Bayesian Analysis of Cell Nucleus Segmentation by a Viterbi Search Based Active Contour

Pascal Bamford and Brian Lovell
Cooperative Research Centre for Sensor Signal and Information Processing,
Department of Computer Science and Electrical Engineering,
University of Queensland, Australia.
Email: {bamford@csee.uq.edu.au,lovell@csee.uq.edu.au}

Abstract

An image segmentation scheme is shown to be exceptionally successful through the application of high-level knowledge of the required image objects (cell nuclei). By tuning the algorithm's single parameter, it is shown that the performance can be maximised for the dataset, but leads to individual failures that may require alternative choices. A second stage is introduced to process each of the resulting segmentations obtained by varying the parameter over the working range. This stage gives a Bayesian interpretation of the results which indicates the probable accuracy of each of the segmentations that can then be used to make a decision upon whether to accept or reject the segmentation.

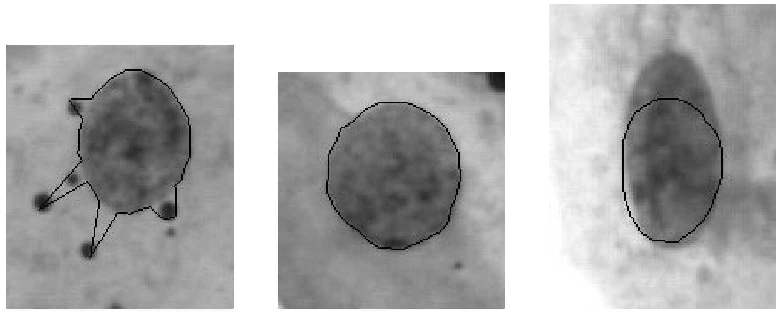
1. Introduction

Since the cervical cancer screening process was identified as a suitable candidate for automation in the early 1950s, many research groups have sought a solution. However, despite the many excellent results in the feature extraction and classification stages, the automatic segmentation of cell images has been uniformly identified as a major hurdle in the development of a working system[2][4][5]. Robust unsupervised segmentation of real images is a difficult problem and often requires a great deal of high-level knowledge in order to produce satisfactory results. This is usually due to the fact that these images can be very noisy, be of poor quality, and can contain many artefacts or unwanted objects. Although the latter case refers to the original purpose of segmentation, the application of knowledge of the required output can greatly reduce its effect. This high-level knowledge can be used in the initial choice of algorithm, through modification of the algorithm (gradient direction, statistical method, etc.) or by parameter tuning for 'optimum' performance. However, the latter optimises performance over the entire dataset without necessarily optimising the parameter for each individual image.

Here, *a priori* knowledge of cell nuclei was applied to all of the stages described above and produced a method that was highly successful (99.64% correct segmentations on a dataset of over 20,000 images). The problem of tuning the algorithm's single parameter, λ, became apparent when it was discovered that many of the failures could have been rectified by a different parameter value. It was therefore proposed to introduce a second process to interpret the output of the segmenter for a range of its parameters. This would provide more information on the image being processed, enabling a further application of high-level knowledge.

2. Segmentation Method

A global minimum searching, active contour based segmentation method was chosen. As cell nuclei borders are smooth, are characterised by areas of relatively large gradient and often contain artefacts and 'noise' (stained cytoplasm) this technique was particularly appealing. Due to space limitations, the reader is referred to [1] for a full description of the procedure. The output of the algorithm was a set of 256 real valued points which described a contour that hopefully delineated the nucleus border. A single parameter, λ, controlled the output of the segmenter and its range of operation was $0 \leq \lambda \leq 1$ [3]. This parameter effectively regularised the two main concerns that the contour should satisfy; smoothness and location of high gradient areas. By choosing λ close to 0, the smoothness constraint was largely ignored, allowing the contour to lie upon the strongest image edges whether they were on the nucleus border or not (figure 1(a)). If λ was set closer to 1 then the gradient information became less important leading to a very smooth contour which did not necessarily lie upon the



(a) (b) (c)

Figure 1. Effect of algorithm parameter, λ : (a) λ close to 0.0 leads to a contour lying upon the areas of greatest gradient in the image (b) correct choice of λ leads to accurate segmentation (c) λ close to 1.0 leads to a smooth contour which does not lie on nucleus border

nucleus border either (figure 1(c)). Therefore a compromise between these two values was sought (figure 1(b)). As the same compromise could not be guaranteed to work for all images, a set of solutions was provided for each image corresponding to the value of λ used.

3. Segmentation Interpretation

For each image, the segmentation results were viewed for each value of λ and either determined as a 'pass' or as a 'fail'. The relationship between the number of successful segmentations and the regularisation parameter, λ, is graphed in figure 2. This graph was obtained on a 772 image subset of the database for λ in the range $\lambda = 0.1k$ where $k = 0, 1, 2, \ldots, 10$.

Figure 2 shows a wide plateau of high success rates between $\lambda = 0.3$ and 0.8, demonstrating that the method is well suited to the application. At the extremities, either the gradient or smoothness constraint dominates and reduces the efficacy of the algorithm. This graph is characteristic of the segmentation procedure and therefore allows decisions to be made on how to interpret the results.

4. Bayesian Smoke Ring Images

It is interesting to combine the *a priori* estimate of probability of segmentation accuracy from figure 2 with the actual segmentations produced for each value of λ. One way to achieve this is to

1. Produce a border image for each value of λ by setting border pixels to 1.0 and all other pixels to 0.0;

2. multiply each border image by the *a priori* estimate of segmentation performance from figure 2; and

3. sum all of the weighted border images.

The resulting images resemble smoke rings, so we have dubbed the images "Bayesian Smoke Rings." Some examples are shown in figure 3 (a),(b),(c), and (d). Such images show the probability of each segmentation result being correct and are therefore helpful in devising methods to reliably determine the accuracy of a segmentation without human intervention (in other words, totally automated accurate segmentation). They may also lead to a quantitative measure of the intrinsic difficulty of the segmentation task for each image in the database.

5. Conclusions and Future Work

An image segmentation method was shown to be highly successful by incorporating high-level knowledge of the desired objects. It was discovered that although tuning the algorithm's parameter leads to exceptionally good success rates, more information could be gleaned from the images being segmented if the results were viewed for a range of parameters. By compiling a graph of segmentation success against the algorithm's parameter, λ, it was possible to demonstrate the suitability of the algorithm to the problem by the presence of a high performance plateau over a wide range of λ. It is envisioned that such a graph would be useful in the initial algorithm selection stage for the comparison of different methods. Finally, the graph was used to produce a 'Bayesian smoke ring' image for each input image. This image could be used in an evaluation stage that classes the results by confidence in the location of the final contour. A system that chose to output only results with a high confidence level would greatly enhance the accuracy of its segmentation stage.

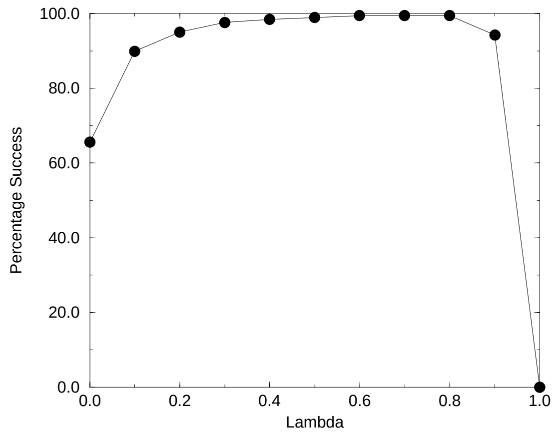


Figure 2. Percent Correct Segmentation versus regularisation parameter

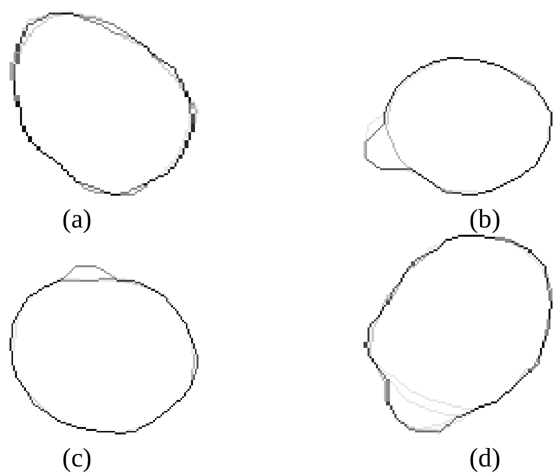


Figure 3. Bayesian smoke ring images of various cells which indicate the probability of each pixel being on the nuclear border as a grayscale value.

6. Acknowledgement

The authors would like to thank Marcus Frean of the Department of Computer Science and Electrical Engineering, University of Queensland, for his suggestion of a Bayesian interpretation of the segmentation results.

References

[1] P. Bamford and B. Lovell. Robust cell nucleus segmentation using a viterbi search based active contour. In *DICTA '97*, pages 89–93, Auckland, New Zealand, December 1997. Technical Keynote 2.

[2] G. L. Garcia. *Feasibility of contextual analysis in an automated cervical prescreening system*. PhD thesis, Worcester Polytechnic Institute, 1986.

[3] K. F. Lai. *Deformable contours: Modeling, extraction, detection and classification*. PhD thesis, University of Wisconsin-Madison, 1994.

[4] C. E. MacAulay. *Development, implementation and evaluation of segmentation algorithms for the automatic classification of cervical cells*. PhD thesis, University of British Columbia, August 1989.

[5] B. Nordin. *The development of an automatic prescreener for the early detection of cervical cancer: Algorithms and implementation*. PhD thesis, Uppsala University, 1989.