



City Research Online

City, University of London Institutional Repository

Citation: Reyes-Aldasoro, C. C., Williams, L. J., Akerman, S., Kanthou, C. & Tozer, G. M. (2011). An automatic algorithm for the segmentation and morphological analysis of microvessels in immunostained histological tumour sections. *Journal Of Microscopy*, 242(3), pp. 262-278. doi: 10.1111/j.1365-2818.2010.03464.x

This is the accepted version of the paper.

This version of the publication may differ from the final published version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/5508/>

Link to published version: <https://doi.org/10.1111/j.1365-2818.2010.03464.x>

Copyright: City Research Online aims to make research outputs of City, University of London available to a wider audience. Copyright and Moral Rights remain with the author(s) and/or copyright holders. URLs from City Research Online may be freely distributed and linked to.

Reuse: Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

City Research Online:

<http://openaccess.city.ac.uk/>

publications@city.ac.uk

Title: An automatic algorithm for the segmentation and morphological analysis of microvessels in immunostained histological tumour sections

Authors: Constantino Carlos Reyes-Aldasoro, Leigh J Williams, Simon Akerman, Chryso Kanthou and Gillian M Tozer

Address: Cancer Research UK Tumour Microcirculation Group, The University of Sheffield, Department of Oncology, K Floor, School of Medicine, Beech Hill Road, Sheffield, S10 2RX, UK

E-mail: c.reyes@sheffield.ac.uk

Telephone: 00 44 (0) 114 271 2850

Fax: 00 44 (0) 114 271 3791

Abstract

A fully automatic segmentation and morphological analysis algorithm for the analysis of microvessels from CD31 immunostained histological tumour sections is presented. Development of the algorithm exploited the distinctive hues of stained vascular endothelial cells, cell nuclei and background, to provide the seeds for a 'region-growing' method for object segmentation in the 3D Hue, Saturation, Value (*HSV*) colour model. The segmented objects, identified as microvessels by CD31 immuno-staining, 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. The automatic segmentation was validated against a hand-segmented set of 44 images from 3 different SW1222 human colorectal carcinomas xenografted into mice. $96.3 \pm 0.9\%$ of pixels were found to be correctly classified. Automated segmentation was carried out on a further 53 images from 3 histologically distinct mouse fibrosarcomas (MFs) for morphological comparison with the SW1222 tumours. Four morphometric measurements were calculated for each segmented vessel: vascular area (*VA*), ratio of lumen area to vascular area (*lu/VA*), eccentricity (*e*), and roundness (*ro*). In addition, the total vascular area relative to tumour tissue area (*rVA*) was calculated. *lu/VA*, *e* and *ro*

were found to be significantly smaller in MF tumours than in SW1222 tumours ($p < 0.05$; unpaired t-test). The algorithm is available through the website <http://www.caiman.org.uk> where images can be uploaded, processed and sent back to users. The output from CAIMAN consists of the original image with boundaries of segmented vessels overlaid, the calculated parameters and a Matlab file, which contains the segmentation that the user can use to derive further results.

Keywords: Vessel segmentation, immunohistochemistry, vessel morphology

Introduction

The immunohistochemical staining of tissue sections for different proteins is a standard method for diagnostic and research purposes (Ramos-Vara, 2005). 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 (Newman, 1997, van Mourik *et al.*, 1985). In oncology, the expression of CD31 by endothelial cells in angiogenic vessels has gained considerable attention (Aroca *et al.*, 1999, Righi *et al.*, 2003, Wang *et al.*, 2008) as the tumour vasculature is emerging as an important therapeutic target for cancer (Tozer *et al.*, 2005, Heath & Bicknell, 2009).

Despite the popularity of the use of immunohistochemistry 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. Manual analysis methods involve qualitative assessment of staining distribution (den Bakker *et al.*, 2003, Duyndam *et al.*, 2002), semi-quantitative assessment of staining intensity via visual observation (Enomoto *et al.*, 2009, Rogers *et al.*, 2007), or quantitative assessment of staining distribution (Hansel *et al.*, 2003, Koukourakis *et al.*, 2000, Lund *et al.*, 2000, West *et al.*, 2005) using direct counting methods with or without applications of stereological principles to obtain morphometric data e.g. Chalkley point counting (Chalkley, 1943, Weibel, 1989).

Numerous algorithms for immunostained image analysis have been developed for different specific tasks: counting objects such as nuclei, cells or microvessels (Van der Laak *et al.*,

1998, Chantrain *et al.*, 2003, Wester *et al.*, 1999, Goddard *et al.*, 2002, Goddard *et al.*, 2003, Zhou & Mao, 2005, Munzenmaier & Greene, 2006), quantifying optical density (Kim *et al.*, 1999, Maximova *et al.*, 2006), measuring the abundance of a stain (Lauronen *et al.*, 2006, Ranefall *et al.*, 1998, Ranefall *et al.*, 1997, Pham *et al.*, 2007) or extracting morphometric measurements such as area, perimeter, ratio or perimeter and area and angle of microvessels (Luukkaa *et al.*, 2007, Virgintino *et al.*, 1997, Maximova *et al.*, 2006, Dagnon *et al.*, 2008, Van der Laak *et al.*, 1998, Laitakari *et al.*, 2003, Laitakari *et al.*, 2004, Sadoun & Reed, 2003, Tsuji *et al.*, 2002). 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 (Kim *et al.*, 1999, Lauronen *et al.*, 2006, Pham *et al.*, 2007), counting of vessels or demarcation of regions of interest or vessels (Tsuji *et al.*, 2002, Lauronen *et al.*, 2006), (b) *post-processing tasks* such as delineation of lumina for segmented objects (Van der Laak *et al.*, 1998), inclusion/exclusion of regions of interest (Chantrain *et al.*, 2003), elimination of artefacts (Van der Laak *et al.*, 1998, Virgintino *et al.*, 1997) or (c) *training of classifiers* (Kim *et al.*, 1999, Ranefall *et al.*, 1997, Ranefall *et al.*, 1998) with supervised data. “Automation” in some cases refers to the automatic selection of a threshold (Kim *et al.*, 1999), quantification and measurement procedures after the user selected a threshold (Dagnon *et al.*, 2008) or selection of objects according to user-defined sizes (Goddard *et al.*, 2002).

Most of these algorithms use either the Red, Green and Blue (*RGB*) channels, a subset of these or the intensity of the grey level conversion from the colour images from a light microscope, and few exploit the higher discrimination that can be achieved when the *RGB* channels, the immediate output of a digital camera, are transformed to the Hue, Saturation and Value (*HSV*), Hue, Saturation and Intensity (*HSI*) or Hue, Saturation and Density (*HSD*) channels (van Der Laak *et al.*, 2000, Maximova *et al.*, 2006). In addition, most algorithms

use single threshold selection as a segmentation technique, which is not always accurate. Moreover, to the best of the authors' knowledge, no segmentation algorithm performs any morphological post-processing except for the deletion of objects according to their size (Kim et al., 1999) or adjacency to the border of the image (Goddard et al., 2003) and splitting of round nuclei (Zhou & Mao, 2005). 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. For this paper, all the processing is done in *MATLAB* (The Mathworks, USA ©), and is deployed as an open web-based application.

The aim of this work was to develop an algorithm for quantitatively analysing stained ECs in tumour tissue, with particular emphasis on producing a series of segmented objects that accurately describe the tumour microvessels from which measurements of stained area, vascular eccentricity and vessel roundness can also be obtained. This algorithm could be extrapolated to aid the analysis and quantification of immunohistochemical staining in general. The algorithm is part of an on-line algorithm repository for Cancer Image Analysis called CAIMAN, which can be accessed at <http://www.caiman.org.uk>. Users can upload immunostained images, which will be processed with the algorithm described below and segmented ECs will be overlaid on the original image and sent back to the user with a series of measurements provided by the algorithm.

Here we describe the theoretical bases of the algorithm (pre-processing, segmentation and post-processing of the images and calculation of vascular parameters from the segmented images), as well as examples of results from a human colorectal carcinoma xenografted into mice and a mouse fibrosarcoma transplanted into mice. We determine

that the carcinomas and fibrosarcomas have similar total vascular areas but that the individual vessels in the fibrosarcomas have relatively smaller lumens than those in the carcinomas, which may have therapeutic significance. Mathematical details of the procedures comprising the algorithm are provided in the appendices.

Materials and Methods

Tumour models and Immunohistochemistry

All animal procedures were carried out in accordance with the UK Animals (Scientific Procedures) Act 1986. Human colorectal carcinoma SW1222 cells (provided by Dr. R. B. Pedley, Department of Oncology, Royal Free and University College Medical School) or mouse fibrosarcoma (MF) cells derived from transformed mouse embryo fibroblasts (Tozer *et al.*, 2008) were grown subcutaneously in the rear dorsum of six severe combined immunodeficiency (SCID) mice (three for each line). Mice were sacrificed and the tumour excised when their geometric mean diameter reached 6-8 mm diameter, 10-14 days post implantation. Zinc-fixed tumour sections (5 μm thick) were deparaffinised and antigen retrieval was performed with 20 $\mu\text{g ml}^{-1}$ Proteinase K (Chemicon International - Millipore, Billerica, USA) for 15 minutes at 37 °C. Sections were blocked with species-specific serum for one hour followed by endogenous avidin, biotin and peroxidase blocks and were then incubated with purified rat anti-mouse CD31 (PECAM-1) monoclonal antibody (1:500, BD Pharmingen – BD Biosciences, San Diego, USA) overnight at 4° C. Sections were washed (PBS + 0.1% Tween-20) and secondary biotinylated rat antibody (1:400, Vector Laboratories – Peterborough, UK) was applied for 1 hour at room temperature. Sections were washed and the immune signal amplified by AB complex (DAKO) for 1 hour at room temperature. The immune reaction was detected by DAB (DAKO), with positive staining

appearing brown, against blue counterstaining, with Mayers Hematoxylin (Fluka – Sigma Aldrich, Gillingham, UK).

Image acquisition

Bright field images were captured with the following acquisition settings: x10 magnification objective (N Plan 10x /0.25) on a Leica DM 1000 light microscope attached to a JVC KY-1030U digital camera (1.5 Megapixel 1/2" CCD). Each image contained 1280 × 960 pixels, which measured $0.93 \times 0.93 \mu\text{m}^2$, thus covering $1.1904 \times 0.8928 \text{ mm}^2$. A stage micrometer (Agar Scientific ©) was used to calibrate the ratio of pixels to μm . Images were acquired across a single central section from each tumour. In total, 44 SW1222 images and 53 MF images (12 to 19 per tumour) were captured. The software used for the acquisition was KY-link (JVC, Japan ©).

Overview of the algorithm

Pre-processing of images was necessary to equalise background inhomogeneity, as published previously (Reyes-Aldasoro 2009). The segmentation method was based on a transformation from the *RGB* (Red, Green, Blue) to the *HSV* (Hue, Saturation, Value) colour model. The *HSV* model describes perceptual colour relationships related to the artistic ideas of hue, tint and shade (Smith, 1978, Gonzalez & Woods, 2008) enabling a three dimensional chromatic histogram to be constructed from the *HSV* images. The saturated brown pixels, representing intense CD31 staining, were selected as the starting point or 'seeds' for a 'region-growing' segmentation method and the bright pixels in the yellow-green-cyan hues were selected as background and a stopping criterion for the growing

process. Since the calculated morphological vascular parameters were obtained from each segmented object i.e. vessel of the image, it was crucial to perform three morphological tasks as post-processing steps to examine the objects; otherwise the statistical analysis on the shape and size measurements of the vessels could be misleading.

First, where segmentation resulted in a series of unconnected objects, which should clearly form a single vessel, these objects were linked into a single vessel, on the basis of a small distance between the objects and their shape. Second, some of the thresholded objects appeared with an elliptical shape, identifying them as vessels, but they were not closed structures. These objects were closed, where the gap between the edges of the object was small. Third, objects with more than one inner lumen were split into separate objects. Since the pixels that described the vessels and their position within the tumour were recorded, statistical measurements such as eccentricity and proportional lumen area were extracted.

Definition of the data and pre-processing

Throughout 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 $[R,G,B]$ red, green and blue. Hue, $h \in [0, 360^\circ]$, is a circular property related to the wavelength of the colour where red corresponded approximately to 0° , yellow to 60° , green to 120° , cyan to 180° , blue to 240° and magenta to 315° . The saturation $s, s \in [0, 1]$, is a measure of the purity of the colour or its departure from white or grey and value $v, v \in [0, 1]$, is a measure of the brightness. Mathematical details of the image and the transformations are described in the appendix. The first step of the algorithm corrected the background inhomogeneity or shading of the images and equalised the histogram of the RGB channels. It was assumed that an original

unbiased image U was corrupted by a slowly-varying shading S so that $I = U + S$. Therefore, the corrected image \hat{U} , which was an estimation of U , was given by: $U \approx \hat{U} = I - S$. The shading component was estimated as the *envelope* of the signal (Reyes-Aldasoro, 2009). The envelope can be understood as the iterative stretching of a thin flexible surface under which a series of objects, that is the cells in this case, are placed. Initially, the surface was identical to the signal intensity but after a series of stretches, the surface adapted to the peaks of the objects, and intermediate values in between them. The process of stretching was repeated iteratively until change was minimal as defined in (Reyes-Aldasoro, 2009). Once shading was removed, the mean value of the three *RGB* channels was equalised to obtain a background with minimum saturation.

Segmentation procedure I: the three-dimensional hue-saturation-value histogram

The next step of the algorithm transformed the *RGB* image to an *HSV* colour model and obtained the three-dimensional hue-saturation-value histogram $m_{HSV}(h,s,v)$. The intensity histogram (Winkler, 1995) of an image is a well-known measurement of the occurrence or relative frequency of elements at certain intensities or grey levels and it is the basis of threshold-based segmentation algorithms, as opposed to those based on clustering or deformable models (Ma *et al.*, 2010). A distribution that varies along three dimensions can be summarised along one dimension, typically through summing the elements with respect to the desired dimension, the result is called a *marginal distribution*. In the same way, marginal distributions of a 2D distribution consist of 1D distributions. Marginal distributions of a three dimensional histogram would correspond to two-dimensional hue-saturation,

hue-value of saturation-value histograms. \bar{I}_{sat} , \bar{I}_{val} denoted the average values of saturation and value. Histograms of colour images have been used as one dimensional histograms of separate channels (*RGB* or *HSV*) for segmentation (Schettini, 1993), histogram equalisation (Tominaga, 1987, Lu *et al.*, 2005), or image enhancement (Bassiou & Kotropoulos, 2007). In some cases, two channels have been used to form a bi-variate histogram (Angulo & Serra, 2007, Hanbury & Serra, 2003). Three dimensional channel description was discussed in (Marcu & Abe, 1995) as a visualising tool. In 3D, the histogram approaches have concentrated on enhancement of the *RGB* channels (Trahanias & Venetsanopoulos, 1992) or motion detection through 3D statistical characteristics obtained by filtering (Iliev *et al.*, 2007). *HSV* segmentation in 3D was proposed in (Cho *et al.*, 2001). For this work, we introduced the *maximum saturation profile* p_{max_s} as measurement of the distribution of the highest value of saturation for every value of hue.

While the marginal distributions revealed the relative frequency of pixels within a certain hue or saturation, the maximum saturation profile revealed important characteristics of the image as it was important to distinguish between the hues with low saturation, which may correspond to background regions, and those highly saturated, which may correspond to regions of interest, that is the endothelial cells and the nuclei. Those hues with very bright pixels only ($p_{max_s} \leq 0.25$) were considered as background (Androutsos *et al.*, 1999, Angulo & Serra, 2005) as the hue did not contain saturated colours. 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 described below. The colour variations inherent to immunohistochemistry (Taylor & Levenson, 2006) resulted in images where the colour of stained nuclei was either blue or light purple and the endothelial cells were brown to light

beige. To compensate for this variation, the amount of dark 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, the range of brown, background and blue were approximately (10° , 50°), (60° , 200°), (220° , 310°) respectively, while for those with a low component it was (350° , 90°), (90° , 190°), (200° , 290°) respectively. These values were established after numerous immunostained images (more than 600) of different experiments were analysed in the *HSV* space.

Segmentation procedure II: region-growing process

The endothelial cells in brown previously defined, were used as seeds in the next 'region-growing' step of the algorithm for image segmentation while bright background and blue nuclei, were also used for the stop criterion. Region-growing is a process in which all the neighbouring pixels of a small region or seed form an outer boundary. If the pixels of the boundary fulfil certain criteria of similarity, they are merged to the region and the procedure continues until no neighbouring pixels comply with the criteria. The criteria of similarity for the images were: (a) distance in rows and columns from the seed pixels, (b) distance in hue between the neighbouring pixels and the hues of the seeds, and (c) a combination of the saturation and value: only dark regions (pixels whose value $v < \bar{I}_{val}$) and saturated regions (for each hue, saturation $s > 0.25 p_{\max s}$) were considered in the region growing, bright and non-saturated regions were not included. The three regions grew iteratively; the seeds were dilated to construct the outer boundary previously mentioned, and the pixels tested for the previous criteria. The region-growing stopped when minimal change (10 pixels) was detected between iterations. Small (1 or 2 pixels) and isolated regions were considered as spurious and removed and the region-growing was repeated. If any pixels remained unassigned to any region, they were considered as background. The regions

corresponding to blue were labelled and those that were in contact with brown areas were reassigned to brown as these could correspond to nuclei of the endothelial cells. The resulting regions were smoothed with a closing (dilation followed by erosion) morphological operator.

The regions corresponding to brown, that is the objects formed by endothelial cells, were labelled and their hues and sizes were analysed. A reliability criterion (*rel*) was defined as the ratio of pixels whose hues belonged to the original definition of brown to their area A , with the idea of retaining the objects that were most likely to be vessels and discarding those that could be formed by dark stained nuclei of cells other than ECs. Objects were discarded, i.e. not considered as a vessel (or part of a vessel), and assigned to “background” class, according to their area and reliability criterion (1) $A < 20$ pixels, (2) $rel < 0.1$, (3) $A < 100$ and $rel < 0.2$. These values were heuristically determined to discard very small objects, objects with a very low ratio of brown, and small objects with a low ratio of brown.

Hand segmentation of images

All forty-four images of the SW1222 tumours were manually segmented by a user without knowledge of the results obtained by the algorithm described above. For each image, the vessels and the lumen were delineated to obtain regions of background and vessels. The delineation was done in Matlab using a routine based on the Matlab command *roipoly* for each vessel and lumen. The final result were 44 images with all the vessels segmented from the background. The automated segmentation was compared with manual segmentation using two methods:

(1) *Correct classification* (sometimes referred as the complementary *misclassification* (Randen & Husøy, 1999)) was calculated as the ratio of pixels that belong to identical classes in the automatic and manual segmentation.

(2) *Positive predictive value* (PPV) and *Negative predictive value* (NPV) (Altman & Bland, 1994) were derived from the number of 'True Positive' (TP) pixels, number of 'True Negative' (TN) pixels, and the corresponding 'False Positive' (FP) and 'False Negative' (FN) pixels in the following way: $PPV = TP / (TP + FP)$ and $NPV = TN / (FN + TN)$.

Post-processing

So far, the algorithm has segmented regions of brown or light beige and surrounding pixels of dark, saturated colours into a series of objects, which corresponded to the endothelial cells stained for CD31. However, before the objects were assigned as vessels, three morphological conditions had to be satisfied: objects were not part of a single vessel which have been segmented as two separate, disjointed objects (Fig. 1a), an object was not an open vessel (one where a small gap prevented the vessel from having a closed lumen) which needed to be closed (Fig. 2a) and an object was not formed by several joined vessels (Fig. 4a). The criteria tested to join separate objects were: closeness to each other, the region in between them not being background, objects not containing lumen and an improvement in the shape, in terms of the complexity of a skeleton, of the joined object. The criteria examined for the open vessels were an open lumen with small gap, which could form a watershed between the inside and the outside of the cell (see below). Finally, for the joined objects the criteria tested were to have multiple lumina larger than a predefined size.

The process of joining separate objects, which could form part of a single vessel, was performed as follows: for every segmented object, the Euclidean distance to all other objects was calculated to find its closest neighbour, with the idea of merging objects to form a single vessel. The dimensions of the objects dictated the maximum distance beyond which the objects would not be linked; larger objects allowed relatively distant objects to be considered while smaller objects considered only close neighbours. For each object, the major axis (MA) of an equivalent ellipse were calculated and combined in the following way: $d_{\max} = \log(MA_1) + \log(MA_2)$. The logarithm was used to prevent large objects from linking to very distant objects while allowing small objects to link with not so distant objects.

A second criterion to link the object was determined by the nature of the pixels between the objects. The region of connection was defined as all the pixels of the line of minimum distance between the objects dilated by two pixels. When the majority of pixels of a region of connection belonged to the background previously defined (60% or more), the objects were not joined. Otherwise, two further criteria were tested before joining the objects into a single one. First, both objects should be solid with no inner holes (Euler number = 1). Second and most important perhaps, is that the shape of the combined objects should be closer to the idealised shape of a vessel than the two separate objects. To test this criterion, both objects were reduced to a series of 1-pixel thick lines that followed the original shape of the objects, commonly known as *skeleton* (Blum, 1967). The skeleton (or medial axis transform) followed the shape of the objects with a thin line and for most objects consisted of several segments connected at branching points. Two objects (cyan), skeleton (yellow) and branching points (brown) are presented in Fig. 1(b, c). The first skeleton had 5 branching points and the second 4. The number of branching points of the skeleton was considered as a measure of the uniformity and smoothness of the shape of

each object, thus if the number of branching points of the single joined object exceeded the number of branching points of the individual objects, the joining was discarded. When all previous criteria were fulfilled, the objects were joined into a single one by changing the class (from background to vessel) of the previously defined region of connection into an object. The skeleton of the combined objects (Fig. 1d) presented nine branching points, thus fulfilling the last criterion.

To close the open objects (Fig. 2a) the following process was carried out: for every segmented object of the image, the Euclidean distance between the pixels of the surrounding background (both inside and outside the object) was calculated to generate a distance map. The distances were defined as a negative value that increased, as the pixels were more distant from the object. In a topographical analogy with the distance corresponding to the altitude of a map, the object constituted a “plateau” while the background was a descending slope. Should this artificial landscape “received rain”; the water would find the path through which it could reach a region of minimum altitude, sometimes called lake or sea. Each point in the map corresponded to a path towards one and only one lake, and the landscape was partitioned into catchment basins or regions of influence of the regional minima (Fig. 2b). Finally, the boundaries defined by the watershed were used to partition the original object into a series of smaller objects. This segmentation process is called a *watershed transformation* (Vicent & Soille, 1991). The boundaries defined by the watershed transformation were added (mathematical *OR*) to the object (Fig. 2c) and the external boundary of the object was calculated. If the external boundary was identical to the original boundary, then the object could not be closed. However, when the external boundary (blue in Fig. 2d) was different from the original boundary of the object (white) through one watershed between the inside and the outside

of the vessel, the pixels that belonged to only the new boundary (and not the original) were added as a bridge, thus closing the open gap of the object (Fig. 2e). A final distance restriction was applied: the object was only close when the gap was relatively small, 10 pixels is a compromise between a small gap that could be due to the segmentation process and a large one that could be due to a region without ECs.

A final and important step was to divide those objects that contained more than one lumen. These objects could be formed by separate vessels lying very close to or even touching each other. Examples include sections cut through a branching vessel (Fig. 3a) or a tortuous vessel (Fig. 3b). Another possibility is that of a vessel with the endothelial cells bridging the lumen. Fig. 4a presents an object with five larger holes and several small holes. By splitting these objects with multiple lumina into individual vessels, the statistics collected will be on a per-lumen basis instead of on a per-vessel basis.

When an object was detected to have more than one inner hole, the holes were labelled and the original object was removed from the analysis, leaving a new image with a series of objects corresponding to the holes. If the holes were smaller than a predefined size, they were considered to be noise and not to be split. Then, a distance map was formed with the Euclidean distance from each pixel of the object to its closest hole (Fig. 4b). The distance map was again segmented with the *watershed transformation* (Fig. 4c). Finally, the boundaries defined by the watershed were used to partition the original object into a series of smaller objects (Fig. 4d).

Calculation of morphological parameters

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

measurement of eccentricity was an estimation that depended on the fitting of the ellipse to the object. For this work, the eccentricity provided by the *regionprops* command of MATLAB and the direct least-square fitting method (Fitzgibbon *et al.*, 1999) were compared; both methods provided equivalent results (data not shown).

The roundness ro of the objects was estimated as the ratio of the perimeter of the object over the perimeter of a circle of the equivalent area:

$$ro = \frac{P}{\sqrt{4\pi VA}}, \quad (2)$$

where P is the actual perimeter of the object and VA is the Vascular Area. This metric described how elongated the object was together with how smooth or crenated the external boundary of each vessel was. A final measurement was calculated per image: 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 relative Vascular Area (rVA) indicates the extent of vascularisation of the image.

The statistical significance of differences in the mean values of the calculated parameters between the two tumour types was tested using an unpaired t-test. Values of $p < 0.05$ were considered significant.

Results and Discussion

Segmentation by simple intensity thresholding of the stained endothelial cells did not produce adequate results (data not shown). Several factors are likely to account for this:

first, the images may have inhomogeneous background intensities, or shading, which required correction. Second, the intensity of the EC staining may not be uniform, and vessels may appear as fragmented objects or several vessels may appear to be linked as a single object. Third, variations in the immunostaining and acquisition process can produce differences in colour and quality of the image and the stained cells may have different shades. Endothelial cells stained from dark brown to light beige. Cell nuclei appeared from blue-purple to light magenta, while the background was white to light grey (Fig. 5a). Optical artefacts were manifest as non-uniform background in the images as can be seen in Fig. 5b, where the intensities of the *RGB* channels of one row of the image in Fig. 5a are presented. The black arrow on the right hand side shows a decrease in the intensity of all three channels. The channels after shading correction are presented in Fig. 5c where the background intensities are uniform within the image.

The segmentation result of Fig. 5a is presented in Fig. 6, first as the boundaries of the segmented objects overlaid with a green line (Fig. 6a) and second with the objects labelled with different colours that allowed visual discrimination (Fig. 6b). Fig. 7 presents three regions of interest from each of three SW1222 tumours (Fig 7a-c) and three MF tumours (Fig 7d-f) at higher magnification. It can be seen that the automated segmentation is effective in both tumour types (green boundaries in Fig 7b and 7e; coloured objects in Fig 7c and 7f), even though there are qualitative differences in the background staining between tumours shown in Fig. 7a and 7d. In order to graphically compare the staining of tumours, the three-dimensional histogram $m_{HSV}(h,s,v)$ obtained from Fig. 7a (centre tumour) is presented in Fig. 8b next to its *RGB* equivalent cloud of points in Fig. 8a. Both plots present a cloud of points located along the axis of red, green and blue or hue, saturation and value with its corresponding colour. It should be noticed how the brown and

purple points are well separated in the *HSV* model while they are difficult to distinguish in the *RGB* model. Fig. 8c presents the marginal distribution hue-saturation $m_{HS}(h,s)$, which has been overlaid on the 2D loci of constant value as an aid to visualise the pixel distribution. A 2D plane where hue and saturation vary along their corresponding axis while value remains constant is called *loci of constant value*. In this loci colours spread from one side to another, while the other axis determines the purity or “whiteness” of the colour. The relative frequency of pixels with a certain hue and saturation determines the height of $m_{HS}(h,s)$, that appears as a mesh of black lines. Notice that the z-axis is in a logarithmic scale. The pixels with high saturation are concentrated in the region of $(250^\circ, 50^\circ)$ in hue (purple to orange), while the region $(50^\circ, 250^\circ)$ (yellow to blue) has very low saturation (tend to white). The hue histogram $m_H(h)$ (solid blue line) and maximum saturation profile $p_{\max S}$ (dashed green line) are presented in Fig. 8d. While the histogram indicates the relative frequency of pixels of different hues, $p_{\max S}$ indicates the maximum saturation for a given hue. $p_{\max S}$ reveals information about the image that is not evident in the histogram such as the distribution of the saturated pixels (peaks at $(0^\circ, 50^\circ)$ and $(270^\circ, 360^\circ)$) and the background (valley between $(90^\circ, 200^\circ)$). The pixels with hues in the low saturation region were used as background, and the seeds for the region-growing algorithm were obtained from the high saturation pixels. Fig. 9 presents the histograms and maximum saturation profile corresponding to the image of Fig. 7a (right-hand tumour). The chromatic difference between the images of Fig. 7a (centre versus right-hand image) can be observed in the blue region (around 250°) both in $m_{HSV}(h,s,v)$ and $m_{HS}(h,s)$ due to a higher presence of blue-purple pixels present in the histograms of Fig. 8 relative to Fig. 9. While there is a striking difference in the relative frequency of pixels with particular characteristics in the 1D histograms (solid blue lines, comparing Figs. 8 and 9), the $p_{\max S}$

profiles (dashed green lines) remain similar, which indicates the robustness of $p_{\max S}$ as an estimator for the seeds of the region-growing process.

The robustness of the segmentation process lies with the chromatic characteristics of the immunostained vessels of the tumours and the intrinsic difference with the haematoxylin-stained nuclei of the SW1222 tumour cells. Although there are differences in the shades obtained from the immunohistochemistry process, the endothelial cells can be better discriminated in the *HSV* channels than in the traditional *RGB* channels as can be seen in Figs. 8,9. Also, the use of an adaptive hue range, rather than a fixed one compensates for variations in the distributions of the colours and resulted in good segmentation of the cells.

All forty-four images of the SW1222 tumours were manually segmented by a user without knowledge of the results obtained by the algorithm described above. Fig. 10a presents the automatic segmentation of one representative image and Fig 10b the manual segmentation. Fig 10c presents the pixel-to-pixel comparison, when pixels were assigned to the same class by both methods, this was labelled in white and the black regions correspond to different classes, that is, the incorrectly classified pixels. The automatic results are presented as green boundaries overlaid in the original image in Fig 10d. For this image 97.12% of the pixels were assigned into identical classes (background/vessels) by both techniques. The 2.88% misclassified pixels correspond to very small vessels not segmented by the algorithm and the edges of larger vessels, where the hand-segmented delineation tended to smooth the boundaries of the cells and the automatic algorithm presented more “serrated” edges. Overall, the *correct classification* of the automatic against the manual segmentation was 0.963 ± 0.009 (mean \pm SD for $n=3$ tumours) of the total pixels from the 44 images. The corresponding *Positive predictive value* (PPV) was

0.991±0.016 and the *Negative predictive value* was 0.978±0.008. Therefore, the automated segmentation method provided an accurate means of identifying tumour blood vessels from which morphological parameters could be calculated.

Fig. 11 shows the histograms of the following morphometric measurements from the SW1222 carcinomas (Fig. 11a) and the MF fibrosarcomas (Fig. 11b): (i) vascular area VA , (ii) ratio of lumen area over vascular area lu/VA , (iii) eccentricity e , (iv) roundness ro , and (v) relative vascular area rVA . The number of vessels segmented were 6,163 from the three SW1222 and 9,920 from the MF tumours. The data for each tumour are shown separately, as indicated in the legend. Some differences in the vascular morphology of the three SW1222 tumours, which are visible in the selected images from each tumour shown in Fig. 7, are also apparent in the histograms for the whole data sets (Fig. 11a) and Table 1. 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. 11a(i)). The larger vessels in tumour 2 were also associated with larger lumina relative to the vascular area, VA (Fig 11a(ii)) and combined with the density of the vessels, the relative vascular area (rVA) i.e. the extent of vascularisation was also larger in tumour 2 than in tumours 1 and 3 (Fig 11a(v)). In the measurements of shape, there was little variation in e , while tumour 3 presented a higher value of ro than all other tumours (Fig 11a(iv)). There were no significant differences between the SW1222 tumours and the MF tumours for vascular areas of individual vessels (VA) or the relative vascular areas (rVA) (Table 1).

There was a significant but small difference in eccentricity (e) between the MF and SW1222 tumours and a borderline significant difference in (ro) between the two tumour types. However, the lumen to vascular area ratio (lu/VA) was significantly smaller in the MF tumours than in the SW1222 tumours (comparison of Figs 11a and b and Table 1). The

larger lumens in the SW1222 tumours may relate to the relatively well-differentiated glandular structure of the SW1222 tumours (El Emir *et al.*, 2007) and our own unpublished results showing that the SW1222 tumours are more resistant to vascular shut-down by the tumour vascular disrupting agent, combretastatin A-4-phosphate (CA-4-P) than the MF tumours. The relationship between vascular morphology and response to vascular targeting agents remains to be fully investigated but the current results illustrate the usefulness of the automated algorithm in providing a rapid means of quantifying vascular characteristics of individual blood vessels, which may have therapeutic significance. In contrast, manual methods are slow and can only provide relatively simple measurements such as vessel counts. In the future we expect to use this algorithm to compare the measurements in different populations, for instance, tumours treated with a vascular-targeted drug against a control group or to measure regional variation according to the position within the tumour.

CAIMAN Website

The segmentation algorithm is provided to users through a website called “CAIMAN”. CAIMAN is an Image Analysis internet-based project (Reyes-Aldasoro *et al.*, 2010) that combines the scripting language *PHP* (The PHP group, free software), the high-level technical computing language *MATLAB* (The Mathworks, USA ©), the *Interactive Object Management Environment* (IOME) Markup Language (Griffiths *et al.*, 2009) and specifically designed image analysis algorithms ((Reyes-Aldasoro *et al.*, 2008, Reyes-Aldasoro, 2009) for instance), to provide a user-friendly web-page where any person can upload images from their experiments and execute analysis algorithms to obtain quantitative measurements. The website and the use of IOME have been presented in

more detail in (Reyes-Aldasoro et al., 2010, Griffiths et al., 2009), but briefly: the user selects the appropriate algorithm from those available through the CAIMAN website and uploads the image to be analysed. Once the images have been uploaded to the web-server, the IOME-PHP toolbox is used to make a web service request to the image analysis service, which resides in a high performance cluster, this request is made using IOME-ML. The image is transferred to the cluster where it is processed in Matlab and the results are sent to the user via e-mail. The email contains a series of morphological measurements together with two attachments: an image file in *jpeg* format with the segmentation overlaid on the original image and a Matlab file with the segmentation as labelled objects. From this second file, the users can extract more measurements, if required.

Conclusions

An algorithm to segment the ECs from images of stained histological sections has been presented. The strength of the algorithm resides with the large number of microvessels that can be analysed by a fully automatic segmentation. In the sample images shown, no user intervention was required for thresholding, adjustments of brightness and contrast or selection of window sizes. 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. The algorithm can be accessed through a user-friendly website where the user can upload immunohistological images and the results are sent back via *e-mail*. 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.

Acknowledgements

The authors would like to thank Cancer Research UK for funding this work, the White Rose Grid and The University of Sheffield for the computing facilities to run the Caiman website and Dr Michael Griffiths, Dr Deniz Savas and Thilini Daranagama for their valuable help. We would also like to thank Dr Barbara Pedley for kindly providing the SW1222 cell line and staff at the University of Sheffield for care of the animals.

Appendices: Mathematical Details

Image Definition and transformation from RGB to HSV

Throughout 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 $L_r = \{1, 2, \dots, r, \dots, N_r\}$, $L_c = \{1, 2, \dots, c, \dots, N_c\}$ be the spatial domains of the data, $x \in (L_r \times L_c)$ be a pixel of the image, and $[R, G, B] = [\{1, 2, \dots, r, \dots, N_i\}, \{1, 2, \dots, g, \dots, N_i\}, \{1, 2, \dots, b, \dots, N_i\}]$ a triplet of values for red, green and blue. An image was represented then as a function that assigned a colour to each pair of coordinates (Haralick *et al.*, 1973):

$$L_r \times L_c; I_{rgb} : L_r \times L_c \rightarrow [R, G, B]. \quad (3)$$

I_{rgb} was also represented by its individual colour channels $I_{rgb} = [I_{red}, I_{green}, I_{blue}]$. An image in the *HSV* model $I_{hsv} = T(I_{rgb})$ was represented by the individual channels: $I_{hsv} = [I_{hue}, I_{sat}, I_{val}]$ corresponding to the new chromatic triplet $[H, S, V] = [\{0, \dots, h, \dots, 1\}, \{0, \dots, s, \dots, 1\}, \{0, \dots, v, \dots, 1\}]$. Both images had the same dimensions.

The conversion from *RGB* to *HSV* (Smith, 1978) required the red, green, and blue channels to be normalised: $\hat{r} = \frac{R}{255}$, $\hat{g} = \frac{G}{255}$, $\hat{b} = \frac{B}{255} \in [0,1]$. Their maximum and minimum values were: $\max_{rgb} = \max\{\hat{r}, \hat{g}, \hat{b}\}$, $\min_{rgb} = \min\{\hat{r}, \hat{g}, \hat{b}\}$. The value v , $v \in [0,1]$, is a measure of the brightness of the element and was defined as $v = \max_{rgb}$. The saturation s , $s \in [0,1]$, is a measure of the purity of the colour or its departure from white or grey and was defined as:

$$s = \begin{cases} 0 & \text{if } \max_{rgb} = 0 \\ \frac{\max_{rgb} - \min_{rgb}}{\max_{rgb}} & \text{otherwise} \end{cases} \quad (4)$$

Hue, $h \in [0,360^\circ]$, is a circular property related to the wavelength of the colour where red corresponded approximately to 0° , yellow to 60° , green to 120° , cyan to 180° , blue to 240° and magenta to 315° . Each hue was obtained by:

$$h = \begin{cases} 0^\circ & \text{if } \max_{rgb} = \min_{rgb} \\ (60^\circ \times \frac{\hat{g} - \hat{b}}{\max_{rgb} - \min_{rgb}} + 0^\circ) \bmod 360^\circ, & \text{if } \max_{rgb} = \hat{r} \\ (60^\circ \times \frac{\hat{b} - \hat{r}}{\max_{rgb} - \min_{rgb}} + 120^\circ), & \text{if } \max_{rgb} = \hat{g} \\ (60^\circ \times \frac{\hat{r} - \hat{g}}{\max_{rgb} - \min_{rgb}} + 240^\circ), & \text{if } \max_{rgb} = \hat{b} \end{cases} \quad (5)$$

Shading Correction

The shading component was estimated as the *envelope* of the signal or as an iterative stretching of a thin flexible surface. Mathematically, each stretch corresponded to a comparison between the intensities of a pixel and the average value of increasingly distant pairs of opposite 8-connectivity neighbours in 4 orientations: $[0^\circ, 45^\circ, 90^\circ, 135^\circ]$. The

distance d_i between the pairs was increased at every iteration i . The intensity of the pixel was replaced with the maximum value of the averages and the pixel itself:

$$S_{\max}^i(r,c) = \max \left\{ \frac{I(r-d_i,c-d_i) + I(r+d_i,c+d_i)}{2}, \frac{I(r+d_i,c-d_i) + I(r-d_i,c+d_i)}{2}, \frac{I(r-d_i,c) + I(r+d_i,c)}{2}, \frac{I(r,c-d_i) + I(r,c+d_i)}{2}, I(r,c) \right\}. \quad (6)$$

The three-dimensional hue-saturation-value histogram

The hue-saturation-value histogram is a tri-variate measurement of the relative frequency of $[H,S,V]$ on I_{hsv} and it was defined as:

$$m_{HSV}(h,s,v) = \frac{\#\{x \in (L_r \times L_c) : I_{hue}(x) = h, I_{sat}(x) = s, I_{val}(x) = v\}}{\#\{L_r \times L_c\}}, \quad h \in H, s \in S, v \in V, \quad (7)$$

where $\#$ denoted the number of elements in the set. The two-dimensional hue-value $m_{HV}(h,v)$, hue-saturation $m_{HS}(h,s)$ or saturation-value $m_{SV}(s,v)$ histograms are marginal distributions of $m_{HSV}(h,s,v)$. A marginal distribution of a n-dimensional distribution corresponds to the distribution when one of the dimensions is averaged over the other dimensions and it is obtained by integrating out the variables that are not of interest for that marginal (Leon-Garcia, 1994). The marginal distributions of $m_{HS}(h,s)$ constituted the individual histograms for hue and saturation $m_H(h) = \sum_s m_{HS}(h,s)$, $m_S(s) = \sum_h m_{HS}(h,s) \cdot \bar{I}_{sat}$,

\bar{I}_{val} denoted the average values of saturation and value: $\bar{I}_{sat} = \frac{1}{N_r \times N_c} \sum_r \sum_c I_{sat}$ and

$$\bar{I}_{val} = \frac{1}{N_r \times N_c} \sum_r \sum_c I_{val} .$$

The *maximum saturation profile* $p_{\max S}$ is measurement of the distribution of the highest value of saturation for every value of hue, mathematically:

$$p_{\max S} = \{y \in (h \times S) : \max(s) \text{ so that } m_{HS}(y) > 0\}, h \in H . \quad (8)$$

References

- Altman, D. G. & Bland, J. M. (1994) Diagnostic-Tests-2 - Predictive Values .4. *Br. Med. J.* **309**, 102-102.
- Androutsos, D., Plataniotis, K. N. & Venetsanopoulos, A. N. (1999) A novel vector-based approach to color image retrieval using a vector angular-based distance measure. *Comput. Vision Image Understanding*, **75**, 46-58.
- Angulo, J. & Serra, J. (2005) Image Color Segmentation using Bi-variate Histograms in Luminance/Saturation/Hue Polar Color Spaces. *Computac. Sist.* **8**, 303-316.
- Angulo, J. & Serra, J. (2007) Modelling and segmentation of colour images in polar representations. *Image and Vision Comput.* **25**, 475-495.
- Aroca, F., Renaud, W., Bartoli, C., Bouvier-Labit, C. & Figarella-Branger, D. (1999) Expression of PECAM-1/CD31 isoforms in human brain gliomas. *J. Neuro-oncol.* **43**, 19-25.
- Bassiou, N. & Kotropoulos, C. (2007) Color image histogram equalization by absolute discounting back-off. *Comput. Vision Image Understanding*, **107**, 108-122.
- Blum, H. (1967) A Transformation for Extracting New Descriptors of Shape. In: *Models for the Perception of Speech and Visual Form* (ed. W. Wathen-Dunn). MIT Press, Cambridge, USA.
- Chalkley, H. W. (1943) Method for the quantitative morphologic analysis of tissues. *J. Natl. Cancer Inst.* **4**, 47-53.
- Chantrain, C. F., DeClerck, Y. A., Groshen, S. & McNamara, G. (2003) Computerized quantification of tissue vascularization using high-resolution slide scanning of whole tumor sections. *J. Histochem. Cytochem.* **51**, 151-158.
- Cho, K. M., Jang, J. H. & Hong, K. S. (2001) Adaptive skin-color filter. *Pattern Recognit.* **34**, 1067-1073.
- Dagnon, K., Heudes, D., Bernaudin, J. F. & Callard, P. (2008) Computerized morphometric analysis of microvasculature in non-small cell lung carcinoma. *Microvasc. Res.* **75**, 112-118.
- den Bakker, M. A., Flood, S. J. & Kliffen, M. (2003) CD31 staining in epithelioid sarcoma. *Virchows Arch.* **443**, 93-97.
- Duyndam, M. C., Hilhorst, M. C., Schluper, H. M., Verheul, H. M., van Diest, P. J., Kraal, G., Pinedo, H. M. & Boven, E. (2002) Vascular endothelial growth factor-165 overexpression stimulates angiogenesis and induces cyst formation and macrophage infiltration in human ovarian cancer xenografts. *Am. J. Pathol.* **160**, 537-548.
- El Emir, E., Qureshi, U., Dearling, J. L. J., Boxer, G. M., Clatworthy, I., Folarin, A. A., Robson, M. P., Nagl, S., Konerding, M. A. & Pedley, R. B. (2007) Predicting response to radioimmunotherapy from the tumor microenvironment of colorectal carcinomas. *Cancer Res.* **67**, 11896-11905.
- Enomoto, S., Sata, M., Fukuda, D., Nakamura, K. & Nagai, R. (2009) Rosuvastatin prevents endothelial cell death and reduces atherosclerotic lesion formation in ApoE-deficient mice. *Biomed. Pharmacother.* **63**, 91-26.
- Fitzgibbon, A. W., Pilu, M. & Fisher, R. B. (1999) Direct least-squares fitting of ellipses. *IEEE Trans. Pattern Anal. Mach. Intell.* **21**, 476-480.

- Goddard, J. C., Sutton, C. D., Furness, P. D., Kockelbergh, R. C. & O'Byrne, K. J. (2002) A computer image analysis system for microvessel density measurement in solid tumours. *Angiogenesis*, **5**, 15-20.
- Goddard, J. C., Sutton, C. D., Furness, P. N., O'Byrne, K. J. & Kockelbergh, R. C. (2003) Microvessel density at presentation predicts subsequent muscle invasion in superficial bladder cancer. *Clin. Cancer Res.* **9**, 2583-2586.
- Gonzalez, R. C. & Woods, R. E. (2008) *Digital image processing*, Pearson Prentice Hall, Upper Saddle River, NJ.
- Griffiths, M. K., Reyes-Aldasoro, C. C., Savas, D. & Greenfield, A. (2009) IOME, A Toolkit for Distributed and Collaborative Computational Science and Engineering. In: *UK e-Science All Hands Meeting* (eds. A. Trefethen & P. Watson). Oxford, UK.
- Hanbury, A. & Serra, J. (2003) Colour image analysis in 3D-polar coordinates. In: *LNCS 2781* (eds. B. Michaelis & G. Krell). Springer-Verlag, Heidelberg.
- Hansel, D. E., Rahman, A., Hermans, J., de Krijger, R. R., Ashfaq, R., Yeo, C. J., Cameron, J. L. & Maitra, A. (2003) Liver metastases arising from well-differentiated pancreatic endocrine neoplasms demonstrate increased VEGF-C expression. *Mod. Pathol.* **16**, 652-659.
- Haralick, R. M., Shanmugam, K. & Dinstein, I. h. (1973) Textural Features for Image Classification. *IEEE Trans. Syst. Man Cybern.* **3**, 610-621.
- Heath, V. L. & Bicknell, R. (2009) Anticancer strategies involving the vasculature. *Nature Reviews Clin. Oncol.* **6**, 395-404.
- Iliev, P., Tzvetkov, P. & Petrov, G. (2007) Motion detection using 3D image histograms sequences analysis. In: *Proceedings of the Third Workshop - 2005 IEEE Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2005*.
- Kim, D., Gregory, C. W., Smith, G. J. & Mohler, J. L. (1999) Immunohistochemical quantitation of androgen receptor expression using color video image analysis. *Cytometry*, **35**, 2-10.
- Koukourakis, M. I., Giatromanolaki, A., Thorpe, P. E., Brekken, R. A., Sivridis, E., Kakolyris, S., Georgoulas, V., Gatter, K. C. & Harris, A. L. (2000) Vascular endothelial growth factor/KDR activated microvessel density versus CD31 standard microvessel density in non-small cell lung cancer. *Cancer Res.* **60**, 3088-3095.
- Laitakari, J., Harrison, D. & Stenback, F. (2003) Automated image analysis of proliferating cells in carcinoma of the larynx. *Acta Oto-Laryngol.* **123**, 759-766.
- Laitakari, J., N y h , V. & Stenback, F. (2004) Size, shape, structure and direction of angiogenesis in laryngeal tumour development. *J. Clin. Pathol.* **57**, 394-401.
- Launonen, J., Hayry, P. & Paavonen, T. (2006) An image analysis-based method for quantification of chronic allograft damage index parameters. *APMIS*, **114**, 440-448.
- Leon-Garcia, L. (1994) *Probability and Random Processes for Electrical Engineering*, Addison-Wesley Publishing, Reading, Mass.
- Lu, J., Gu, Q., Plataniotis, K. N. & Wang, J. (2005) A comparative study of skin-color models. In: *LNCS 3656* (eds. M. Kamel & A. Campilho). Springer-Verlag, Heidelberg.
- Lund, E. L., Thorsen, C., Pedersen, M. W., Junker, N. & Kristjansen, P. E. (2000) Relationship between vessel density and expression of vascular endothelial growth factor and basic fibroblast growth factor in small cell lung cancer in vivo and in vitro. *Clin Cancer Res*, **6**, 4287-4291.
- Luukkaa, H., Laitakari, J., Vahlberg, T., Klemi, P. & Grenman, R. (2007) Morphometric analysis using automated image analysis of CD34-positive vessels in salivary gland acinic cell carcinoma. *Acta Oto-Laryngol.* **127**, 869-873.
- Ma, Z., Tavares, J. M. R. S., Jorge, R. N. & Mascarenhas, T. (2010) A review of algorithms for medical image segmentation and their applications to the female pelvic cavity. *Comput. Meth. Biomech. Biomed. Eng.* **13**, 235-246.

- Marcu, G. G. & Abe, S. (1995) Three-dimensional histogram visualization in different color spaces and applications. *J. Electron. Imaging*, **4**, 330-346.
- Maximova, O. A., Taffs, R. E., Pomeroy, K. L., Piccardo, P. & Asher, D. M. (2006) Computerized morphometric analysis of pathological prion protein deposition in scrapie-infected hamster brain. *J. Histochem. Cytochem.* **54**, 97-107.
- Munzenmaier, D. H. & Greene, A. S. (2006) Chronic angiotensin II AT(1) receptor blockade increases cerebral cortical microvessel density. *Am. J. Physiol-Heart C*, **290**, H512-H516.
- Newman, P. J. (1997) The biology of PECAM-1. *J. Clin. Invest.* **99**, 3-8.
- Pham, N.-A., Morrison, A., Schwock, J., Aviel-Ronen, S., Iakovlev, V., Tsao, M.-S., Ho, J. & Hedley, D. W. (2007) Quantitative image analysis of immunohistochemical stains using a CMYK color model. *Diagn. Pathol.* **2**, 8-18.
- Ramos-Vara, J. A. (2005) Technical aspects of immunohistochemistry. *Vet. Pathol.* **42**, 405-426.
- Randen, T. & Husøy, J. H. (1999) Filtering for Texture Classification: A Comparative Study. *IEEE Trans. Patt. Anal. Machine Intel.* **21**, 291-310.
- Ranefall, P., Egevad, L., Nordin, B. & Bengtsson, E. (1997) A new method for segmentation of colour images applied to immunohistochemically stained cell nuclei. *Anal. Cell. Pathol.* **15**, 145-156.
- Ranefall, P., Wester, K., Andersson, A. C., Busch, C. & Bengtsson, E. (1998) Automatic quantification of immunohistochemically stained cell nuclei based on standard reference cells. *Anal. Cell. Pathol.* **17**, 111-123.
- Reyes-Aldasoro, C. C. (2009) Retrospective shading correction algorithm based on signal envelope estimation. *Electron. Lett.* **45**, 454-456.
- Reyes-Aldasoro, C. C., Biram, D., Tozer, G. M. & Kanthou, C. (2008) Measuring cellular migration with image processing. *Electron. Lett.* **44**, 791-793.
- Reyes-Aldasoro, C. C., Griffiths, M. K., Savas, D. & Tozer, G. M. (2010) CAIMAN: An online algorithm repository for Cancer Image Analysis. *Comput. Meth. Prog. Bio.* **in press**, doi:10.1016/j.cmpb.2010.07.007.
- Righi, L., Deaglio, S., Pecchioni, C., Gregorini, A., Horenstein, A. L., Bussolati, G., Sapino, A. & Malavasi, F. (2003) Role of CD31/platelet endothelial cell adhesion molecule-1 expression in in vitro and in vivo growth and differentiation of human breast cancer cells. *Am. J. Pathol.* **162**, 1163-1174.
- Rogers, P. A., Donoghue, J. F. & Girling, J. E. (2007) Endometrial Lymphangiogenesis. *Placenta*, **29**, 48-54.
- Sadoun, E. & Reed, M. J. (2003) Impaired angiogenesis in aging is associated with alterations in vessel density, matrix composition, inflammatory response, and growth factor expression. *J. Histochem. Cytochem.* **51**, 1119-1130.
- Schettini, R. (1993) A segmentation algorithm for color images. *Pattern Recognit. Lett.* **14**, 499-506.
- Smith, A. R. (1978) Color gamut transform pairs. *ACM Siggraph Computer Graphics*, **12**, 12-19.
- Taylor, C. R. & Levenson, R. M. (2006) Quantification of immunohistochemistry - issues concerning methods, utility and semiquantitative assessment II. *Histopathology*, **49**, 411-424.
- Tominaga, S. (1987) Expansion of color images using three perceptual attributes. *Pattern Recognit. Lett.* **6**, 77-85.
- Tozer, G. M., Akerman, S., Cross, N., Barber, P. R., Bjorndahl, M., Greco, O., Harris, S., Hill, S. A., Honess, D. J., Ireson, C. R., Pettyjohn, K. L., Prise, V., Reyes-Aldasoro, C. C., Ruhrberg, C., Shima, D. T. & Kanthou, C. (2008) Blood Vessel Maturation and Response to Vascular-Disrupting Therapy in Single Vascular Endothelial Growth Factor-A Isoform-Producing Tumors. *Cancer Res.* **68**, 2301-2311.

- Tozer, G. M., Kanthou, C. & Baguley, B. C. (2005) Disrupting tumour blood vessels. *Nat. Rev. Cancer*, **5**, 423-435.
- Trahanias, P. E. & Venetsanopoulos, A. N. (1992) Color image enhancement through 3-D histogram equalization. In: *Pattern Recognit., 11th IAPR Int. Conf.*, The Hague, Netherlands.
- Tsuji, T., Sasaki, Y., Tanaka, M., Hanabata, N., Hada, R. & Munakata, A. (2002) Microvessel morphology and vascular endothelial growth factor expression in human colonic carcinoma with or without metastasis. *Lab. Invest.* **82**, 555-562.
- van Der Laak, J. A., Pahlplatz, M. M., Hanselaar, A. G. & de Wilde, P. C. (2000) Hue-saturation-density (HSD) model for stain recognition in digital images from transmitted light microscopy. *Cytometry*, **39**, 275-284.
- Van der Laak, J. A. W. M., Westphal, J. R., Schalkwijk, L. J. M., Pahlplatz, M. M. M., Ruiter, D. J., de Waal, R. M. W. & de Wilde, P. C. M. (1998) An improved procedure to quantify tumour vascularity using true colour image analysis. Comparison with the manual hot-spot procedure in a human melanoma xenograft model. *J. Pathol.* **184**, 136-143.
- van Mourik, J. A., Leeksa, O. C., Reinders, J. H., de Groot, P. G. & Zandbergen-Spaargaren, J. (1985) Vascular endothelial cells synthesize a plasma membrane protein indistinguishable from the platelet membrane glycoprotein IIa. *J. Biol. Chem.* **260**, 11300-11306.
- Vicent, L. & Soille, P. (1991) Watersheds in digital spaces - an efficient algorithm based on immersion simulations. *IEEE Trans. Patt. Anal. Machine Intel.* **13**, 583-598.
- Virgintino, D., Monaghan, P., Robertson, D., Errede, M., Bertossi, M., Ambrosi, G. & Roncali, L. (1997) An immunohistochemical and morphometric study on astrocytes and microvasculature in the human cerebral cortex. *Histochem. J.* **29**, 655-660.
- Wang, D., Stockard, C. R., Harkins, L., Lott, P., Salih, C., Yuan, K., Buchsbaum, D., Hashim, A., Zayzafoon, M., Hardy, R. W., Hameed, O., Grizzle, W. & Siegal, G. P. (2008) Immunohistochemistry in the evaluation of neovascularization in tumor xenografts. *Biotech. Histochem.* **83**, 179-189.
- Weibel, E. R. (1989) Measuring through the Microscope - Development and Evolution of Stereological Methods. *J. Microsc.-Oxford*, **155**, 393-403.
- West, C. C., Brown, N. J., Mangham, D. C., Grimer, R. J. & Reed, M. W. (2005) Microvessel density does not predict outcome in high grade soft tissue sarcoma. *Eur. J. Surg. Oncol.* **31**, 1198-1205.
- Wester, K., Ranefall, P., Bengtsson, E., Busch, C. & Malmstrom, P. U. (1999) Automatic quantification of microvessel density in urinary bladder carcinoma. *Brit. J. Cancer*, **81**, 1363-1370.
- Winkler, G. (1995) *Image Analysis, Random Fields and Dynamic Monte Carlo Methods*, Springer, Berlin, Germany.
- Zhou, H. & Mao, K. Z. (2005) Adaptive Successive Erosion-based Cell Image Segmentation for p53 Immunohistochemistry in Bladder Inverted Papilloma. In: *Engineering in Medicine and Biology*. IEEE, Shanghai.

Figures

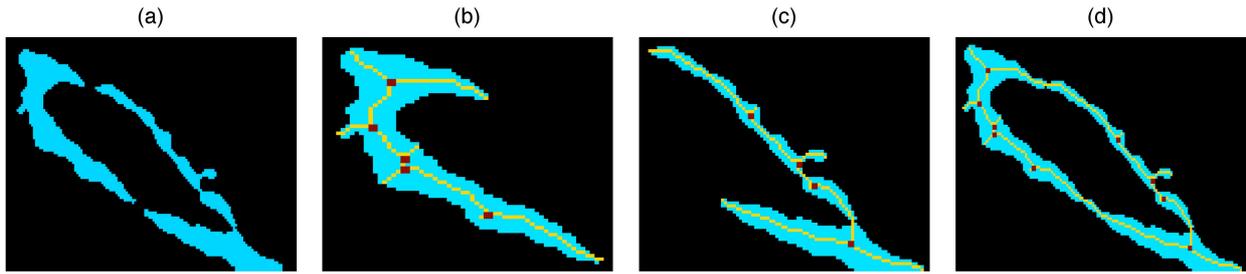


Fig. 1. Morphological task of joining two separate objects. (a) Two objects considered for joining. (b, c) Skeleton (yellow) and branching points (brown) of individual objects; objects present 5 and 4 branching points respectively. (d) Skeleton and branching points (nine) of the objects joined into a single object. When the number of branching points of the joined object exceeds the number of points of the individual objects, joining is not considered (see text for all criteria). In this case, the objects were joined.

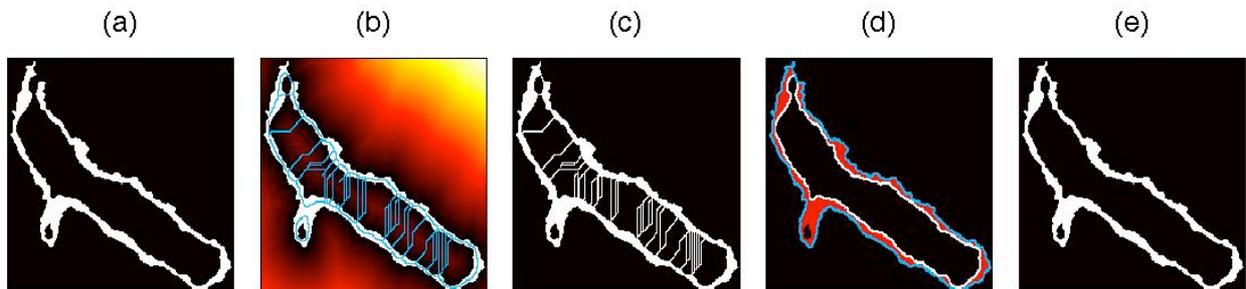


Fig. 2. Morphological task of closing of open objects: (a) A single object with an opening in the upper left hand side. (b) Distance map (brightness denotes distance from object) and segmentation boundaries of a watershed transformation (blue lines). (c) Combination of the object with the boundaries. (d) External boundaries of the original object (white) and the new object (blue) (e) Final closed object.

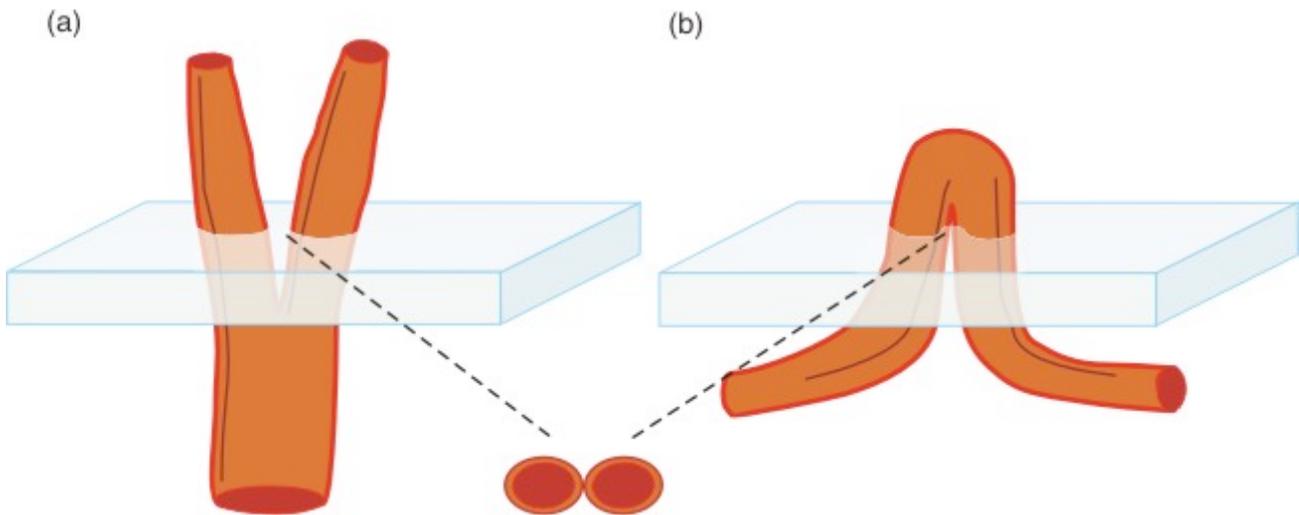


Fig. 3. Two examples of vessels that may appear as a single object when thin sectioned. (a) A branching vessel, (b) a tortuous vessel crossing the plane of sectioning.

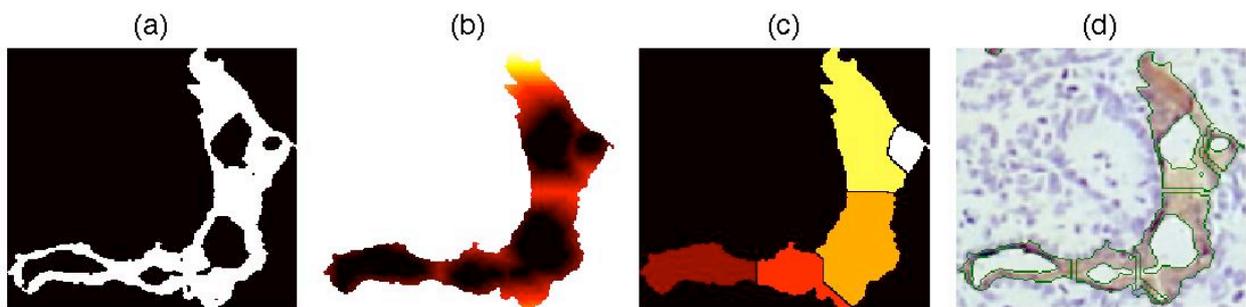


Fig. 4. Morphological task of splitting objects. Any object with more than one inner area or “hole” or lumen is considered as a group of vessels and it will be split into several new objects: (a) One object with more than one hole. (b) A distance map (the minimum Euclidean distance from every pixel of the object to its nearest hole, brightness denotes distance from the holes) corresponding to the larger holes of (a). (c) Watershed segmentation of the distance map. The segmentation is analogous to the catchment basins of a topological map (see text). For the example, 5 regions were generated. (d) The original object split into 5 new objects, the segmentation boundaries (green) are overlaid on the original image.

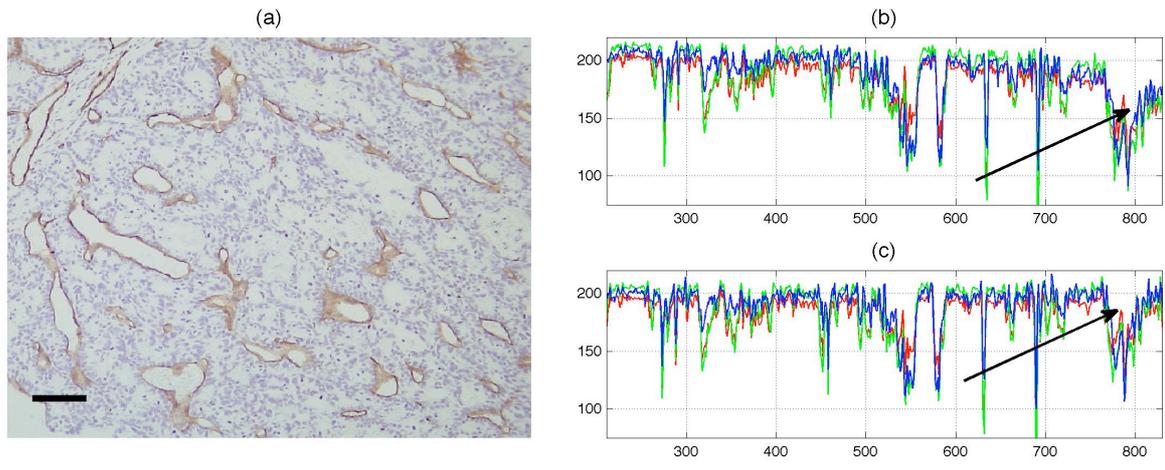


Fig. 5. (a) One representative immunostained image. SW1222 carcinoma cancer cell nuclei appear blue-purple, endothelial cells in brown-beige and the background in white-grey. Bar = 80 μ m. (b) The intensity profile of a line inside the image; red, green and blue lines correspond to the values of each of the colour channels. Notice the general decrease in intensity of the three channels on the right side (black arrow). These inhomogeneities are due to optical artefacts. (c) After shading correction, the intensities of the same profile remain constant.

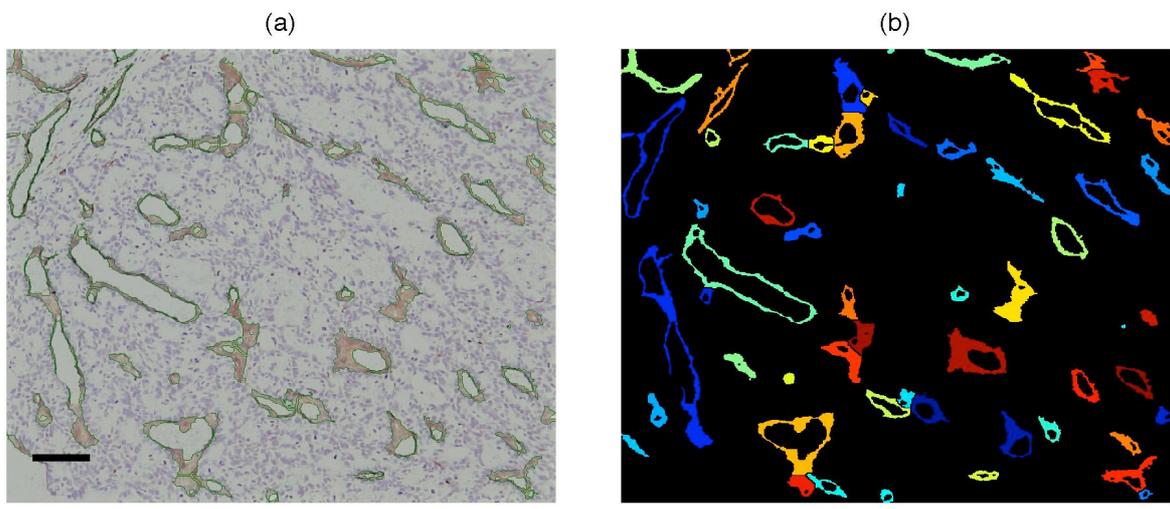


Fig. 6. Vessel segmentation with the proposed algorithm. (a) Immunostained images with object boundaries overlaid with green lines. Bar = 80 μ m. (b) Segmented objects corresponding to the vessels on the left column, each object was labelled with a different colour for visual discrimination.

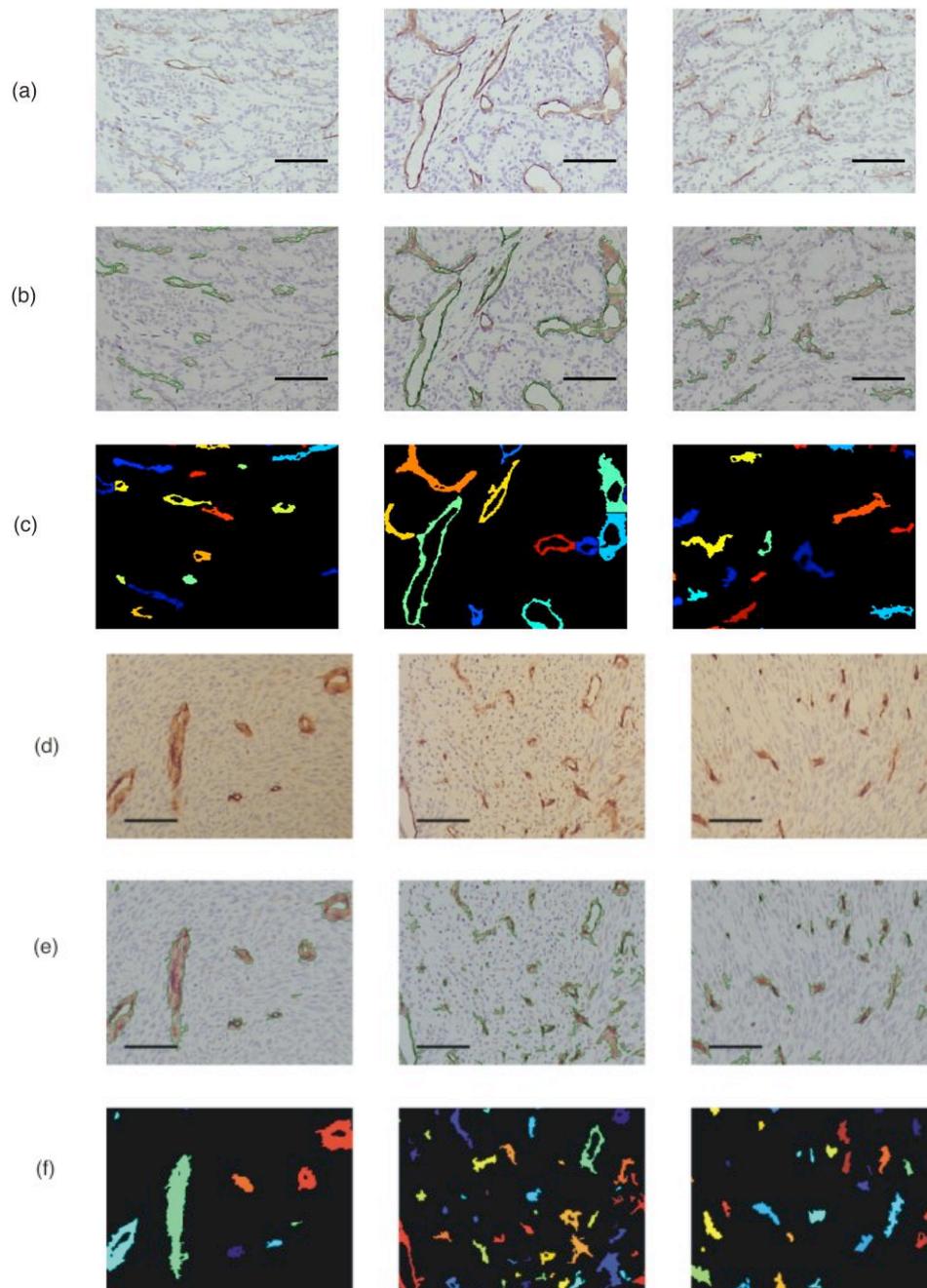


Fig. 7 Six regions of interest of images, one per tumour, with different colour shades and quality due to variations in the immunostaining and acquisition process. Bar = 80 μm . (a) Original immunostained images for the SW 1222 tumours, (b) Segmented objects overlaid with green lines, (c) Segmented objects labelled with different colours. (d) Original immunostained images for the MF tumours, (e) Segmented objects overlaid with green lines, (f) Segmented objects labelled with different colours.

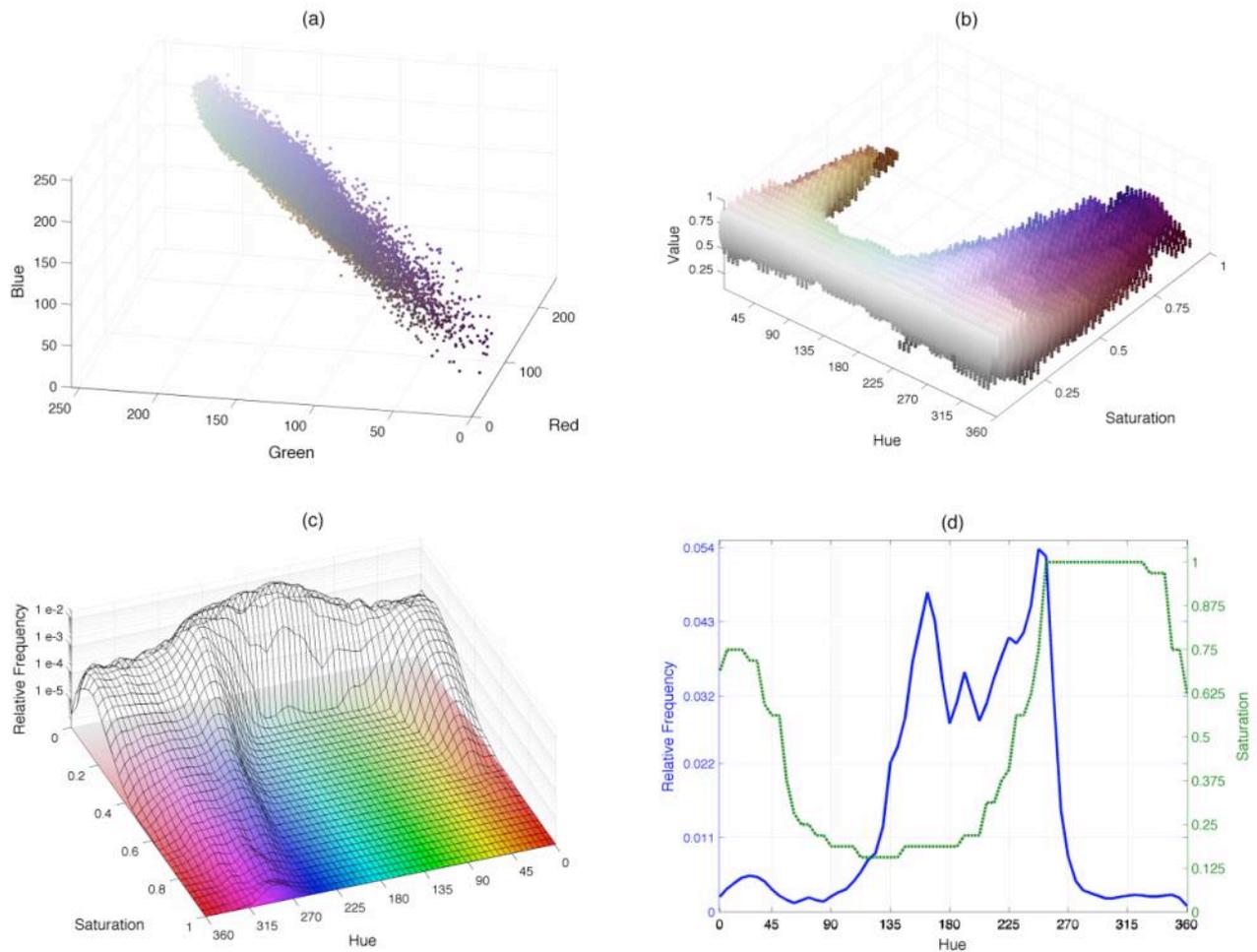


Fig. 8. Graphical depictions of the coloured pixels from an immuno-stained section. (a) A cloud of coloured points describes the distribution of pixels within the *RGB* domain corresponding to Fig. 7a (centre). (b) Three-dimensional histogram $m_{HSV}(h,s,v)$ for the same image. A cloud of coloured points describes the distribution of pixels within the *HSV* domain. (c) Two-dimensional histogram $m_{HS}(h,s)$ overlaid on loci of constant value, notice the logarithmic z-axis. The higher peaks at the low saturation ($s < 0.2$) correspond to the background pixels while the lower peaks in the foreground correspond to the pixels stained by immunohistochemistry. (d) One-dimensional hue histogram (solid blue line) and maximum saturation profile $p_{\max S}$ (green dotted line). While the histogram describes the relative frequency of the ranges of hue, $p_{\max S}$ indicates those hues that reach high saturations (intense colours, far from white or grey) within the images.

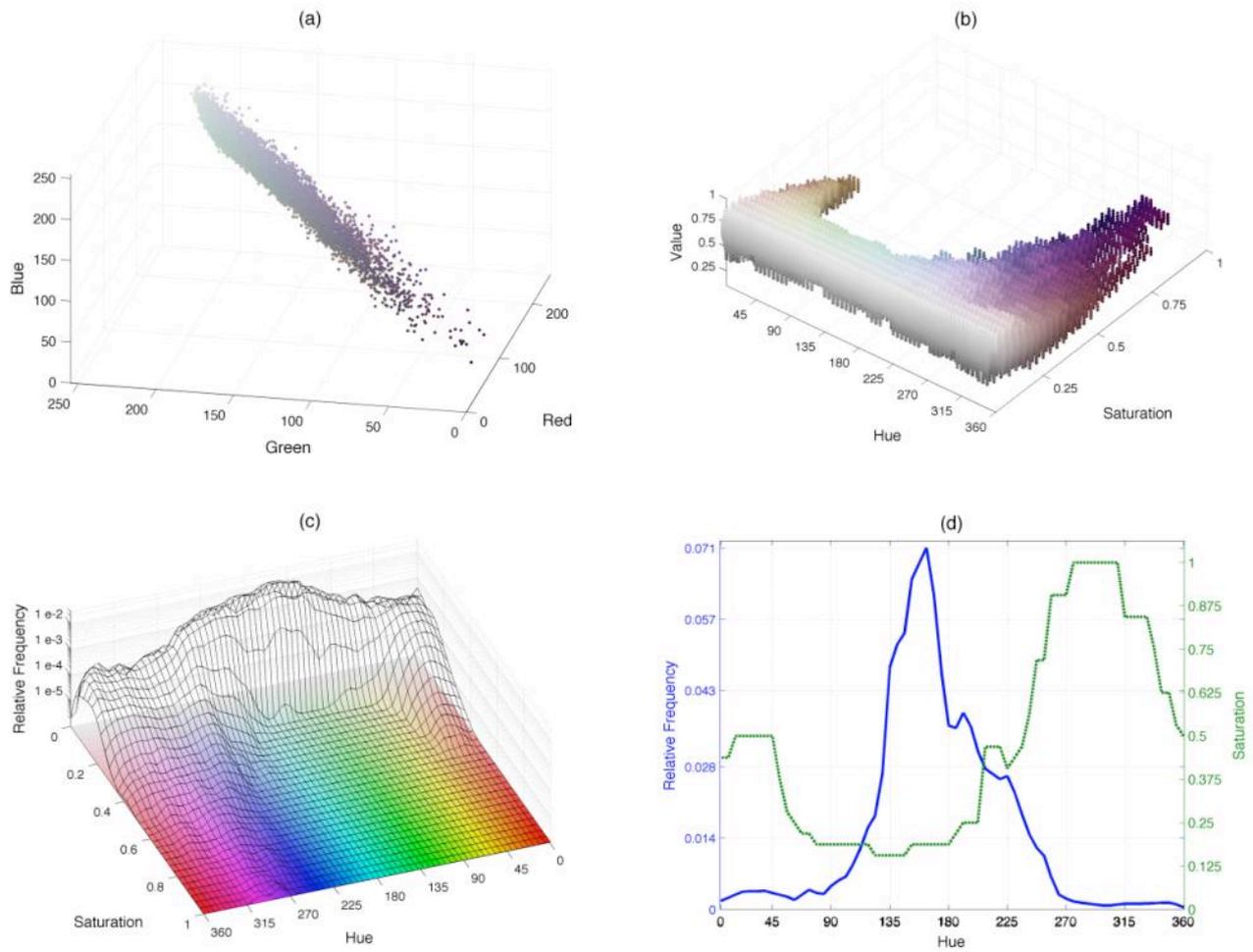


Fig. 9. (a) A cloud of coloured points describes the distribution of pixels within the RGB domain corresponding to Fig. 7a (right). (b, c, d) Three-, two- and one-dimensional histograms $m_{HSV}(h, s, v)$, $m_{HS}(h, s)$, $m_H(h)$ and maximum saturation profile $p_{\max s}$. See Fig. 8 legend for full details.

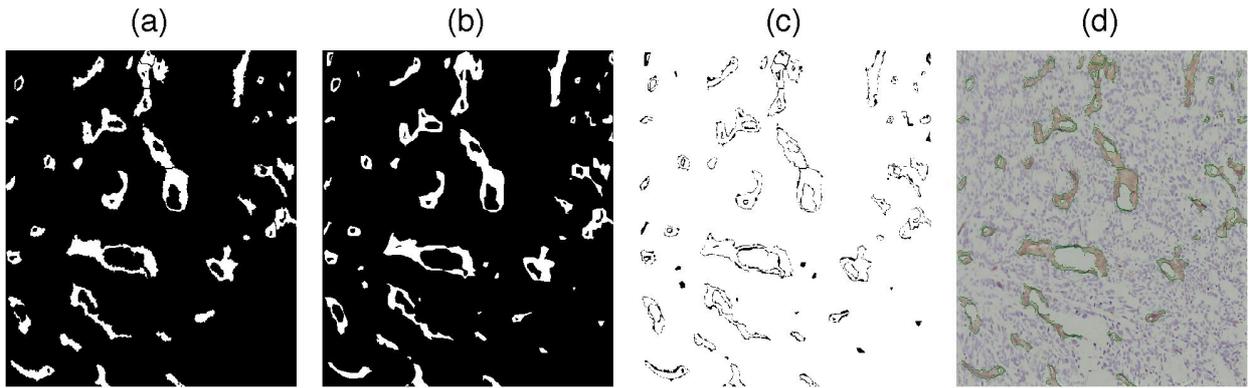


Fig. 10 Manual validation of the algorithm. (a) The automatic segmentation produced by the algorithm, vessels are white and background is black. (b) Manual segmentation produced by hand-delineation. (c) Comparison of the two segmentations; pixels that have identical classes in both segmentations appear in white (97.12%), while incorrect classifications (2.88%) appear in black. (d) Original image with the automatic classification overlaid with green lines.

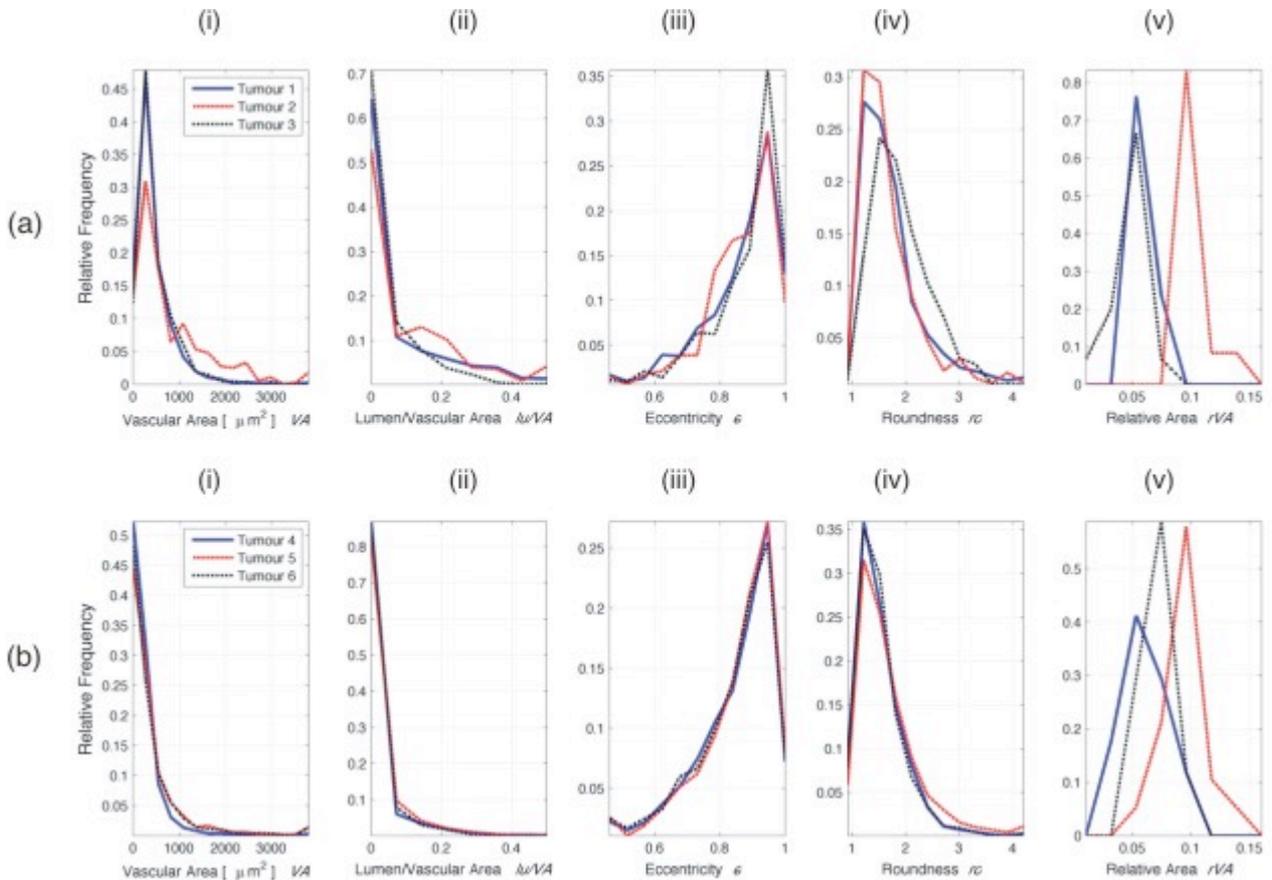


Fig. 11 Histograms for the morphometric measurements of microvessels from 3 SW1222 tumours (a) and 3 MF tumours (b). (i) vascular area VA , (ii) ratio of lumen area over vascular area lu/VA , (iii) eccentricity e , (iv) roundness ro , and (v) relative area rVA . Tumours 1,4 (Fig. 7a,d left) are represented by a solid blue line, tumours 2, 5 (Fig. 7a,d centre) by a dashed red line and tumours 3, 6 (Fig. 7a,d right) by a dash-dot black line. It should be noted that there is some variation in the measurements between the different tumours (see text for details). Measurements in (i-iv) are acquired per object from 12-19 images per tumour, whereas a single value per image is acquired for rVA in (v).

Table 1. Statistical descriptors (Mean, Standard Error of the Mean (sem), Median, Minimum and Maximum) for five measurements (Vascular Area, eccentricity, Lumen/Vascular area, roundness and relative vascular area) obtained from 6 tumours. *p*-values of an unpaired t-test were calculated between the SW 1222 (tumours 1-3) and MF (tumours 4-6) lines. Values of *p*<0.05 are highlighted in bold.

	Metric	Vascular Area [μm^2]	Lumen/Vessel Area	Eccentricity	Roundness	Relative vascular Area
Tumour 1 SW 1222	mean	441.6	0.071	0.858	1.77	0.059
	sem	9.3	0.002	0.002	0.01	0.001
	median	310.0	0	0.897	1.58	0.059
	minimum	25.9	0	0.167	0.90	0.048
	maximum	7976.9	0.651	0.997	5.97	0.070
Tumour 2 SW 1222	mean	833.0	0.098	0.856	1.71	0.102
	sem	20.4	0.002	0.002	0.01	0.004
	median	477.4	0.022	0.889	1.52	0.103
	minimum	25.9	0	0.207	0.93	0.085
	maximum	9153.2	0.705	0.996	7.08	0.144
Tumour 3 SW 1222	mean	472.9	0.041	0.875	1.93	0.046
	sem	13.5	0.002	0.003	0.01	0.004
	median	325.6	0	0.919	1.79	0.051
	minimum	25.9	0	0.164	0.93	0.001
	maximum	8893.7	0.683	0.997	5.54	0.066
Tumour 4 MF	mean	266.8	0.019	0.839	1.53	0.061
	sem	10.4	0.001	0.002	0.01	0.004
	median	127.1	0	0.878	1.41	0.057
	minimum	22.4	0	0.109	0.89	0.034
	maximum	22067.9	0.678	0.997	5.85	0.092
Tumour 5 MF	mean	431.3	0.023	0.844	1.67	0.095
	sem	14.6	0.001	0.002	0.01	0.004
	median	161.7	0	0.882	1.49	0.093
	minimum	22.4	0	0.180	0.88	0.062
	maximum	17810.8	0.573	0.998	8.01	0.146
Tumour 6 MF	mean	420.5	0.017	0.835	1.52	0.068
	sem	19.3	0.002	0.001	0.01	0.002
	median	149.1	0	0.875	1.41	0.068
	minimum	22.4	0	0.206	0.89	0.050
	maximum	16032.6	0.502	0.998	5.76	0.099
<i>t</i> -test SW vs MF	<i>p</i> -value	0.199	0.039	0.019	0.046	0.791