

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/229327296>

A fully automated approach to prostate biopsy segmentation based on level-set and mean filtering

Article in *Journal of Pathology Informatics* · January 2011

DOI: 10.4103/2153-3539.92032 · Source: PubMed

CITATIONS

13

READS

73

6 authors, including:



John Galeotti

Carnegie Mellon University

76 PUBLICATIONS 406 CITATIONS

[SEE PROFILE](#)



Marcial Garcia-Rojo

Hospital Universitario Puerta del Mar, Cádiz

196 PUBLICATIONS 2,287 CITATIONS

[SEE PROFILE](#)



Oscar Deniz

University of Castilla-La Mancha

202 PUBLICATIONS 3,012 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Aqualitas-retos Identificación y Clasificación Automática de Diatomeas [View project](#)



Digital pathology [View project](#)

Symposium - Original Research

A fully automated approach to prostate biopsy segmentation based on level-set and mean filtering

Juan Vidal, Gloria Bueno, John Galeotti¹, Marcial García-Rojo², Fernanda Relea², and Oscar Déniz

VISILAB - Intelligent Systems and Computer Vision Group, University of Castilla la Mancha, Ciudad Real, Spain, ¹Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA, ²Department of Anatomic Pathology, University General Hospital of Ciudad Real, Ciudad Real, Spain

E-mail: *Gloria Bueno - gloria.bueno@uclm.es

*Corresponding author

Received: 25 October 11

Accepted: 25 October 11

Published: 19 January 12

This article may be cited as:

Vidal J, Bueno G, Galeotti J, García-Rojo M, Relea F, Déniz O. A fully automated approach to prostate biopsy segmentation based on level-set and mean filtering. *J Pathol Inform* 2011;2:5.

Available FREE in open access from: <http://www.jpathinformatics.org/text.asp?2011/2/2/5/92032>

Copyright: © 2011 Vidal J. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

With modern automated microscopes and digital cameras, pathologists no longer have to examine samples looking through microscope binoculars. Instead, the slide is digitized to an image, which can then be examined on a screen. This creates the possibility for computers to analyze the image. In this work, a fully automated approach to region of interest (ROI) segmentation in prostate biopsy images is proposed. This will allow the pathologists to focus on the most important areas of the image. The method proposed is based on level-set and mean filtering techniques for lumen centered expansion and cell density localization respectively. The novelty of the technique lies in the ability to detect complete ROIs, where a ROI is composed by the conjunction of three different structures, that is, lumen, cytoplasm, and cells, as well as regions with a high density of cells. The method is capable of dealing with full biopsies digitized at different magnifications. In this paper, results are shown with a set of 100 H and E slides, digitized at 5×, and ranging from 12 MB to 500 MB. The tests carried out show an average specificity above 99% across the board and average sensitivities of 95% and 80%, respectively, for the lumen centered expansion and cell density localization. The algorithms were also tested with images at 10× magnification (up to 1228 MB) obtaining similar results.

Key words: Histological segmentation, level set, mean filtering, prostate cancer, whole-slide imaging

Access this article online

Website:

www.jpathinformatics.org

DOI: 10.4103/2153-3539.92032

Quick Response Code:



INTRODUCTION

Prostate cancer is currently the most common cancer type among men in the United States.^[1] Screening protocols aimed at cancer prognosis include digital rectal examinations (DRE), prostate-specific antigen (PSA) measurements, and *in vivo* imaging techniques (CT, MRI, or ultrasound). However, these tests allow the pathologists to know that there is a reason to be suspicious, rather than the actual cause of the problem, so if the doctor finds any irregularity in these examinations, he usually

requires the patient to undergo a biopsy in order to make an accurate diagnosis. Early detection is essential to overcome the disease, but a third of the affected patients present a highly developed scenario when they are initially diagnosed.^[2]

Once the biopsy is performed, tissue samples are stained with some biomarker, usually Hematoxylin-Eosin (H&E), and placed onto a transparent slide. Then, doctors place that slide under a microscope and examine it looking for suspicious areas, gradually increasing the level of detail (changing the objective to obtain higher magnification)

if they find any clue of cancer. Samples are classified using the Gleason Score,^[3] which quantifies both cancer spreading grade and its aggressiveness from 1 to 5. Robotic microscopes equipped with digital cameras, along with whole slide imaging (WSI) methods, allow pathologists to watch the sample on a screen rather than using a microscope directly,^[4] and also provide the starting point for computer-based biopsy processing, providing a complete toolkit for computer-aided diagnosis (CAD).

One of the problems working with WSI is the image size, which usually takes several gigabytes depending on the magnification of the objective used in the digitization process. Despite fast improvement in image processing techniques, there are few software tools that are able to analyze prostate biopsy images in a fully automated way, in order to find ROI in those images.

In recent years, there have been several studies that focus on H&E prostate biopsy image processing. However, these systems are focused on the extraction of descriptors that may be useful for automatic tissue classification, rather than in ROI segmentation. Jafari-Khouzani *et al.*^[5] used an algorithm that calculates multiwavelet coefficients from 100× magnification images, and then computes energy and entropy from them, using those values for image classification into grades 2-5 of the Gleason Score.

Farjam *et al.*^[2,6] aim at malignancy detection in the same type of images. They first preprocess images using wavelets or simply transforming them into grayscale images. Texture features are extracted from the preprocessed images and clustered using a K-means algorithm. Doyle *et al.*^[7,8] have developed a classification system that is able to differentiate between 40× magnification images of benign epithelium, benign stroma, Gleason grade 3, and Gleason grade 4. They use first and second order statistics, as well as wavelet features, and support vector machines (SVM) for classification. Recently, they have improved their system to use multiresolution classification.^[9] Other novel classification methods use fractal dimension to perform Gleason classification.^[10]

Naik *et al.*^[11-13] have worked on the integration of high-level a-priori information (such as size and structure of the glands) with the image features computed. Using this approach, they are able to detect and accurately segment glands in H&E images using Bayesian probability and level sets. Their system is also able to differentiate images of benign tissue, Gleason grade 3, and Gleason grade 4. Xu *et al.* also propose a way to improve robustness of level set segmentation which is suitable for histopathological images.^[14]

Hafiane *et al.*^[15,16] perform prostate biopsy segmentation using a variation of fuzzy C-means that includes a spatial constraint to the traditional clustering to initially classify the image pixels into four classes (two classes for epithelial nuclei, one for lumen, and one for cytoplasm).

Then, multiphase vector-based level sets are used to refine the initial segmentation. Their main aim is the correct segmentation of nuclei, and they obtain an 85% accuracy versus manual segmentation.

There are two key differences between our method presented here and the previous ones. First, none of the systems uses complete mosaics but rather fragments of them (with the exception of Ref.^[9]), and they were tested with high magnification (up to 100×) images, which are slow to acquire and extremely large to be processed without a dedicated cluster. Our proposed algorithm works with images acquired with low magnification (5×, 10×), and thus requires much lower computational resources. Second, the previous algorithms aim at the segmentation of the different structures within the images, that is, lumen, cytoplasm, and nuclei, whereas our method, presented here, is aimed at separating ROIs from non interesting areas.

The method that will be described here is able to work with mosaics created at different magnification. Here, the method is illustrated with image digitized at 5× magnification, whose typical sizes take from 12 up to 500 MB. The aim of this system is the segmentation of ROI from these images in a way that mimics the method used by the doctors, that is, identifying at low magnification the regions with high concentration of cells or where the architectural distribution between lumen and cells is relevant. Thus, in order to fulfill the criteria used by pathologists to identify ROIs, two techniques have been implemented. The methods are called *Lumen Centered Expansion* and *Cell Density Localization*, and they are based on level-set segmentation and mean filtering respectively.

Methods and materials used are detailed in section 2. Experimental work carried out so far, as well as discussion and future work, is presented in sections 3 and 4 respectively.

MATERIALS AND METHODS

The images that we are working with were digitized using an ALIAS II motorized microscope from LifeSpan Biosciences Inc. This microscope acquires tiles with a size of 2000×2000 pixels and 24 bits per pixel (RGB). These tiles are stored in an uncompressed RAW file, with no header, and planar format (first part of the file is used to store the red channel, second one to store the green channel, and last part to store blue channel). Each tile requires 11.4 MB. The ALIAS microscope is equipped with five different objectives, whose magnifications are 2.5×, 5×, 10×, 20×, and 40×. Although the method is magnification independent, in this work the results are provided on samples digitized at 5× magnification.

The slides have been provided by the Anatomic Pathology Department of the Hospital General Universitario de Ciudad Real (HGUCR). This group of pathologists has

also specified for us the most relevant features that should be considered when analyzing prostate biopsies at these magnifications, such as minimum size of relevant regions, or their architectural distribution. H and E stained prostate biopsies have three types of well-differentiated structures of interest: lumen, cytoplasm, and cells [Figure 1]. Apart from those, crystals or necrosed areas might as well appear. Furthermore, if the biopsied tissue is not perfectly extracted and placed onto the slide, folded tissue and/or tissue cuts are likely to be present.

For pathological purposes, the most relevant structures are cells.^[17] Their density, morphology, distribution between them, and relationship with lumen and cytoplasm are the most relevant features that pathologists consider to elaborate a diagnosis. Thus, groups of cells are especially important, and they may appear either surrounding a lumen area, or packed very closely in isolated clusters.

As mentioned above, the method takes into account the pathologists criteria. The technique is able to detect complete ROIs, where a ROI is composed by the conjunction of three different structures, that is, lumen, cytoplasm, and cells. Therefore, we focus on the following: (a) Regions where the three structures of interest appear concentrically: lumen is surrounded by cytoplasm, which is surrounded by cells. It is not rare that the last two structures appear mixed. (b) Regions where several cells lie together. The concentric regions are detected using our novel *Lumen Centered Expansion* approach, while the latter are detected using another novel approach, which we term *Cell Density Localization*.

In the first case, since there are three different structures of interest, we have developed our algorithm to sequentially segment each type of structure. First of all, potential lumen areas are segmented. Then, for each of them, two different level sets are used to segment cytoplasm and cells. Finally, the outputs of all three steps are merged into one single ROI.

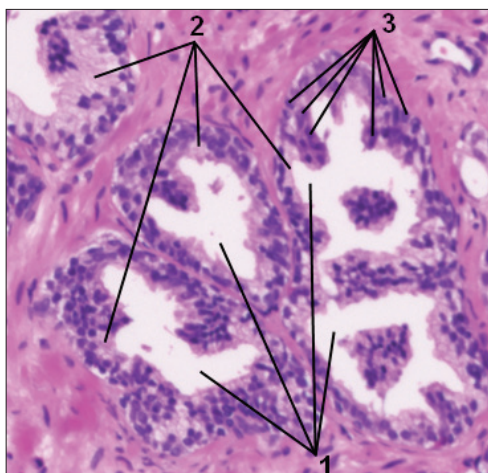


Figure 1: Examples of structures of interest: lumen (1), cytoplasm (2), and cells (3)

These techniques are described as follows.

Lumen centered expansion

Although H and E images are typically digitized using RGB format, this approach only uses the green channel of the original RGB image. As shown in Figure 2b, the green channel provides better contrast between the lumen areas and the other types of structures of interest than either the red [Figure 2a] or the blue [Figure 2c] channels. This decision is also aimed at lowering the

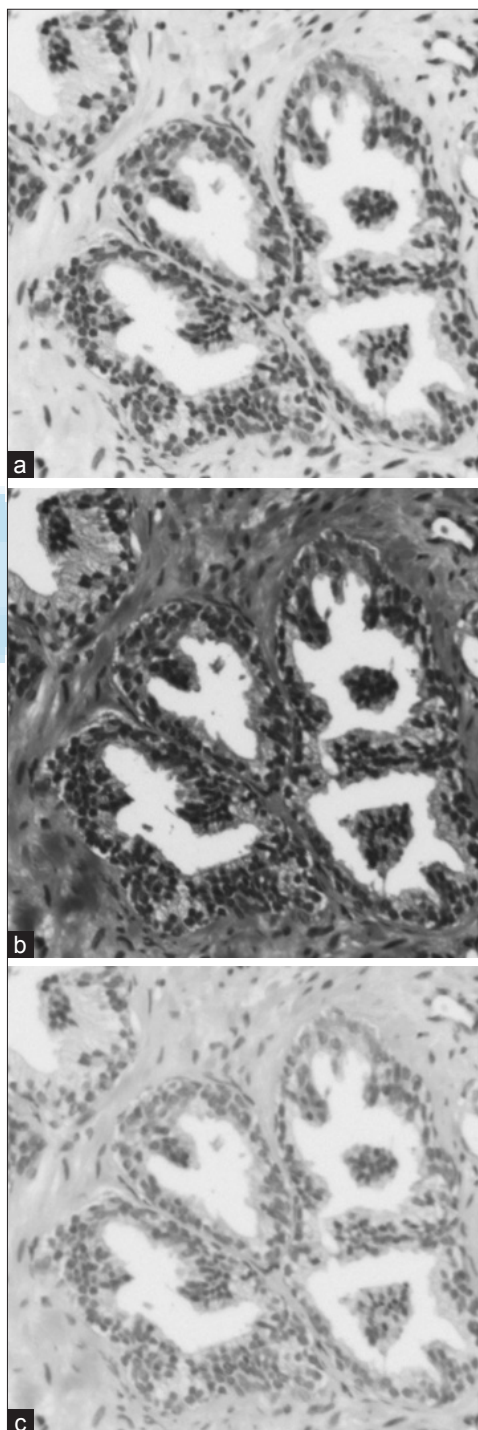


Figure 2: (a) Red; (b) Green; (c) Blue channels of Figure 1

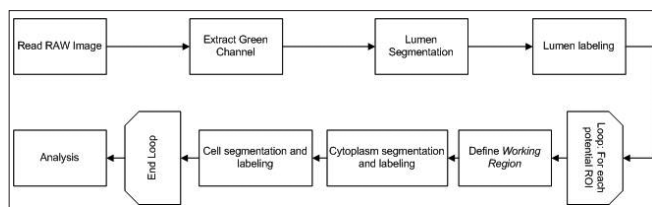


Figure 3: Flowchart of the lumen centered expansion algorithm

memory requirements of the processing, which may be huge if the input image is large. A flow chart of the complete algorithm is shown in Figure 3.

The first step of the segmentation, once the green channel is extracted, consists in smoothing the image to remove noise. Since the acquisition conditions are very well controlled, there is little noise on the images. However, smoothing with an anisotropic diffusion filter reduces pixel intensity variance, which improves the result of subsequent processing. After smoothing, the image is binarized with high threshold, so that only high intensity areas such as lumen (and “empty” areas containing no tissue at all) appear at foreground. We are currently using a fixed threshold, because all images in our test set have similar intensities. This threshold has been set to 200 (the range of possible intensities using 8 bits is 0-255).

Each set of touching foreground pixels (i.e., each connected component) from the binarization output is considered a *blob*. The physical size of each blob is measured and those which are too small or too large to be considered lumen areas are removed, since they probably represent noise in the former case, and areas lacking any tissue in the latter. Minimum blob size has been set to 500 nm², and maximum size to 20,000 nm².

Thresholding usually produces blobs that are not compact and contain small holes, so a voting technique is used to fill in those holes. The voting method is similar to morphologic dilation, but it is more restrictive. In traditional dilation, a structuring element (kernel) is moved along the image and all the background pixels that lie in the region covered by the kernel when it is centered in a foreground pixel are automatically promoted to the foreground set. In comparison, when using voting, instead of directly promoting the pixels, they receive a *vote*, and only the pixels that receive a certain number of votes are subsequently moved to the foreground set. This filter is used iteratively five times, using a squared 3×3 kernel. Each of the resulting blobs is considered a *potential lumen area* and labeled individually.

Potential lumen areas are processed one by one, in a sequential loop, in order to expand them outward in search of concentric regions. For each of them, a *working region* (a square with its sides separated 120 pixels in each direction from the potential lumen) is defined. Then,

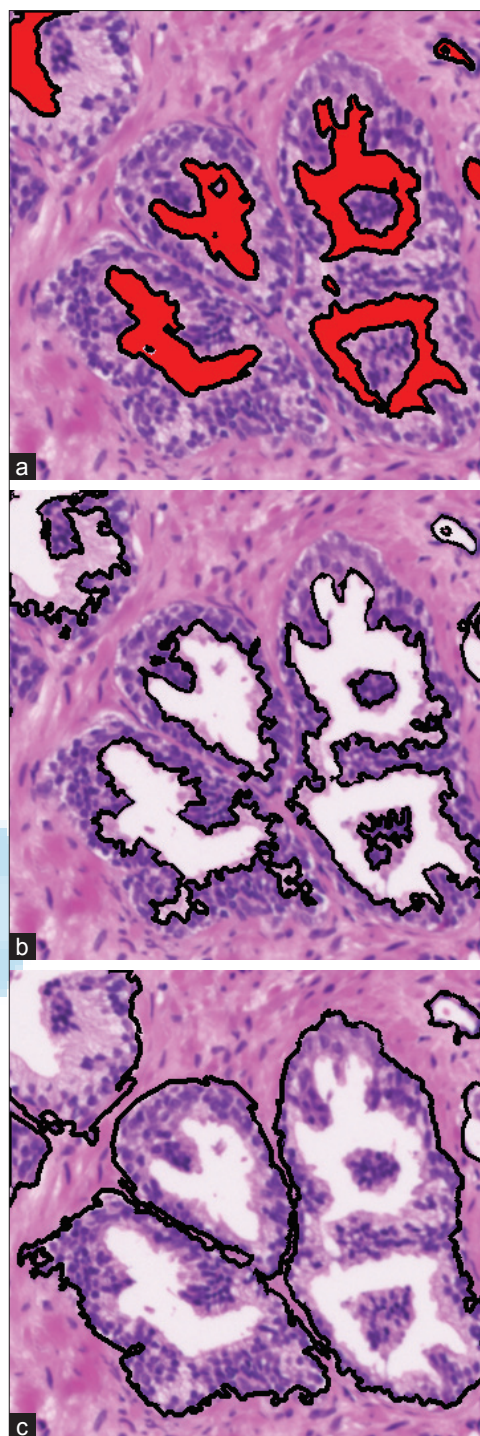


Figure 4: Partial and final results of the lumen centered expansion algorithm. (a) Initialization of the level set with lumen segmentation. (b) First level set with cytoplasm segmentation. (c) Final result with cell segmentation

the contour of the potential lumen area is used as initial zero set for a geodesic active contours level set.^[18] This is illustrated in Figure 4. Figure 4a shows the lumen areas in pseudocolor (red), and their contours as initial zero set. The evolution function of this level set is established so that it can grow freely almost everywhere except on

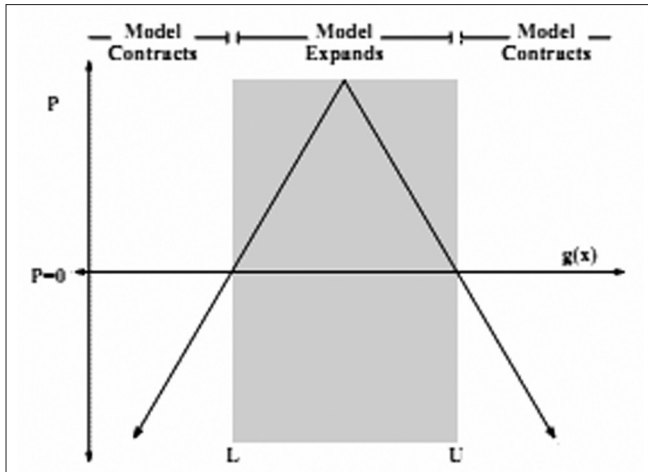


Figure 5: Threshold segmentation propagation term^[19]

cell regions, where it gets stuck and, as stated in Ref.^[19], it can be represented as

$$\frac{\partial u}{\partial t} = \alpha A(x) + \beta P(x) + \gamma Z(x) \quad (1)$$

where A is the advection term, P is the propagation term, and Z the curvature term. The first term attracts the curve to desired areas, the second term regulates the speed with which the curve moves outward (always in a normal direction of the curve), and the third term controls how much the curve may bend itself, preventing uncontrolled growth. It should be noted that the curve may split and merge at any time. Since the objective is to make the curve grow from the contour of the lumen to the inner contour of the cells, the advection term is constructed so that the curve is attracted by cells. The curvature and propagation terms are calculated using a distance map where the lumen contour is zero valued, points inside the contour are negative, and points outside it are positive. In this scenario, the curve grows outward, and such behavior allows it to include the glandular border into the segmented region.

Propagation and curvature coefficients are set to low values compared to the advection coefficient, so that the curve grows slowly and is able to bend (creating high-curvature edges if necessary). These coefficients are set to 1.0 for the advection term, 0.1 for the propagation term, and 0.3 for the curvature term. It should be remarked that, for each potential lumen area, the presence of other lumen areas does not affect the level set evolution, because the distance map is calculated based only on the working region, and thus only one potential lumen area contributes for each distance map at any time.

When the geodesic level set has finished its evolution, the curve should be touching the inner border of the cells that surround the lumen and cytoplasm. This partial result is illustrated in Figure 4b.

Since the region of interest also includes the cells, a second level set is used to obtain the ROI. The initialization of this second level set is a dilation of the output of the previous one, so that the initial zero set goes across the cells. This time, the curve may either grow or contract, but the evolution function is simpler, and it just depends on the intensity values of the image. An expansion interval is determined by a lower and upper threshold (L, U). The mean between L and U is called maximum expansion value (MEV). For each pixel belonging to the curve, if its intensity is inside the interval (L, U), then the level set will expand with a speed proportional to the distance to the MEV as shown in Figure 5.

The evolution function for this level set does not consider an advection term. Rather, it just considers propagation and curvature, and propagation does not depend on the distance from the initial contour. According to Ref.^[19], the evolution of this level set is

$$\frac{\partial u}{\partial t} = \beta P(x) + \gamma Z(x) \quad (2)$$

$$P(x) = \begin{cases} g(x)-L & \text{if } g(x) < (U-L)/2+L \\ U-g(x) & \text{otherwise} \end{cases} \quad (3)$$

It has been observed that this level set works better if the MEV is set close to the mean value of the pixels that are to be segmented. The targets are the cells which have medium-low intensity values in the green channel of the image, so the lower and upper thresholds were set to 60 and 160, respectively. The weights for the propagation and advection terms were the same as in the geodesic level set.

Finally, outputs of both level set segmentations are merged into just one, which represents the segmented region. If the potential lumen area was really a true one (composed of the three structures of interest in the predicted fashion) the output of the segmentation should be the full ROI. This is illustrated in Figure 4c.

Cell density localization

As with lumen centered expansion, this approach just uses one of the channels, in this case, the red channel of the original RGB image. As Figure 2 shows, the red channel provides better contrast between the cells and the other types of structures of interest than any other channel. As said before, using just one channel also reduces memory requirements, which is a nice side effect. A flowchart depicting this algorithm is shown in Figure 6.

The first step of the segmentation, once the red channel has been extracted, is thresholding that channel to separate the cells from the rest of the tissue. The cells are darker than the rest of the tissue, which is almost white

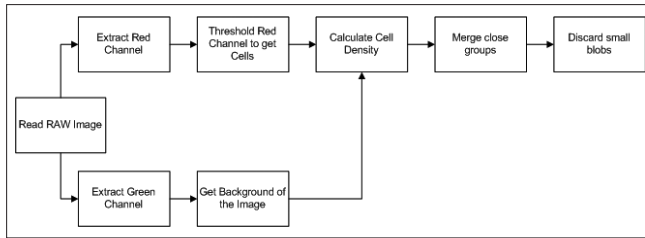


Figure 6: Flowchart of the cell density localization algorithm

(Figure 2, red channel). The threshold has been set to the 80% of the maximum intensity of the image (around 200 when using 8-bit images), so that any pixel with an intensity lower than the threshold is considered a cell.

The result of the thresholding is used to compute the cell density of the image. For each pixel in the image, a circular neighborhood of radius 7 pixels is used to calculate its cell density. This radius is valid for both 5×

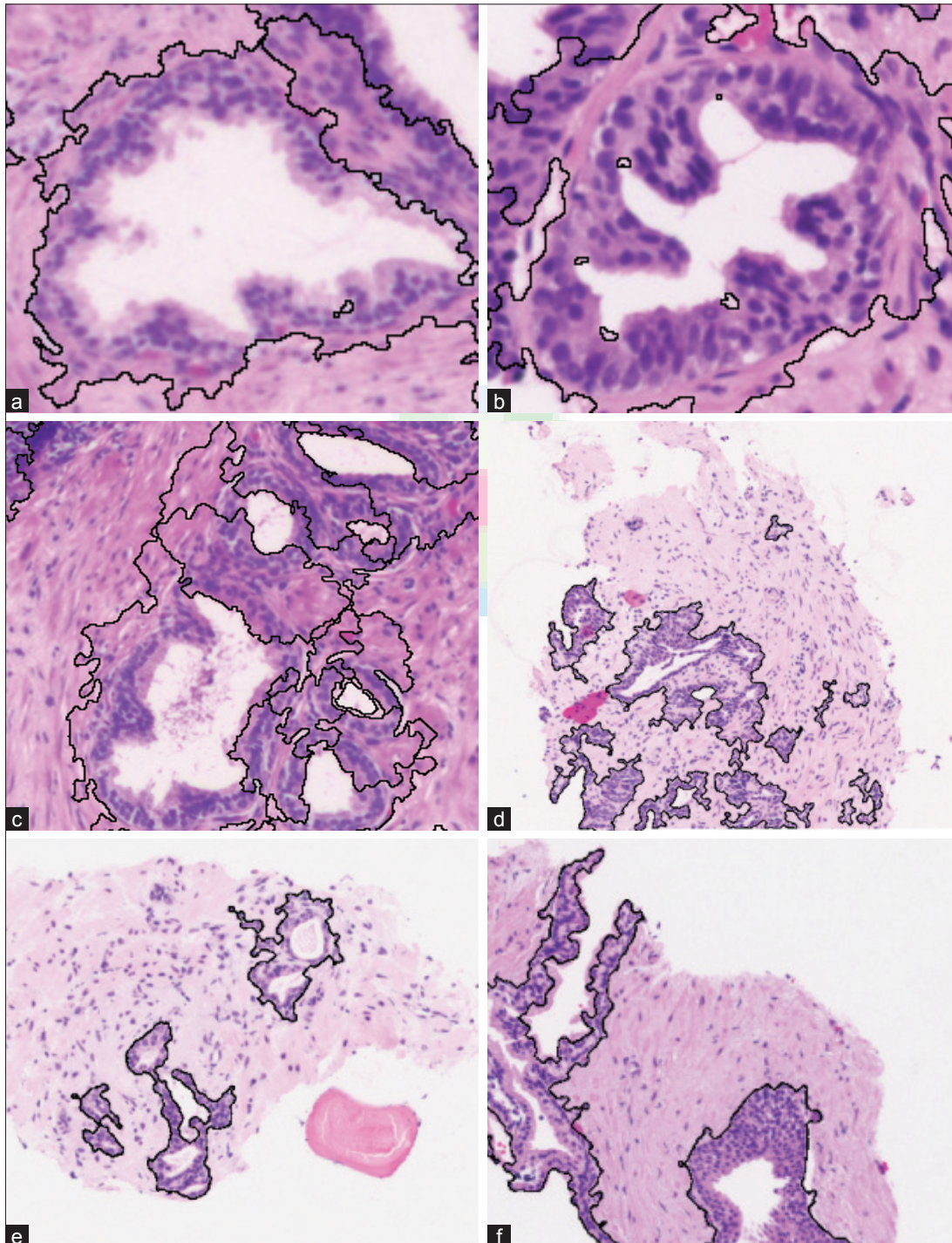


Figure 7: Segmented ROIs. (a)-(c) Results of the lumen centered expansion algorithm. (d)-(f) Results of the cell density localization algorithm

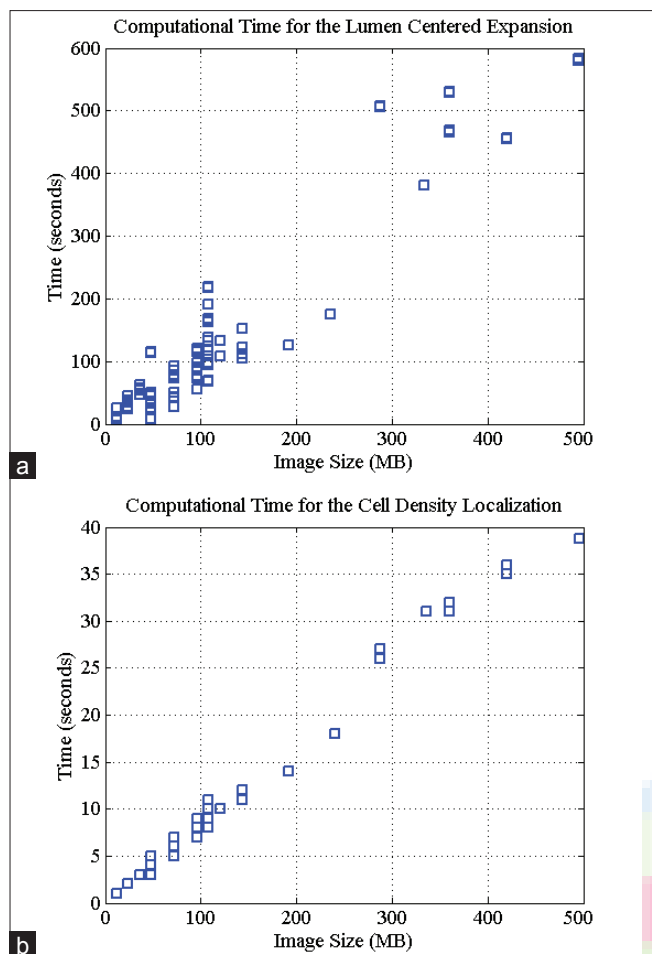


Figure 8: Computational times. (a) Times of the lumen centered expansion algorithm. (b) Times of the cell density localization algorithm

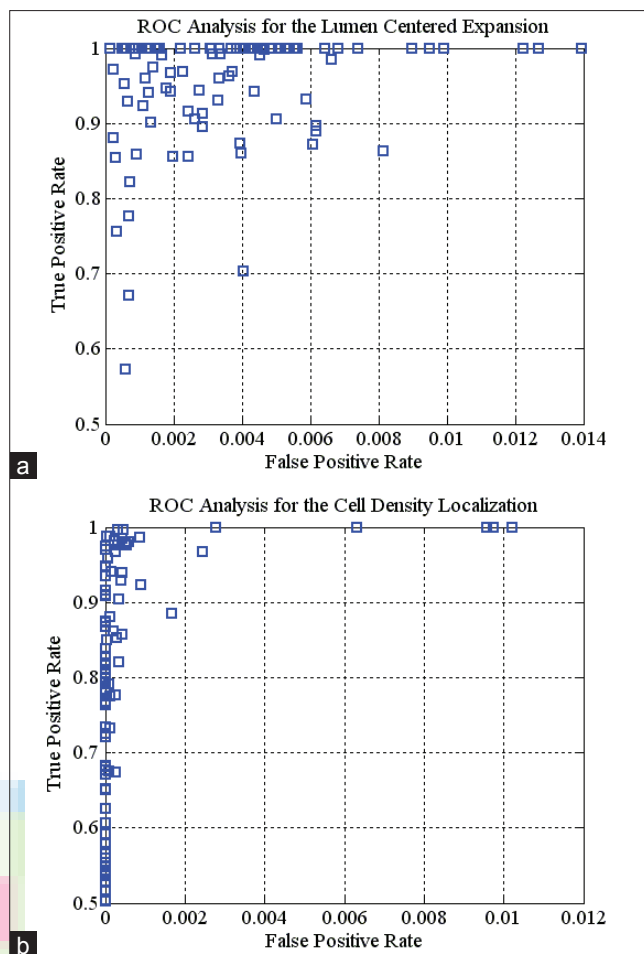


Figure 9: ROC analysis. (a) ROC of the lumen centered expansion algorithm. (b) ROC of the cell density localization algorithm

and 10× images. The pixel density is easily computed using a mean filter:

$$\sum_i^N \frac{p_i}{N} \quad (4)$$

where N is the number of pixels in the neighborhood, and p_i are each one of those pixels. In order to keep the regions that feature a higher cell density, only pixels with a cell density higher than 30% are considered relevant. Radius 7 was chosen because it is small enough to keep the calculations localized, and large enough to let each pixel be influenced by more than one cell.

In order to obtain a smooth and compact final result, several operations are performed. First, a dilation using a 3×3 kernel (4-neighbors, 5 iterations) is executed. This aims at closing small gaps between cells, so that groups of cells are merged into big blobs. Next, in case that any of those blobs is not compact (i.e., it has holes inside it), it is filled. Then erosion using the same kernel and number of iterations is used to restore the size of the blobs previously dilated. Finally the blobs that are not big enough (at least 110 μm in perimeter) to be relevant are discarded.

All these steps are only performed in the regions of the image where there is tissue present. In order to know where the tissue is present, the background of the image (which is almost white, due to illumination used in the microscope), is extracted by thresholding the green channel of the image. Pixels with an intensity higher than 90% and not surrounded by any tissue are considered background. The application of the algorithm only in tissue regions speeds up the execution.

RESULTS

A dataset of 100 complete prostate H and E stained biopsy images has been used to test the algorithms. All the image were acquired with 5× magnification, with memory requirements ranging from 12 MB (2000×2000 pixels) to 500 MB (14,000×12,000 pixels). Some selected fragments that exemplify the Lumen Centered Expansion and the Cell Density Localization algorithms are shown in Figures 7a-c, and Figures 7d-f respectively.

Although the images used to test the algorithm were large, computational times were not deemed excessive. Figure 8 shows a scatter plot with the computational

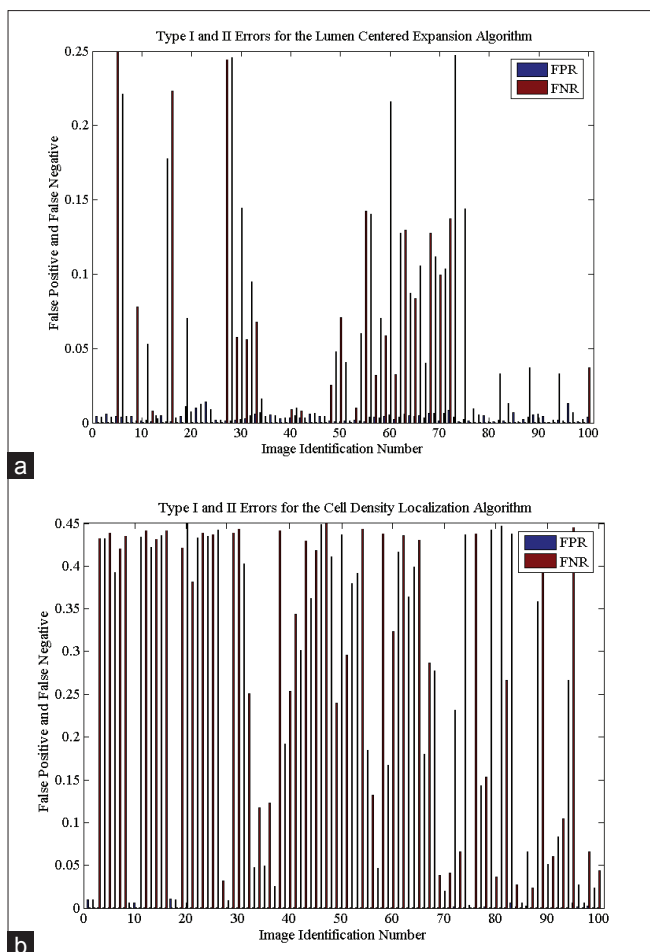


Figure 10: Type I and II errors (FP and FN). (a) Errors for the lumen centered expansion algorithm. (b) Errors for the cell density localization algorithm

time for both algorithms run with the 100 images. The level set algorithm takes from 6 seconds to 9 minutes for 12 MB and 500 MB images respectively [Figure 8a], and mean filtering method takes from 1 second to 38 seconds [Figure 8b]. The testing machine was equipped with an Intel Core i7 950 (3.07 GHz) processor and 12 GB RAM.

A quantitative validation based on ROC analysis was carried out with our set of 100 different tissue samples of WSI, stained with a variety of H and E dyes (weak and dark). The samples were both benign and malign samples of prostate biopsy. The results of both algorithms were compared to the manual selection of ROIs done by pathologists from the local Hospital (HGUCR). Thus, the rates of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) detections were calculated. The ROC analysis for the algorithms of lumen centered expansion and cell density localization are shown in Figure 9.

In the case of the lumen centered expansion algorithm, the rate of both FP and TP is higher than that for the cell density localization algorithm. However, the errors (FP and FN) are kept quite low for both methods. Thus, in the case of the lumen expansion [Figure 9a], an average

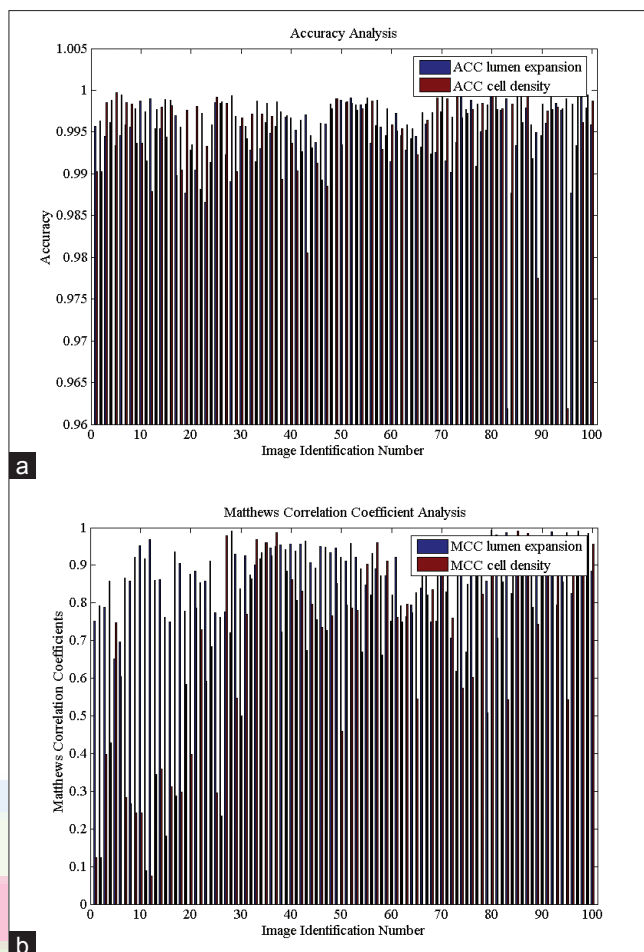


Figure 11: Validation metrics. (a) Accuracy. (b) Matthews correlation coefficient

of 5% detections were FN, 0.39% were FP, 95% were TP, and 99.61% were TN. In the case of the cell density technique [Figure 9b], an average of 20% detections were FN, 0.08 were FP, 80% were TP, and 99.92 were TN. The results show an average sensitivity of 95% and 80% with specificity above 99% for the lumen centered expansion and cell density localization respectively. The type I and II errors, that is, the FP and FN for the 100 images are illustrated in Figure 10a for the Lumen Centered Expansion and Figure 10b for the Cell Density Localization algorithms. Most of the FP errors occur in those samples stained with weak H&E dye, above all for the lumen centered algorithm.

The accuracy (ACC) and Matthews correlation coefficient (MCC) were also obtained. These quantitative metrics are defined by

$$ACC = \frac{(TP + TN)}{(TP + FN + FP + TN)} \quad (5)$$

$$MCC = \frac{(TP * TN) - (FP * FN)}{\sqrt{(TP + FP)(TP + FN) + (TN + FP)(TN + FN)}} \quad (6)$$

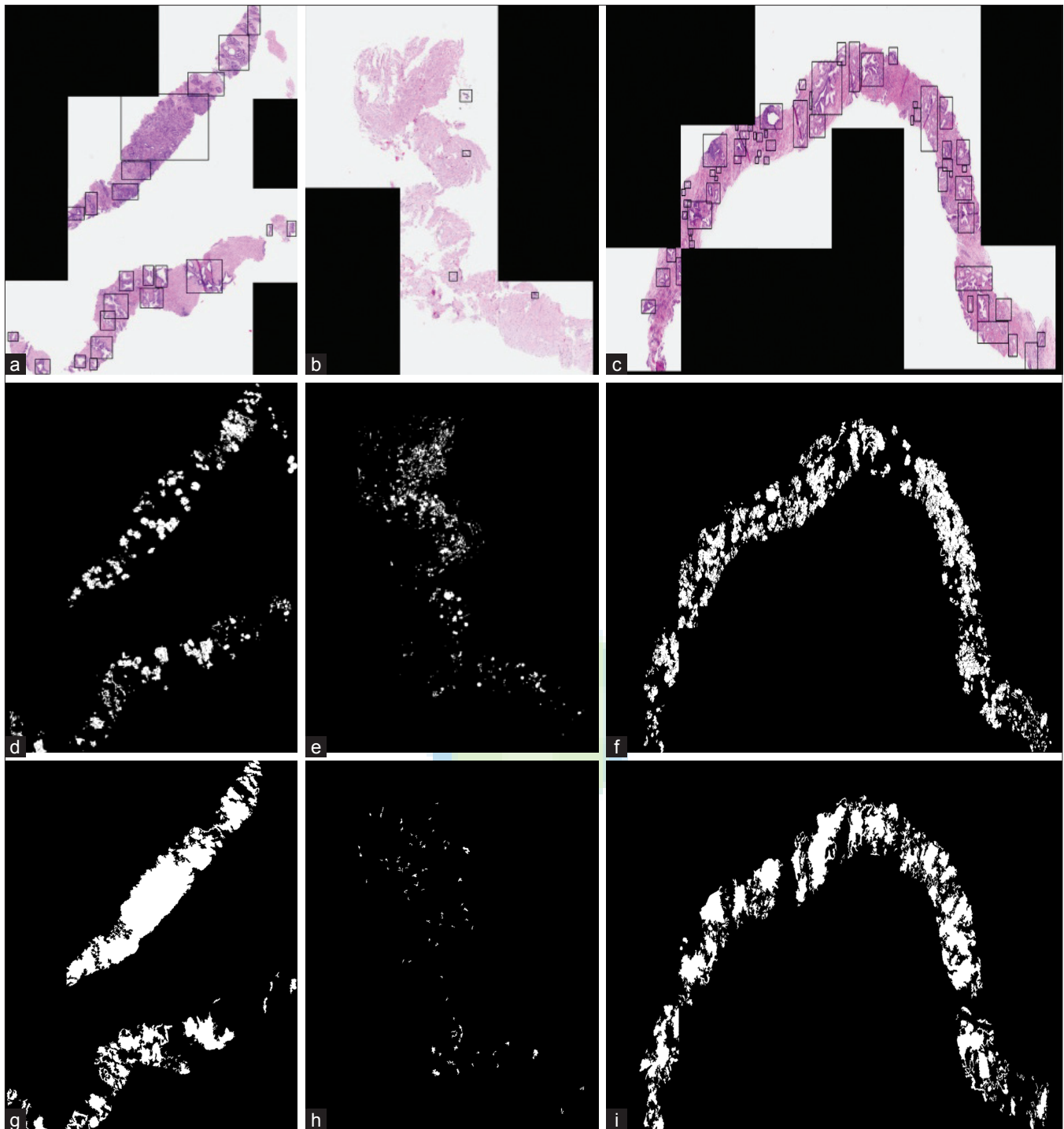


Figure 12: Results for the quantitative validation. (a-c) Manual ROI selection. (d-f) ROIs obtained with the lumen centered expansion algorithm. (g-i) ROIs obtained with the cell density localization algorithm

Both the ACC [Figure 11a] and the MCC [Figure 11b] give good results. An average of 99% ACC with 0.87 MCC is obtained for the lumen expansion and 99% ACC with 0.71 MCC for the cell density algorithm. The MCC is a correlation coefficient between the truth values and detected ones; it returns a value between -1 and +1. A coefficient of +1 represents a perfect prediction, 0 an

average random prediction, and -1 an inverse prediction. These metrics are illustrated in Figure 11.

The results on the WSI for both algorithms are illustrated in Figure 12. Figure 12a-c are the manual ROI selection, (Figure 12d-f) show the ROIs obtained with the lumen centered expansion segmentation, and (Figure 12g-i) show the ROIs obtained with the cell density localization

segmentation.

The algorithms were also tested with images at 10× magnification and up to 1228 MB (24,000×16,000 pixels) obtaining similar results.

Further improvements shall consider the fusion of information in the two algorithms together with texture analysis^[2,8,20,21] and methods based on invariant color information.

CONCLUSIONS

In this paper, an approach to ROI segmentation in whole-slide images of prostate biopsies has been described. The method proposed is based on level set and mean filtering techniques for lumen centered expansion and cell density localization respectively. The approach followed in this paper is different from previous work that attempts to segment all significant regions such as nuclei, lumen, and epithelial cytoplasm. The novelty of the technique lies in the ability to detect complete ROIs, where a ROI is composed by the conjunction of three different structures, that is, lumen, cytoplasm, and cells with a high density of cells and the architectural distribution between lumen and cells. The method is capable of dealing with full biopsies digitized at different magnification. The proposed algorithm is also original because it works on large images acquired with low magnification, thus being different from other algorithms that require higher magnification and have been tested only on small samples. In this way, the method tries to mimic the manual procedure of expert clinicians.

The proposed system is also useful because it can be used for different purposes. It could be integrated into a slide visualization environment to highlight the ROIs for the pathologists, either for slide analysis or even with teaching purposes. Another possible use of the ROI segmentation is virtual microscopy systems. In order to avoid the full digitization of all samples, they could be first digitized at low magnification (5×, or 10×), and then processed to locate the ROIs, which would be the only regions to be subsequently digitized at higher magnification. The system could also be used as a previous step in classification applications, since it could reduce the amount of information to be processed, and probably speed up the whole classification process.

A dataset of 100 prostate biopsies WSI stained with a variety of H and E dyes (weak and dark) has been used to test the algorithms. All the images were acquired with 5× magnification, with memory requirements ranging from 12 MB to 500 MB. The tests carried out show that the algorithms are both fast and accurate. The segmentation accuracy by means of ROC analysis and the Matthews correlation coefficient give good results. An average 99%

ACC with 0.87 MCC, sensitivity of 95% and specificity of 99.61% is obtained for the lumen expansion, and 99.92% ACC with 0.71 MCC, sensitivity of 80%, and specificity of 99% is obtained for the cell density algorithm.

Although segmentation accuracy is not high enough to be used in a medical environment in the short term, we consider that our results are promising and we are confident that future enhancements to the system will improve the results. Further improvements shall consider the fusion of information in the two algorithms together with texture analysis and methods based on invariant color information.

ACKNOWLEDGMENTS

This work has been carried out with the support of COST Action IC0604 and the research projects DPI2008-06071 of the Spanish Research Ministry, and PI-2010/040 of the FISCAM.

REFERENCES

- Jemal A, Siegel R, Xu J, Ward E. Cancer Statistics 2010. *CA Cancer J Clin* 2010;60:277-300.
- Farjam R, Soltanian-Zadeh H, Jafari-Khouzani K, Zoroofi RA. An image analysis approach for automatic malignancy determination of prostate pathological images. *Cytometry Part B: Clinical Cytometry* 2007;72:227-40
- Gleason DF. The veteran's administration cooperative urologic research group: Histologic grading and clinical staging of prostate carcinoma. In Tannenbaum, M. *Urologic Pathology: The Prostate*. Philadelphia: Lea & Febiger, 1977. p. 171-98.
- García M, Bueno G, PecesC, González J, Carbajo M. Critical comparison of 31 commercially available digital slide systems in pathology. *Int J Surg Pathol* 2006;14:285-305.
- Jafari-Khouzani K, Soltanian-Zadeh H. Multiwavelet grading of pathological images of prostate. *IEEE Trans Biomed Engg* 2003;50:697-704.
- Farjam R, Soltanian-Zadeh H, Zoroofi RA. Wavelet-based determination of malignancy of the pathological images of the prostate. *WSEAS Trans Electronics* 2004;3:476-82.
- Doyle S, Madabhushi A, Feldman M, Tomaszewski J. A boosting cascade for automated detection of prostate cancer from digitized histology. In: Larsen R, Nielsen M, Sporring J. (editors.) *Medical Image Computing and Computer-Assisted Intervention MICCAI 2006, Lecture Notes in Computer Science*. Heidelberg: Springer Berlin; 2006;4191:504-11.
- Doyle S, Hwang M, Shah K, Madabhushi A, Feldman M, Tomaszewski J. Automated grading of prostate cancer using architectural and textural image features. In: *Biomedical Imaging: From Nano to Macro*, 2007. ISBI 2007. 4th IEEE International Symposium on. 2007. p. 1284-7.
- Doyle S, Feldman M, Tomaszewski J, Madabhushi A. A boosted Bayesian multi-resolution classifier for prostate cancer detection from digitized needle biopsies. *IEEE Trans Biomed Engg (Preprint)* (2010) doi: 10.1109/TBME.2010.2053540.
- Huang PW, Lee CH. Automatic classification for pathological prostate images based on fractal analysis. *IEEE Trans Med Imaging* 2009;28:1037-50.
- Naik S, Doyle S, Feldman M, Tomaszewski J, Madabhushi A. Gland segmentation and computerized Gleason grading of prostate histology by integrating low, high-level and domain specific information. In: Metaxas DN, Rittscher J, Lockett S, Sebastian TB. (editors.) *Proceedings of 2nd Workshop on Microscopic Image Analysis with Applications in Biology*, Piscataway, NJ, USA: 2007.
- Naik S, Madabhushi A, Tomaszewski J, Feldman M. A quantitative exploration of efficacy of gland morphology in prostate cancer grading. In: *Bioengineering Conference, 2007. NEBC '07. IEEE 33rd Annual Northeast*. 2007. p. 58-9.
- Naik S, Doyle S, Agner S, Madabhushi A, Feldman M, Tomaszewski J. Automated gland and nuclei segmentation for grading of prostate and breast cancer

- histopathology. In: Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on. 2008. p. 284-7.
14. Xu J, Madabhushi A, Janowczyk A, Chandran S. A weighted mean shift, normalized cuts initialized color gradient based geodesic active contour model: Applications to histopathology image segmentation. In: Proceedings of SPIE. 2010;7623:76230Y.
 15. Hafiane A, Bunyak F, Palaniappan K. Fuzzy clustering and active contours for histopathology image segmentation and nuclei detection. In: ACIVS. 2008;5259:903-14.
 16. Hafiane A, Bunyak F, Palaniappan K. Level set-based histology image segmentation with region-based comparison. In: Proc. Microscopic Image Analysis with Applications in Biology 2008.
 17. Epstein JI. Biopsy Interpretation of the Prostate. 4th ed. Philadelphia: Lippincott Williams and Wilkins; 2007.
 18. Caselles V, Kimmel R, Sapiro G. Geodesic active contours. Int J Comput Vision 1995;22:61-79.
 19. Ibañez L, Schroeder W, Ng L, Cates J. The ITK Software Guide. 2nd ed. New York: Kitware, Inc; ISBN 1-930934-15-7., 2005.
 20. Jafari-Khouzani K, Soltanian-Zadeh H. Rotation-invariant multiresolution texture analysis using radon and wavelet transforms. IEEE Trans Image Process 2005;14:783-95.
 21. Tabesh A, Teverovskiy M. Tumor classification in histological images of prostate using color texture. In: Proc. Asilomar Conf. Signals, Systems, and Computers 2006, p. 841-5.

