Automated optimisation of cell segmentation parameters in phase contrast using discrete mereotopology

Flight, Rachel; Landini, Gabriel; Styles, Iain; Shelton, Richard; Milward, Michael; Cooper, Paul

Automated optimisation of cell segmentation parameters in phase contrast microscopy using discrete mereotopology

Rachel Flight[1] rxf144@bham.ac.uk

Gabriel Landini[2] g.landini@bham.ac.uk

Iain Styles[3] i.b.styles@cs.bham.ac.uk

Richard Shelton[2] r.m.shelton@bham.ac.uk

Michael Milward[2] m.r.milward@bham.ac.uk

Paul Cooper[2] p.r.cooper@bham.ac.uk

[1] PSIBS DTC, University of Birmingham, UK

[2] School of Dentistry, University of Birmingham, UK

[3] School of Computer Science, University of Birmingham, UK

Abstract

It has been shown previously that the number of epithelial cells in a monolayer can be determined *in vitro* using phase contrast microscopy by subtracting images mean-filtered with two different kernel radii and then thresholding to segment cells. Careful selection of filter sizes was essential to ensure the number of segmented regions corresponded accurately with the number of cells in the image, however manual parameter selection and verification is time-consuming and prone to human error. We propose an intelligent imaging approach for evaluating the success of filter size combinations for cell detection using discrete mereotopology to compare segmentations with ground truth binary images of stained cell nuclei. Applying this approach to phase contrast images of H400 epithelial monolayers with varying levels of confluency, a region in the parameter space could be identified where more than 90% of cells were correctly detected.

1 Introduction

Epithelial cells cultured *in vitro* form adherent monolayers and enable the study of cell behaviour in a simplified environment. These 2D epithelial monolayer models are used for a wide variety of purposes including toxicity assays and assessing potential wound healing treatments [1, 2]. One important metric in such studies is the change in cell growth rate in response to stimuli and this may be assessed by establishing growth curves from cell counts obtained at multiple time-points during the culture period. Cell numbers are frequently determined by counting cells in a haemocytometer chamber, a technique which involves enzymatic detachment of adherent monolayers from the culture surface to perform manual cell counts using a microscope. The destructive nature of this technique means it is unsuitable for observing cell growth longitudinally in the same cell population thus large numbers of cultures are required to generate growth curves, consuming time and laboratory resources and introducing error through potential inaccuracies in seeding density across different cultures. The accuracy of this technique is further limited by inter-operator variability [3].

Phase contrast (PC) microscopy is a non-invasive method for imaging unstained cells using small phase shifts in light transmitted through the sample and provides a potential approach for non-destructively determining cell number using images of cell monolayers. However, PC images suffer from intrinsic "halo" artefacts – regions of bright pixels at object edges (Figure 1B) – which are a barrier to the application of simple thresholding techniques for segmenting cells.

A method for determining cell number from PC images has previously been developed using morphological filters to overcome the halo artefact and create an image that may be segmented with a single threshold value [4]. This method involves applying two mean filters to the image – one with a kernel radius, r_k, larger than the average cell and the other smaller than the average cell. Subtracting the two resultant images and then applying a threshold yields a binary image where the number of connected regions represents the cells (Figure 2A). The radii of the two filters – r_{k_small} and r_{k_large} – must be selected carefully to avoid separate cells from being merged, split or entirely undetected resulting in an erroneous cell count.

In previous work, r_k values were selected by applying a large range of r_k pairs to a test image and comparing the number of regions segmented by each pair with the total number of cells in the image as determined by a manual count. However, this approach is inadequate in a number of ways. Firstly, manually determining cell number from a PC image is a laborious and error-prone process, particularly for images of confluent monolayers where cell numbers can be in the thousands. Furthermore, considering the total number of segmented regions as the sole measure of cell detection success does not guarantee that the regions have any spatial correspondence with the true cells.

Discrete Mereotopology (DM) is a version of the Region Connection Calculus (RCC) logic used to describe the parthood relationships between discrete regions [5]. RCC5D is an implementation of DM which describes a set of five possible relationships between two regions in discrete 2D space (Figure 1A). RCCD has been used previously to encode information about cellular structure into an automated evaluation of segmentation success in histological images [6].
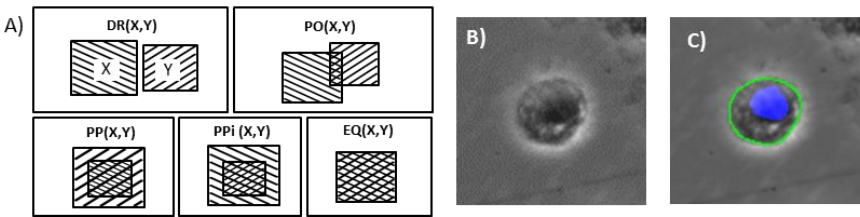
Figure 1 : A) The 5 spatial relationships as described by the RCCD5 relation set and may be read as follows: DR(X,Y) – "X is disconnected from Y", PO(X,Y) – "X partially overlaps with Y", PP(X,Y) – "X is a proper part of Y", PPi(X,Y) – "X is a proper part inverted of Y", EQ(X,Y) – "X is equal to Y". B) PC image of a H400 cell showing the "halo" artefact – bright pixels at the cell edge. C) Image B overlaid with cell nucleus (blue) and manual cell segmentation (green). The relation shown is PP(nucleus, segmentation).

We propose to use DM to incorporate biological structure information about epithelial cells in an automated procedure to assess the success of pairs of r_k values for detecting H400 epithelial cells in PC images, through comparison with a ground truth image of cell

nuclei. First we define a desired relationship between cell segmentation and nucleus in terms of RCC5D relations. Subsequently, the optimum values of r_k are identified by searching through the parameter space to find the r_k pairs which maximise the number of cells satisfying the defined relationship conditions. The cell count accuracy is then further improved by using properties of segmented regions identified as being correct detections to set a minimum cell size for removing small noise regions.

2 Optimisation of filter pairs

2.1 Defining successful cell detection from DM relations

Epithelial cells generally have a single nucleus, N, entirely contained within the cell cytoplasm, C (Figure 1C). In the ideal case, the segmented region corresponds with the cell cytoplasm, thus the following conditions may be applied to define a correctly detected cell:

  i.   A nucleus should be a proper part, PP(N,C) of, or equal to EQ(N,C) the segmented cell region. PP is the expected case but the restriction is relaxed to also include PO(N,C) and EQ(N,C) relations to account for cases when the cell is a slightly under-segmented region but still detected.

  ii. There should be an exclusive one-to-one relationship between segmented cell regions and nuclei (i.e. our definition does not include multi-nucleated cells).

2.2 Identification of optimal parameters

An image of fluorescently stained cell nuclei is acquired alongside the PC images to act as the ground truth for cell locations. By performing the segmentation on the PC image multiple times with a range of r_k pairs, optimal values may be defined as the combinations which return the highest number of correctly detected cells (Figure 2B).
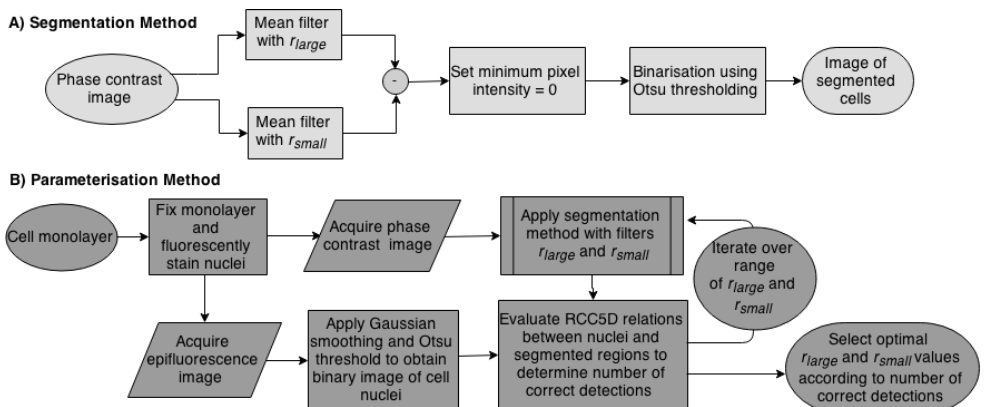

Figure 2 : A) Workflow for the segmentation method described in [4]. B) Workflow for the parameterisation method described in this paper.

2.3 Removing small noise regions

Image noise causes some small clusters of pixels to fall above the threshold value resulting in incorrectly segmented regions (Figure 3C). Once parameters have been selected

according to the method described in section 2.2, the empirical cumulative distribution function (CDF) of areas of segmentations meeting the correct detection criteria, CDF_{cells}, is compared with the CDF of regions that do not, CDF_{noise}. An appropriate size threshold for removing noise which maximises noise removal whilst minimising removal of true cell segmentations is chosen by finding the object area size for which CDF_{noise} - CDF_{cells} is maximised.
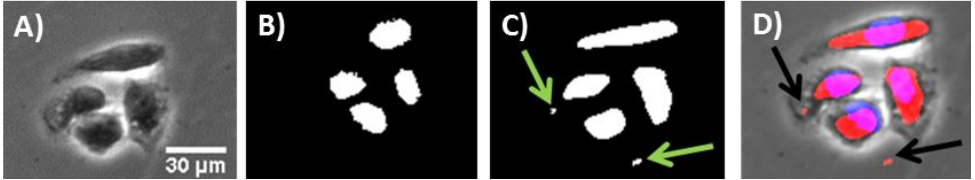

Figure 3 – A) Phase contrast image of DAPI-stained H400 cells. B) Thresholded fluorescence image of nuclei. C) Cells segmented using $r_{k\_small} = 7$ pixels and $r_{k\_large} = 18$ pixels. Arrows indicate noise regions. D) Cell nuclei (blue) and segmentations (red) overlaid on original image. It can be seen that all regions segmented with these parameters (except noise regions indicated with arrows) meet the correct cell detection criteria.

3 Experimental materials and methods

3.1 Cell culture and fluorescent staining

Cells used were H400 keratinocytes derived from a human oral squamous carcinoma [7]. Cultures were maintained at 37°C in a humidified atmosphere with 5% CO_2 in Dulbecco's MEM / nutrient mix supplemented with 10% foetal calf serum (FCS) (Biosera, UK), 0.6 µg/mL L-glutamine (Sigma, UK) and 0.4 mg/mL hydrocortisone (Sigma, UK).

For the imaging steps, cells were seeded in a 6-well plate (Sarstedt, UK) at 5×10^4, 7.5×10^4 and 1×10^5 cells/mL to achieve a range of cell densities and were incubated as described above for 48 hours. Cultures were washed three times with phosphate buffered saline (PBS) and fixed by overnight incubation with 10% formalin (Leica, UK) and finally washed 3 times with PBS, ProLong Gold antifade mountant with DAPI (Life Technologies, UK) was delivered dropwise onto the cell monolayer before placing a coverslip and then incubated at room temperature in the dark overnight prior to imaging.

3.2 Image acquisition and analysis

Cells were imaged concurrently with PC and fluorescence microscopy using a Nikon TE300 microscope with a x10 objective and a Retiga-2000R CCD camera (Qimaging, UK) using Micro-Manager software for ImageJ [8, 9]. 1600 x 1200 pixel images were acquired and calibrated using a stage micrometer (image size was of 1.19 x 0.89 mm).

Epifluorescence images were pre-processed with a Gaussian filter with radius of 2 pixels to reduce noise and a rolling ball algorithm [10] to account for uneven illumination, and then segmented using manual threshold selection. Images were visually inspected to ensure there were no merged or unsegmented nuclei. PC images were segmented using the method described in [4] using parameters in the range $2 \le r_{k\_small} \le 30$ and $r_{k\_small} <$

$r_{k\_large} \leq 70$. In total, the success of 1566 r_k combinations in detecting 1425 cells were considered.

Macros to analyse DM relations between nuclear and PC segmentations were developed in-house with ImageJ [9] utilising freely available RCCD plugins [6]. Graphs representing correct detections for parameter combinations and empirical CDFs were generated using Matlab.

4 Results

Figure 4A shows graphical data of correct detections over the range of r_k pairs tested. A region in the parameter space was identified inside which more than 90% of cells were successfully detected in all images. The values of the centroid of this region were $r_{k\_small} = 7$ pixels = 6 µm and $r_{k\_large} = 18$ pixels = 16 µm, which were used for further analysis of segmented region properties.

Figure 4B shows the CDFs of the correct segmentations and noise regions from all three images as segmented with these parameter values. When CDF_{cells} was subtracted from CDF_{noise}, there was a maximum at an object area of 130 pixels. Setting this as the minimum cell area condition removed 82% of noise regions at the expense of 14% of correct segmentations.
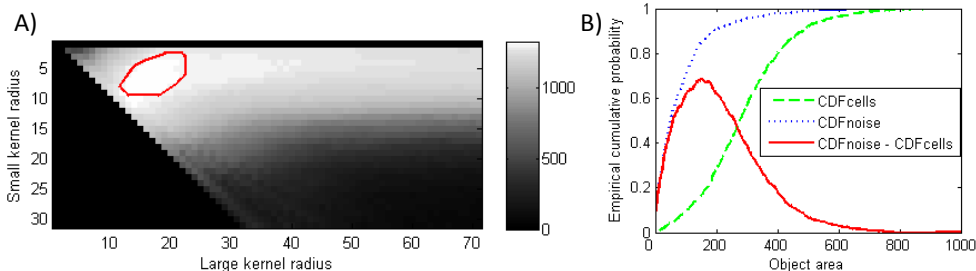

Figure 4 : A) Graph showing the number of correctly detected cells as a percentage of total cells in all of the three images, for all r_{k_small} and r_{k_large} combinations tested. The area circled in red shows regions where at least 90% of cells were correctly detected, the centroid of which was $(r_{k\_small}, r_{k\_large}) = (18,7) = (16$ µm,6 µm) B) Empirical cumulative distribution functions of correct segmentations, noise segmentations and the difference between the two. The maxima of the difference in CDFs occurred at 130 pixels (106 µm^2).

5 Discussion and conclusion

Defining cell detection in terms of the spatial relationships between segmented and ground truth images provides a quantitative and objective method for assessing the success of parameters for optimally segmenting cells and removes the requirement for tedious and error prone manual counting. We have applied this approach to three images each representing a different level of confluence to select a pair of r_k values which, when applied to PC images of H400 cells, correctly detected more than 90% of cells. The parameterisation results detailed here were valid for images of H400 cells acquired with the experimental set up described, but must be repeated if any part of the experimental setup changes, for example, when using other cell types. However, this approach needs to be performed only once to determine the optimal filter parameters. H400 cells are

generally rounded in shape, relatively homogeneously sized and form a single non-overlapping layer until extremely high density is reached. Other cell lines with similar morphological properties have potential to achieve comparable cell detection success rate, but further validation will need to be undertaken to confirm this.

By using the DM definition of correct cell detections in terms of the nuclear ground truth, it was possible to compare the properties of cell regions and incorrect noise regions separately. By considering the empirical cumulative density functions of segmentation areas a minimum cell size threshold could be determined which removed small incorrect segmentations whilst retaining the maximum number of true segmentations. This provides a simple and effective way of removing noise however it is envisaged that the DM conditions defining correct cell detection in terms of the nuclear ground truth may be used as a training set to apply more sophisticated supervised learning techniques for noise removal.

A further advantage of the parameterisation method is that by verifying cell location rather than only considering the total number of cells detected, there is scope for studying the distribution of cells in monolayer as well as the population size - information which is lost if cell number is determined using the current gold standard of haemocytometer counting.

References

[1] M. N. M. Walter, K. T. Wright, H. R. Fuller, S. MacNeil, and W. E. B. Johnson, "Mesenchymal stem cell-conditioned medium accelerates skin wound healing: an in vitro study of fibroblast and keratinocyte scratch assays.," *Experimental cell research*, vol. 316, no. 7, pp. 1271–81, Apr. 2010.

[2] S. Arora, J. Jain, J. M. Rajwade, and K. M. Paknikar, "Cellular responses induced by silver nanoparticles: In vitro studies.," *Toxicology letters*, vol. 179, no. 2, pp. 93–100, Jul. 2008.

[3] R. Biggs and R. L. Macmillan, "The error of the red cell count," *Journal of clinical pathology*, vol. 1, no. 5, pp. 288–91, Nov. 1948.

[4] R. Flight, G. Landini, I. Styles, R. Shelton, M. Milward, and P. Cooper, "Semi-automated cell counting in phase contrast images of epithelial monolayers," in *Proceedings of the 18th conference on Medical Image Understanding and Analysis*, 2014, pp. 241–246.

[5] D. A. Randell, G. Landini, and A. Galton, "Discrete Mereotopology for Spatial Reasoning in Automated Histological Image Analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 3, pp. 568–581, 2013.

[6] G. Landini, D. A. Randell, and A. P. Galton, "Discrete Mereotopology in Histological Imaging," in *Proceedings of the 17th conference on Medical Image Understanding and Analysis*, 2013, pp. 101–106.

[7] S. Prime, S. Nixon, and I. Crane, "The behaviour of human oral squamous cell carcinoma in cell culture," *The Journal of Pathology*, vol. 160, no. 3, pp. 259–269, 1990.

[8] A. Edelstein, N. Amodaj, K. Hoover, R. Vale, and N. Stuurman, "Computer control of microscopes using µManager.," *Current protocols in molecular biology*, vol. Chapter 14, no. October, p. Unit14.20, Oct. 2010.

[9] W. S. Rasband, "ImageJ." U. S. National Institutes of Health, Bethesda, Maryland, USA.

[10] S. Sternberg, "Biomedical Image Processing," *Computer*, vol. 16, no. 1, pp. 22–34, 1983.