Estimation of Average Cell Shape from Digital Images of Cellular Surfaces

Peter J.W. Iles, David A. Clausi and G. Wayne Brodland
University of Waterloo, Waterloo, ON, Canada, N2L 3G1
(pjwiles, dclausi, brodland)@uwaterloo.ca

***Abstract*— A method is proposed to determine the average cellular geometry in high-resolution images of embryonic epithelia. The concept of a 'composite cell' is used to represent the average cell shape. This composite cell can provide evidence of stresses present in the epithelia. A new adaptive contrast enhancement routine is applied to the input image first, followed by an iterative watershed segmentation. The composite cell is calculated from the segmentation results. Qualitative results show that this computationally inexpensive algorithm produces accurate results for a variety of image sizes, contrast levels, cell shapes and appearances.**

***Index Terms*— composite cell, adaptive contrast enhancement, watershed segmentation, shape detection.**

I. Introduction

THERE are many applications for shape recognition, such as automated product inspection and image data mining. This paper proposes an approach to one such application: determining the average cellular geometry in high-resolution images of embryonic epithelia. This approach is shown to be computationally inexpensive and robust to different image sizes, contrast levels, cell shapes and appearances. An example embryonic epithelium image is shown in Figure 1.

This technique is part of a larger research project headed by G.W. Brodland at the University of Waterloo who is modelling the biomechanical causes of embryonic development. This is being done to determine the cause of birth defects such as spina bifida, a condition which occurs when the spine forms incorrectly early in the development process. It is believed that physical stresses occurring in the epithelium form the basis of the occurrence of this condition. It has been shown that these stresses are related to the shape of the epithelium cells [1].

To describe the average cellular geometry of a given patch of the epithelium, Brodland and Veldhuis developed the idea of an elliptical 'composite cell' [2]. The composite cell represents the average shape of cells in the patch. The shape of this composite cell gives information on the overall direction and magnitude of stresses in the patch. The composite cell is completely quantified by three parameters: its orientation (θ, the angle of its major axis from the horizontal) and the lengths of its major and minor axes, L_{major} and L_{minor}, respectively as shown in Figure 2. The goal of the algorithm described in this paper is to calculate the composite cell parameters from a digital image of a patch of an embryonic epithelium.

Puddister approached this problem by first transforming the image into the frequency domain, then using the shape of the spectrum to infer the average cell shape and size [3]. This, however, suffered because many cell pattern repetitions (and hence high quality, large images) are required to produce a usable frequency spectrum. As shown in the example input image, this cannot be assumed. Most other multiple shape recognition algorithms use *a priori* shape information to detect a desired shape in an image [4], [5]. This does not work for the given problem, however, since the cells need to be detected *whatever* their shape. These two aspects (large, high-quality images, and *a priori* shape information) are not required in the proposed algorithm. The algorithm can detect any convex cell shape, including the case where only a small number of cells can be detected.

There are three main stages in this method, as shown in Table I. First, local and global contrast problems are addressed through the application of a new adaptive contrast enhancement routine. Next, an iterative watershed segmentation technique is used to detect individual cells. Finally, a st

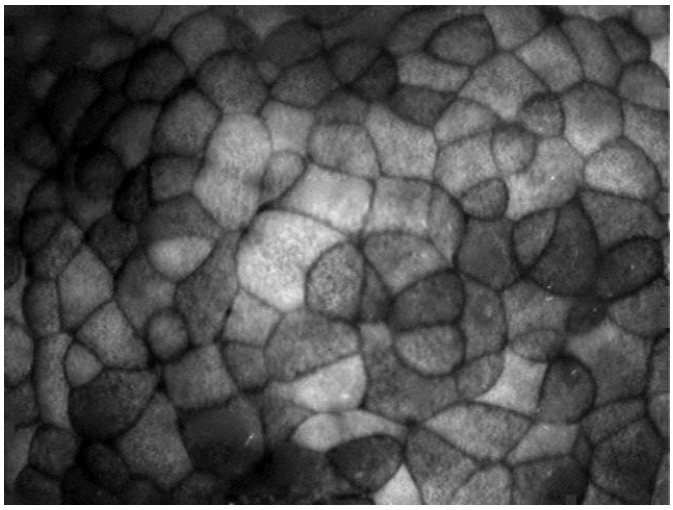

Fig. 1. Example input embryonic epithelium image

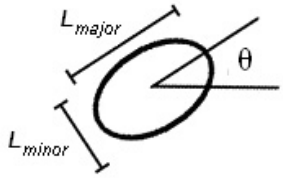

Fig. 2. The composite cell is defined by its orientation θ and its major and minor axes, L_{major} and L_{minor}

Proceedings of the First Canadian Conference on Computer and Robot Vision (CRV'04)
0-7695-2127-4/04 $20.00 © 2004 IEEE

TABLE I

ALGORITHM SUMMARY. INPUT IS EMBRYONIC EPITHELIUM IMAGE, OUTPUT IS θ, L_{major} AND L_{minor} THAT DEFINE COMPOSITE CELL

Step	Description	Parameters
1	*Adaptive contrast enhancement*	
1.1	Calculate local averages	r
1.2	Apply local sigmoidal transfer fcns	m
2	*Iterative watershed segmentation*	
2.1	Take image complement	
	loop	
2.2	Apply extended-minima transform	d
2.3	Apply imposed-minima transform	
2.4	Perform watershed segmentation	
2.5	Decrement d	d_{dec}
2.6	exit if min # segments found	n_{min}
	end loop	
2.7	Remove boundary segments	
3	*Calculation of composite cell*	
3.1	Find centroid of each cell	
3.2	Find median 16 ray directions	
3.3	Fit ellipse to find θ, L_{major} and L_{minor}	

calculation routine is applied to calculate the composite cell. The organisation of this paper is as follows: in Sections II-IV, the three major stages outlined above are presented. Results are given in Section V, future work is outlined in Section VI and conclusions are drawn in Section VII.

II. ADAPTIVE CONTRAST ENHANCEMENT

The purpose of this step is to increase the local contrast of the input image, $f_{in}(x, y)$, to make the image appropriate for watershed segmentation. Global contrast problems, such as illumination variation (due to microscope lighting) as well as local contrast problems, introduced during image acquisition or due to the pigmentation of the cells, are corrected. For example, the patch of cells in Figure 1 shows a severe illumination variation and contains cells of varying pigmentations. To correct for these problems, an adaptive contrast enhancement algorithm is used.

There are many adaptive contrast enhancement algorithms in the literature from which to choose. An adaptive version of unsharp masking is suggested by Polesel et. al [6] and a multiscale contrast enhancement method suggested by Boccignone [7] both use measures of local contrast to decide on the amount of enhancement needed. However, the extended-minima transform (see description in the next section) in the watershed segmentation algorithm needs only a *minimum* level of contrast to detect the locations of cells. For this reason the local contrast does not need to be measured, and contrast can simply be enhanced everywhere in the image. An adaptive histogram equalisation algorithm [8] was also considered. This algorithm is computationally expensive, however, because it requires the calculation of the ordered set of local intensities. It was therefore decided to develop a computationally-inexpensive algorithm that enhances local contrast everywhere in the image.

The developed algorithm involves spreading local pixel intensities away from a local average through use of a sliding sigmoidal transfer function. This achieves the desired contrast enhancement. The algorithm is described in the following subsections.

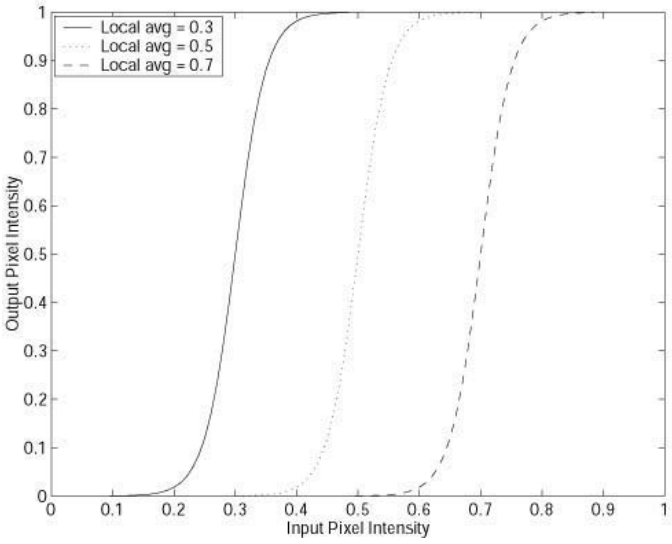

Fig. 3. Sigmoid transfer functions for average local intensities of 0.3, 0.5 and 0.7. ($m = 10$)

A. Local Average Calculation

A calculation of local average intensities is done by masking the input image $f_{in}(x, y)$ with a uniform circular local-averaging mask, $h_r(m, n)$, where r is the mask radius in pixels. This local average is used to position the sigmoidal transfer function as described in the next subsection. The resulting image is simply:

$$f_{avg}(x, y) = h_r(m, n) * f_{in}(x, y).$$

B. Sigmoidal Transfer Function Point Operation

Using the the local average image, the local contrast of the original image is enhanced with this point operation. For each pixel, a sigmoidal transfer function is created which is centred on the local average. When a maximum slope of $m > 1$ is used, this expands the pixel intensity range around the local average, increasing the local contrast. This entire operation can be tidily performed with the equation

$$f_{out}(x, y) = \frac{1}{1 + e^{4m[f_{avg}(x,y) - f_{in}(x,y)]}}.$$

where $f_{out}(x, y)$ is the resulting image. Note that all equations given in this paper are calculated for real-valued images (intensity values from 0 to 1).

The sigmoidal function was chosen because it is monotonically increasing (to provide a sensible pixel intensity mapping), and the translation and maximum slope can be set as single parameters in the equation. Examples of the sigmoidal transfer function for different local averages are shown in Figure 3.

The computation time for the entire contrast enhancement algorithm is < 1 second in MATLAB® for a 256x256 image (Pentium 4 @2.4GHz). The result of performing this algorithm on the image shown in Figure 1 is shown in Figure 4. Note the significant enhancement that has been done to the dark upper-left region.

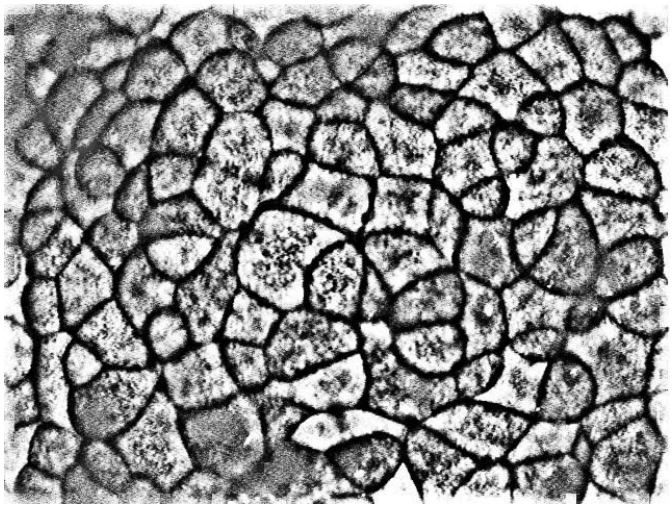

Fig. 4. Result of applying adaptive contrast enhancement method to image in Figure 1. ($m = 10$; $r = 15$)

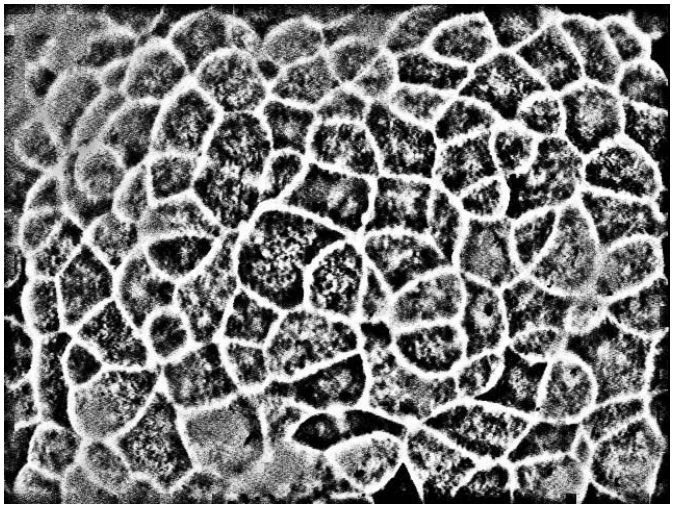

Fig. 5. Image complement taken of contrast enhancement result shown in Figure 4

III. Iterative Watershed Segmentation

In order to calculate individual statistics of the cells, it is necessary to first segment the image, giving each cell a unique label. Once this is done, each cell can be analysed in isolation. The proposed iterative watershed method segments the output image of the adaptive contrast enhancement method, $f_{out}(x, y)$, into the individual cells wherever possible.

The classic watershed segmentation method treats an image as a three-dimensional landscape (where light areas are 'hills' and dark areas 'valleys') and segments it into its component watershed areas [9]. Every point in a given watershed segment can be thought of as draining into a common area, just as every point in the Lake Superior watershed drains into Lake Superior. It is up to the algorithm developer to choose the most appropriate network of drainage areas for the image. The watershed segmentation method is used for this application because it is fast and does not require *a priori* information of the number of segments.

In this context, the contrast-enhanced epithelium in Figure 4 contains many round hills (the cells) surrounded by a network of narrow valleys (the cell boundaries). The proposed iterative watershed segmentation method involves first taking the complement of this image,

$$f_{compl}(x, y) = 1 - f_{out}(x, y)$$

so that the cells become basins and the boundaries become ridges, as shown in Figure 5. This is done because the watershed segmentation method will now recognise the cell basins as individual segments.

After the image complement is taken, there are three further steps involved in the proposed segmentation method. These are iterated over until the number of segments (cells) detected exceeds a user-defined minimum, n_{min}. As described below, the number of segments detected increases for each iteration. It is up to the user to decide the minimum number of segments necessary in order to make up a representative number of all the cells in the image. The reason for choosing n_{min} as a parameter is explained below, after some necessary explanations.

The first of the three steps is to take the extended-minima transform F_1 (*imextendedmin* function in MATLAB®) of the contrast-enhanced image which locates areas that are of a minimum depth d below (darker) than their neighbours [10].

$$f_{extended}(x, y) = F_1(f_{compl}(x, y), d)$$

An 8-connected neighbourhood (3x3 window) is used in this transform. These areas will become the 'drainage areas' used by the watershed algorithm; the number of areas detected here is the same as the number of segments detected by the watershed method. The minimum depth rule is employed to make the extended-minima transform robust against minor intensity variations. It is desired that these 'drainage areas' correspond to the interiors of all the cells on a one-to-one basis. The second of the three steps is the imposed-minima transform F_2 (*imimposemin* function in MATLAB®) that imposes local minima on the contrast-enhanced image only where the extended-minima transform detected the location of cells.

$$f_{imposed}(x, y) = F_2(f_{compl}(x, y), f_{extended}(x, y))$$

Each local minima now coincides with a detected cell location, and is seen as a drainage area for the watershed method. Finally, the watershed segmentation method (*watershed* function in MATLAB®) is performed. The imposed-minima transform and the watershed segmentation are performed using an 8-connected neighbourhood as well.

The iteration is done over the minimum depth d. A large d_o is initially chosen which will result in only a few cells being detected for most types of input images. As d is lowered by decrement d_{dec}, more cells are detected since the intensity threshold criteria is being relaxed. When the number of cells detected, n, is greater than n_{min}, the iterations stop. The necessity of setting d as a hard threshold is removed through the use of the iterations, making this stage more rol[illegible] different input image types. The hard threshold n_{min}

Proceedings of the First Canadian Conference on Computer and Robot Vision (CRV'04)
0-7695-2127-4/04 $20.00 © 2004 IEEE

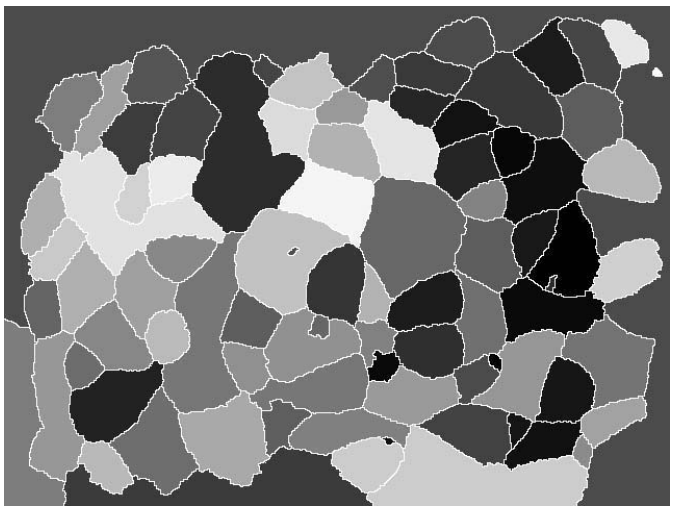

Fig. 6. Watershed segmentation performed on image shown in Figure 5

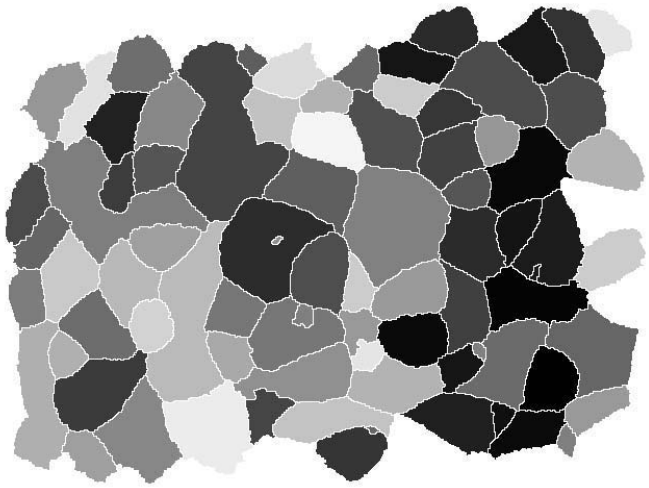

Fig. 7. Boundary cells removed from segmentation shown in Figure 6

more confidently set by the user than d, since it represents only the number of cells that must be detected to be confident in the results.

An image with well-defined cells and edges requires few iterations since the intensity depressions are already deep. An image with poorly-defined cells and edges requires additional iterations. An example segmentation is shown in Figure 6.

A single additional step is then performed on the segmentation result. Any cell that touches an edge of the image is removed. This is done because the overall shape of these cells is unknown, and hence cannot be used in the calculation of the composite cell. An example result of this procedure performed on the Figure 6 segmentation is shown in Figure 7.

IV. Calculation of Composite Cell

Given the segmented image, measurements of the individual cells are made. These measurements are then combined in order to calculate the composite cell. It is not, however, a straightforward task to find the 'average' shape the individual cells. It was decided to quantify a cell by taking ray measurements in multiple directions from its centroid to its edge. The medians of these measurements are then found over all n cells.

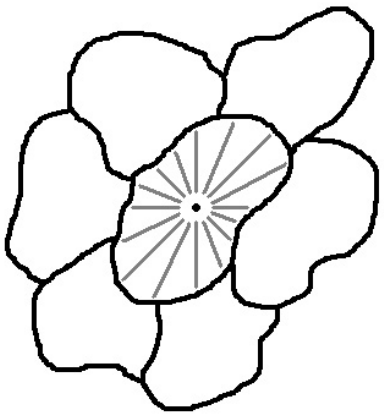

Fig. 8. 16 ray measurements taken from centroid to edge of cell

The location of the centroid of each cell is calculated using the *regionprops* function and the *centroid* statistic in MATLAB®. Ray measurements are then taken from the centroid to the cell boundary in 16 uniformly-spaced directions:

$$\theta = 0, \frac{\pi}{8}, ..., \frac{15\pi}{8} \ [rads]$$

The ray measurement of the i^{th} cell in the θ direction is:

$$ray_{i,\theta} = dist(centroid_i, edge_{i,\theta})$$

where $dist(a, b)$ is the Euclidean distance between a and b, $centroid_i$ is the location of the centroid of the i^{th} cell and $edge_{i,\theta}$ is the location of the cell boundary in the θ direction from the centroid of the i^{th} cell. These ray measurements are shown for a typical cell in Figure 8. The median measure (over all n cells) is calculated for each of the 16 ray directions. This results in 16 ray measurements:

$$ray_\theta = median(ray_{1,\theta}, ray_{2,\theta}, ..., ray_{n,\theta}).$$

The median measure was chosen over the mean due to its robust nature. The median measure prevents spuriously detected cells (both large and small) from having a major impact on the final result.

The ray_θ measurements define a sixteen-sided polygon that represent the average cell in the image. It is difficult, however, to quantify results that are in this form. For this reason, an ellipse is fit to this polygon. A best-fit ellipse (in the least-squares sense) is calculated using an algorithm developed by Fitzgibbon et al. [11]. This ellipse is the final goal of the algorithm: the elliptical composite cell defined by the three parameters θ, L_{major} and L_{minor}. These three parameters of the composite cell allow for direct comparisons of the stresses detected in different cell patches.

This completes the discussion of the major steps in the algorithm. A summary of the algorithm steps is shown in Table I. The next section outlines the experimental results.

Proceedings of the First Canadian Conference on Computer and Robot Vision (CRV'04)
0-7695-2127-4/04 $20.00 © 2004 IEEE

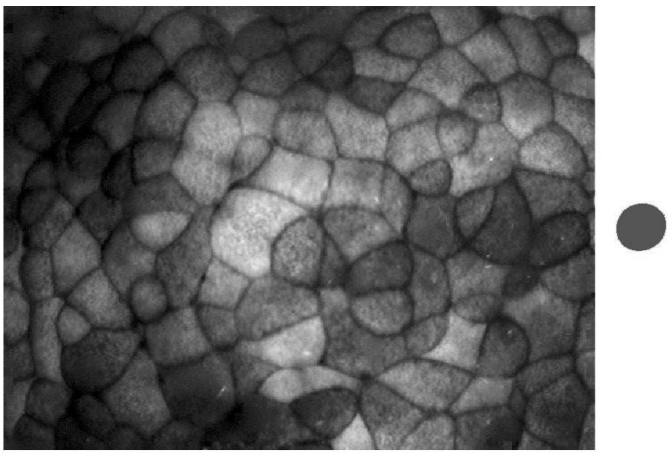

Fig. 9. Algorithm result for a 640x480 input image with varying contrast, large cells and well-defined edges. Parameters: $m = 10,\ r = 15,\ n_{min} = 25,\ d_o = 0.55$ with $d_{dec} = 0.04$. Results: $\theta = 27.8^o,\ L_{major} = 26.7,\ L_{minor} = 24.4$

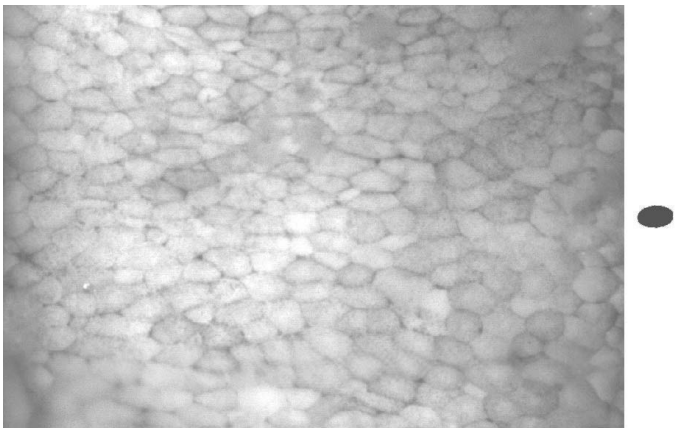

Fig. 10. Algorithm result for a 976x672 input image with lower contrast, small cells and less well-defined edges. Parameters: $m = 10,\ r = 15,\ n_{min} = 25,\ d_o = 0.55$ with $d_{dec} = 0.04$. Results: $\theta = 3.6^o,\ L_{major} = 27.5,\ L_{minor} = 16.2$

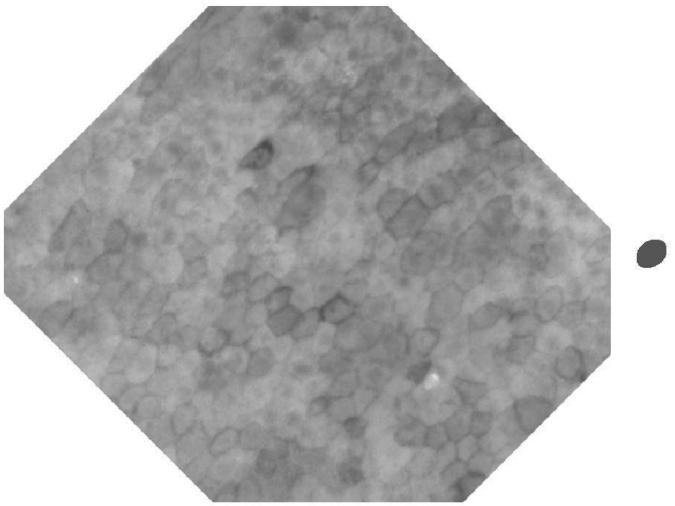

Fig. 11. Algorithm result for a 738x603 input image with low contrast, small cells and regions with undetectable edge information. Parameters: $m = 10,\ r = 15,\ n_{min} = 25,\ d_o = 0.55$ with $d_{dec} = 0.04$. Results: $\theta = 38.2^o,\ L_{major} = 18.7,\ L_{minor} = 13.5$

V. Experimental Results

The experimental results presented in this section are qualitatively evaluated. The composite cells generated by the algorithm are visually compared to the input images. Quantitative testing will be included in future work.

As shown in Figures 9-11, the results of the proposed technique accurately detect the average shape of cells in the image. The composite cell calculated by the algorithm is shown on the right in each of the figures. Each result was produced using identical parameters: sigmoidal slope $m = 10$, disc radius $r = 15$, minimum number of segments $n_{min} = 25$ and initial minimum depth $d_o = 0.55$ (measured in pixel intensity) with an decrement $d_{dec} = 0.04$. Computation times given are for MATLAB® running on a Pentium 4 @2.4GHz.

Figure 9 shows the technique performed on an epithelium image with high contrast in some areas, but also with significant illumination variation. The cells are large and their edges are distinct. The composite cell is defined by: $\theta = 27.8^o,\ L_{major} = 26.7,\ L_{minor} = 24.4$ (axes lengths given in pixels). This result captures the size and lack of significant orientation of the cells. Computation time for this 640x480 image is about 6 seconds.

The image shown in Figure 10 has lower contrast, less distinct edges and also a significant illumination variation. The cells are much smaller than in the previous image. The composite cell is defined by: $\theta = 3.6^o,\ L_{major} = 27.5,\ L_{minor} = 16.2$. This result accurately captures the size and horizontal orientation of the cells. Computation time for this larger image (976x672) is about 15 seconds.

Figure 11 shows a result for a low-contrast image with little cell detail in some areas (see upper left area). The composite cell is defined by: $\theta = 38.2^o,\ L_{major} = 18.7,\ L_{minor} = 13.5$. This result demonstrates that the algorithm is able to detect a sufficient number of cells to capture their noticeable diagonal orientation. Computation time for this 738x603 image is about 13 seconds.

As shown in the results presented, this algorithm detects the composite cell accurately for many types of embryonic epithelium images. These include images of varying size and those with high, low or variable contrast. Images containing regions of undetectable cell information are also shown to be analysed accurately by the proposed algorithm. Finally, the algorithm has been shown to be robust to cells of different sizes and at various orientations. This is all done with identical input parameters.

VI. Future Work

As mentioned above, rigorous quantitative testing is required to further validate this algorithm. Images with known composite cells will form a benchmark for results. Sensitivity testing to incremental rotation and stretching will also be completed.

It is secondly desired to make the proposed algorithm more robust to noisy images. The adaptive contrast enhancement algorithm enhances noise as well as edge detail. The approach to noise reduction must handle all expected sources of noise such as due to the optical characteristics of the microscope

Finally, a further investigation into the chosen parameters is necessary to make this algorithm more widely applicable. It is desired to remove any dependance on any hard thresholds. One of these is the local-averaging disc radius, r. The selection of this can have a significant effect on the algorithm results. If the cells are much larger than the chosen r an unrealistic contrast enhancement would result. This would occur because the algorithm would significantly increase contrast in the interior of the cells. The radius would optimally have a direct dependance on the average cell size, but finding the size is one of the purposes of the algorithm. Hence, some kind of iteration could be used here.

VII. Conclusions

This paper has presented a robust approach to finding an elliptical 'composite cell' which represents the average cell shape in a embryonic epithelium image. The composite cell, which is described by its orientation and the lengths of its major and minor axes, provides quantified information about stresses present in the epithelium.

A novel contrast enhancement method has been suggested, which is followed by an iterative watershed segmentation. The shapes of the individual segments are averaged together, and an ellipse is fit to this shape to determine the composite cell.

The proposed approach has been shown to be applicable for many input image sizes and appearances, cell sizes and orientations, and illumination scenarios. The approach is also computationally inexpensive; needing only 6 seconds for even large images (640x480) in MATLAB® on a Pentium 4 computer @2.4GHz.

References

[1] G. Brodland, J. Veldhuis, and D.-L. Chen, "How cell shape affects the stresses in a cell sheet," in *ASME IMECE*, Orlando FL, Nov. 5-10 2000, pp. 79–80.

[2] G. Brodland and J. Veldhuis, "Computer simulations of mitosis and interdependencies between mitosis orientation, cell shape and epithelia reshaping," *Journal of Biomechanics*, vol. 32, pp. 673–681, 2002.

[3] S. M. Puddister, "Estimating bulk geometrical properties of cellular structures," Master's thesis, University of Waterloo, Waterloo, ON, 2003.

[4] A. Ashbrook *et al.*, "Multiple shape recognition using pairwise geometric histogram based algorithms," in *Proc. Fifth International Conference on Image Processing and its Applications*, Edinburgh , UK, July 1995, pp. 90–94.

[5] H. Moon *et al.*, "Optimal edge-based shape detection," *IEEE Trans. Image Processing*, vol. 11, no. 11, pp. 1209–1227, Nov. 2002.

[6] A. Polesel, G. Ramponi, and V. Mathews, "Image enhancement via adaptive unsharp masking," *IEEE Trans. Image Processing*, vol. 9, no. 3, pp. 505–510, 2000.

[7] G. Boccignone, "A multiscale contrast enhancement method," in *Proceedings., International Conference on Image Processing*, vol. 1, Santa Barbara, CA, Oct. 1997, pp. 306–309.

[8] J. A. Stark, "Adaptive image contrast enhancement using generalizations of histogram equalization," *IEEE Trans. Image Processing*, vol. 9, no. 5, pp. 889–896, May 2000.

[9] K. R. Castleman, *Digital Image Processing*. Englewood Cliffs, NJ: Prentice Hall, 1996.

[10] *Image Processing Toolbox: for use with Matlab*, The Mathworks Inc., 2001.

[11] A. Fitzgibbon, M. Pilu, and R. Fisher, "Direct least-squares fitting of ellipses," vol. 21, no. 5, pp. 476–480, May 1999.