Proc. 29th International Conference on Systems, Signals and Image Processing "IWSSIP 2022", June 01 - 03, 2022, Sofia, Bulgaria

A semi-automatic approach for quantitative analysis of histological images

Sergio A. da Silva Junior
Departamento de Informática
Universidade Estadual de Maringá
Maringá, Brazil
ra115735@uem.br

Mateus F. T. Carvalho
Departamento de Informática
Universidade Estadual de Maringá
Maringá, Brazil
pg404059@uem.br

Celso V. A. Q. Calomeno
Departamento de Biotecnlogia,
Genética e Biologia Celular
Universidade Estadual de Maringá
Maringá, Brazil
ra110688@uem.br

Julio C. Polonio
Departamento de Biotecnlogia,
Genética e Biologia Celular
Universidade Estadual de Maringá
Maringá, Brazil
jcpolonio2@uem.br

Franklin C. Flores
Departamento de Informática
Universidade Estadual
de Maringá
Maringá, Brazil
fcflores@din.uem.br

Jaqueline Rinaldi
Programa de pós-graduação em
Biociências e Fisiopatologia
Universidade Estadual de Maringá
Maringá, Brazil
jak.rinaldi@gmail.com

Yandre M. G. Costa
Departamento de Informática
Universidade Estadual
de Maringá
Maringá, Brazil
yandre@din.uem.br

***Abstract*—The use of a computational approach in histological analysis can help in the acquisition and interpretation of results. In this work, we present a computational method to automate the morphometrical protocol proposed by Ewald R. Weibel in 1966. The method proposed here was based on the use of widely known image processing algorithms aiming to identify regions of interest (i.e. image segmentation, thresholding and flood fill). The method also uses mathematical morphology, such as the opening operator to filter the noise left by the thresholding. As a contribution of this work, we also created a software tool, named AutoGrid, to support reasearchers of biology labs apply the method. The proposal was evaluated on a set of images taken from rat prostate. In this scenario, we extracted data to perform histological analysis both using the software tool and following the manual protocol. When the tool was used (i.e. semi-automatic mode), the user was asked to perform a little manual intervention, and the most time consuming and tiring parts of the process runs automatically. Finally, the computational approach presented a faster (i.e. $\approx$ 45% less time) and less monotonous histological analysis, with statistically equivalent results to those obtained on the manual mode.**

***Index Terms*—Weibel's Grid, Biological tissue segmentation, Mathematical Morphology.**

I. Introduction

In this work, we propose a semi-automatic approach for quantitative analysis of histological images taken from the rat ventral prostate. The quantitative analysis of histological images is defined by the application of mathematical axioms. By means of these axioms, it is possible to reconstitute and quantify, in a three-dimensional way, the analyzed biological tissue, observing its volume density, surface, length and number of different tissue structures.

The literature presents different authors who have already introduced several techniques of quantitative methods [1]–[3]. Quantification allows the researchers to evaluate and compare parameters that allow to apply statistical analysis. Here, we propose the implementation of a semi-automatic method to support researchers to perform analysis in biological tissues using the grid introduced by Weibel [1], in 1966.

This type of methodology is normally a repetitive and tiring job, the researcher has to overlap the tissue image with the grid image, which is part of the protocol proposed by Weibel, and count which tissue is overlapping the grid points. This process can be very demanding, and especially susceptible to errors.

In this sense, the automation of such type of task would be of great value, in order to increase the productivity of the researchers. The approach proposed here is capable of segmenting the image, allowing some interaction with the researcher. Moreover, it allows the automatically counting of the overlapping structures.

As a model, we used images of rat prostate from the GEBIOREP research group of the State University of Maringá. The morphology of the prostate is organized in three compartments segmented in Epithelium, Stroma and Lumen. In our approach, the Epithelium is segmented by applying binarization to the red or blue channel of the original image, a morphological filter is applied afterwards to remove the noise. The Lumen is segmented with help of the researcher, by clicking inside the borders of the Epithelium the flood fill algorithm is applied in that region to segment it. The remaining of the image is considered the stroma.

The remaining of this work is organized as follows: Section II describes Digital Image Processing techniques used here (i.e. segmenters and mathematical morphology), the Weibel Grid, and also introduces some details about the dataset used for experimentation. In Section III we describe the Study Design. Section IV presents the results and discussion. And, finally, in Section V, the concluding remarks are described.

II. Materials and Methods

In this section we describe the main image processing techniques used for the implementation of method proposed here. In addition, we also introduce some details of the dataset used for experimentation.

A. Threshold

Threshold is one of the simplest segmentation techniques, it is used to segment areas of the image that are in the same grayscale value range [4], [5]. If the Threshold is utilized for segmentation two regions, it is denominated binarization. The Threshold, is defined with:

$$f(x,y) = \begin{cases} 0, & \text{if} \quad f(x,y) \leq T \\ 255, & \text{if} \quad f(x,y) > T, \end{cases} \tag{1}$$

978-1-7281-4387-3/22/$31.00 © 2022 IEEE

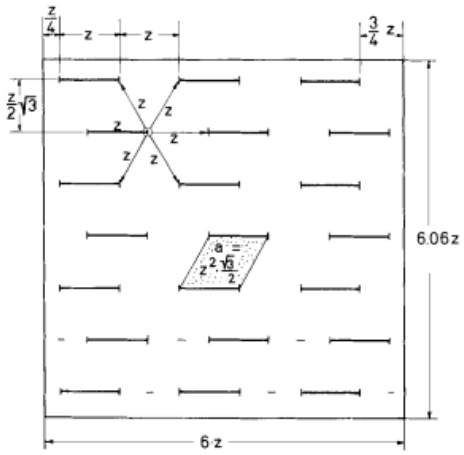


Fig. 1: Weibel Grid with 42 points [1].

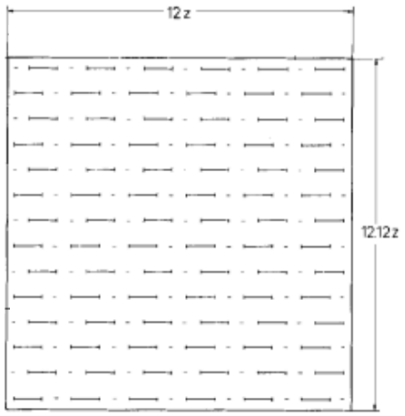


Fig. 2: Weibel Grid with 168 points [1].

where $f(x, y)$ is the pixel value in position x, y of the image in analysis and T is the chosen Threshold value. The Threshold value is an integer, i.e., $T = 0, \ldots, 255$.

B. Mathematical Morphology

The study of Mathematical Morphology has begun in the 60's, in the *École des Mines de Paris*, introduced by Matheron and Serra [6]. Its main objective is to extract information regarding the geometry and topology of an unknown set, by transforming it through a completely defined set, called Structuring Element (SE) [7]. An SE can assume different types of form, cross, disk and line are some examples. Here, we used SE disk, because the structures in the image are round.

The basic operations of Mathematical Morphology (MM) are *erosion* and *dilation*. By combining these operations, other MM operations can be done. These operations were obtained from [4]. Let A be a binary digital image and B a Structuring Element (SE). Binary dilation $A \oplus B$, is defined as $A \oplus B = \bigcup_{b \in B}(A + b)$ and Binary Erosion $A \ominus B$, is defined as $A \ominus B = \bigcap_{b \in B}(A - b)$. Note that $A \oplus B$ and $A \ominus B$ are binary digital images.

With the operations defined above, it is possible to define *opening* operations. Binary Opening is defined as $A \circ B = (A \ominus B) \oplus B$. This operation causes the image to be smoothed around the edges and eliminate small preeminences. For further details, see [4], [8], [9]. The implementation of the described filter was made with assist of the Scikit-image morphology library [10] and the OpenCV image processing library [11].

C. The Weibel grid

In the morphometric analysis, the methods *test lines*, *test point* e *test area* are used. The *test lines* method estimates the surface, however, it is necessary to have a grid composed of vertically oriented parallel line segments; The *test point* estimates the volumetric work, however, it is necessary to have a grid composed of point to do it; and the *test area* performs the counting of the structures. Therefore, depending on the analysis, it is necessary to use different grids.

With this in mind, Weibel published in 1966 [1] a method consisting of a grid capable of performing the three analyzes mentioned above. This method received Weibel's name, and became known as *Weibel Grid*. Fig. 1 shows a Weibel grid with 42 points.

By carefully analyzing the Fig. 1, we can observe the mathematical spacial relation between the components of the grid. In this case, the Weibel grid is composed of 21 line segments of length z, separated in a vertically equidistant manner, by a distance, $\frac{z}{2}\sqrt{3}$. The horizontal lines are also separated in a equidistant manner, by a distance z. Consequently, the grid has 42 points, two for each line segment, one at the start of the segment of size z, and another at the end. It is possible to expand the 42 points grid to a grid of 168 points, as shown in Fig. 2.

In our experiments, the line segments were created according to Fig. 2, that is, formed by 84 segments of length z, separated in a vertically equidistant manner, by a distance $\frac{z}{2}\sqrt{3}$, and in a horizontally equidistant manner, by a distance z.

Finally, it should be noted that the value z depends on the height of the image and also on the number of points the grid has. For example, in case the image is 500 pixels high, by applying a grid of 42 points, z will assume the value $z = 500/6.06 = 82.5 \simeq 83$ pixels of length, if a grid of 168 were to be applied in the image, z will assume the value $z = 500/12.12 = 41.25 \simeq 41$ pixels of length. In an image with 1000 pixels of height, if grid with 42 and 168 points were applied, z would assume $z = 1000/6.06 \simeq 165$ and $z = 1000/12.12 = 82.5 \simeq 83$ pixels of length, respectively.

D. The dataset

The method proposed here was tested on images taken from a dataset created by the GEBIOREP research group of the State University of Maringá. The dataset is composed of ten images obtained from Wistar rat prostate using a digital microscope. The images were stained with Hematoxylin and eosin, which helps to enhance some regions of interest. An example of such image is shown in Fig. 3.

III. Study Design and Execution

In this section, we describe the rationale behind the strategy proposed here aiming to automate the use of the Weibel technique to perform the quantitative analysis of a given biological tissue. Basically, the automation of the use of Weibel technique on digital images is based on counting the occurrences of intersection between a predefined grid of points and regions of interest present in the image. For this purpose, the first step consists of performing the image segmentation, to get the different regions of interest clearly identified.

The images used to test our approach were taken from a Rat ventral prostate (more details in subsection II-D). In this type of image, there are three different regions of interest: Stroma, Epithelium

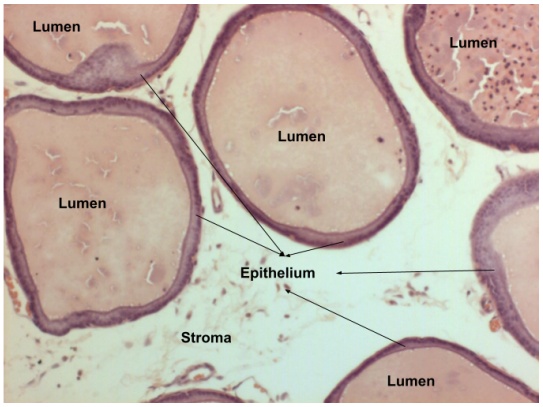


Fig. 3: Photomicrography of ventral rat prostate stained by Hematoxylin and eosin illustrating the glandular compartments measured through Weibel's grid.

and Lumen. These images were stained with Hematoxylin and eosin, which makes easier to separate some regions by exploring the color channels of the digital images (i.e. red, green and blue) in isolation. For example, when a thresholding is applied to the red or blue channel, the Epithelium is easily segmented.

In Fig. 3, it is possible to observe that the Epithelium presents a very striking purple color, which in the RGB color space (**R**ed, **G**reen, **B**lue) can be obtained by using values of R and B greater than G, for example $R = 186$, $G = 85$ and $B = 211$, represented by the tuple $(186, 85, 211)$. However, the Epithelium color can vary a little, its coloration depends on dye time exposure, histological cut thickness, sample quality and other environment factors. Nonetheless, for all images tested in this work, we observed that the colors of the Epithelium always have values like the example presented above, that is, R and B values greater than the value of G. For others types of biological tissue, new parameters must be found and can be adjusted in the interface.

Once the image is loaded, the color channels are separated and three new grayscale images are created, one for each channel. After separating the channels, its possible to apply the thresholdsegmenter, equation 1, which returns three images formed only by the colors white (255) and black (0). In these images, it is noticed that the green channel is not interesting for segmenting the Epithelium, however, the red and blue channels segment the Epithelium clearly, but with some noise.

These noises are treated with the Bynary Opening morphological filter, $A \circ B$, using a disk as the structuring element. To aggregate in the filtering, a brush tool was implemented, it allows the user to manually adjust any type of impression. After the noise is correctly handled, the researcher have to mark manually a point inside the Epithelium, the *flood fill* segmenter is then applied in these regions, to mark the Lumen (interior of the cell).

The choice for the development of a semi-automatic approach is justified by the search for a good performance in the relation time consumed *versus* quality of segmentation. By semi-automatic mode, we mean that the process will have a minimal human intervention. It is illustrated in Fig. 4.

At last, the Weibel grid with 168 points, described in the subsection II-C, and the function responsible for counting the points that are overlapping each image structure are applied.

The counting function operates as follows: if the endpoint of the line segment, represented in red in Fig. 5 is over the ephitelium, then

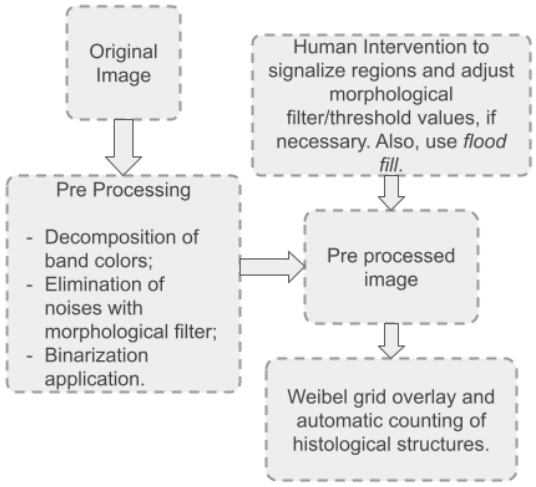


Fig. 4: Flow diagram of the semi-automatic proposed method.

the counter[1] $x \leftarrow x + 1$, or, if the line segment is over the Lumen, then the counter $y \leftarrow y + 1$, otherwise, the counter $w \leftarrow w + 1$, such that w is the Stroma counter.

IV. Results and Discussion

The main goal in this work was to reduce the amount of manual labor done by technicians from biology labs, that usually deals with the application of the Weibel grid to perform analysis of biological tissues. Aiming to check whether the semi-automatic approach proposed here is faster than the manual mode, we performed an experiment using ten photomicrograph images of Rat ventral prostate. As a result, we also developed a software named AutoGrid[2] to serve as a tool to support the analysis of images of biological tissues in biology laboratories operating practices. Fig. 5 shows an example of segmentation using this software tool. Here, the Epithelium is segmented as black, the Lumen is segmented as gray and the Stroma as white. The result of the Weibel grid superimposed to the image can also be seen.

As can be seen in Table I, the mean time to count the histological structures was decreased from 1 minutes and 16 seconds to 42

[1]The counters x, y, z all start at 0, i.e. $x = y = z = 0$

[2]This software tool will be made available in the final version of the paper, if it is accepted.

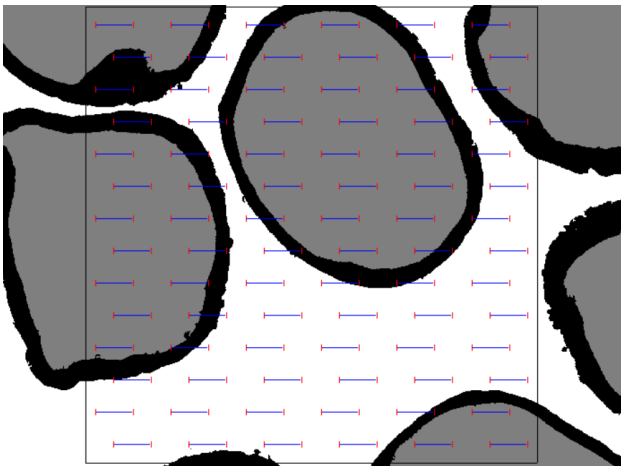
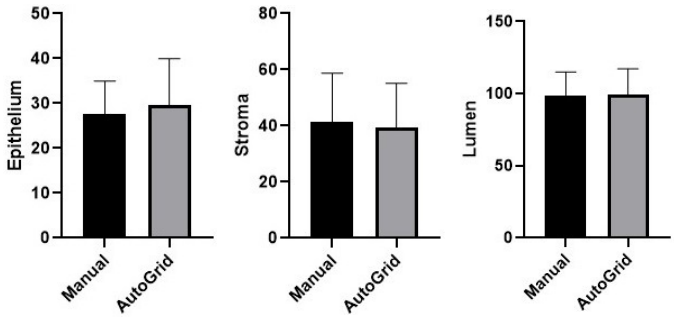


Fig. 5: Resulting segmentation using AutoGrid.



(a) Epithelium (b) Stroma (c) Lumen

Fig. 6: Comparison between AutoGrid and Manual Analysis.

seconds. Other than the time, the software is advantageous because it avoids mistakes when counting and it is less monotonous when compared to the manual counting. In this sense, it is important to point that if a larger number of images need to be processed, the manual operation tends to be affected by other undesirable factors, like fatigue and loss of attention by the annotator.

TABLE I: Comparison between the manual analysis and AutoGrid.

	Manual				AutoGrid			
Image	Stroma	Epithelium	Lumen	Time	Stroma	Epithelium	Lumen	Time
I1	56	27	85	01:44	52	30	86	00:41
I2	44	16	108	01:44	39	13	116	00:55
I3	53	33	82	01:43	48	41	79	00:49
I4	40	39	89	01:33	27	48	93	01:20
I5	74	19	75	01:26	70	26	72	00:30
I6	28	25	115	01:24	30	23	115	00:27
I7	30	22	116	01:26	30	20	118	00:33
I8	36	34	98	01:12	38	37	93	00:31
I9	43	27	98	01:06	46	26	96	00:41
I10	11	24	123	01:05	12	32	124	00:36
Mean	41.5	27.6	98.9	1:16	39.2	29.6	99.2	0:42

A comparative analysis was made using the dataset presented in II-D, to check if there is a statistical significant difference between AutoGrid implementation and the manual approach. For this comparison, we used a two tailed Unpaired T-test [12] with a confidence interval of 95%. In this case, if the returned P-value is lower than 0.05, then the methods are significantly different, otherwise, they are equivalents. In Fig. 6, we can see the mean values obtained for all the regions of interest, both on the AutoGrid and on the manual mode. Using AutoGrid, the Epithelium presented a mean value of 29.6 and the manual method presented a mean value of 27.6. In this case, the P-value obtained in the comparison between these methods was 0.625. The Stroma presented a mean value of 39.2 using AutoGrid and 41.5 in the manual mode, and the P value obtained was 0.970. Finally, the Lumen presented a mean value of 99.2 using AutoGrid and 98.9 in the manual method, and the P-value obtained was 0.761. Observing the P-values obtained, it is possible to conclude that the AutoGrid implementation of the Weibel Grid technique obtain values equivalent to the manual approach, that is, there are no significant difference between both results. Regarding the time consumption, the semi-automatic method consumed $\approx 45\%$ less time than the manual approach.

V. Conlcluding Remarks

In this work, we presented a semi-automatic method for automating the Weibel Grid technique using rats prostate images. For this, the threshold algorithm combined with area opening filter algorithms was used to segment the Epithelium and the flood fill algorithm was used to segment the Lumen.

Experiments performed on a dataset containing rat prostate images showed a decrease of $\approx 45\%$ in the time required to apply the Weibel Grid protocol, and presented a statistically equivalent result when compared to the manual technique.

As future work, we intend to expand our approach also considering eventual peculiarities of other types of biological tissues. We also plan to check whether the difference of performances between AutoGrid and the manual operation is even greater for larger sets of images, both in terms of time and quality of annotation.

Acknowledgment

We thank the Brazilian research support agencies: Coordination for the Improvement of Higher Education Personnel (CAPES), and National Council for Scientific and Technological Development (CNPq) for their financial support.

References

[1] E. R. Weibel, G. S. Kistler, and W. F. Scherle, "Practical stereological methods for morphometric cytology," *The Journal of cell biology*, vol. 30, no. 1, pp. 23–38, 1966.

[2] H. Rohr, M. Oberholzer, G. Bartsch, and M. Keller, "Morphometry in experimental pathology: methods, baseline data and applications." *Int. Review of Experimental Pathology*, vol. 15, pp. 233–325, 1976.

[3] E. Huttunen, T. Romppanen, and H. Helminen, "A histoquantitative study on the effects of castration on the rat ventral prostate lobe." *Journal of Anatomy*, vol. 132, no. Pt 3, p. 357, 1981.

[4] R. C. Gonzalez and R. E. Woods, *Digital image processing*. Upper Saddle River, N.J.: Prentice Hall, 2008.

[5] A. S. Abutaleb, "Automatic thresholding of gray-level pictures using two-dimensional entropy," *Computer vision, graphics, and image processing*, vol. 47, no. 1, pp. 22–32, 1989.

[6] J. Serra, "Introduction to mathematical morphology," *Computer vision, graphics, and image processing*, vol. 35, no. 3, pp. 283–305, 1986.

[7] J. FACON, *Morfologia Matemática: teoria e exemplos*. Curitiba: Editora Universitária Champagnat da PUC-PR, 1996.

[8] R. M. Haralick, S. R. Sternberg, and X. Zhuang, "Image analysis using mathematical morphology," *IEEE transactions on pattern analysis and machine intelligence*, no. 4, pp. 532–550, 1987.

[9] L. Najman and H. Talbot, *Mathematical morphology: from theory to applications*. John Wiley & Sons, 2013.

[10] S. Van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, "scikit-image: image processing in python," *PeerJ*, vol. 2, p. e453, 2014.

[11] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

[12] R. E. Walpole, *Probability & Statistics for Engineers & Scientists*. Prentice Hall, 2006.