Systematic characterisation of cellular spatial arrangement in tissues.

Gabriel Landini [1]

Oral Pathology Unit, School of Dentistry, The University of Birmingham, England.

ABSTRACT

This work presents a number of methods to facilitate the characterisation of the spatial arrangement of objects (cells) in a tissue as observed in 2D histological sections. The approach is based on an algorithmic segmentation of tissue space into discrete binary cells associated with the position of the cell nuclei. The segmentation is done by means of greyscale reconstruction to extract cell nuclei markers (as regional minima in the image) followed by the so-called watershed transform to partition the rest of the tissue into exclusive basins of attraction (the cells) associated with those markers. Once the cell elements have been identified, user-defined local neighbourhoods (a local set of cells) are used to determine the extent of the relations to be sought. The relations between the cells within these neighbourhoods are characterised using several Graph Theory networks constructed over nodes defined by the centres of mass of the cells. The cell segmentation also allows to quantify other spatial features in the tissues which are difficult to measure – for example defining and labelling tissue layers and the cell orientation within a layer. These features can conveniently extracted after a minimal amount of user-defined landmarks are provided by the operator. The methodology is illustrated with examples of normal, precancerous and cancerous epithelia of the mouth, and with the epithelial lining in different types of odontogenic cysts.

Keywords: tissue architecture, Graph Theory, networks

1. INTRODUCTION

Cell nuclear shape and size are features that have generated a great deal of interest in image analysis as applied to tissue pathology. However, the statistical description of tissue spatial architecture has received much less attention, perhaps because it is difficult to analyse. Since cells are embedded in a 3-D space, the analysis of their spatial relationships derived from the analysis of 2-D histological sections is a non-trivial task. Stereology provides some tools to overcome these limitations, but most often these solutions require serial or random sectioning which may not always be feasible when dealing with biopsies and diagnostic material. Nevertheless, since histopathological diagnosis is usually performed on 2-D images of single sections, it seems reasonable to assume that some non-negligible amount of information about cell spatial arrangement may be retrievable from such images (otherwise histopathological diagnosis would not be possible!). Quantification of the architecture of tissues may therefore provide additional information of diagnostic interest. Here are presented some approaches to quantify the arrangement of cells within a tissue as measures of tissue architecture. The examples given in this presentation concentrate on stratified epithelia from normal, premalignant and malignant oral mucosa as well as from epithelial linings of odontogenic cysts, but the procedures should be also applicable to other types of tissues.

2. MATERIALS AND METHODS

2.1 Segmentation

A two step procedure was developed to segment and analyse epithelial tissue sections based on models of random tessellations from stochastic geometry [1]. The first step consists of the partition of the epithelial compartment into its constitutional units (epithelial cells). These boundaries are difficult, and often impossible, to identify using routine staining methods (for instance Haematoxylin and Eosin) and light microscopy. Instead, the epithelial compartment is subdivided into theoretical cell areas defined as the exclusive areas relative to each epithelial cell nucleus. This is achieved by localising the epithelial nuclei based on the optical density of the histological stain and using the nuclei extents to divide the remainder of the tissue space. For sections stained with Haematoxylin and Eosin, stain separation is

[1] E-mail: G.Landini@bham.ac.uk, Telephone: +44 121 237 2885
Address: Oral Pathology Unit, School of Dentistry, The University of Birmingham, St. Chad's Queensway, B4 6NN, Birmingham, England.

first performed using colour deconvolution [2] to preserve only the Haematoxylin stain (for nuclear localisation). The spatial partition of the epithelial compartment into the exclusive areas of influence of each nucleus can be achieved by two methods that produce very similar results: 1) using the so-called watershed transform [3-5]. This is similar to a Voronoi tessellation, but has the additional advantage that guarantees that the segmentation boundaries do not pass through the nuclear spaces (see figure 1). 2) Alternatively, similar results can be achieved using a skeletonization of the background (also called skiz) to obtain the so-called "influence zones". This is preformed on binary images after "dome" or "basin" extraction [4]. These are grey morphology procedures that preserve regional minima (i.e. dark basins) or regional maxima (bright domes) of up to a certain size (the depth or height of the domes/basins). Regional minima are those connected areas of pixels with a given heigh or depth value for which every pixel around them has a strictly higher value [4]. The skeletonization of the background can be achieved with a hit-and-miss transform by performing consecutive thinnings until *idempotence* (i.e. until no more pixel changes take place), followed by pruning (deletion of free ends in the resulting skeleton) taking care that the resulting skeleton must be a 4-connected construct if the segmented cells are expected to be 8-connected.

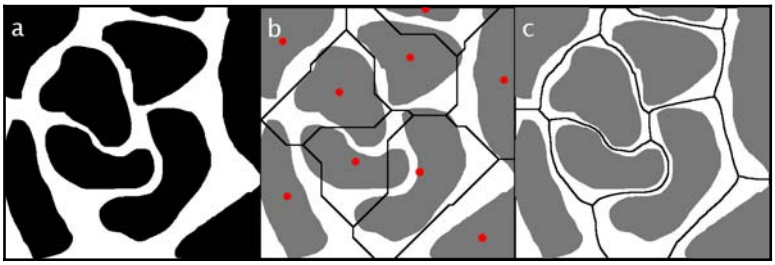

figure 1. A diagram to show the problem of segmentation using a single point (typically the centre of mass or centroid) of a particle (here the cell nucleus) for defining exclusive domain areas (i.e. cells). a) Original image with nuclei (in black), b) a morphological approximation to the Voronoi tessellation by "dilation without merging" of the seed (in red). The dark lines indicate the limits between cells. c) Results using the watershed transform. Note that in this case the boundaries between cells avoid passing through any nuclear part.

The second step consists of the quantification of the relations of cells within local neighbourhoods. Two types of architectural relationships will be shown: local neighbourhoods and layers.

2.2 Local neighbourhoods

Here, a local neighbourhood is considered to be an arbitrarily sized group of contiguous cells at a particular location in the tissue. These can be defined for every cell, C, in the tissue with centroid coordinates C(x,y) so that the neighbourhood N of radius R is centred on C. For sampling consistency, neighbourhoods N(C,R) are considered only if they are completely populated by cells (i.e. the neighbourhoods are not partially empty because of border effects at the edges of the tissue or the image). Those neighbourhoods correspond to portions of the epithelial profiles that contain a complete circular mask and therefore the centres of such neighbourhoods lie in cells that remain after and erosion of the epithelial compartment with a kernel of radius R. As an example, neighbourhoods of radii R=30 pixels and R=60 pixels will be investigated (37.5 and 75µm respectively).

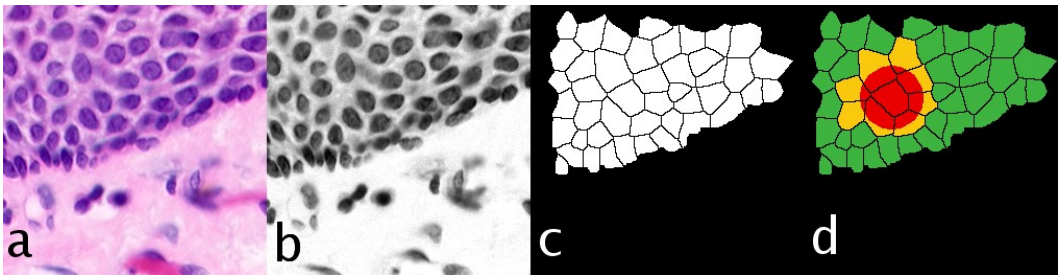

figure 2. In (a) is shown a section of a fragment of oral mucosa stained with Haematoxylin and Eosin. The mucosa is composed of stratified squamous epithelium (top) and the underlying connective tissue (bottom). (b) the Haematoxylin stain contribution obtained after colour deconvolution. In (c) is shown the segmented epithelial cells (white). The local neighbourhoods - orange and red in (d) - are defined by binary reconstructing the cells intersections between image (c) and a sampling disc (in red) of radius R centred in a particular cell C. The neighbourhood is N(C,R) is therefore associated to the cell C and can be used to select and classify all cells in the tissue according to the data collected about their neighbourhoods.

2.2.1. Graph networks constructs

Once the neighbourhoods are constructed, a number of quantitative measurements can be extracted from them. Apart from some obvious morphometrical measures of the cells (and their statistics within the neighbourhood) other descriptors of the spatial arrangement of the cells can be obtained through the construction of planar graphs. Graph Theory provides the principles to construct such graphs. The graphs consist of geometrical constructs based on the cell positions (the centre of mass or centroids of the cells, called "nodes") interconnected by means of lines or "edges" while respecting certain constrains [8-10]. Below are introduced five specific types of graphs that can be defined in algorithmic terms and implemented as automated procedures (see figure 3):

1. the *minimal spanning tree*: the shortest path that connects all the nodes in the set.
2. the *relative neighbour graph*: two points A and B are considered relative neighbours, if for all other points X in the set: distance(A, B) <max[distance(A, X), distance(B, X)]
3. the *Gabriel graph*: two points A and B are considered G-neighbours if their diametral circle (i.e. such that AB is the diameter) does not contain any other points of the set. The computation consists in finding whether for every potential pair of neighbours A and B, any other point X is contained in their diametral circle: distance2 (A, X) + distance2 (B, X) < distance2 (A, B)
4. the *Delaunay triangulation*: the graph created by identifying sets of three points for which the osculating circle (i.e. the circle that passes through the three points) contains no other points inside. A triangle is formed with those points as vertices and their sides are added to the graph.
5. the *nearest neighbour graph*: connects each node to its nearest neighbour. The same principle can be extended to the *k*-nearest neighbours to produce a *k-nearest neighbours graph*.

A number of parameters derived from the distribution of the edge lengths can be derived (for instance total graph length, number of nodes and edges, mean, standard deviation, skewness and kurtosis of the edge distribution). As examples of these principles, the methods were applied to 2 histopathological problems: 1) an analysis of the architectural characteristics of oral mucosa samples and 2) the analysis of the epithelial lining of 3 types of jaw cysts.

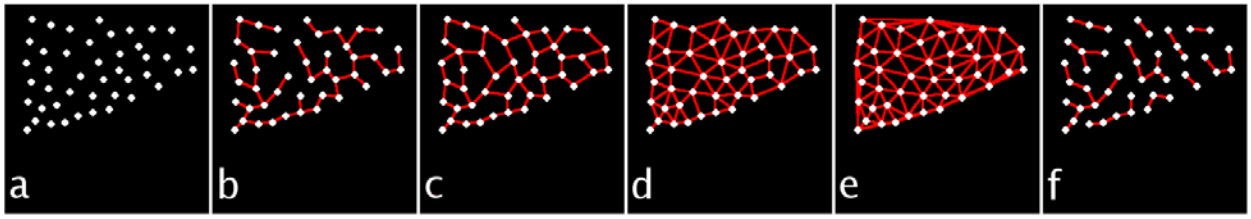

figure 3. Some of the constrained networks that can be used to characterise the spatial distribution of the nodes in the space. (a) a set of nodes, (b) the minimal spanning tree, (c) the relative neighbour graph, (d) the Gabriel graph, (e) the Delaunay triangulation, (f) the nearest neighbour graph.

2.2.2. Tissue layers

Some tissues (e.g. stratified epithelia) are arranged in layers. Despite that it seems easy to visually determine the layers in an appropriately oriented section, it is not straightforward to consistently determine layer membership in an unequivocal manner. This becomes more complicated in tissues that do not maintain a constant thickness. Here is therefore proposed that layer structure can be revealed systematically by means of a distance transform for non-regular lattices [6,7]. Using this method, the distance, in layers, between cells can be estimated from an arbitrary reference point.

The methodology relies on the fact that the cell segmentation procedure described above produces an irregular lattice of cells, which are separated by 1 pixel wide gaps. This makes it possible to determine the next-neighbours of a cell by the intersection (logical AND operation) of the twice-dilated cell of interest C with the rest of the set (one dilation to reach the gap and a second dilation to intersect the neighbour cell). The intersection of the twice-dilated cell and its neighbours lies on the neighbouring cells themselves, so those intersections are used as seeds for binary reconstruction. The reconstructed cells are the neighbours that form the nearest layer to C. If this layer is merged with C and the whole procedure repeated, the result is the 2nd nearest layer of cells to C, and so on until all the cells have been labelled in the image. In figure 4, is show various possible layer labelling depending on where the reference is located.

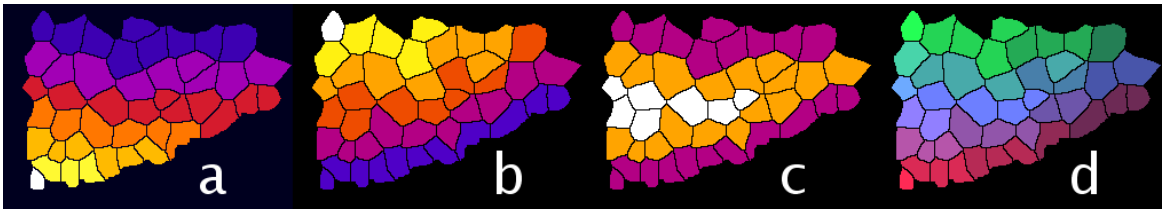

figure 4. Layer labelling. Images (a) to (c) show in pseudocolour the different tissue layers (each individual layer in the same colour) referenced from the top layer (a), from the bottom layer (b) and from either top or bottom reference layers (c). In (d) is shown an RGB composite in which the intensity of the RGB components correspond to the layer each cell belongs to (red= from the top, green=from the bottom and blue= from either). Therefore the colour defines the 3 distances.

2.2.3. Cell orientation in a layer

The angle of orientation of objects within an image can be used to asses whether the distribution of directions (angles) follows a particular trend. However, the angle returned by traditional morphometrical procedures is in most cases relative to the image frame orientation. It the tissues are not aligned, or if there are changes of direction it is not possible to compare angle orientations across images.

To effectively characterise the orientation of a cell major axis with respect to the tissue orientation, it is necessary to define the local direction of the tissue based on the position of the neighbouring cells that share the same layer and use this direction to offset the cell major axis. This can be achieved using a similar procedure to the layering algorithm once the layers can be identified: 1) select a layer, 2) then select a cell within that layer (the current cell), and 3 find its nearest neighbours by dilating the cell twice and performing a logical AND with the current layer. The intersection of the double dilation and the layer provides the seeds that allow reconstruction of the nearest neighbours. A further iteration of the procedure provides the next-to-nearest neighbours. After these cells have been reconstructed and their centres of mass determined (for instance using ImageJ's Particle Analyzer) the direction of the neighbourhood can be determined using some statistical method. In figure 5, this direction was determined by linear regression of the centre of mass coordinates. Since 2 regressions are possible (*x on y* and *y on x*), the one that returned the smallest standard deviation of the residuals was chosen, however, other methods may be also appropriate. Figure 5 shows the local layer directions for each cell in the tissue. Because the direction has been determined for every particular cell, the cell orientation with respect to its neighbours in the layer can be expressed in terms of a relative angle.

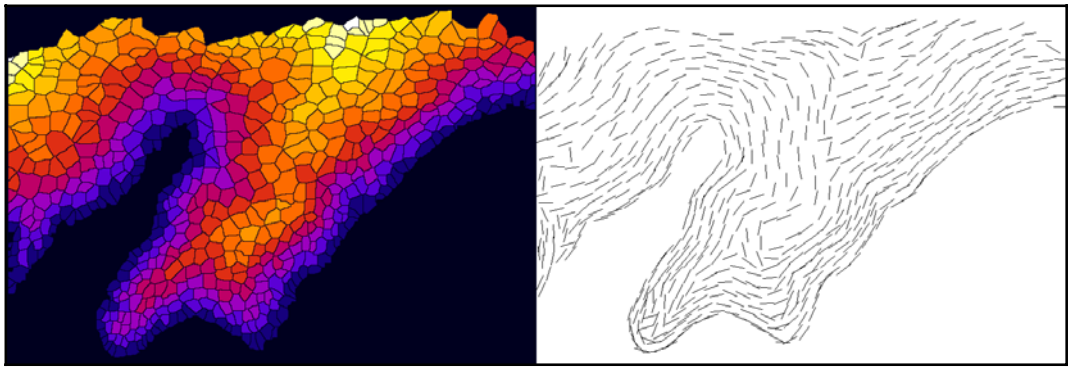

figure 5. Left: Image with labelled layers (in pseudocolour) from the basal layer (nearest to the bottom of the image). **Right**: vectors centred at each cell indicating the direction of the local layer neighbourhood (the local layer neighbourhood is formed by the cell of interest, its nearest neighbours and the next-to-nearest neighbours within the layer).

3. RESULTS

3.1 Oral mucosa analysis.

The graph network characterisation described above was applied to 441 digital images (x20 magnification images, 1 pixel =0.67µm) from 9 normal samples (surgical margins of non-epithelial disease), 7 premalignant oral lesions with epithelial dysplasia and 5 well differentiated oral squamous cell carcinoma. The local neighbourhood data (104,627 neighbourhoods for R=30 and 67,590 for R=60) was analysed using a general linear model and post-hoc Tukey tests and

revealed that for both neighbourhood sizes, the mean of the majority of the parameters investigated were significantly different ($p<0.01$) across the three diagnostic classes. The discrimination power of the parameters was assessed using hierarchical stepwise discriminant analysis and it was found that the best discrimination was always achieved for neighbourhoods sized R=60 (75μm). The pooled data (cell-wise) discrimination was always higher than what would be obtained by random choice (33%) indicating that, in principle, this type of analysis could be useful for quantitative for diagnostic purposes (75% correct between normal and malignant groups, 55% between normal, premalignant and malignant groups) [7].

3.2 Cyst lining analysis.

An analysis of 150 digital images (at x40 magnification) of the epithelial linings cysts of the jaws (9 cases radicular cysts, 13 from solitary keratocysts and 8 from keratocysts associated with the Basal Cell Naevus Syndrome; 5 images for each case) were segmented using the methodology presented earlier an several morphometrical and architectural parameters were determined (cell morphology, number of layers, cell orientation).

Statistically significant differences were observed across the 3 cyst types both at the morphological and architectural levels of the lining. The case-wise discrimination between radicular cysts and keratocyst (disregarding whether they were solitary or syndromic) was highly accurate (with an error of just 3.3%). However, the odontogenic keratocyst subtypes could not be reliably separated into the original classes, achieving discrimination rates slightly above random allocations (60%) [11].

4. CONCLUSIONS

The methodology presented is able to provide some formal descriptions of epithelial architecture. With respect to histopathology, the approaches described here are able to contribute to quantifying systematically some aspects of tissue spatial organisation. Traditionally, histopathological terms like “cell palisading”, “loss of orientation”, “herring bone pattern” or “storiform pattern” have been (and are still) used. They may elicit mental images and facilitate recognition of tissue arrangements, but they are largely subjective and therefore cannot be subjected to quantification. Instead, systematic approaches such as those outlined here have the advantage of being reproducible and quantifiable. In addition, most steps in the analyses shown can be automated and require little operator intervention. One down side is that like for many other imaging procedures, image quality and standardised conditions are necessary to guarantee the reliability of the segmentation.

The *Influence_zones*, *Colour_Deconvolution*, *Binary_Reconstruct* and *Domes* plugins are part of the Morphology Collection of plugins for ImageJ (providing more than 80 new commands) written by the author and freely available for download [12].

Acknowledgement: The author wishes to thank Wayne Rasband for developing and maintaining ImageJ.

REFERENCES

1. D. Stoyan, D, W, S. Kendall, J. Mecke, *Stochastic Geometry and its Applications,* 2nd edition, Wiley & Sons, Chichester, 1995.
2. A. C. Ruifrok, D. A. Johnston, “Quantification of histological staining by color deconvolution,” *Anal. Quant. Cytol. Histol.* **23**, pp. 291-299, 2001.
3. L. Vincent, P. Soille, “Watersheds in digital spaces: an efficient algorithm based on immersion simulations,” *IEEE Trans. Patt. Anal. & Mach. Intell.*, **13**(6):583-598, 1991.
4. L. Vincent, “Morphological greyscale reconstruction in Image Analysis: Applications and efficient algorithms,” *IEEE Trans. Image Proc.* **2**(2), pp. 176-201, 1993.
5. D. Sage, Watershed segmentation plugin. http://bigwww.epfl.ch/sage/soft/watershed/
6. G. Landini, I.E. Othman, “Estimation of tissue layer level by sequential morphological reconstruction.” *J. Microsc.* **209**(2): 118-125, 2003.
7. G. Landini, I.E. Othman, “Architectural analysis of oral cancer, dysplastic and normal epithelia,” *Cytometry A*, **61A**: 45-55, 2004.
8. R. Sedgwick, *Algorithms*. 2nd ed. Addison-Wesley, Reading, MA, 1988.
9. B. K. Bhattacharya, R. S. Poulsen, G. T. Toussaint , “Application of proximity graphs to editing nearest neighbor decision rule,” *International Symposium on Information Theory,* Santa Monica, California; 1981.
10. G. T. Toussaint, ”The relative neighbourhood graph of a finite planar set,” *Patt. Recogn.* **12**:261–268, 1980.
11. G. Landini, “Quantitative analysis of the epithelial lining architecture in odontogenic cysts,” *Head Face Med* **2** (Feb): 4, 2006.
12. G. Landini. Morphology plugins Collection for ImageJ. http://www.dentistry.bham.ac.uk/landinig