Segmentation of Bone Marrow Stromal Cells in Phase Contrast Microscopy Images

Christopher J. Bradhurst¹, Wageeh Boles¹, Yin Xiao¹,²

¹ School of Engineering Systems, Queensland University of Technology, Brisbane, Australia.
² Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Australia.

Email: (christopher.bradhurst|w.boles|yin.xiao)@qut.edu.au

Abstract

The morphology of bone marrow stromal cells (BMSCs) and the nature of phase contrast images make segmentation of such images challenging as many standard segmentation approaches do not work. This presents an obstacle to the development of systems that could use pattern recognition (PR) techniques to assess culture quality, since successful segmentation is an important precursor to successful pattern recognition. A method is presented for image normalisation and segmentation of cell regions within sub-confluent cell cultures of human bone marrow stromal cells, including a novel method of dealing with the halo associated with phase contrast images. The proposed method was evaluated by measuring its effect on the accuracy of a subsequent PR stage that was trained to discriminate between two BMSC cultures of differing quality. The accuracy achieved averaged 93% across four commonly used PR algorithms, corresponding to an overall accuracy gain of 17% compared to non-normalised, unsegmented images.

Keywords: segmentation, phase contrast, bone marrow stromal cells, mesenchymal stem cells, pattern recognition

1 Introduction

Bone marrow stromal cells (BMSCs), also called mesenchymal stem cells, are the subject of intense research in the biomedical community because of their potential for use in regenerative medicine. Although samples of these cells can be taken relatively easily from human donors and cultured in vitro, there is presently no quick and easy way to assess the quality of such samples.

With this in mind, a system is being developed with the aim of assessing BMSC culture quality from phase contrast images via the use of pattern recognition techniques. As part of this process, statistical measures (potentially indicative of cell culture quality) must be extracted from the image regions that contain cells. This requires such cell regions be segmented from the background. This is a crucial step since the success of pattern recognition techniques in image processing systems is known to be very dependent upon the effectiveness of the preceding segmentation [1].

However, segmentation of living cell images presents a number of problems. Firstly, the images must be taken in a non-invasive manner so that the cells may continue to grow unharmed. This precludes the use of fluorescent markers and histological stains that can make segmentation much simpler. Furthermore, the nature of images taken using phase contrast (figure 1) renders standard segmentation approaches ineffective. For example, the background gray levels are a subset of the cell gray levels and so intensity thresholding cannot be used. Furthermore, the bright halo that appears around many cell borders creates false edges when standard edge detection methods are applied.

Figure 1: Phase contrast images of BMSCs taken at 10x magnification – note the irregular cell morphology and bright halo around many of the cell borders. The images are typical of those taken from (a) sample A and (b) sample B

A variance based method of segmenting living cells has been proposed by Wu et. al. [2], however being designed for normal bright field (non phase contrast) images, it does not deal adequately with halo regions. Nevertheless, the variance based approach is still useful in locating the approximate cell regions, and
has been employed in a modified form in this paper. A probabilistic living cell segmentation model [3] has been developed for segmenting hematopoietic stem cells under phase contrast, however the method takes advantage of the regular circular morphology of such cells. Unfortunately, the irregular morphology of BMSCs renders this approach ineffective. Debeir et. al. [4] have recently suggested the use of marked weakened watersheds, a method used to deal with the problem of oversegmentation typical of the watershed approach. This method appears to deal well with the irregular morphology and phase contrast halo, however it requires that cell centroids be initially marked. Other applications (for example [5]) designed to track the movement of living cells over multiple phase contrast image frames also involve segmentation steps. However, the techniques employed are concerned primarily with tracking cell centroids rather than producing an accurate contour of the cell regions themselves and are thus not considered suitable for the present application.

The segmentation approach presented here, involves three processing steps. Firstly, the images are normalised to reduce the effect of different lighting and camera exposure settings. Two common approaches are evaluated, histogram equalisation and contrast stretching. Secondly, approximate cell regions that include the halo areas are located in a manner similar to that proposed in [2]. This step is termed rough segmentation. The third phase (termed refined segmentation) refines this approximate cell region by excluding the halo and thus providing a much closer fit to the real cell contour. As the halo is likely dependent in some manner on the underlying cell structural features, it is important to determine if the halo contains information useful to the assessment of cell culture quality. For this reason, both rough and refined segmentation methods are evaluated.

In order to determine the most effective normalisation and segmentation approach (i.e. the one that maximises PR performance), the algorithms are evaluated by measuring their effect on the accuracy of a subsequent pattern recognition system’s ability to discriminate between images from two BMSC cultures of differing quality. As part of this process, statistical features are extracted from the images and segmented regions and then input to the pattern recognition system. During training the statistical features are provided to the PR system along with the classification (i.e. sample A or sample B). During testing, the PR system is provided with only the statistical features and must guess the correct sample that the image belongs to.

Below, in section 2, we describe technical details of the images followed by the image normalisation and segmentation algorithms and the method of evaluating their performance. In section 3 we present and discuss results quantifying the effect of image normalisation, rough segmentation (which includes the halo regions) and refined segmentation (excludes halo regions) and determine the generalisation error of the PR system when used with the optimal parameters for normalisation and segmentation suggested. We conclude this section with a discussion of the limitations of the segmentation method. Finally, conclusions are drawn in section 4.

2 Materials and Methods

2.1 Images

The subject of the images for this study were samples of human bone marrow stromal cells (BMSCs) taken from two different adults and are labelled sample A and B. The cells of sample A (figure 1a) were considered by expert opinion to be of a higher quality than those of sample B (figure 1b), as they had a higher proliferation rate.

Images were taken using a PixelLINK PL-B686CU colour USB microscopy camera with DC10NN C mount Nikon 38mm adapter attached to a Nikon Eclipse TS100-F inverted microscope with a 10x phase-contrast objective lens.

Original images were taken at the camera’s maximum resolution of 2208x3000 pixels. Images were then converted to 8-bit greyscale and rescaled to a size of 552x750 (using ImageJ software version 1.40e [6]) to allow for a more manageable speed of processing and more economical use of storage space. To increase the amount of data available, each rescaled image was split into 4 tiles (following normalisation, if applicable) resulting in images with a final size of 276x375. The size of each pixel in these tile images corresponded to an area of 1.3µm x 1.3µm.

2.2 Image Normalisation

Image normalisation attempts to reduce the effect of variation in the input images due to differing lighting, camera exposure and other settings. Two common methods of normalisation are contrast stretching and histogram equalisation [7].

Contrast stretching applies a linear transformation to the input image so that the intensity histogram is stretched across the full range of possible pixel intensity values (e.g. 0-255 for the 8-bit gray level images used here). In order to prevent a relatively small number of outlier pixel intensities adversely affecting the result, it is normal to allow a certain percentage (say, 0.5%-3%) of pixels to become saturated (set to 0 or 255). For example, allowing a saturation of 2% would imply that pixels with an intensity in the 1st percentile of the histogram would be set to 0, pixels in the 99th percentile would be set to 255, and the pixels in between would be linearly stretched between the intensities of 1 and 254.
Histogram equalisation, applies a non-linear transformation to the image histogram, so that the output image histogram is flattened and approaches a uniform distribution. An advantage of this method over contrast stretching is that it does not require selection of a parameter value.

2.3 Rough Segmentation

A rough segmentation was performed by taking the input image tile (figure 2a) and calculating the standard deviation at each pixel, computed over a circular neighbourhood, or mask, centred on the pixel. A mask size of radius 3, equivalent to a diameter of approximately 9µm, was found empirically to provide the most suitable result. A smaller mask resulted in more cell areas with little intensity variation being incorrectly considered background. A larger mask resulted in more background areas being included in the segmented cell area, as well as a noticeably reduced performance in the subsequent refined segmentation process. The result of applying this standard deviation mask can be seen in figure 2b.

In order to separate cell regions from the background, the standard deviation image was automatically thresholded using minimum error thresholding [8]. The result of this process is the binary image shown in figure 2c. It can be observed from this image that foreground objects include not only cells, but also small particles, which can be seen as small white blobs. Small holes can also be seen in the centre of some of the cell regions in figure 2c. These holes are caused by contiguous dark cell regions within the input image, where the standard deviation may be very low or even zero, causing them to be considered part of the background. Finally, image areas that contain the white halo are considered by this algorithm to be part of the foreground since they have a high local intensity variation.

To clean up the image, the small particles (area <= 1024 pixels / ~1700 µm²) and holes (area <= 256 pixels / ~430µm²) were removed by applying a particle filter to the thresholded image of figure 2c with the results shown in figure 2d and figure 2e. The area of the particles and holes to be removed was determined manually by measuring their average sizes across a range of typical images.

An outline of the rough segmentation result of figure 2e is shown superimposed on the input image in figure 2f. It can be seen that the rough segmentation result produces a reasonably accurate contour for cell areas where there is no halo (point A). However, in areas where the halo is present (point B), the contour balloons outwards to encompass it, resulting in non-cell regions being included within the contour. The next section details a simple method of dealing with this effect and producing a refined contour that more closely matches the actual contour of the cell regions.

![Figure 2: Rough segmentation (a) normalised image tile (b) standard deviation filter applied (c) thresholded image (d) small objects removed (e) small holes removed (f) roughly segmented region superimposed on input image](image)

2.4 Refined Segmentation

An examination of the pixel intensity profiles of cross-sections taken across roughly segmented areas (figure 3a) indicates that there is a general increase in pixel intensity in halo regions as one moves in the direction from the rough segmentation contour toward the actual cell boundary (figure 3b). In contrast, areas where the rough segmentation contour is already following the cell boundary closely, there is a decrease in intensity in the direction of the cell. The following algorithm was developed that takes advantage of this, by moving the contour toward the cell until the intensity gradient becomes negative.

As a first step, a mean filter of radius 1 is applied to the image to reduce the effect of noise on the intensity gradient as shown in figure 3c. Next, the portions of
the smoothed image that lie within the roughly segmented areas are scanned in four directions (left to right, right to left, top to bottom and bottom to top) with the rough segmentation contour being moved toward the cell for as long as the smoothed intensity gradient in that direction is non-negative. This produces a more refined binary mask image of the cell areas as shown in figure 4b. Finally, a median filter of radius 4 is applied to the refined mask image to remove spurs caused by noise not fully removed by the initial mean filter. This gives the final result of the refined segmentation process as shown in figure 4c. The outline of this mask is shown superimposed on the original input image in figure 4d. It can be seen that the refined segmentation mask produces a much closer fit to the actual cell boundaries compared to the initial rough segmentation contour of figure 2f. In particular, the edge halo areas no longer cause a false ballooning out of the cell boundary contour.

2.5 Performance Evaluation

To evaluate the normalisation and segmentation algorithms, a database of 128 images (64 of sample A, 64 of sample B) was split using two thirds (84 images) for training and one third (44 images) for testing. Equal numbers of both samples were maintained in the two sets. The effect of the different normalisation and segmentation methods was evaluated by testing the performance of the system on the training set (i.e. empirical performance) using stratified 10-fold cross validation [9]. Once the methods that maximised the empirical performance were determined and applied, the generalisation performance of the resulting system was determined by evaluation on the (previously unseen) test set.

Feature extraction: Statistical measurements of pixel intensity values (mean, mode, median, standard deviation, skew and kurtosis) were extracted from either the entire image (when testing normalisation methods) or the segmented regions only (when testing segmentation methods). The extracted measurements were used as input to the pattern recognition system, along with the class (i.e. sample A or B). The reason the above statistical features have been used, rather than morphological measures, such as cell size/length and so on, is that they only require the segmentation of cell regions, rather than individual cells. This is an important consideration, since cells often touch and cannot be segmented separately by the proposed method. Furthermore, in many cases, only a portion of a cell is visible, making measurements of cell size and shape unreliable.

Pattern Recognition: Pattern recognition performance was evaluated using four PR algorithms previously employed in the literature relating to the analysis of cell images, namely, the naïve Bayes [10], multi-layer perceptron (MLP) [11], radial basis function (RBF) network [12] and the support vector machine (SVM) [13]. The Weka system (version 3.5.7) [9] was used for the implementation of these algorithms. Default parameters provided by Weka were used and no optimisation of PR parameters was performed. The default parameters used by Weka are designed to give reasonable performance in many applications and this was found to hold true for this project. A brief description of these classifiers, along with the default parameters, are now given below, with the reader being referred to [1] and [9] for a more in-depth coverage.

The naïve Bayes classifier is a simple probabilistic classifier based on the application of Bayes’ theorem. It is termed naïve because it assumes that the input
features are conditionally independent given the class. Despite this (often incorrect) assumption, the method often works well in practice. The classifier was used with input features assumed to follow a normal distribution.

The multi-layer perceptron is an artificial neural network consisting of multiple layers of artificial neurons with a typically non-linear (e.g. sigmoid) activation function. The connection weights are trained using backpropagation. The network used here consisted of one hidden layer containing four neurons, sigmoid activation function, a learning rate of 0.3 and a momentum of 0.2.

The radial basis function network is an artificial neural network incorporating a weighted sum of (typically Gaussian) radial basis functions. The Weka implementation uses a k-means clustering algorithm to determine the centres and widths of the Gaussian RBFs and the weights are determined by logistic regression. The default parameters use two clusters (per class) with a minimum standard deviation of 0.1.

The support vector machine is a learning method developed from statistical learning theory. It finds the hyperplane that separates the two classes with the maximum margin. The Weka implementation uses the sequential minimal optimisation algorithm for training. The default configuration uses a polynomial kernel of degree one without the use of lower order terms. The complexity parameter (C) is set to 1.0.

### 3 Results and Discussion

#### 3.1 Effect of Normalisation

It can be seen from table 1 that the method of normalisation had a very significant impact on the performance of all the classifiers, with accuracy improvements ranging from 7% to 12% compared to images that were not normalised. Contrast stretching was found to be the best method of normalisation, with a 1.5% saturation producing the best results for 3 out of the 4 learning algorithms (see figure 5).

Interestingly, histogram equalisation resulted in a significant loss of accuracy for all classifiers even compared to the non-normalised images, and therefore is best avoided. The reason for the poor performance is that histogram equalisation flattens the image histogram toward a uniform distribution. This makes the histograms (and hence the extracted statistical features) of both classes more alike.

It is important to note that the absolute accuracies don’t necessarily tell us which classifier is the best, since only the default parameters were used. In fact, we would expect them all to improve performance when their individual parameters are optimised. However, it can be seen from the flatter curve in figure 5 that the MLP was the most robust technique for dealing with variations in the normalisation process. It maintained a more consistent degree of accuracy, compared to other methods, when the normalisation process was not ideal.

#### 3.2 Effect of Segmentation

Table 2 shows the effect of segmentation on classifier performance (% images correctly classified) using images normalised using contrast stretching at 1.5% saturation. Accuracy was improved significantly when using refined segmentation compared to both rough segmentation and no segmentation. This indicates that the halo is best excluded from segmented cells regions for best pattern recognition performance.

#### 3.3 Performance on Test Data

Table 3 shows the performance of the different learning algorithms on previously unseen test data when using the combination of contrast stretching at 1.5% saturation and refined segmentation. The results show that all classifiers generalised very well, achieving similar or higher accuracy than that achieved on the training set.

---

**Table 1:** Effect of normalisation method vs PR performance (% images correctly classified)

<table>
<thead>
<tr>
<th>Norm. Method</th>
<th>MLP</th>
<th>Naïve Bayes</th>
<th>RBF Net</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hist. Eq.</td>
<td>75.00</td>
<td>69.52</td>
<td>59.52</td>
<td>85.33</td>
</tr>
<tr>
<td>None</td>
<td>83.33</td>
<td>70.24</td>
<td>66.19</td>
<td>73.81</td>
</tr>
<tr>
<td>CS (0.5% sat)</td>
<td>86.90</td>
<td>70.24</td>
<td>77.38</td>
<td>63.10</td>
</tr>
<tr>
<td>CS (1% sat)</td>
<td>89.29</td>
<td>79.76</td>
<td>88.10</td>
<td>78.57</td>
</tr>
<tr>
<td>CS (1.5% sat)</td>
<td>90.48</td>
<td>80.95</td>
<td>88.10</td>
<td>80.95</td>
</tr>
<tr>
<td>CS (2% sat)</td>
<td>90.48</td>
<td>79.76</td>
<td>89.29</td>
<td>79.16</td>
</tr>
<tr>
<td>CS (2.5% sat)</td>
<td>90.48</td>
<td>76.19</td>
<td>86.90</td>
<td>77.38</td>
</tr>
<tr>
<td>CS (3.0% sat)</td>
<td>86.90</td>
<td>75.00</td>
<td>80.95</td>
<td>73.81</td>
</tr>
</tbody>
</table>

**Figure 5:** Classifier performance vs % saturated pixels (contrast stretching)

**Table 2:** Effect of segmentation on classifier performance (% images correctly classified)

<table>
<thead>
<tr>
<th>Segmentation Method</th>
<th>MLP</th>
<th>Naïve Bayes</th>
<th>RBF Net</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>90.48</td>
<td>69.52</td>
<td>68.10</td>
<td>80.95</td>
</tr>
<tr>
<td>Rough (inc. halo)</td>
<td>93.98</td>
<td>83.13</td>
<td>84.34</td>
<td>91.57</td>
</tr>
<tr>
<td>Refined (excl. halo)</td>
<td>96.34</td>
<td>87.80</td>
<td>92.68</td>
<td>95.12</td>
</tr>
</tbody>
</table>

**Table 3:** Performance on previously unseen test data

<table>
<thead>
<tr>
<th>Performance</th>
<th>MLP</th>
<th>Naïve Bayes</th>
<th>RBF Net</th>
<th>SVM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training Set</td>
<td>96.34</td>
<td>77.80</td>
<td>92.68</td>
<td>95.12</td>
</tr>
<tr>
<td>Test Set</td>
<td>95.45</td>
<td>93.18</td>
<td>95.12</td>
<td>93.18</td>
</tr>
</tbody>
</table>
3.4 Limitations

Our method does have its limitations. Firstly, although the refined segmentation method generally forms accurate contours around cell regions, it does not attempt to segment touching cells. It may thus be of limited use, at least in its present form, in applications that require such separation. Secondly, since the method uses an automatic thresholding algorithm [8] to segment regions of high intensity variance (cell regions) from regions of low intensity variance (background), the method fails if an image has insufficient background area, as is the case when cells approach confluence. When this happens, cell body areas that are flat and spread out (low variance) become treated as background regions, producing an unsatisfactory result. However, this is not considered to be a major problem for cell quality assessment, since it is intended that images will be taken early in the culture process, well before confluence is reached.

4 Conclusions

Both image normalisation and segmentation was shown to have a large impact on the ability of a subsequent pattern recognition system to correctly discriminate between two BMSC cultures of differing quality. Best results were obtained when images were first normalised using contrast stretching (at 1.5% saturation) and then segmented using the refined segmentation method described in section 2.4.

Using the above methods the system achieved an average image classification accuracy of 93% on the training data and 94% on previously unseen test images. This represents a large (~17%) improvement in accuracy compared to that achieved with non-normalised, unsegmented images.

Furthermore, it was found that including the halo region in the segmented regions hinders pattern recognition performance. This leads to the conclusion that it does not contain information indicative of cell culture quality.

Having now developed a suitable image normalisation and segmentation algorithm, further research will focus on the feature selection and pattern recognition components along with training and testing the system on a wider range of BMSC culture images.

References