

Correspondence

Robust Quantification of *In Vitro* Angiogenesis Through Image Analysis

Antti Niemistö*, Valerie Dunmire, Olli Yli-Harja, Wei Zhang, and Ilya Shmulevich

Abstract—An automated image analysis method for quantification of *in vitro* angiogenesis is presented. The method is designed for *in vitro* angiogenesis assays that are based on co-culturing endothelial cells with fibroblasts. Such assays are used in many current studies in which anti-angiogenic agents for the treatment of cancer are being sought. This search requires accurate quantification of the stimulatory and inhibitory effects of the different agents. The quantification method gives lengths and sizes of the tubule complexes as well as the numbers of junctions in each of them. The method is tested with a set of test images obtained with a commercially available *in vitro* angiogenesis assay. The results correctly indicate the inhibitory effect of suramin and the stimulatory effect of vascular endothelial growth factor. Moreover, the image analysis method is shown to be robust against variations in illumination. We have implemented a software package that utilizes the methods. The software as well as a set of test images are available at <http://www.cs.tut.fi/sgn/csb/angioquant/>.

Index Terms—Angiogenesis, image analysis, quantification, segmentation.

I. INTRODUCTION

Angiogenesis, the formation of new capillary blood vessels, has become an important area of scientific research. It can occur abnormally around malignant tumors, and anti-angiogenic agents for the treatment of cancer are currently being sought [1], [2]. This search requires the quantification of the stimulatory and inhibitory effects of the different agents. In order for the quantification to be accurate, objective, and consistent, it should be obtained with automated image analysis. In fact, since quantification of angiogenesis requires the measurement of tubule lengths, it is never done manually.

Numerous *in vivo* angiogenesis assays have been developed [3]. However, *in vivo* assays often use nonhuman tissue and are time consuming, expensive, and difficult to interpret. *In vitro* angiogenesis assays are usually more rapid, less expensive, and easier to interpret; they have, therefore, gained popularity [4]. They operate on the principle that endothelial cells form tubule-like structures when cultured on a supportive matrix. Light microscopy is commonly used to obtain the images to be analyzed.

In this paper, a quantitative image analysis method is developed for *in vitro* angiogenesis assays that are based on co-culturing endothelial cells with fibroblasts in wells arranged in a plate [5]. Each well can be

x_4	x_3	x_2
x_5	p	x_1
x_6	x_7	x_8

Fig. 1. The 3×3 neighborhood of a point p .

used to test a condition. Our approach is to first detect the tubule complexes by binarizing the angiogenesis images. The sizes of the tubule complexes can be readily measured from the binary image. After binarization, the tubule complexes are thinned, and the resulting skeletons are used to measure the total lengths of the tubule complexes as well as to detect junctions, i.e., branching points, of the skeletons. The image processing methods are described in detail in Section II. Section III describes the quantification results obtained with a set of angiogenesis images obtained with the TCS Cellworks Angiokit (Buckingham, UK), a commercially available *in vitro* angiogenesis assay. Finally, some concluding remarks are given in Section IV.

II. METHODS

A. Definitions

The following definitions will be adopted in the discussion below. We assume the labeling of points (pixels) depicted in Fig. 1. The points x_1, x_2, \dots, x_8 are the eight-neighbors of p , and are said to be eight-adjacent to p . Similarly, the points x_1, x_3, x_5, x_7 are the four-neighbors of p , and are said to be four-adjacent to p .

A sequence of points y_1, y_2, \dots, y_n is an eight-path (four-path), if y_{i+1} is an eight-neighbor (four-neighbor) of y_i for all $i < n - 1$. A subset Q of an image P is eight-connected (four-connected) if for every pair q_i, q_j of points in Q there is an eight-path (four-path) from q_i to q_j consisting of points in Q . If P is a binary image and all points in Q have the same value (zero or one), Q is said to be an eight-connected (four-connected) component of P . An algorithm for labeling connected components in a binary image can be found in [6].

An end point is a point whose eight-neighbors form at most one eight-connected component. A junction point is a point whose eight-neighbors form more than two four-connected components.

B. Segmentation

The objective of segmentation is to distinguish objects of interest (foreground) from the background. In our case, the aim is to reduce a color or grayscale image into a binary image that accurately represents the objects of interest in the image. The objects of interest are naturally the tubule complexes, i.e., the networks of connected tubules that are darker than the background. In the binary image, the background will be represented by zeros and the foreground by ones.

As an example, a test image that will be used throughout this section is presented in Fig. 2(a). This is just a small portion of an image representing a complete well. The quality of the eventual results is dependent on the accuracy of the binary representation of tubule complexes, and segmentation is, therefore, the most crucial step of the overall quantification procedure.

Global thresholding is a simple approach to segmentation. It creates a binary image based on the intensity values of the grayscale image. Pixels with intensity values falling within a given range are marked as foreground (ones), while all other pixels get marked as background

Manuscript received May 12, 2004; revised August 31, 2004. This work was supported in part by the Academy of Finland. *Asterisk indicates corresponding author.*

*A. Niemistö is with the Institute of Signal Processing, Tampere University of Technology, 33101 Tampere, Finland, and also with the Department of Pathology, The University of Texas M. D. Anderson Cancer Center, Houston, TX 77030 USA (e-mail: antti.niemisto@tut.fi).

V. Dunmire, W. Zhang, and I. Shmulevich are with the Department of Pathology, the University of Texas M. D. Anderson Cancer Center, Houston, TX 77030 USA.

O. Yli-Harja is with the Institute of Signal Processing, Tampere University of Technology, 33101 Tampere, Finland.

Digital Object Identifier 10.1109/TMI.2004.837339

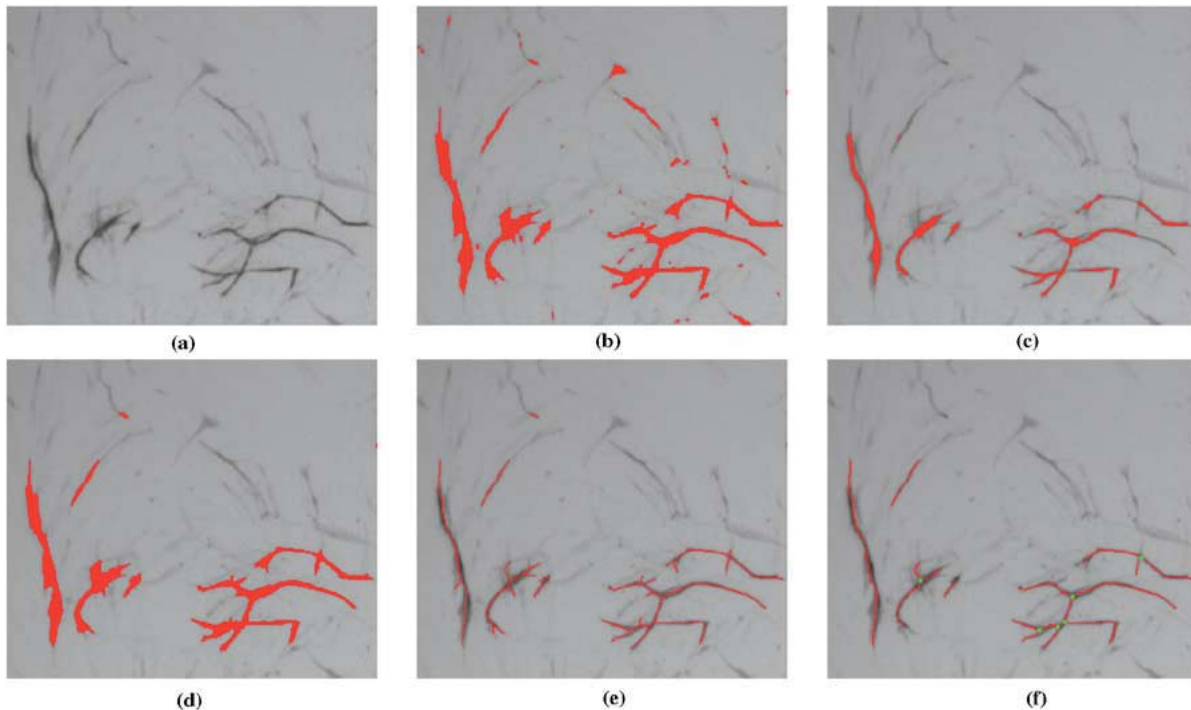


Fig. 2. A test image at different image processing stages. (a) Original. (b) High thresholding. (c) Low thresholding. (d) Segmentation. (e) Skeleton. (f) Pruned skeleton.

(zeros). However, global thresholding does not typically provide good results. This is in part because illumination is usually uneven across an image obtained with a microscope. For example, in many illumination arrangements the intensity is highest near the center of the image, and gets lower when one moves further away from the center. The corners, thus, become darker than the center, and clearly the same thresholds cannot be used in different parts of the image.

An obvious solution in the case of uneven illumination is to design better illumination. However, this is not always feasible. Therefore, we correct illumination digitally by fitting, in the least squares sense, a second degree polynomial surface on the image. This polynomial gives us an estimate of the intensity of the background in different image regions and allows us to make corrections if the intensity varies between these regions.

In addition to actual tubules, angiogenesis images obtained using assays based on a fibroblast matrix generally contain fibroblasts. In Fig. 2(a), fibroblasts can be seen in the upper part of the image. They have a similar appearance to tubules, but are in general lighter (have higher intensity). However, some parts of tubules can be as light as fibroblasts and, therefore, segmentation cannot be achieved by simple thresholding even after illumination is corrected. Instead, we need to look at each complex in the image, and decide if it is a fibroblast or a tubule complex. This can be achieved with a modification of the hysteresis thresholding approach [7]. The implementation uses connected component analysis in the same way as suggested in [8]. If the input image is a color image, it is converted into a grayscale image by eliminating the hue and saturation information while retaining the luminance.

We perform thresholding in such a way that pixel values below the threshold are taken as objects of interest. First, we choose a high threshold T_H to recognize both tubule and fibroblast complexes [see Fig. 2(b)]. Second, we choose a low threshold $T_L < T_H$ with which only the darker parts of tubule complexes are recognized as foreground [see Fig. 2(c)]. Let us call the obtained images I_H and I_L , respectively. Each eight-connected component in I_H is a candidate for

a tubule complex. A candidate complex is a tubule complex if at least a fraction t of pixels belonging to the candidate complex are recognized as foreground pixels in I_L . Thus, the resulting binary image represents the tubule complexes accurately if

- 1) all pixels belonging to a tubule complex are recognized as foreground pixels in I_H ;
- 2) at least a fraction t of pixels belonging to each tubule complex in I_H are recognized as foreground pixels in I_L ; and
- 3) less than a fraction t of pixels belonging to each fibroblast complex in I_H are recognized as foreground pixels in I_L .

The choice of t depends in general on the thresholds, i.e., the lower T_L is with respect to T_H , the smaller t should be.

Because the tubule and fibroblast complexes are typically well defined, the classic method by Otsu [9] can be used to determine the high threshold T_H . The method is nonparametric and unsupervised in the sense that no *a priori* knowledge about the image is needed. It maximizes the separability of the resultant classes (foreground and background) in gray levels, and is insensitive to variability of the overall intensity of the images. However, some parts of tubule complexes often have a high intensity and get falsely recognized as background. The resulting small holes in the recognized tubule complexes are filled using a morphological closing by a 3×3 square structuring element [6]. Note also that it is enough to make sure that all tubule complexes are recognized, i.e., it does not matter if fibroblast complexes are not recognized.

The choice of the low threshold T_L can be made based on the histogram of the original grayscale image. Because after correcting illumination the background of the image is homogeneous in angiogenesis images, most of the unique intensity values represent pixels that belong to a tubule or fibroblast complex. Thus, it can be safely assumed that those intensity values that are below the midpoint of the histogram range represent tubule pixels, and T_L is selected accordingly. Most likely some intensity values higher than T_L also represent tubule pixels, but it does not matter as long as some pixels in each tubule complex get recognized as foreground pixels, and t can be selected such that our segmentation algorithm recognizes all tubule complexes.

With the test image of Fig. 2(a), Otsu's algorithm gives $T_H = 146$, and the resulting binary image is shown superimposed in red color on the original image in Fig. 2(b). Similarly, the low threshold is $T_L = 118$, and Fig. 2(c) shows the resulting binary image. It can be seen that the low threshold results in disconnections in the recognized tubule complexes. If the threshold is made higher, the tubule complexes become connected, but at the same time some pixels representing fibroblasts get recognized. In general, there is no single threshold that can recognize all tubule complex pixels and only them, pointing out the need for a hysteresis thresholding approach.

The final segmentation result obtained with $t = 0.005$ is presented in Fig. 2(d). It can be seen that the tubule complexes are recognized accurately, while the lighter fibroblast complexes are marked as background. In this case, the same result could have been obtained by applying hysteresis thresholding as suggested by Canny [7], i.e., by simply requiring that for each candidate complex there must be at least one corresponding foreground pixel in I_L . However, a single foreground pixel in I_L might result from noise or other artefacts. In our method, the number of pixels that is required depends on the size of the candidate complex: the larger the complex, the more pixels are required. This makes our method more robust.

C. Thinning

In order to be able to measure the lengths of tubules, they must be reduced to arcs that are one pixel in width. The reduction process is generally referred to as thinning, and the result of thinning an object is the skeleton of the object. Thinning is a fundamental preprocessing step in many image processing and pattern recognition algorithms. A wide range of different thinning algorithms can be found in the literature. They produce slightly different skeletons, and are often designed with a particular application in mind. A comprehensive review of thinning methods can be found in [10]. Some of the more recent studies, such as the ones suggested in [11] and [12], have concentrated on thinning in the presence of different types of noise.

In our case the requirements for thinning are: 1) the resulting skeleton must be topologically equivalent to the object; 2) the resulting skeleton must be one pixel thick; and 3) the algorithm must preserve the connectivity of the object. Insensitivity to boundary noise, i.e., to small protrusions and indentations, is not crucial. Instead, spurious false arcs can be pruned after thinning (to be discussed later).

A thinning algorithm suggested by Guo and Hall [13] is suitable for our application. Specifically, we use Algorithm 1 from that paper. It is a parallel two-subiteration thinning algorithm that is applied iteratively until further thinning is not possible. It examines in each iteration the 3×3 neighborhood of each pixel to determine if that pixel should be deleted. The algorithm can be implemented efficiently using two look-up tables, one for each subiteration. The result of applying the thinning algorithm to the segmentation result of Fig. 2(d) can be seen in Fig. 2(e). Alternatively, the algorithm suggested by Shih and Wong [11] could be used, but the pruning phase would still be required.

D. Pruning

The skeletons that are obtained with the thinning algorithm generally contain spurious arcs, i.e., short arcs that are not topologically equivalent to the corresponding tubule complex. As can be seen in Fig. 2(e), this is also the case with our test image. The spurious arcs result from small protrusions and indentations in the binary image obtained after segmentation, and should be removed.

The detection and removal of spurious arcs is based on their size: they are short arcs between an end point and a junction point. If a skeleton is thought of as a network of roads, an end point is a dead

TABLE I
QUANTIFICATION OF THE IMAGE OF FIG. 2

complex	1	2	3	4	5	6
length	15.7	41.7	82.9	110.7	151.2	298.7
size	84	94	537	378	925	1139
junctions	0	0	1	1	0	4

end and a junction point is an intersection. The following algorithm removes from a skeleton image I those spurious arcs whose sizes are less than or equal to the pruning size N_P pixels.

- 1) Delete from I all pixels in the 3×3 neighborhood of each junction point. Store the result in J .
- 2) Delete from J all eight-connected components that do not contain an end point of an eight-connected component in I . Store the result in J .
- 3) Delete from I all pixels that correspond to an eight-connected component in J with size smaller than or equal to N_P . Store the result in I .
- 4) Delete from I all end points in the 3×3 neighborhood of a junction point. Store the result in I .

The fourth step is needed to remove the remaining spurious arcs of size one pixel. In some cases, the removal of spurious arcs may create new spurious arcs. To remove all of these, the algorithm should be applied iteratively until there is no change in the image between two consecutive iterations.

The result of applying the algorithm with $N_P = 10$ to the image in Fig. 2(e) is shown in Fig. 2(f). The image also shows the junction points as green squares. It can be seen that the skeletons are topologically equivalent to the respective tubule complexes. Also, the small skeleton near the top left corner of the image has been removed. This is desired, since skeletons that are very small (in this case less than 10 pixels) are likely to result from segmentation errors and should, thus, be removed.

E. Data Generation

The data generation phase gives the lengths and sizes of each tubule complex as well as the numbers of junctions in each complex. The lengths and sizes are measured in pixels.

The number of pixels in a tubule complex is a coarse estimate of its length. However, it should be taken into account that the distance between two diagonally connected pixels is greater than the distance between two four-connected pixels. To be exact, if the number of pixels in the skeleton of a tubule complex is N and the number of diagonally connected pairs of pixels is N_d , the length of the skeleton is given by $L = N + (\sqrt{2} - 1)N_d - 1$.

The size of a tubule complex is given by the total number of pixels representing the complex. The size is calculated from the segmentation result and, thus, takes into account the thickness of the tubules. The number of junctions in each tubule complex can be obtained readily using the definition of a junction point given in Section II-A. The lengths, sizes, and numbers of junctions for each tubule complex in the image of Fig. 2 are shown in Table I.

III. RESULTS

We have developed a software package called *AngioQuant* that utilizes the image processing methods described in the previous section. *AngioQuant* is a fully automated tool that can be used to analyze a set of angiogenesis images as a batch without any user intervention. However, the software also allows the user to set some of the parameters interactively. *AngioQuant* can be downloaded at <http://www.cs.tut.fi/sgn/csb/angioquant/>. A set of 24 test images is

TABLE II
QUANTIFICATION OF THE TEST IMAGES

treatment	complexes	length		size		junctions	
		total	mean	total	mean	total	mean
none	190	15431.4	81.2	61851	325.5	118	0.62
none	208	16347.1	78.6	63633	305.9	130	0.63
suramin	121	3769.1	31.2	17073	141.1	5	0.041
suramin	92	3109.9	33.8	14489	157.5	9	0.098
VEGF	164	18550.2	113.1	80249	489.3	171	1.04
VEGF	199	26959.5	135.5	117909	592.5	329	1.65

also available at the same location. The TCS Cellworks Angiokit, a commercially available *in vitro* angiogenesis assay that utilizes a fibroblast matrix seeded with endothelial cells, was used. The images were then obtained with a light microscope.

At the imaging stage it is important to make sure that the quality of the images is sufficient for the AngioQuant software. First, it is typically possible to get the whole well in good focus to avoid blurring. Also, the imaging should be done soon after the angiogenesis experiment is finished in the laboratory. This makes sure that the staining is strong enough to provide a sufficient contrast between the tubules and the rest of the image.

Seven of our test images are of wells that were treated with 20 μ M suramin, an anti-angiogenic drug [14]. Visual inspection of the images reveals that the tubule complexes are short and have few junctions. Nine of the images are of wells that were treated with 2 ng/ml vascular endothelial growth factor (VEGF). The effects of the growth factor can be seen in the images, as the tubule complexes are large and have many junctions. Finally, eight of the images are of wells with no treatment.

A. Quantification

We applied AngioQuant to each of the test images. Due to the limited space, Table II presents the results for only six of the test images. The complete results can be found on the supplementary web site. The lengths and sizes of tubule complexes are measured in pixels, since this is sufficient for the purpose of comparing the results obtained with different treatments. Since the number of tubule complexes varies significantly between the images, we compare the mean lengths of tubule complexes and the mean numbers of junctions per tubule complex. The sizes of the tubule complexes are omitted from the analysis, because they provide little additional information.

It is seen that the quantification is similar between duplicate (same treatment) wells. Moreover, we observe the expected results when the wells with suramin and VEGF treatments are compared against the wells with no treatment. In the wells treated with suramin, the number of tubule complexes is smaller than in the wells with no treatment. Further, the tubule complexes are much shorter and the numbers of junctions are much smaller than in the wells with no treatment. In fact, the quantification shows that the number of junctions per tubule complex is close to zero. For the wells treated with VEGF, the quantification shows increased mean tubule complex lengths and junction counts as compared with wells with no treatment.

Fig. 3 shows each of the 24 wells in a two-dimensional feature space. The horizontal axis is the mean length of the tubule complexes, and the vertical axis is the mean number of junctions per tubule complex. The downward triangles represent the wells with no treatment, the squares represent the wells with suramin treatment, and the upward triangles represent the wells with VEGF treatment. The three classes are mainly well separated in the feature space.

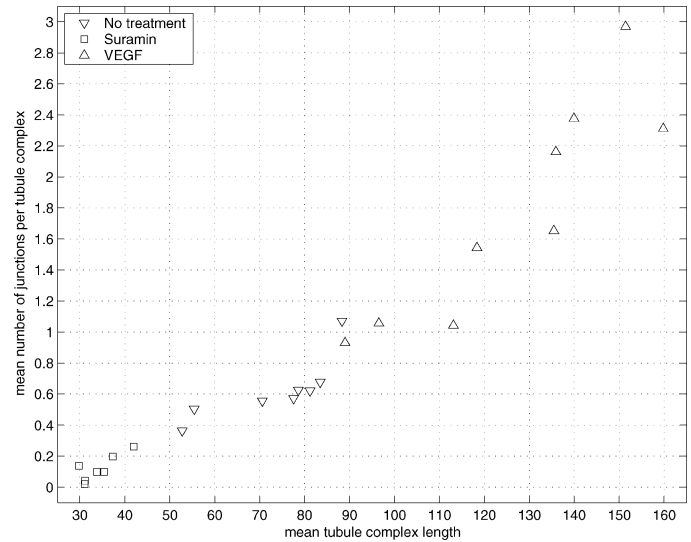


Fig. 3. The 24 wells in a two-dimensional feature space.

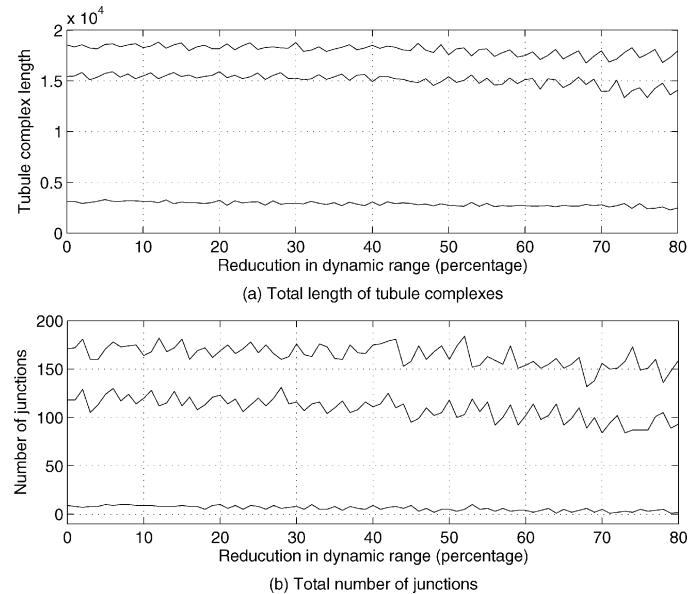


Fig. 4. The changes in the quantification caused by reducing the dynamic range of the original image. In both figures, the topmost curve corresponds to the first VEGF image from Table II, the middle curve corresponds to the first no-treatment image, and the lowest curve corresponds to the first suramin image.

B. Robustness

It is important that the image analysis method is robust against variations in the images. This is a requirement for the AngioQuant software to be applicable for a wide range of different *in vitro* angiogenesis assay images. In particular, illumination can vary significantly in different imaging systems and the methods must be robust against these changes. The segmentation method is the most critical in this respect, since as long as the binary image obtained with the segmentation method represents the tubule complexes accurately, the subsequent image processing methods will give the desired results. Our approach in testing robustness is to simulate changes in illumination digitally, and to examine the changes in the quantification results.

The overall intensity of the illumination of each image was varied digitally by adding an integer constant $k \in [-255, 255]$ to all the pixel values of each image. Pixels whose values became negative were replaced with zeros, and pixels whose values became greater than 255

were replaced by 255. This is called clipping. Clearly, clipping destroys some of the information contained in the original image and in the extreme case, i.e., with very small or large values of k , clipping destroys all the information. It is, thus, not fair to require robustness against heavy clipping. However, as long as there is no clipping, the quantification turns out to be the same as for the original image.

Illumination also affects the dynamic range, i.e., the range of intensity values, of the obtained image. This can be simulated digitally by mapping the intensity values in the original image to a smaller range. The result is an image with a histogram that has approximately the same shape as that of the original image, only narrower.

Fig. 4 shows how the quantification of three of the test images is affected by reducing the dynamic range of the original image. Since reducing the dynamic range destroys some of the information contained in the original image, the quantification is affected. However, even when the dynamic range is reduced by up to 80%, the quantification of each image remains close to the original quantification. Most importantly, the relative quantification is not affected by a reduction in the dynamic range, i.e., when the quantifications obtained with different treatments are compared, the conclusions are not affected by a reasonable reduction in the dynamic range.

IV. CONCLUSION

An automated image analysis method was developed for quantifying the stimulatory and inhibitory effects of the different agents in *in vitro* angiogenesis assays. The method is designed for assays that are based on co-culturing endothelial cells with fibroblasts. A software package that employs the image analysis method was implemented and is publicly available. It can be used to measure the lengths and sizes of tubule complexes as well as the numbers of junctions in each tubule complex. Our experiments suggest that the image processing methods employed by the software are robust, which makes the software applicable to a wide range of microscopic images taken of *in vitro* angiogenesis assays. The quantification results obtained for a set of test images correctly

indicate the inhibitory effect of suramin and the stimulatory effect of VEGF. The software enables a move from qualitative studies to quantitative studies.

REFERENCES

- [1] J. Folkman, "Role of angiogenesis in tumor growth and metastasis," *Semin. Oncol.*, vol. 29, no. 6, pp. 15–18, Dec. 2002.
- [2] R. Kerbel and J. Folkman, "Clinical translation of angiogenesis inhibitors," *Nat. Rev. Cancer*, vol. 2, no. 10, pp. 727–739, Oct. 2002.
- [3] R. K. Jain, K. Schlenger, M. Höckel, and F. Yuan, "Quantitative angiogenesis assays: Progress and problems," *Nat. Med.*, vol. 3, no. 11, pp. 1203–1208, Nov. 1997.
- [4] D. Donovan, N. J. Brown, E. T. Bishop, and C. E. Lewis, "Comparison of three *in vitro* human 'angiogenesis' assays with capillaries formed *in vivo*," *Angiogenesis*, vol. 4, no. 2, pp. 113–121, 2001.
- [5] E. T. Bishop, G. T. Bell, S. Bloor, I. J. Broom, N. F. K. Hendry, and D. N. Wheatley, "An *in vitro* model of angiogenesis: Basic features," *Angiogenesis*, vol. 3, no. 4, pp. 335–344, 1999.
- [6] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Upper Saddle River, NJ: Prentice-Hall, 2002.
- [7] J. F. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.
- [8] G. Liu and R. M. Haralick, "Two practical issues in Canny's edge detector implementation," in *Proc. 15th Int. Conf. Pattern Recognition*, 2000, pp. 676–678.
- [9] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.
- [10] L. Lam, S.-W. Lee, and C. Y. Suen, "Thinning methodologies—a comprehensive survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 9, pp. 869–885, Sep. 1992.
- [11] F. Y. Shih and W.-T. Wong, "Fully parallel thinning with tolerance to boundary noise," *Pattern Recogn.*, vol. 27, no. 12, pp. 1677–1695, Dec. 1994.
- [12] Y.-S. Chen and Y.-T. Yu, "Thinning approach for noisy digital patterns," *Pattern Recogn.*, vol. 29, no. 11, pp. 1847–1862, Nov. 1996.
- [13] Z. Guo and R. W. Hall, "Parallel thinning with two-subiteration algorithms," *Comm. ACM*, vol. 32, no. 3, pp. 359–373, Mar. 1989.
- [14] A. Gagliardi, H. Hadd, and D. C. Collins, "Inhibition of angiogenesis by suramin," *Cancer Res.*, vol. 52, no. 18, pp. 5073–5075, Sept. 1992.