumes, the results of which subsequently reflect structural and surface detail not visible in standard two-dimensional (2-D) projections [1]. The confocal configuration (i.e., point-like illumination and a detector with a pinhole mask), demonstrates increased ability for *optical sectioning* (improved depth resolution) over conventional methods (i.e., using a large-area-detector). As such, it is ideal for obtaining highly resolved image slices across the specimen volume, which can then be recombined to form a 3-D image representation.

Manuscript received June 2, 2001. The work of M. J. Kyan was supported in part by the Chancellor's Scholarships in Engineering (CSIE) program at the University of Sydney. *Asterisk indicates corresponding author.*

M. J. Kyan is with the Signal and Multimedia Processing Group, School of Electrical and & Information Systems Engineering, University of Sydney, NSW 2006, Australia.

*L. Guan is with the Department of Electrical and Computer Engineering, Ryerson Polytechnic University, Toronto, ON M5B 2K3, Canada (e-mail: lguan@ee.ryerson.ca).

M. R. Arnison and C. J. Cogswell are with the Physical Optics Department, School of Physics, University of Sydney, NSW 2006, Australia.

Publisher Item Identifier S 0018-9294(01)09143-1.

Confocal transmission Nomarski differential interference contrast (DIC) is a mode that images changes in the refractive index of the specimen and is, thus, suited to the imaging of translucent subjects, such as chromosomes [2]. Standard 3-D image processing and visualization techniques such as those employed in other 3-D modalities (e.g., confocal fluorescence) are ineffective when applied to DIC due to their inability to deal with the inherent differential shading embedded in the imaged specimens. The problem is that each DIC image slice shows a differential *bas-relief* effect, depicting highlights for positive phase gradients in the specimen, shadows for negative gradients, and grey for regions of zero phase gradient (or a level surface). As a result, it becomes difficult to distinguish the level surfaces on the chromosomes from the background, prompting the need for removal of the differential effect before the original image may be reconstructed.

In the past, various techniques have been proposed to reverse the differential effect, so that image specimens may be isolated from their background. Isolation would then allow for clear 3-D volume visualization of the data specimens, by rendering the background transparent. Quantitative approaches outlined in the literature use integration and iterative deconvolution [3], [4] to reconstruct the DIC image. However both methods have a smoothing effect, reducing the clarity of 3-D visualization. Other, more advanced techniques, such as Weiner filtered deconvolution [5], attempt to model the DIC imaging system to accurately restore phase, however, in doing so make simplifying assumptions about the DIC process that are inconsistent when applied to complex biological specimens. The Hilbert Transform has been shown to be a useful qualitative alternative [6] in its ability to reconstruct phase information from DIC images. The transform allows features of interest (those exhibiting high spatial frequencies) as well as contrast between specimen objects and background information to be retained.

The majority of effort addressing DIC images, is centerd on direct reconstruction, however, little work has been conducted in terms of further image processing to resolve important features within the data. Further *image segmentation* (the process of subdividing or categorising an image into its different constituent parts) for example, can resolve greater surface detail *within* the objects themselves than can be obtained through a simple background removal. The motivation of this paper is, thus, based on formulating improved techniques that extend the Hilbert processing of DIC images to allow for better extraction of chromosome *features* during visualization. The aim is twofold: to represent the dataset through advanced, meaningful features; and to classify and segment into regions exhibiting similar features, allowing for more distinct isolation of chromosome regions from

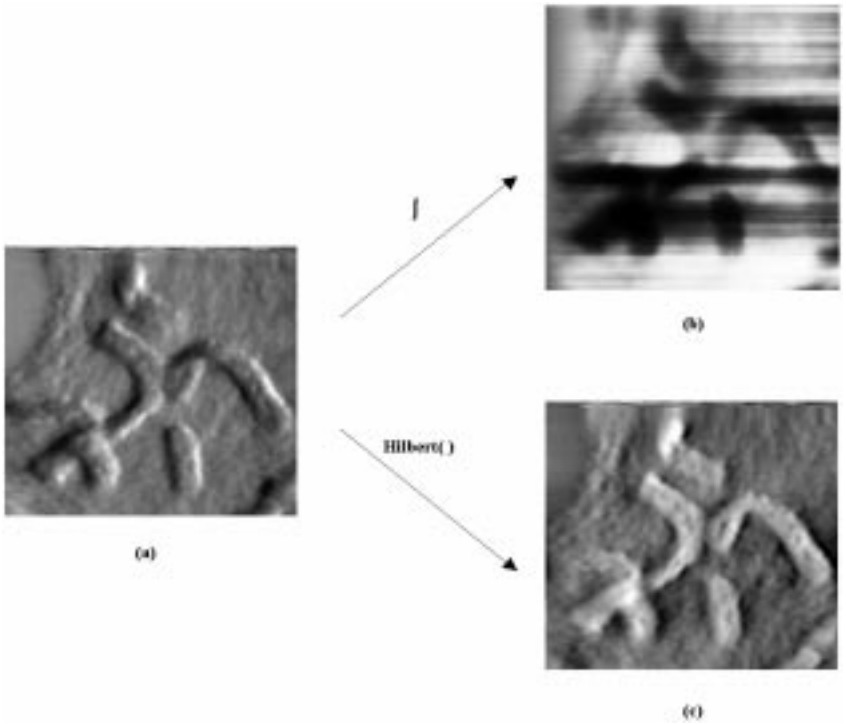


Fig. 1. Image reconstruction (images courtesy of Arnison *et al.* [6]). (a) One of the original 2-D transmission DIC images of orchid chromosomes showing *bas-relief* (differential shading gradient). (b) Directional integration of the DIC image showing blurred result and streaking across the lateral axis. (c) Hilbert transform of the DIC image shown in (a). The differential shading is removed, providing an initial stage in image reconstruction.

background. Thereby simplifying the visualization process with the view for future automation.

Kohonen's self organizing feature map (SOFM) is an established neural network algorithm that works to cluster an input set of statistics using unsupervised learning. It has been used across multiple disciplines for classifying and segmenting data of which there is no prior knowledge. It is especially useful in image processing as it allows for a high dimensional feature space to be adaptively mapped onto a lower dimensional, neural manifold. Achieving what is essentially, an optimal vector quantization of the input feature space [7].

The concept of feature space simply refers to the vectorial representation of a group of image statistics defining each pixel in the original image. In past work conducted by Nguyen *et al.* [8], pixels were redefined using 3-D vectors incorporating classical image statistics, namely: pixel intensity, localized mean and localized variance. This new, feature space description of the image offers a more detailed description of each voxel, reflecting a voxel *topology* rather than a simple intensity level. Such a description, when applied to the SOFM, has met with favorable segmentation results in that the chromosome bodies were adequately distinguished and, thus, extracted from their surroundings for volume visualization. The problem however, was that there was very little resolution of the internal structure of these chromosomes, and the results presented no significant improvements over thresholding, except that a segmented dataset would aid in automatic rejection of background.

In this paper, a biologically inspired feature known as *local energy* is investigated as an alternative image statistic. Predominantly local energy has been investigated for its ability as a 2-D feature detector [9], based on phase congruency (PC) in the image. More recently, preliminary 3-D applications in confocal microscopy focus on surface detection [10], wherein a ridge tracing algorithm is used to form an opaque, skeleton image of the specimen. In our investigation, we propose a novel scheme that exploits the propensity of the SOFM for unsupervised categorization, with the heightened feature sensitivity of 3-D local energy detection: incorporating these techniques in order to more completely describe the internal and external features existing in the chromosome specimens. As such, we extend beyond describing just the outer skeletal shell of a specimen.

Due to the qualitative nature of the original data (a nonlinear mixture of DIC phase and amplitude information) [6], visual inspection remains the only form of evaluation. This is conducted via *volume visualization* of the resultant datasets. For this purpose, we use the software package *VoxelView*,[1] in which the background can be made transparent, while the 3-D segmented features are rendered to various degrees of opacity in order to reveal the remaining chromosome structures.

The following presentation is structured as follows: Section II will outline the initial reconstruction techniques, Section III will discuss the SOFM algorithm, Section IV will introduce the concept PC while also describing the local energy detection process. Finally, results are presented in Section V with conclusions drawn in Section VI.

II. Image Reconstruction

The first stage in processing the DIC image data is to remove the differential *bas-relief* effect present in the image [Fig. 1(a)]. The technique used is based on spatial frequency analysis, and is

[1]VoxelView: Vital Images, Inc. USA

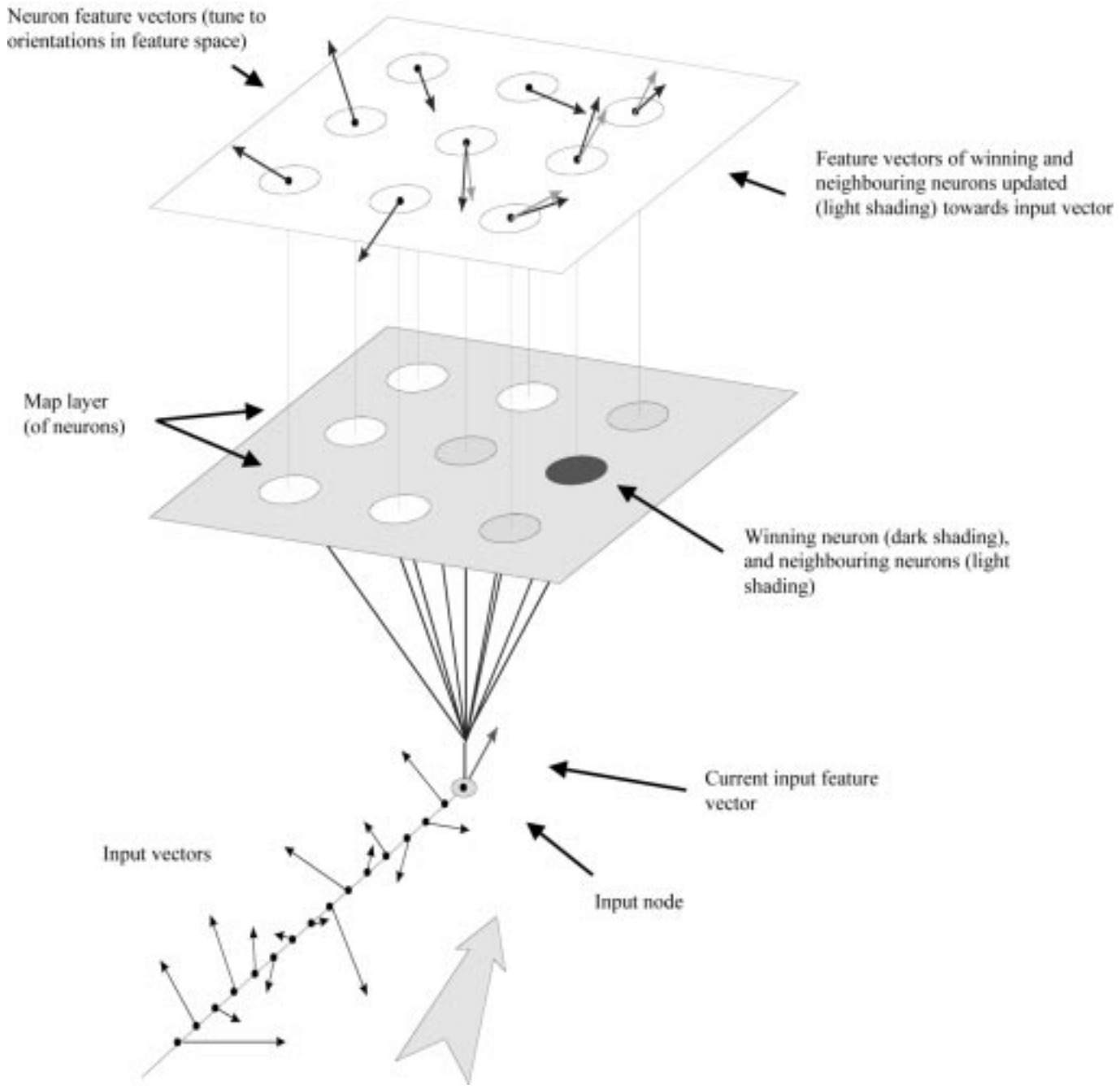


Fig. 2. Kohonen's SOFM algorithm.

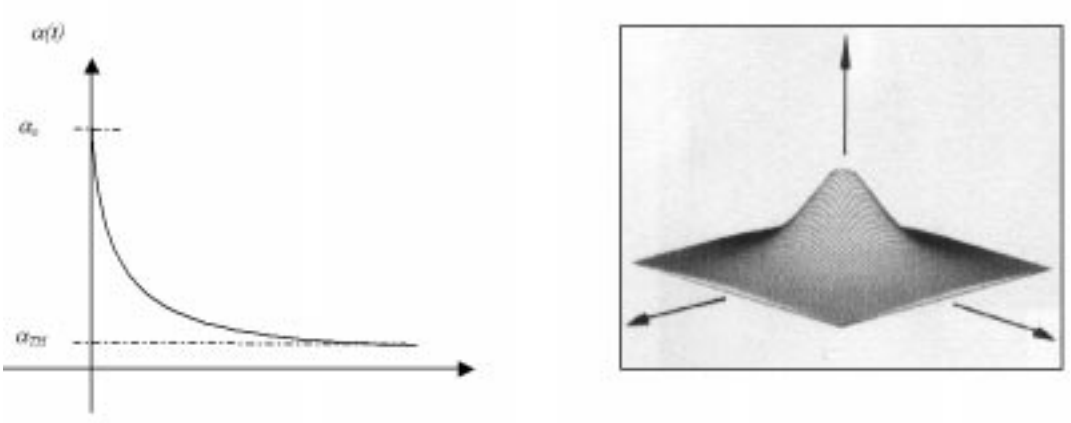
known as the Hilbert transform (developed by Oppenheim and Schafer) [6], [11].

In the real space domain, differentiation corresponds to a multiplication by the spatial frequency, while integration corresponds to division. Straight integration alone induces a division process that attenuates higher frequency components, often resulting in a blurred result [Fig. 1(b)]. The Hilbert transform differs in its ability to maintain the relative balance of frequency components in the image. It achieves this by keeping all positive frequency components and reversing the negative frequency components [Fig. 1(c)]. This makes the image symmetric, removing the *bas-relief* effect.

III. Image Segmentation Using SOFM

A. SOFM Architecture

Kohonen's self-organizing map algorithm can be visualised in terms of the structure shown in Fig. 2. The map consists of a 2-D *lattice* of neurons (memory elements). Input feature vectors are applied to the input node of the SOFM individually, each imparting information to the neural map about a particular characteristic in the image. Each neuron in the map has a feature vector of its own, initially randomly oriented in feature space. Over time, these neurons learn from the input vectors, according to a competitive learning algorithm, selectively *tuning*



(a) (b)

Fig. 3. Control parameters describing proportion of information learnt from each new input vector. (a) Graph describing the learning rate over time. When $\alpha(t)$ falls below the threshold α_{TH}, learning ceases. (b) Gaussian neighborhood smoothing function. The peak of which corresponds to a value of one, decaying radially to zero.

themselves to a particular orientation in feature space. The final orientations of the neurons then represent particular feature categories into which the individual image pixels (denoted by the input feature vectors) can be grouped.

The real power of the SOFM, therefore, lies in its ability to form a *topographical* map of the input patterns where the spatial locations (coordinates) of the neurons in the lattice represent *intrinsic* features contained in the set of input vectors [12]. This architecture, thus, aids in the categorization of pixels reflecting an overall segmentation of the original image.

B. Training Vectors—The Learning Process

When each input vector is applied to the network, a single neuron known as the *winning* neuron is activated. This neuron is the one with a feature vector closest to the input vector. The *winning* neuron is found by calculating the minimum euclidean distance of all the neurons from the input vector as follows, where $m_i(t)$ is the feature vector of neuron i, $x(t)$ is the current input vector, and c is the winning neuron

$$\begin{aligned}\|x(t)-m_c(t)\| &= \min_i \{\|x(t)-m_i(t)\|\} \\ &= \min_i \left\{ \sqrt{\sum_{j=1}^{n} (x_j(t)-m_{i,j}(t))^2} \right\}. \end{aligned} \tag{1}$$

The winning neuron and its neighboring neurons, each learn information by updating their orientations toward the input vector by some proportion. The proportion is highest for the winning neuron, reducing somewhat for the surrounding neurons, depending on how close they are to the winning neuron. Closer neurons are updated more dramatically, thus, learn more information from the input. The learning process is iterative, converging over time

$$m_i(t+1) = m_i(t) + \alpha(t) h_{ci}(t) [x(t) - m_i(t)] \tag{2}$$

where $\alpha(t)$ is the learning rate and $h_{ci}(t) =$ is a 2-D neighborhood smoothing function.

Together, $\alpha(t)$ and $h_{ci}(t)$ control the proportion by which a neuron is updated.

C. Control Parameters

Learning Rate $\alpha(t)$: The learning rate is defined as

$$\alpha(t) = \alpha_o e^{-t/\tau_\alpha} \tag{3}$$

where α_o is the starting rate and $\tau\alpha$ is the decay rate.

This function decays over time, depending on adjustable time constants, and controls the *global* proportion of information that is to be learnt at each iteration in the SOFM algorithm. When it falls below a defined tolerance (α_{TH}), this signifies the algorithm's convergence, and learning ceases.

Neighborhood Smoothing Function $h_{ci}(t)$: The neighborhood smoothing function controls the relative proportion of change occurring in the winning and neighboring neuron feature vectors. It is modeled on a 2-D Gaussian normalized between zero and one [Fig. 3(b)], centerd on the position of the winning neuron in the map lattice. The smoothing function is described by

$$h_{ci}(t) = e^{\left(-\|r_c - r_i\|^2 / 2\sigma^2(t)\right)} \tag{4}$$

where r_c and r_i are the position vectors of the winning neuron (c) and the neuron being updated (i), while σ represents a radius function controlling the extent over which the neighborhood function is effective.

Thus, $h_{ci}(t) = 1$ at the winning neuron, and its feature vector is updated by the full proportion of $\alpha(t)$, while surrounding neuron vectors are updated by a fraction of $\alpha(t)$.

The set of neurons to be updated at an instant in time, is controlled by the radius function as mentioned previously. This radius $\sigma(t)$ also decays exponentially

$$\sigma(t) = (\sigma_o - 1) e^{-t/\tau_\sigma} \tag{5}$$

where σ_o is the initial value and $\tau\sigma$ is an arbitrary fixed time constant.

D. Clustering and Reassigning Pixel Intensities

After the neuron feature vectors have been *tuned* to particular orientations in feature space (i.e., representing particular surface types in the original image), the original pixels are then *pooled* into categories, each of which is assigned a new intensity value (intensity values divided equally across the greyscale range).

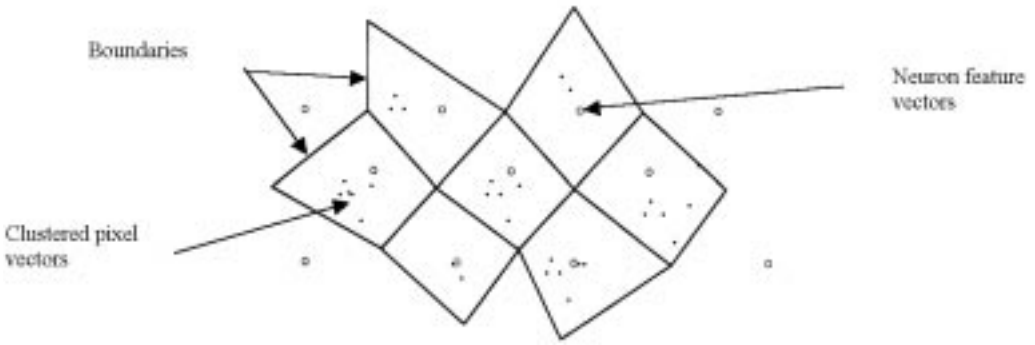


Fig. 4. A Voronoi tessellation showing the formation of category boundaries about the neurons, and the pixels clustered into these categories.

Each category is defined by a neuron. Pixels are assigned to the neurons they are closest to in feature space. This process forms *clusters* of pixels, using an algorithm known as a *Voronoi tessellation*. A Voronoi tessellation defines the region in feature space surrounding each neuron vector such that the boundary dividing two clusters is perpendicular to the line joining the center of two clusters (see Fig. 4).

The pixel values are then reassigned according to their categorised intensity, thus producing an image segmented by intensity into its primary surface types. Intensities representing unwanted surface types can then be removed during visualization.

E. SOFM Algorithm

The SOFM segmentation process was applied to the 3-D datasets in a number of separate stages. These steps are as outlined here.

- Step 1) Build 3-D dataset: The individual image slice files (in '*FITS*' file format) are combined into a single 3-D dataset in memory. This dataset forms the input to steps 2 and 4.
- Step 2) Generate features: The appropriate set of three statistical features are generated for each voxel in the 3-D dataset. A three-tuple feature *vector* for each voxel is then constructed from the generated features. The output is a file of voxels represented in feature space.
- Step 3) SOFM: The feature space vectors describing the image, produced in Step 2, are then used as input to the SOFM algorithm, which builds and trains a neural map of chosen dimension $(n \times n)$. The output of this step is a map file of n^2 neural vectors.
- Step 4) Clustering: The clustering step takes input from steps 1 and 3, where the voronoi algorithm is used to tag each voxel in the original 3-D dataset, with a label classifying it into one of the groupings defined by the characteristic neural vectors in the map file.
- Step 5) Assign intensities and reslice dataset: In the final step, the labeled voxels are assigned an intensity value to enable viewing, before being split into separate image slices for saving. The separation of intensity labeling from the clustering process allows for different assignment techniques to be explored in future experiments.

IV. Local Energy Detection

A. Perceiving Features in an Image

The problem of extracting features from within 3-D images produced using a confocal microscope is very similar to the notion of edge detection within 2-D images. Many techniques used today are simply 3-D extensions of 2-D gradient-based edge detectors.

Gradient-based methods work on the principle that people perceive features wherever there is a sharp difference in intensity between two regions (i.e., an intensity gradient). However, these methods only locate features with a step-like intensity profile and do not account for the variety of other possible feature profiles such as lines, roofs or ramps—feature profiles which all exist in confocal microscope images. For this reason, simple gradient-based methods fail to adequately locate much of the surface detail present in an image.

Rather than searching for points where there are sharp changes in intensity, the local energy or PC model of detection searches for patterns of *order* in the phase component of the FT of the image. Physiological evidence [13] indicates that the human visual system responds strongly to points in an image where phase information is highly ordered, redefining how *features* are actually perceived.

B. Phase Congruency

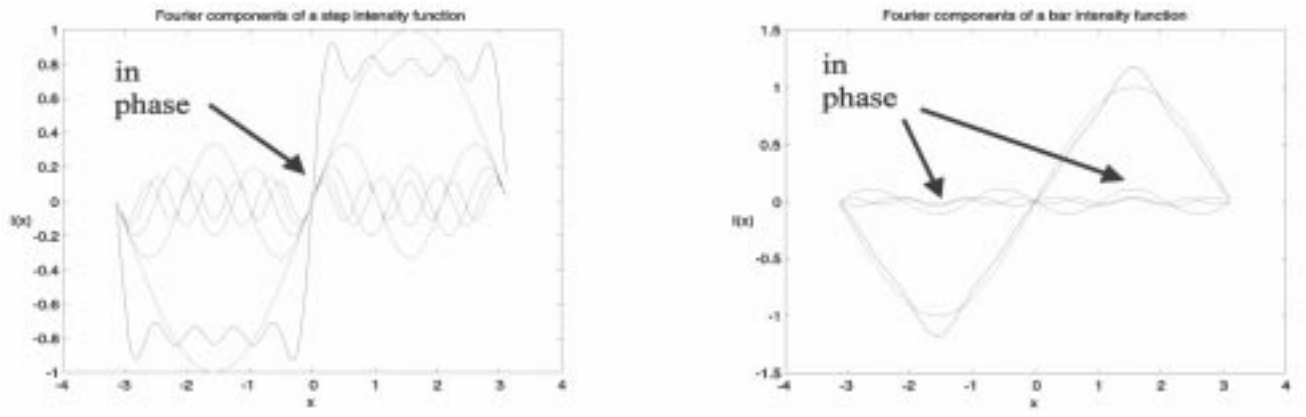
The local energy model of feature detection [14] is based on the principle that features are perceived at points of maximal PC within an image. The first step in understanding PC is to consider the Fourier expansion of a one dimensional (1-D) intensity profile into its component sinusoidal waveforms

$$F(x) = \sum_{n=0}^{\infty} (A_n \sin(n\gamma x + \phi_n)) \tag{6}$$

where

- γ constant (usually 2π);
- ϕ_n phase offset of the nth component;
- A_n magnitude of the n the sine component.

Generally, the Fourier components will be in phase in regions where features are visible. For example, the Fourier components of a square wave (step-like intensity profile) are in phase at the step, independent of whether the step represents a positive or negative transition in intensity [see Fig. 5(a)]. Alternatively, the Fourier components of a bar feature, or triangular wave pro-



(a) (b)

Fig. 5. Fourier components of (a) a step intensity and (b) a bar intensity function. The in-phase components correspond to regions in which features are detected.

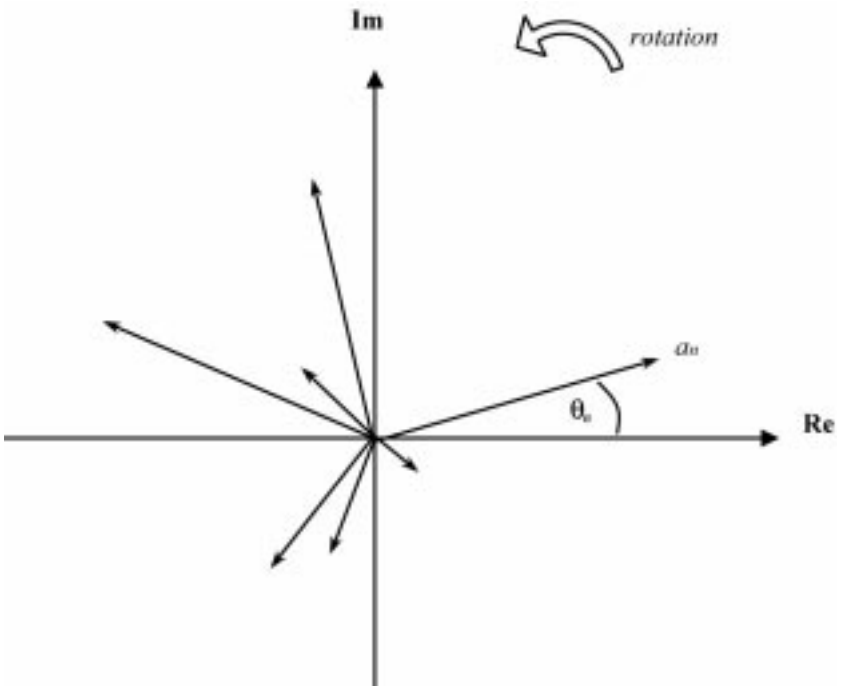


Fig. 6. Phasor representation of the Fourier components at a single point x in the signal $F(x)$.

file, are in phase at each peak and trough [see Fig. 5(b)], corresponding to points that are responsible for perceived features.

The PC function is defined by Morrone and Owens [14] as

$$PC(x) = \underset{\theta\in[0,2\pi]}{\text{Max}} \frac{\sum_n (A_n \cos(n\gamma x + \phi_n - \theta))}{\sum_n A_n} \tag{7}$$

where parameters are as in (6).

The signal $F(x)$, at any point x in the image, can be thought of in terms of a sum of vectors (rotating phasors at different angular frequencies in the complex domain which represent individual Fourier sine components, each with a different amplitude a_n and phase θ_n). This is depicted in the Fig. 6.

A relatively simple method for obtaining a measure of PC would be to find the mean phase angle of the vectors, and calculate the standard deviation of all the individual phase angles about this mean. This method breaks down, however, since it gives a large deviation when comparing vectors at 355° with those located at 1°, where it should be small.

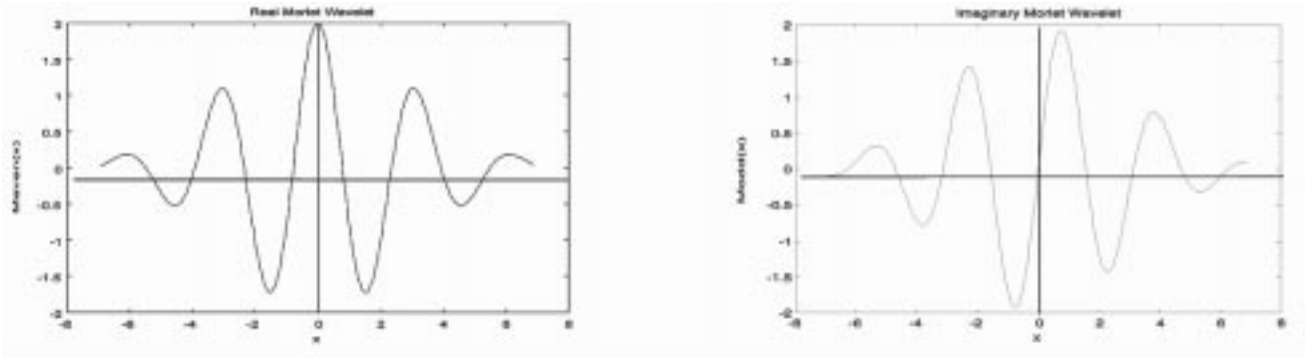
The cosine term was subsequently added to the PC equation in order to capture the proximity of these two angles. Thus, the PC is equal to one when the Fourier component vectors are all aligned (such as at the discontinuity in a step-function), and falls between zero and one otherwise.

The PC function can be approximated using the Taylor expansion

$$\cos(x) \approx 1 - \frac{x^2}{2} \tag{8}$$

so that

$$PC(x) = \underset{\theta\in[0,2\pi]}{\text{Max}} \left[1 - \frac{\sum_n \left(A_n(n\gamma x + \phi_n - \theta)^2\right)}{2\sum_n A_n}\right]. \tag{9}$$



(a) (b)

Fig. 7. (a) Real and (b) imaginary Morlet wavelet functions forming a complex Gabor function.

This approximation is maximized, as mentioned previously, for a value of theta equal to the mean phase angle $\bar{\theta}$, thus

$$PC(x) \approx 1 - \frac{\sum_n \left(A_n(n\gamma x + \phi_n - \bar{\theta})^2\right)}{2\sum_n A_n}. \tag{10}$$

C. Local Energy

Phase congruency as defined in the Section IV-B is an awkward value to calculate. It was subsequently shown by Venkatesh and Owens [15], that the PC function is directly proportional to a function known as the local energy, which is a popular computation used in modeling biological vision.

The local energy function is defined for a 1-D signal as:

$$E(x) = \sqrt{F^2(x) + H^2(x)} \tag{11}$$

where $F(x)$ is the integral over the real parts of the phasors defining the signal at x(luminance profile); while $H(x)$ is the integral over the imaginary parts of these phasors [Hilbert transform of $F(x)$].

This energy $E(x)$, defined at the point x in the signal, is the magnitude of the vector sum of all the Fourier components of the signal at that point (i.e., the vector sum of all the phasors as depicted in Fig. 6). $F(x)$ then becomes the sum (integral) of all the real parts of these phasors (luminance profile), while $H(x)$ is the sum of all the imaginary parts of these phasors.

The local energy function shows peaks at points in the signal where there is a high degree of alignment in the Fourier component phasors: producing a larger resultant vector. The local energy function is, thus, related to the PC function. The relation can be described by

$$E(x) = PC(x)\sum_n A_n. \tag{12}$$

D. Morlet Wavelets

The preferred method for obtaining local frequency information, used in calculating local energy, is known as the wavelet transform [16]. This method uses a bank of filters to analyze the signal. The filters are created from rescalings of one type of wave shape, with each scaling designed to "pick out" particular frequencies of the signal being analyzed [10]. This allows for a more *localized* description of frequency content than can be obtained through the simple use of a FT. The range of frequencies can also be easily controlled in this manner.

Morlet wavelets [16] are used to define the wave shape for the filter banks. They are nonorthogonal wavelets based on complex Gabor functions: cosine waves (even) and sine waves (odd), each modulated by a Gaussian, forming the real and imaginary parts respectively (see Fig. 7).

The two components $F(x)$ and $H(x)$, of the local energy are obtained by convolving the signal with a bank of Morlet wavelets as follows:

$$F(x) = \sum_{f=\mathrm{Freq}} I(x) \otimes M_f^{\mathrm{even}}(x) \tag{13}$$

$$H(x) = \sum_{f=\mathrm{Freq}} I(x) \otimes M_f^{\mathrm{odd}}(x) \tag{14}$$

where

Freq	set of frequencies over which the bank of filters is defined [2];
$M_f^{\mathrm{even/odd}}$	real and imaginary Morlet wavelet functions at frequency f;
$I(x)$	one-dimensional intensity profile at position x;
$\otimes$	one-dimensional convolution.

E. Extension to Higher Dimensions

The inherent problem in moving to higher dimensions is that the Hilbert transform is only defined in one dimension. Furthermore, it is an odd function and, thus, exhibits an *imposed orientation* in higher dimensions.

This prompts a more general definition of the Morlet wavelet functions

$$M_{f,\hat{v},x}^{\mathrm{even}} = \frac{1}{\left(\sqrt{2\pi}\sigma\right)^3} e^{-\|x\|^2/2\sigma^2} \omega \cos(\omega \hat{v}^T \cdot x) \tag{15}$$

$$M_{f,\hat{v},x}^{\mathrm{odd}} = \frac{1}{\left(\sqrt{2\pi}\sigma\right)^3} e^{-\|x\|^2/2\sigma^2} \omega \sin(\omega \hat{v}^T \cdot x) \tag{16}$$

where $\sigma = k/f$, k is the controlling factor between the filter's width σ and frequency f, and $\omega = 2\pi f$.

These two filters constitute a quadrature pair, where $\hat{v}$ represents the unit orientation vector of the filter. This vector can be defined in two or three dimensional space, where $\|x\|^2$ also extends to become the distance of the pixel position from the origin, within the chosen dimension. The origin is generally chosen to

[2]The frequency range (spatial) can be adjusted to make the local energy function sensitive to particular sized features in the image.

be the center of a 2-D image space, or the center of a 3-D image volume (rather than a corner as is defined in standard images).

Local energy in higher dimensions is, thus, computed over a set of orientations of these quadrature filter pairs in either 2-D or 3-D. This can be summarized by the following transformation [10]:

$$E(x) = \sum_{\hat{v} \in \text{Orient}} E_{\hat{v}}(x). \tag{17}$$

Equation (17) can then be broken up into the following directional set of calculations:

$$E_{\hat{v}}(x) = \sqrt{F_{\hat{v}}^2(x) + H_{\hat{v}}^2(x)} \tag{18}$$

$$F_{\hat{v}}(x) = I(x) \otimes G_{\hat{v}}^{\text{even}}(x) \tag{19}$$

$$H_{\hat{v}}(x) = I(x) \otimes G_{\hat{v}}^{\text{odd}}(x) \tag{20}$$

$$G_{\hat{v}}^{\text{even}}(x) = \sum_{f \in \text{Freq}} M_{f,\hat{v}}^{\text{even}} \tag{21}$$

$$G_{\hat{v}}^{\text{odd}}(x) = \sum_{f \in \text{Freq}} M_{f,\hat{v}}^{\text{odd}}. \tag{22}$$

Alternatively, the calculation may be performed in the frequency domain where real and imaginary processing may be achieved simultaneously through

$$E_{\hat{v}}(x) = \|\text{Inverse}\, FT\,(\mathbf{I}(u) \times \mathbf{G}_{\hat{v}}(u))\| \tag{23}$$

where

$$\mathbf{I}(u) = FT(I(x)) \tag{24}$$

$$\mathbf{G}_{\hat{v}}(u) = \mathbf{G}_{\hat{v}}^{\text{even}}(u) + i\mathbf{G}_{\hat{v}}^{\text{odd}}(u) \tag{25}$$

$$= FT\left(G_{\hat{v}}^{\text{even}}(x) + iG_{\hat{v}}^{\text{odd}}(x)\right). \tag{26}$$

Calculation in the frequency domain can save vast amounts of time, especially when working in 3-D with the number of voxels in the order of 10^6. This vast amount of data would make convolutional computation in real space extremely time consuming.

F. Morlet Filter Construction

The even and odd Morlet filters $G_{\hat{v}}^{\text{odd}}(x)$ and $G_{\hat{v}}^{\text{even}}(x)$, as defined in (21) and (22), are combined to form a single *complex* Morlet filter $G_{\hat{v}}(x)$ for each orientation in either 2-D or 3-D space.

The filter is made to be equal in size to the original image. A 2-D image of size 256×256 pixels, for example, requires a Morlet filter of size 256×256 pixels to be built. This size is preserved when the FT is taken of each, allowing the two resulting frequency domain matrices to be multiplied directly.[3] This also applies to the 3-D case.

The difference between Morlet wavelets created in 2-D space and 3-D space lies in both the definition of the direction vector $\hat{v}$, and the distance of the vector x (measured from the center of the filter to each pixel in the filter).

In both the 2-D and 3-D case, a loop traverses through the filter (after it has been allocated memory and its locations initialized to zero) and iterates once for each frequency in the desired filter bank. At each pixel location in a traverse, the real and imaginary Morlet values at the current frequency are calculated, based on (15) and (16), and added to the complex Morlet value at that location. When the loop has been traversed for every frequency, the resulting filter represents the sum of scaled Morlet wavelets forming $G_{\hat{v}}(x)$.

Frequencies are chosen according to the parameters $fsep$ (frequency separation) and $fnum$ (number of frequencies). They are used to form a set of *scaled* frequencies, where each frequency is given by

$$f = \frac{1}{4 \times (fsep)^n} \tag{27}$$

where $1 < n < fnum$.

Manual tuning of these parameters was necessary so that the Morlet filters could be chosen to be sensitive to the size of expected features in the image. This involved adequate choice of k, $fsep$, and $fnum$. If k or $fsep$ were too small (i.e., $k < 0.3$ or $fsep < 1$), aliasing would occur in the filtering process. This is analogous to insufficient sampling of frequencies in the original image, and produces a distorted version of the original image rather than a local energy distribution as the result.

V. Results

A. Simulation Resources

Simulations conducted in this experiment were performed on an 18-slice 256×256 pixel image volume of chromosomes taken from a *Paphiopedilum* (orchid root tip) specimen. The chromosomes were imaged in mitotic metaphase, one of the substages during cell division. The results from this 18-slice dataset will be presented in this section. The original data was prepared and supplied by the Physical Optics Department, Sydney University.

The local energy testing process began with consideration of 2-D single orientation local energy calculations performed on individual slices in the image volume. The results are presented with representations of associated Morlet filters, giving a more intuitive look at how the individual filters interact with the image. The 2-D calculations were also conducted in order to determine the set of *tuning* parameters necessary to achieve the best sensitivity to features in the image slices. Tuning these parameters involved adequate choice of $fsep$ (the separation between frequencies in the filter bank), $fnum$ (the number of frequencies or scaled wavelets in the filter bank), and k (the ratio between the filter width and each frequency).

The simulations were then extended to calculation in multiple 2-D orientations, and finally, to three dimensions. The 3-D cases were then compared, on the slice level, with the 2-D results, in order to view any improvements offered by considering axial information in the image volume. Finally, the 3-D local energy results were used as features in four SOFM simulations. Each simulation using a different combination of image features describing the 3-D dataset.

B. Local Energy—Single Orientation Results

Fig. 8 shows a series of single orientation calculations made on the eighth slice in the 18-slice dataset, along with the associated Morlet filter in their calculation. The tuning parameters described in Fig. 8 were used in subsequent calculations performed in multiple directions, in both the 2-D and 3-D cases.

[3]Multiplication is not matrix multiplication, but rather, point-to-point. For example, in two dimensions, $I(u,v)G(u,v)$; in three dimensions, $I(u,v,w)G(u,v,w)$.

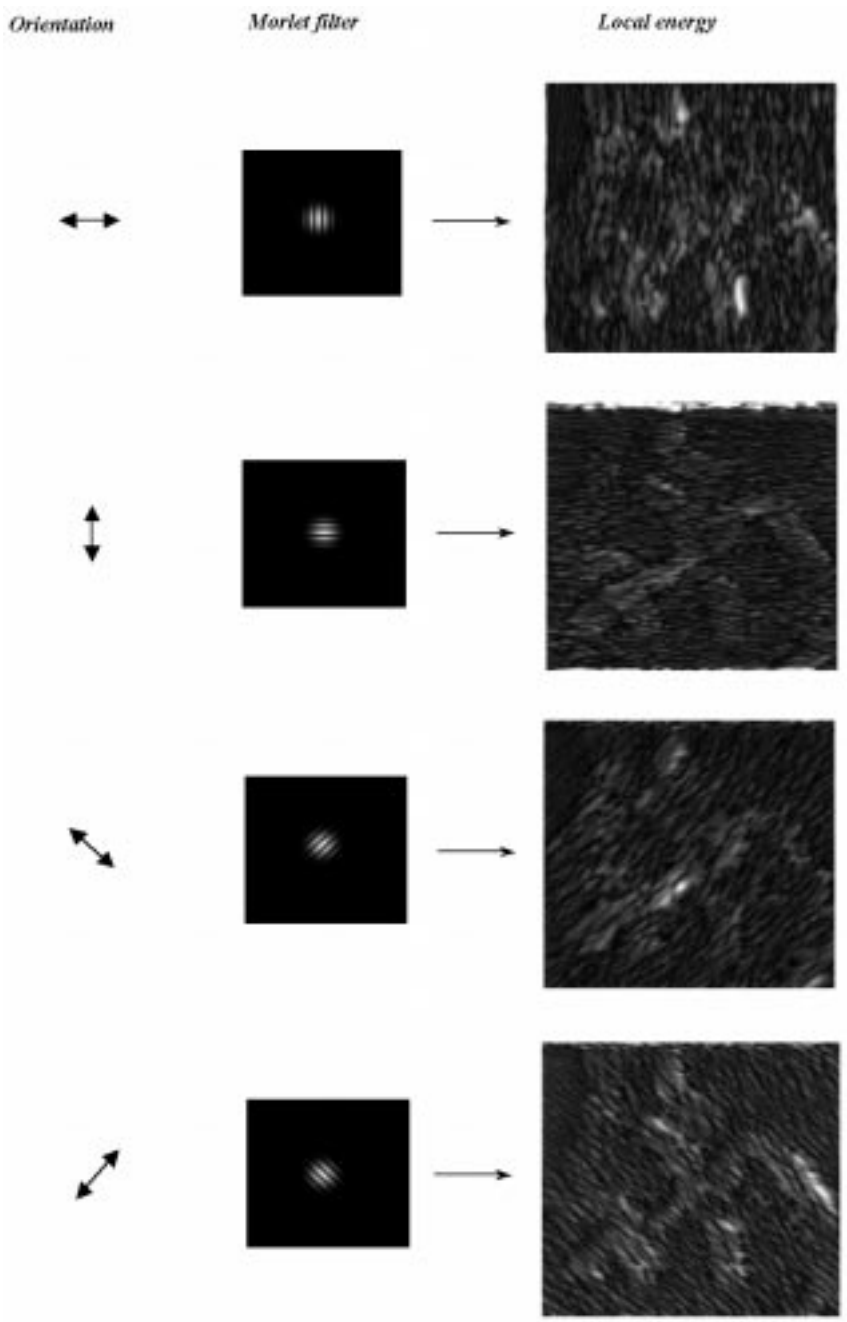
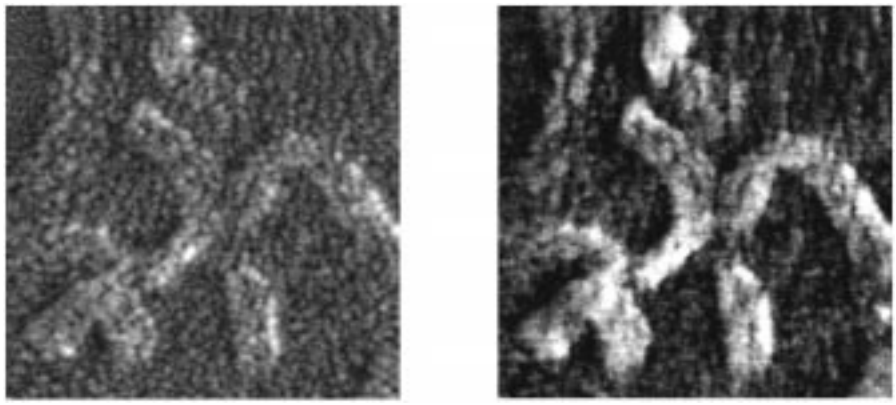


Fig. 8. Single orientation 2-D calculations of local energy on the eighth slice of the 18-slice dataset. The Morlet filters used in each local energy calculation are shown in the middle column, with their associated orientation in the 1st column. It is evident that particular directions pick out more information than others. The associated tuning parameters were: $fsep = 1.2$, $fnum = 5$, and $k = 0.5$.

C. Local Energy—Multiple Orientation Results

Fig. 8 shows a comparison between two local energy calculations made in multiple orientations. Fig. 9(a) shows the 2-D case, where local energy was calculated across 20 equally spaced orientations in 2-D space. Fig. 9(b) shows the same slice in the image volume, however, it is now the result of having performed a local energy calculation in 20 orientations in 3-D space, thereby considering information across image planes (slices). By considering the extra information available in the axial direction of the image volume, the 3-D case was expected to show greater ability to reject spurious changes in intensity in the original image than may have been picked up in any single 2-D filtered direction.

As is evident in these results, the 3-D calculation in Fig. 9(b) gives far superior extraction of the local energy distribution. The *noise-like* signal present in the 2-D result [Fig. 9(a)], due to slight intensity variations in the background, is effectively rejected in the 3-D case. This occurs as a direct result of the fact that the energy defining the chromosome locations is much greater in the 3-D case. As a result, immediate improvements in the subtleties of surface variations can be seen over the initial Hilbert transformed slice of Fig. 1(c), including variations that could only be detected in the axial (third) dimension.



(a) (b)

Fig. 9. Comparison of two 20-orientation calculations of local energy. The first (a) uses 20 orientations in 2-D space, while the second (b) uses 20 orientations in 3-D space. The 3-D case shows a clear improvement in the definition of local energy as distributed through the image volume.

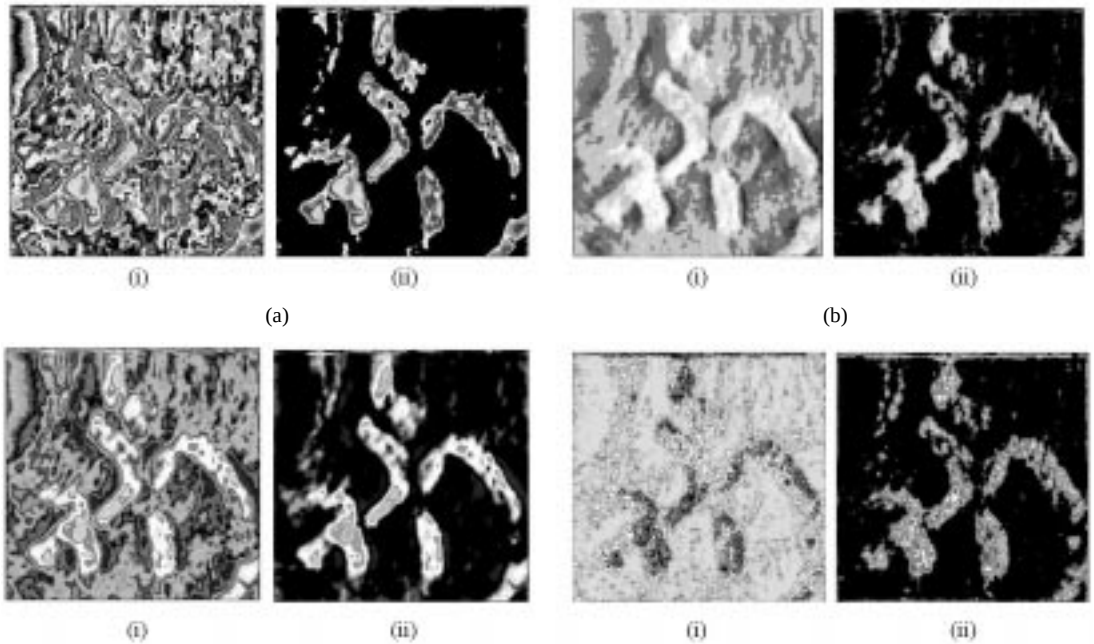


Fig. 10. Segmentation results of the ninth slice of the 18-slice dataset for four test cases. 3-D segmentation was performed using a 4×4 neural map. The two images for each test case are shown: (i) shows the raw output of the SOFM, showing $4 \times 4 = 16$ different colors each representing a surface feature tuned by a single neuron; (ii) shows the same segmentation, with the unwanted intensities (surfaces) removed, revealing the chromosome structures. (a) Case 1—uses pixel intensity, mean, and variance. (b) Case 3—uses pixel intensity, local energy, and variance. (c) Case 2—uses pixel intensity, mean, and local energy. (d) Case 4—uses local energy (LE), mean of LE, and variance of LE.

D. SOFM Segmentation

Primary simulations were conducted on four test cases of the 18-slice dataset for the purpose of examining the effects on the segmented result due to different combinations of input features. The four test cases were created. The sets of features (each three-fold) used in each test case were as follows.

Case 1: {pixel intensity (PI), localized mean (M), localized variance (V)};
Case 2: {PI, M, local energy (LE)};
Case 3: {PI, LE, V};
Case 4: {LE, localized mean of LE, localized variance of LE}.

At this point, we make note that case 1 represents a resimulation of previous work conducted by Nguyen [8]. This is done for the purpose of evaluating the advantages of the different methods. Due to the nature of 3-D visualization (performed later), this is an essential step for unbiased visual inspection. As such it is possible to make comparisons across all simulations viewing each from a similar aspect with similar rendering conditions.

The SOFM was run in two consecutive stages according to the following parameters:

First phase:

$$\alpha_o = 0.9$$
$$\alpha_{\mathrm{TH}} = 0.05$$
$$\text{no. Cycles} = 11$$

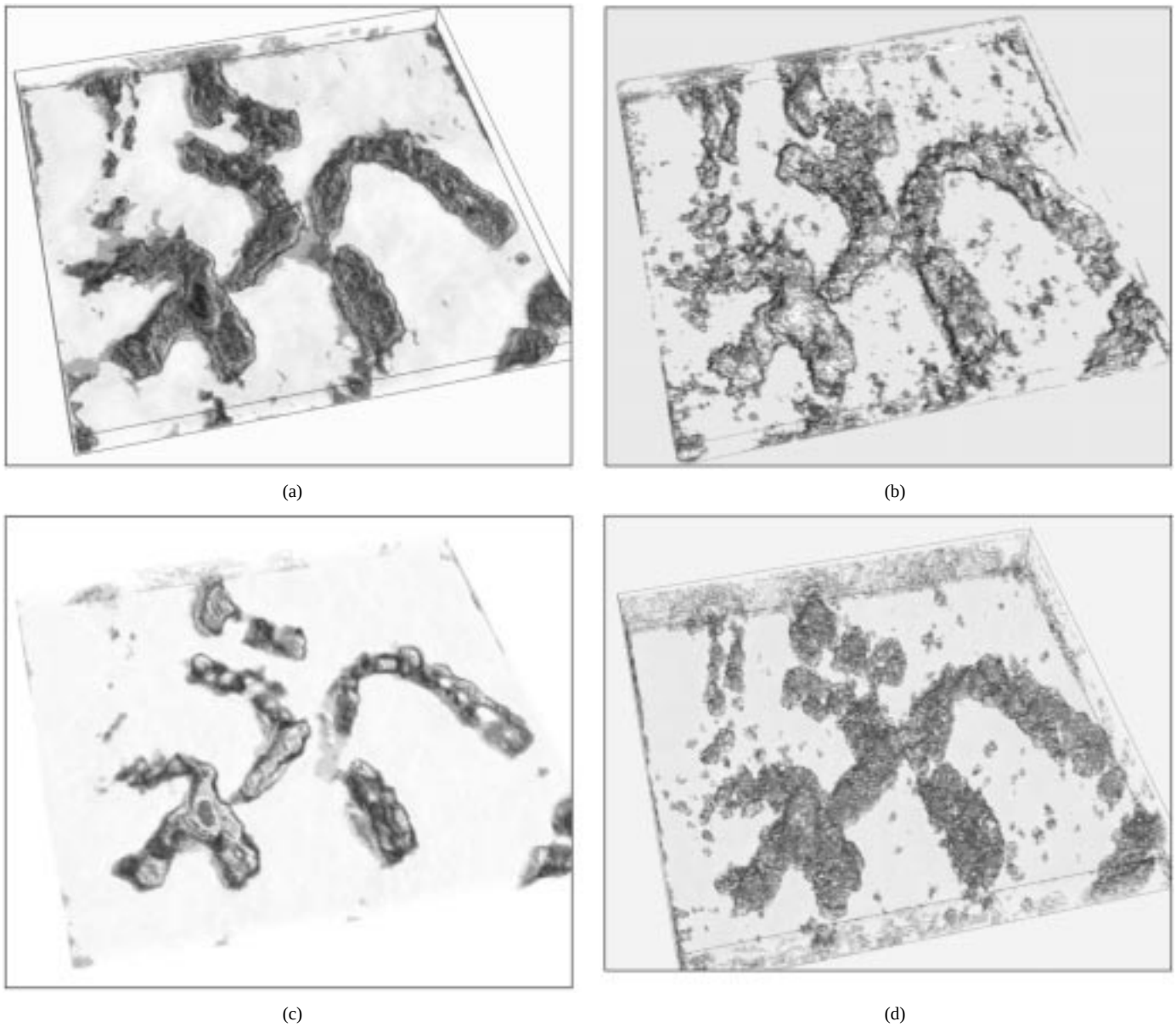


Fig. 11. (a) Case1—Volume visualization of an 18-slice data set of orchid root tip chromosomes after segmentation using a 4×4 SOFM and case 1 image statistics. (b) Case 2—Volume visualization of an 18-slice data set of orchid root tip chromosomes after segmentation using a 4×4 SOFM and case 2 image statistics. (c) Case 3—Volume visualization of an 18-slice data set of orchid root tip chromosomes after segmentation using a 4×4 SOFM and case 3 image statistics. (d) Case 4—Volume visualization of an 18-slice data set of orchid root tip chromosomes after segmentation using a 4×4 SOFM and case 4 image statistics.

Second phase:

$$\alpha_o = 0.4$$
$$\alpha_{\mathrm{TH}} = 0.01$$
$$\text{no. Cycles} = 9$$

where $\tau\alpha = 3$, $\tau\sigma = 5$.

Fig. 10 shows the ninth slice of the thin dataset, after each application of the SOFM process using the different definitions of feature space (denoted by cases 1–4 as defined above).

The SOFM structure for these simulations uses a 4×4 neural map. The map, thus, categorizes pixels into 16 different categories, each representing a characteristic surface type in the original image. The image slices, as presented in Fig. 10, were examined using the graphical program xv. This preliminary form of evaluation involved manually removing the surface types that were not associated with the chromosome bodies. This was achieved by visual inspection, i.e., by setting the intensity values in the image corresponding to nonchromosome surfaces to black, while keeping the others. The resulting image slices with unwanted surfaces removed, are also shown in Fig. 10, and give some indication as to the success of each particular feature set in influencing the final segmentation.

There are two main points that can be drawn from a comparison of these 2-D slices. The first is that when local energy is used as a feature, more greylevels are mapped to the chromosome regions. This means that more of the neurons are tuning themselves to orientations in feature space that represent characteristic features of the chromosomes. The second point is that when local energy is used, there is higher correlation between

the *number* of greylevels assigned to the different surface features on the chromosomes [Fig. 10(b)–(d) (i)], as opposed to surface features in the background. Thus, revealing an image in which the chromosome regions are immediately more distinguishable.

All of these results, as viewed on a 2-D basis, however, fail to adequately portray the 3-D nature of the resulting chromosome structures. The ultimate aim of this experiment was to perform the segmentation in order that these resultant images could be viewed in their 3-D state. This is achieved through use of the software package *VoxelView*, which provides an environment for volume rendering three-dimensional data.

Fig. 11(a)–(d) shows 3-D rendered results of each of the four simulations conducted using the 4×4 neural map. In each case, a front view of the rendered result is presented. These images form the basis for sets of stereo pair images, each taken from a different angle of view, which can be viewed through stereo-vision spectacles, highlighting the true 3-D nature of the chromosomes within.

In Fig. 11(a) and (b), which incorporates the mean of the original pixel intensities, the segmentation tends to blur or reveal somewhat coagulated chromosome bodies, with Fig. 11(b), (using local energy) revealing slightly more surface detail. Fig. 11(d), using only LE based statistics, produced an extremely highly detailed result, showing very dramatic (perhaps exaggerated) variations across the chromosome surfaces. This result reflects the hypersensitivity of the local energy function to different feature profiles in the image.

Finally, Fig. 11(c), using local energy and variance, depicts a dramatic improvement in clarity of the fine structure of the chromosomes, to the point where the two components making up a chromosome (known as *sister chromatids*) are visibly twisted or coiled together, in some places bunched up—a property consistent with the stage of cell mitosis in which the chromosomes were imaged.

VI. Conclusion

The use of the confocal microscope to image translucent specimens such as chromosomes, allows for increased optical sectioning across a specimen volume, aiding in the formation of 3-D image data. The DIC mode used to produce the image slices, however, requires further image processing in order that the objects within the specimen volume can be extracted from the image background. Kohonen's SOFM is a neural network approach to segmenting the chromosomes within the image volume, resulting in adequate extraction of the chromosome bodies, as well as resolving surface detail of their structure. Local energy has been explored as a possible feature to be used in directing the SOFM during segmentation, along with other standard features such as original, mean and variance of pixel intensities. Local energy has shown to be an extremely effective feature for resolving surface details of chromosomes from confocal DIC images beyond levels achieved in the past. Visualization of chromosome images produced in this fashion is possible, greatly enhancing the ability for biologists to observe, in three dimensions, the intricate nature of plant chromosomes.

Acknowledgment

The authors would like to thank the Physical Optics Department, University of Sydney, and the staff of Sydney Vislab[4] for technical assistance and resources used in the visualization of datasets.

References

[1] C. J. Cogswell, K. G. Larkin, M. R. Arnison, and J. W. O'Byrne, "3D fourier analysis methods for digital processing and 3D visualization of confocal transmission images," *SPIE*, vol. 2412, pp. 230–235, 1995.

[2] C. J. Cogswell and C. J. R. Steppard, "Confocal differential interference contrast (DIC) microscopy: Including a theoretical analysis of conventional and confocal DIC imaging," *J. Microscopy*, vol. 165, no. 1, pp. 81–101, 1992.

[3] P. A. Feineigle, A. P. Witkin, and V. L. Stonick, "Processing of 3-D DIC microscopy images for data visualization," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 4, 1996, pp. 2160–2163.

[4] Z. Kam, "Microscopic differential interference contrast image processing by line integration (LID) and deconvolution," *Bioimaging*, vol. 6, pp. 166–176, 1998.

[5] E. B. van Munster, L. J. van Vliet, and J. A. Aten, "Reconstruction of optical pathlength distributions from images obtained by a wide-field differential interference contrast microscope," *J. Microscopy*, vol. 188, pp. 149–157, 1997.

[6] M. R. Arnison, C. J. Cogswell, N. I. Smith, P. W. Fekete, and K. G. Larkin, "Using the Hilbert transform for 3-D visualization of differential interference contrast microscope images," *J. Microscopy*, vol. 199, pp. 79–84, 2000.

[7] H. Yin and N. M. Allinson, "On the distribution and convergence of feature space in self-organizing maps," *Neural Computation*, vol. 7, pp. 1178–1187, 1995.

[8] P. Nguyen, L. Guan, L. Cinque, R. Romagnoli, and S. Levialdi, "Self organizing map for segmenting 3-D biological images," in *Proc. 14th Int. Conf. Pattern Recognition*, vol. 1, Brisbane, Australia, 1998, pp. 471–473.

[9] B. Robbins and R. A. Owens, "2D feature detection via local energy," *Image Vis. Computing*, vol. 15, pp. 353–368, 1997.

[10] C. Pudney, M. Robins, B. Robbins, and P. Kovesi, "Surface detection in 3-D confocal microscope images via local energy and ridge tracing," *J. Comput. Assist. Microscopy*, vol. 8, no. 1, pp. 5–20, 1996.

[11] A. Oppenheim and R. Schafer, *Discrete-Time Signal Processing*. Englewood Cliffs, NJ: Prentice Hall, 1989.

[12] S. Haykin, *Neural Networks a Comprehensive Foundation*. New York: Macmillan College, 1994.

[13] M. C. Morrone and D. C. Burr, "Feature detection in human vision: A phase-dependent energy model," in *Proc. Roy. Soc. London B*, vol. 235, 1988, pp. 221–245.

[14] M. C. Morrone and R. A. Owens, "Feature detection from local energy," *Pattern Recogn. Lett.*, vol. 6, pp. 303–313, 1987.

[15] S. Venkatesh and R. A. Owens, "An energy feature detection scheme," in *Proc. IEEE Int. Conf. Image Processing*, Singapore, 1989, pp. 553–557.

[16] J. Morlet, G. Avens, E. Fourgeau, and D. Giard, "Wave propagation and sampling theory—part ii: Sampling theory and complex waves," *Geophysics*, vol. 47, no. 2, pp. 222–236, 1982.

Matthew J. Kyan received the B.Sc. degree in computer science in 1998 and the B.Eng. degree in electrical engineering in 1999 from the University of Sydney, Australia. Since 1999, he has been working toward the Ph.D. degree in electrical engineering, with a thesis in dynamic magnetic resonance image analysis.

His research interests include biomedical and multimedia signal processing, computational intelligence, neural networks, and 3-D image visualization.

In May 1999, Mr. Kyan was awarded the Inaugural Siemens National Prize for Innovation for his work with 3-D confocal image analysis, a prize awarded for the most innovative undergraduate thesis in Australia for the fields of computer, electrical, electronic, or telecommunications engineering.

[4]http://www.vislab.usyd.edu.au/

Ling Guan (S'88–M'90–SM'96) received the Ph.D. degree in electrical engineering from University of British Columbia, Canada, in 1989.

From 1989 to 1992, he was a Research Engineer of Array Systems Computing, Inc., Toronto, Canada, in machine vision and signal processing. From October 1992 to April 2001, he was a Faculty Member at University of Sydney, Australia. In May 2001, he joined the Ryerson Polytechnic University, Toronto, Canada, where he is currently a Professor in Department of Electrical and Computer Engineering. In 1994, he was a visiting fellow at British Telecom. In 1999, he was awarded a visiting fellowship by Australia Academy of Science/Japan Society for the Promotion of Science and worked in Tokyo Institute of Technology. In 2000, he was on sabbatical leave and a Visiting Professor at Princeton University, Princeton, NJ. His research interests include multimedia processing and systems, optimal information search engine, signal processing for wireless multimedia communications, computational intelligence and machine learning, and adaptive image and signal processing. He has published more than 130 technical articles, and is the editor/author of two books *"Multimedia Image and Video Processing"* and *"Adaptive Image Processing: A Computational Intelligence Perspective"*. He is an associate editor of IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION, *IS&T/SPIE Journal of Electronic Imaging* and *Journal of Real-Time Imaging*. In 1999, he co-guest-edited the special issues on Computational Intelligence for PROCEEDINGS OF THE IEEE. He also serves on the editorial board of CRC Press' Book Series on Image Processing. He has been/is involved in organizing numerous international conferences. He played the leading role in the inauguration of the IEEE Pacific-Rim Conference on Multimedia, and served as the Founding General Chair in Sydney 2000.

Dr. Guan is a member of IAPR and SPIE. He is currently serving on IEEE Signal Processing Society Technical Committee on Multimedia Signal Processing.

Matthew R. Arnison received the B.Sc. (Hons) degree in physics, from the University of Sydney, Sydney, Australia, in 1993 and is Ph.D.degree student in physics at the same university

His current research project is extending the depth of field in high-resolution microscopy using wavefront coding. This is an international collaboration with team members in Oxford University, Oxford, U.K., and Colorado University, Boulder. His research history includes adaptive optics for confocal microscopy, and 3-D microscope image reconstruction.

Carol J. Cogswell has over 30 years of experience in optical microscopy and image processing. She has recently taken up a position as an Adjunct Professor at the Univeresity of Colorado, Boulder, in the Department of Electrical and Computer Engineering where she continues her research in developing new microscopy imaging techniques. She was previously on the faculty of the University of Sydney, School of Physics and Deputy Director of the Australian Key Centre for Microscopy and Microanalysis.

Dr. Cogswell continues to chair the annual SPIE conference, "Three-Dimensional Microscopy: Image Acquisitjon and Processing," at Photonics West.