City Research Online

City, University of London Institutional Repository

Segmentation and morphological analysis of microvessels in immunostained histological tumour sections

Constantino Carlos Reyes-Aldasoro, Leigh J. Williams, Simon Akerman, Chryso Kanthou, Gillian M. Tozer
http://tmg.group.shef.ac.uk

Cancer Research UK Tumour Microcirculation Group, Department of Oncology, The University of Sheffield, Sheffield, S10 2RX, UK.

Abstract

A fully automatic segmentation and morphological analysis algorithm for the analysis of microvessels from CD31 immunostained histological tumour sections is presented. The algorithm exploited the distinctive hues of stained vascular endothelial cells, cell nuclei and background, which provided the seeds for a region-growing algorithm in the 3D Hue, Saturation, Value (*HSV*) colour model. The segmented objects, identified as microvessels by CD31 immunostaining, were post-processed with three morphological tasks: joining separate objects that were likely to belong to a single vessel, closing objects that had a narrow gap around their periphery, and splitting objects with multiple lumina into individual vessels. **Keywords:** vessel segmentation and morphology, IHC.

1   Introduction

The immunohistochemistry (IHC) for staining of tissue sections for different proteins is a standard method for diagnostic and research purposes. Staining for platelet endothelial cell (EC) adhesion molecule (PECAM-1 / CD31) with tagged antibodies is an effective method for identifying and localising the ECs that line blood vessels, as CD31 is expressed constitutively on the surface of adult, embryonic and tumour ECs. In oncology, the expression of CD31 by endothelial cells in angiogenic vessels has gained considerable attention as the tumour vasculature is emerging as an important therapeutic target for cancer. Despite the popularity of the use of IHC to stain for different proteins and the growth and power of computer and image analysis algorithms, manual procedures are still the most common method for assessing the presence, absence, distribution or intensity of staining[1, 2].

Numerous algorithms for IHC image analysis have been developed for different specific tasks: counting objects such as nuclei, cells or microvessels, quantifying optical density, measuring the abundance of a stain, or extracting morphometric measurements such as area, perimeter, ratio or perimeter and area and angle of microvessels [3, 4, 5]. Many of these algorithms are semi-automated, but most require a certain degree of user interaction, either for (a) pre-processing tasks such as adjustments of brightness and contrast or selection of window sizes, counting of vessels or demarcation of regions of interest or vessels, (b) post-processing tasks such as delineation of lumina for segmented objects, inclusion/exclusion of regions of interest, elimination of artefacts or (c) training of classifiers with supervised data.

Most of these algorithms use either the Red, Green and Blue (*RGB*) channels, a subset of these or the grey level intesity equivalent from the colour images, and few exploit the higher discrimination that can be achieved when the *RGB* channels are transformed to the Hue, Saturation and Value (*HSV*) channels. Post-processing of the segmented objects, as presented in this paper, is desirable to introduce a higher reliability of any morphometric measurements and therefore all the statistical measurements derived from these.

2 Segmentation Algorithm

The segmentation algorithm was based on a transformation from the *RGB* to the *HSV* colour model. *HSV* describes perceptual colour relationships related to the artistic ideas of hue, tint and shade [7] enabling a 3D chromatic histogram to be constructed from the *HSV* images.

In this work, we considered that a colour image I_{rgb} had dimensions $N_r \times N_c \times 3$ for rows, columns and three colour channels that were quantised to N_i levels, which are usually 256. Let $L_r = \{1, 2, ..., r, ..., N_r\}$, $L_c = \{1, 2, ..., c, ..., N_c\}$ be the spatial domains of the data, $x \in (L_r \times L_c)$ be a pixel of the image, and $[R, G, B] = [\{1, 2, ..., r, ..., N_i\}, \{1, 2, ..., g, ..., N_i\}, \{1, 2, ..., b, ..., N_i\}]$ a triplet of *RGB* values. An image was represented then as a function that assigned a colour to each pair of co-ordinates: $L_r \times L_c; I_{rgb} : L_r \times L_c \rightarrow [R, G, B]$.

The shading of an original unbiased image U, which was corrupted by a slowly-varying shading S so that $I = U + S$, was corrected by estimating S as the envelope of the signal [6] and removing it from the biased image $\tilde{I}_{rgb} \approx U = I - S$. Then, the mean value of the three *RGB* unbiased channels was equalised to obtain a background with minimum saturation.

The unshaded and equalised image *RGB* was converted to an *HSV* colour model: $I_{hsv} = T(\tilde{I}_{rgb})$, $I_{hsv} = [I_{hue}, I_{sat}, I_{val}]$. The hue-saturation-value histogram $m_{HSV}(h, s, v)$ is a tri-variate measurement of the occurrence of $[H, S, V]$ on I_{hsv} and it was defined as:

$$m_{HSV}(h, s, v) = \frac{\#\{x \in (L_r \times L_c) : I_{hue}(x) = h, I_{sat}(x) = s, I_{val}(x) = v\}}{\#\{L_r \times L_c\}},$$

where # denoted the number of elements and $h \in H$, $s \in S$, $v \in V$. The 2D hue-value $m_{HV}(h, v)$, hue-saturation $m_{HS}(h, s)$ or saturation-value $m_{SV}(s, v)$ histograms are marginal distributions of $m_{HSV}(h, s, v)$. For this work, we introduced the *maximum saturation profile* as measurement of the distribution of the highest value of saturation for every value of hue, mathematically: $p_{\max S} = \{y \in (h \times S) : \max(s) \text{ so that } m_{HS}(y) > 0\}, h \in H$. While the marginal distributions revealed the frequency of pixels within a certain hue or saturation, $p_{\max S}$ revealed the distribution of saturated regions. Fig. 1 shows a representative IHC image and its final segmentation, m_{HSV} as a cloud of coloured points and an *RGB* cloud to compare the separability of the *HSV* model, m_{HS} as a mesh overlaid on the 2D loci of constant value and $p_{\max S}$ together with the histogram m_H.

Three criteria defined the brown colour of the endothelial cells, low value ($v < \bar{I}_{val}$), high saturation ($s > 0.25$) and an adaptive hue range. The colour variations inherent to IHC [8] resulted in images where the colour of stained nuclei was either blue or light purple and the ECs were brown to light beige. To compensate for this variation, the amount of brown ($0°, 67°$) was measured and used to determine the range of hues that were selected as brown. For those images with a strong component of brown, ranges of brown, background and blue were approximately ($10°, 50°$), ($60°, 200°$), ($220°, 310°$) while for those with a low component it was ($350°, 90°$), ($90°, 190°$), ($200°, 290°$) respectively.

The segmented ECs were used as seeds in a region growing algorithm, with the following criteria of similarity: (a) a combination of the s and v, distances in (b) position and (c)

hues from the pixels to the seeds. Stop criterion was minimal change (10 pixels) between iterations. Small (1 or 2 pixels) and isolated regions were discarded. Unassigned pixels were considered as background.

Three morphological conditions were used to assign objects as vessels: (a) **join** objects which could form part of a single vessel (Fig. 2a). Solid objects close to each other were considered to be joined. A skeleton of the individual objects was compared with one of the joined ones and only when the combined had lower number of branching points the objects were joined. (b) **Close** the open objects (Fig. 2b), Small gaps were closed by using the combination of the external edge of the object with the watersheds of a distance map of the background. (c) **Split** joined objects, e.g. a single vessel that turned in the *up-down* dimension (Fig. 2c). The larger holes of the lumina were used to generate a distance map that was split with the watershed transform to partition the original object into smaller ones.

3   Morphological Analysis

The segmentation algorithm described above provided a series of objects that described the endothelial cells of a tumour that were stained for CD31 and as such the objects described closely the shape of the microvessels of the tumour, as they appear in 2-dimensions. The last stage of the segmentation algorithm was to obtain a series of measurements that provided morphological information about the vasculature of the tumour. The following measurements were extracted from each individual vessel (segmented objects):

(1) *SA* - Area of segmented object excluding lumen (Stained Area [μm^2]) (2) *VA* - Area of segmented object including lumen (Vascular Area [μm^2]), (3) lu/VA - Ratio of lumen to vascular area, (4) *e* - Eccentricity of the vessel: $e = \sqrt{1 - \frac{ma^2}{MA^2}}$ where (MA, ma) are the major and minor axes of an equivalent ellipse, (5) *ro* -Roundness of the external boundary of the vessel: $ro = P/\sqrt{4\pi VA}$ where P is the actual perimeter of the object and *VA* is the vessel area, (6) *rVA* - relative Vascular Area as the ratio of the total Vascular Area (sum of the areas of all objects in the image) relative to the total area of the image. This metric indicates the extent of vascularisation of the tumour.

4   Results and Discussion

The robustness of the segmentation algorithm lies with the chromatic characteristics of the immunostained vessels of the tumours and the intrinsic difference with the haematoxylin-stained nuclei of tumour cells. Although there are differences in the shades obtained from the IHC process, the endothelial cells can be better discriminated in the *HSV* channels than in the traditional *RGB* channels.

Fig. 4 shows the histograms of the following morphometric measurements: (a) *VA*, (b) lu/VA, (c) *e*, (d) *ro*, and (e) *rVA*. Forty-four images were acquired from the three different tumours and 6,163 vessels were segmented in total. The data for each tumour are shown separately, as indicated in the legend. Some differences in the vascular morphology of the three tumours are apparent in the histograms. For instance, a higher proportion of larger-sized vessels (as measured by *VA*) are shown in tumour 2 compared with tumours 1 and 3 (Fig. 4a). The larger vessels in tumour 2 were also associated with larger lumina relative to *VA* (Fig 4b) and combined with the density of the vessels, the *rVA* i.e. the extent of vascularisation was also larger in tumour 2 than in tumours 1 and 3 (Fig 4e). In the future

we expect to use this algorithm to compare the measurements in different populations, for instance, tumours treated with a certain drug against a control or regional variation according to the position within the tumour.

The strength of the algorithm resides with the large number of microvessels that can be analysed by a fully automatic segmentation. The automatic segmentation also ensures consistent criteria for object identification and allows the process to be run on a set of images in the background. In addition, the algorithm provides morphometric analysis of microvessels from which general population statistics can be calculated and has general applicability for a range of different tissues and therapeutic interventions.
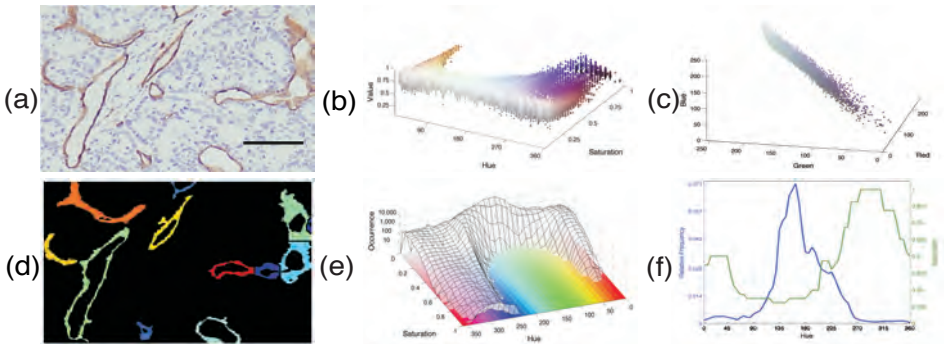
Figure 1: (a) Immunostained tumour section, nuclei appear blue-purple, ECs brown-beige and background in white-grey (bar $= 80 \, \mu$m). (b) A cloud of coloured points describes the chromatic distribution of (a) in the *RGB* domain. (c) A cloud describing m_{HSV} within the *HSV* domain. Notice the discrimination between hues in (c) that is not visible in (b).(d) Final segmentation of (a). (e) m_{HS} overlaid on the loci of constant value. (f) $p_{\max S}$ and m_H.
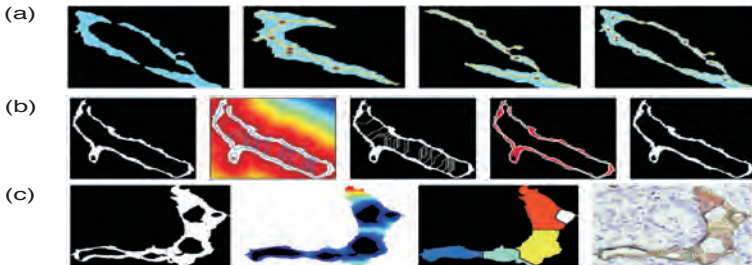

Figure 2: Morphological analysis: (a) joining 2 separate objects, skeleton (yellow) and branching points (brown) of individual objects (blue). (b) Closing of open objects, watershed transformation (blue lines), external boundaries of the original object (white) and the new object (blue) (c) Splitting of objects that may correspond to more than one vessel.
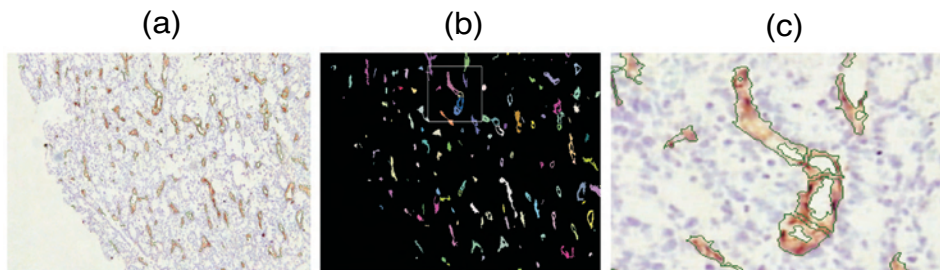
(a)   (b)   (c)

Figure 3: (a) IHC image with the boundaries of the segmented vessels overlaid with a green line. (b) Segmented objects. The region of interest denoted by the white box is shown in detail in (c).
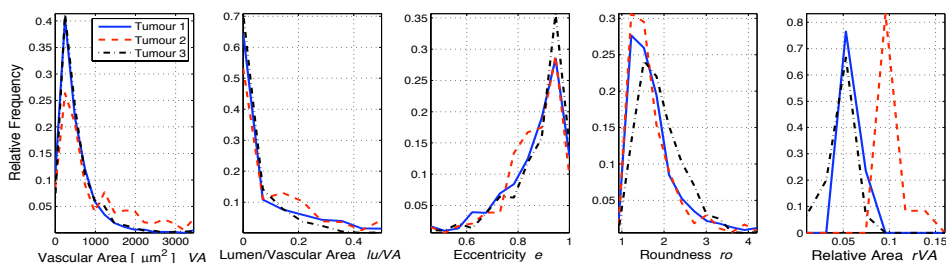

Figure 4: Histograms for the morphometric measurements of microvessels from 3 tumours. (a) *VA*, (b) *lu/VA*, (c) *e*, (d) *ro*, and (e) *rVA*. Solid blue, dashed red line and dash-dot black lines correspond to tumours 1-3 respectively. Parameters in (a-d) are acquired per object from 12-17 images per tumour, whereas a single value per image is acquired for *rVA* in (e).

References

[1] M. A. den Bakker, S. J. Flood, and M. Kliffen. CD31 staining in epithelioid sarcoma. *Virchows Arch*, 443 (1):93–7, 2003.

[2] S. Enomoto, M. Sata, D. Fukuda, K. Nakamura, and R. Nagai. Rosuvastatin prevents endothelial cell death and reduces atherosclerotic lesion formation in apoe-deficient mice. *Biomed Pharmacother*, 63: 19–26, 2009.

[3] J. C. Goddard, C. D. Sutton, P. D. Furness, R. C. Kockelbergh, and K. J. O'Byrne. A computer image analysis system for microvessel density measurement in solid tumours. *Angiogenesis*, 5:15–20, 2002.

[4] D. Kim, C. W. Gregory, G. J. Smith, and J. L. Mohler. Immunohistochemical quantitation of androgen receptor expression using color video image analysis. *Cytometry*, 35(1):2–10, 1999.

[5] P. Ranefall, L. Egevad, B. Nordin, and E. Bengtsson. A new method for segmentation of colour images applied to immunohistochemically stained cell nuclei. *Analytical Cellular Pathology*, 15(3):145–156, 1997.

[6] C. C. Reyes-Aldasoro. Retrospective shading correction algorithm based on signal envelope estimation. *Electronics Letters*, 45(9):454–455, 2009.

[7] A. R. Smith. Color gamut transform pairs. *ACM Siggraph Computer Graphics*, 12(3):12–19, 1978.

[8] C. R. Taylor and R. M. Levenson. Quantification of immunohistochemistry - issues concerning methods, utility and semiquantitative assessment ii. *Histopathology*, 49(4):411–424, 2006.