

An Overlapping and Integral Blurry Evaluation Method to Optimize Tissues Scanning from Microscope

Jaishir Bayuelo¹, Javier Sanjuan², Julián Yepes-Martinez², Wilson Tovar¹, Fabio Zapata¹, Giselle Peñaloza²

¹Department of Systems Engineering, Universidad del Norte, Barranquilla, Colombia

²Department of Mechanical Engineering, Universidad del Norte, Barranquilla, Colombia

ABSTRACT. Virtual microscopes are devices that employ an automated XYZ mechanism to scan a sample, leading to the obtention of a series of small pictures that, when merged, compose a high-quality representation of the specimen. Due to the assembly tolerances, these devices may suffer from zones out of focus, reducing the quality of the final image. To solve this problem, researchers employ evaluation methods to calculate the blurriness of the image, and when an out of focus picture is located, performs the process of autofocus. Because of the variation on the types of samples, especially in pathology, the existing evaluation methods may fail to deliver a proper blur detection. This article proposes an optimized algorithm for the detection of the blurriness while conducting the sample scan in real time, ensuring that every scanned picture will be in focus. For this purpose, the algorithm relies on two functions, the comparison of the overlapping zones of two consecutive images, and the multivariate linear regression of a series of focus functions. The algorithm proved to be a reliable tool when applied in different pathology samples.

1 Introduction

Whole-slide-Image (WSI) is a procedure that permits the digitalization of samples of pathological purpose to obtain high-resolution images. Due to the various applications of this technology, it has become a key factor for the innovation of digital pathology [1], unleashing in the construction of different microscopes based on this approach [2-11]. Unfortunately, there are limitations for the implementations of WSI for clinical use, among them is its high cost due to the use of specialized software, hardware, and maintenance.

Various microscopes have been proposed to tackle the mentioned drawback. Notwithstanding, the proposed low budget devices decrease the image quality of the sample. Due to the tilt of the mechanical stage produced by the increasing assemble tolerances, that causes blurriness in small zones of the WSI. This inconvenient, although on a lesser scale, is still an issue of expensive microscopes [12]. Different researchers have stated algorithms to detect the blurry zones and to conduct the autofocus of the microscope. The detection of the blurry zones is accomplished by the selection of an adequate focus measure. Many subroutines were found in the literature to calculate the focus, such as Tenengrad function, energy gradient function, Brenner function, and Entropy function [13-21]. The evaluation method should be selected based on its performance, unbiasedness, higher signal to noise ratio [11][18][22]. Unfortunately, the wide variety of tissues make the selection between

those functions difficult, as for some samples the functions can detect false blurriness.

To increase the generality of the evaluation function we propose a new algorithm, overlapping and integral blurry evaluation method (BIEMO). This method integrates the contrast measure function, variance function, the lens distance, the Laplacian variance, and the Brenner function, to calculate the focus of the microscope. Furthermore, as the WSI is based on the capture of a series of small pictures and then in the stitching of them, we also use this stitching process as a second evaluation function. The stitching process merges two subsequent images calculating the feature difference between the overlapping regions of the two pictures. Thus, we propose the use of this overlapping region as a mean to evaluate the degree of blurriness of two images.

2 Functions and correlations

Let's consider the WSI image, which is divided by a set of small pictures $C(i)$, each figure being in gray-scale. Then, for each pair of adjacent photos, a series factors are considered to determine if an image is out focus. The mentioned factors are the key-points matcher K_m , the data variance σ , the Brenner's High B_h and low B_l quality scores, the contrast measure function C_m , the Laplacian variance of the matrix L_V , and the distance between the camera and the sample F_d . The definition of these functions is in the following section.

2.1 Functions definition

The key-points matcher Km is the function utilized for the overlapping. Given two images, the function searches for the interest points to locate the area that is coincident in both pictures, returning a number that quantifies the percentage of coincidence. The algorithm used in this article for the calculation of the key-points is the AKAZE key-point detector and the Flann Based Matcher. Due to the complexity of those functions, the reader is encouraged to look for the reference [23].

The mean and the variance functions are both expressions used to evaluate the blurriness of the image (in the next sections we will be using N and M as the dimensions of the matrix, where N is number of columns and M number of rows). They are presented in the following:

$$u(I) = \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M I(i, j) \quad (1)$$

$$\sigma(I) = \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M [I(i, j) - u(I)]^2 \quad (2)$$

The Brenner's high (B_{hq}) and low (B_{lq}) quality scores calculate how much of an image has high and low quality, we calculate the proportion of the two quantities B_{hq}/B_{lq} . The following equation defines both quality scores:

$$B_{hq}(I) = \sum_{i=1}^N \sum_{j=1}^{M-4} \frac{1}{4} \left[\sum_{k=j}^{j+1} I(i, k) - \sum_{k=j+2}^{j+3} I(i, k) \right]^2 \quad (3)$$

$$B_{lq}(I) = \sum_{i=1}^N \sum_{j=1}^{M-6} \frac{1}{4} \left[\sum_{k=j}^{j+1} I(i, k) - \sum_{k=j+3}^{j+5} I(i, k) \right]^2 \quad (4)$$

The main feature of this Brenner coefficient is that when the image is focused, the value is lesser than 1 and when is not focused is greater or equal to 1. Figure 1 has a Brenner coefficient of 0,59.

- The Laplacian variance (equation 5) highlights the edges of the image, using the gradients in both directions of the Cartesian plane. We rely on this measure to calculate how blurry is the photo. Then, if we cannot differentiate the edges of the image, the picture is out of focus. Figure 2, which has a Laplacian variance of 118,68 and a normal variance of 1004,65. Thus, due to the difference between the two measures is huge, we conclude that the picture is out of focus.

$$LP(I) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2}{\partial y^2} \quad (5)$$

- To calculate the contrast measure C_m , we calculated The Sobel operators. The Sobel operators are a measure of the robustness of the image to noise. To calculate them we calculate two gradients on the image, G_x and G_y , by applying the following kernels, respectively.

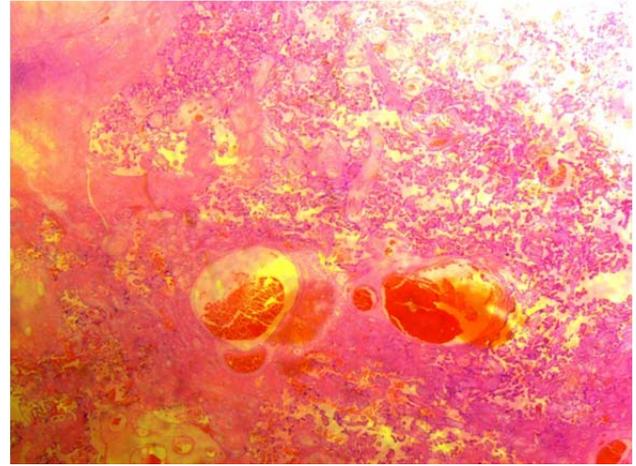


Fig. 1. Kidney tissue image taken from 68,001 mm of the lens.

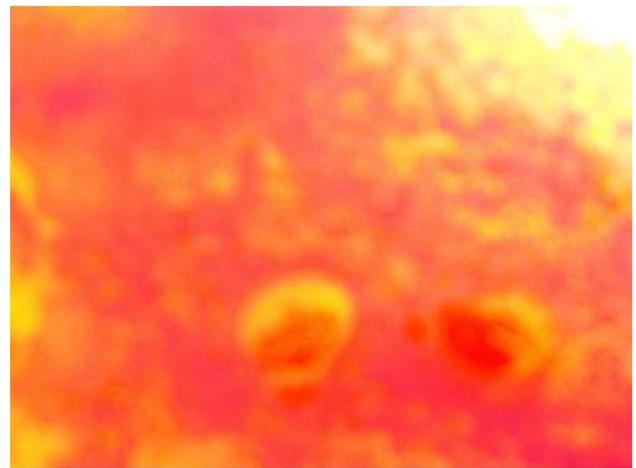


Fig. 2. Kidney tissue image taken from 73,049 mm of the lens.

$$\text{Kernels: } x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad (6)$$

Then, by applying the norm to the gradients, we obtain the contrast measure utilizing the next equation:

$$C_m = \frac{\sqrt{|G_x|^2 + |G_y|^2}}{NM} + 10^{-6} \quad (7)$$

2.2 Multivariate linear regression

We utilize the functions defined in section 2 to obtain a multivariate linear regression model to predict the focus of the picture. We picked those functions due to the results of the ANOVA analysis, which stated that these evaluation methods had the greatest correlation with the blurriness of the image. The equation of the model is the following:

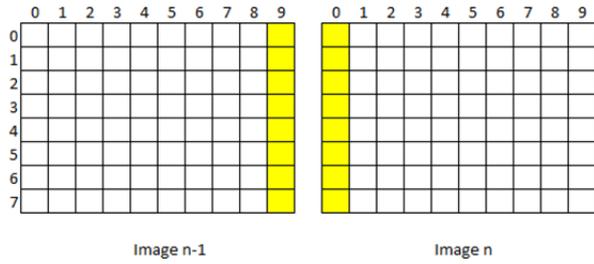


Fig. 1. Image matrices comparison

$$y = \frac{2,81C_m}{10^4} - \frac{5,46\sigma}{10^5} - \frac{1,09F_d+2,49L_v}{10^3} + 1,8 \frac{B_{hq}}{B_{lq}} + 2,14 \quad (8)$$

Where a value of y close to 1 means that the image is focused, otherwise it means blur. We obtained the previous equation employing a dataset of 100 images for the linear regression. Moreover, to prove the effectiveness of the model, we calculated the determination coefficient R , which yielded a value of 0,98, implying fitness of the equation to describe the blurriness.

3 Biemo algorithm

Figure 4 presents pseudo-code of the BIEMO algorithm. The algorithm starts at the first position, C_{i-1} , then goes to the next point, C_i (with i starting in 1). With both of the images in gray-scale, the algorithm divides each image into ten equally spaced parts. Then, the algorithm compares the last division area (LDA) of the C_{i-1} image with the first division area (FDA) of the image C_i . Figure 3 describes the process. The FDA and LDA are used to calculate K_m if this value is below to 0.001 the microscope must conduct autofocus; otherwise, the microscope goes to the next image C_{i+1} . If the value of i is less than the total of images n , the algorithm calculates $i \bmod 6$. If this operation is equal to 0 the algorithm employs equation 8, then if this value is less or equal to 0.5 the microscope conduct autofocus. In the opposite case, the algorithm returns to step 2. The $i \bmod 6$ function was implemented to reduce the use of the regression (decreasing the computational cost) and to avoid the mix between different zones of tissue, in the case of pathological samples.

4 Experiment and analysis

We captured six sequential images to test the validity and feasibility of the algorithm. To analyze the different scenarios in the analysis some of the pictures were out of focus, and others in the transition zones of different tissues.

For the analysis we use two matrices, a state matrix (Tables 1,3) and a comparison matrix (Tables 2,4). In the analysis, we replace a focused image by a blurred one in order to see how much the key-points matches decreased

(we did it on purpose). Two different results were found (Tables 1-4).

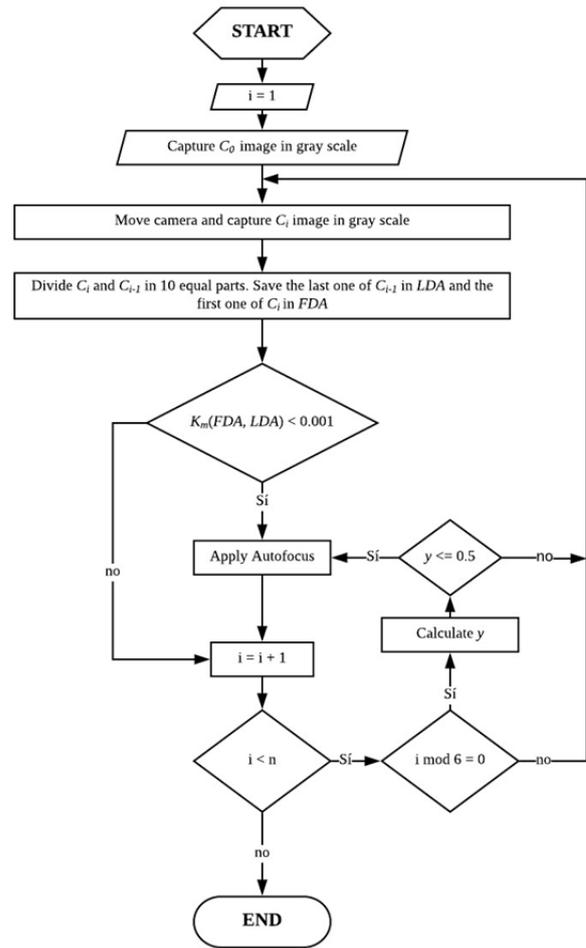


Fig. 2. Algorithm flow chart

- **State matrix:** This matrix consists of 6 columns, image number, focused, first division, last division, first division area and last division area.

- (i) Image number: unique id for the image in process.
- (ii) focused: YES/NO field, describe if the image is focused or not.
- (iii) first division: describe if the first division of the image is dense or has dense tissue in the area, i.e compact.
- (iv) last division: describe if the last division of the image is dense or has dense tissue in the area, i.e compact.
- (v) first division area: total area of the first division image in square pixels.
- (vi) last division area: total area of the last division image in square pixels.

- **Comparison matrix**

- (i) image 1: id of the first image of the pair to compare key-point matches.

Table 1. State matrix 1

State Matrix 1					
Image	Focused	First Division	Last Division	FDA	LDA
0	YES	DENSE	DENSE	128271	128271
1	NO	DENSE	DENSE	128271	128271
2	YES	NO DENSE	DENSE	128271	128271
3	YES	DENSE	DENSE	128271	128271
4	YES	DENSE	DENSE	128271	128271
5	YES	DENSE	DENSE	128271	128271

Table 2. Comparison matrix 1

Comparison Matrix 1					
Img 1	Division	Img 2	Division	Km	Proportion
0	First	1	Last	9	3,5082E-05
1	Last	2	First	4	1,5592E-05
2	Last	3	First	821	0,003200256
3	Last	4	First	1463	0,00570277
4	Last	5	First	891	0,003473116

Table 3. State matrix 2

State Matrix 2					
Image	Focused	First Division	Last Division	FDA	LDA
0	YES	DENSE	DENSE	128271	128271
1	YES	DENSE	DENSE	128271	128271
2	YES	NO DENSE	DENSE	128271	128271
3	YES	DENSE	DENSE	128271	128271
4	YES	DENSE	DENSE	128271	128271
5	YES	DENSE	DENSE	128271	128271

(ii) *division*: division of the image in process (first, second, .. ,tenth or last).

(iii) *image 2*: id of the second image of the pair to compare key-point matches.

(iv) *Km*: number of key-point matches between last division of image 1 and first division of image 2.

(v) *proportion*: the relation between the number of key-point matches and the sum of the areas of the last division area of image 1 and the first division area of image 2.

From the results obtained in Tables 1-4, we conclude that any change in the blur level of the study area will make the algorithm to do the auto-focus process. In Table 2, we can see a proportion colored in red, as well as in Table 4, that means that in Comparison matrix 1, the proportion of key-point matches decreased by 94% with respect to Comparison matrix 2, due to the blur image that we introduce in the first experiment (replacing the focus version by the blur version). In this way, if we continue analyzing in group of six images, we will get a high performance and less computational cost of processing, because we will use the regression model each six times instead of 1 and detect in real time when the camera get out of focus based on this two factors.

Table 4. Comparison matrix 2

Comparison Matrix 2					
Img 1	Division	Img 2	Division	Km	Proportion
0	First	1	Last	152	0,000592496
1	Last	2	First	4	1,5592E-05
2	Last	3	First	821	0,003200256
3	Last	4	First	1463	0,00570277
4	Last	5	First	891	0,003473116

5 Conclusions

In this article was introduced a new algorithm for focusing tissue's scanning in real time with low computational cost and high performance. According to the results shown above, the two factors used to determine whether or not doing auto-focus are highly relevant, they explain the behavior of the focus based on the data of two adjacent images and groups of six ones. There is not a particular restriction in the use of the algorithm, the implementation is quite easy, and the precision is first class.

Acknowledgements

We appreciate the Center for Teaching Excellence CEDU Uninorte because of they bring spaces for the development of meetings and facilitate the creation of the interdisciplinary group. Also, We appreciate Dr. Maria Jose Herrera for her time to evaluate and propose consideration for optimizing the BIEMO method as well as Andrés Deulofeu, Miguel Padilla and Miguel Segrera for their constant support in the development of this article.

References

1. Hossain MS, Nakamura T, Kimura F, Yagi Y, Yamaguchi M. Practical image quality evaluation for whole slide imaging scanner. In Biomedical Imaging and Sensing Conference 2018 Apr 24 (Vol. 10711, p. 107111S). International Society for Optics and Photonics.
2. Andrew J Evans et al. "Implementation of whole slide imaging for clinical purposes: Issues to consider from the perspective of early adopters". In: Archives of pathology & laboratory medicine 141.7 (2017), pp. 944–959.
3. Vosoughi A, Smith PT, Zeitouni JA, Sodeman GM, Jorda M, Gomez-Fernandez C, Garcia-Buitrago M, Petit CK, Chapman JR, Campuzano-Zuluaga G, Rosenberg AE. Frozen section evaluation via dynamic real-time non-robotic Telepathology system in a university Cancer center by resident/faculty cooperation team. Human pathology. 2018 May 1.
4. Tabata K, Mori I, Sasaki T, Itoh T, Shiraiishi T, Yoshimi N, Maeda I, Harada O, Taniyama K, Taniyama D, Watanabe M. Whole-slide imaging at

- primary pathological diagnosis: Validation of whole-slide imaging-based primary pathological diagnosis at twelve Japanese academic institutes. *Pathology international*. 2017 Nov 1;67(11):547-54.
5. Farahani N, Liu Z, Jutt D, Fine JL. Pathologists' Computer-Assisted Diagnosis: A Mock-up of a Prototype Information System to Facilitate Automation of Pathology Sign-out. *Archives of pathology & laboratory medicine*. 2017 Jul 7;141(10):1413-20.
 6. Farahani N, Liu Z, Jutt D, Fine JL. Pathologists' Computer-Assisted Diagnosis: A Mock-up of a Prototype Information System to Facilitate Automation of Pathology Sign-out. *Archives of pathology & laboratory medicine*. 2017 Jul 7;141(10):1413-20.
 7. Lu Q, Liu G, Xiao C, Hu C, Zhang S, Xu RX, Chu K, Xu Q, Smith ZJ. A modular, open-source, slide-scanning microscope for diagnostic applications in resource-constrained settings. *PloS one*. 2018 Mar 15;13(3):e0194063.
 8. Chantziantoniou N, Mukherjee M, Donnelly AD, Pantanowitz L, Austin RM. Digital Applications in Cytopathology: Problems, Rationalizations, and Alternative Approaches. *Acta cytologica*. 2018;62:68-76.
 9. Farahani N, Liu Z, Jutt D, Fine JL. Pathologists' Computer-Assisted Diagnosis: A Mock-up of a Prototype Information System to Facilitate Automation of Pathology Sign-out. *Archives of pathology & laboratory medicine*. 2017 Jul 7;141(10):1413-20.
 10. Jafarian AH, Tasbandi A, Mohamadian Roshan N. Evaluation of photoshop based image analysis in cytologic diagnosis of pleural fluid in comparison with conventional modalities. *Diagnostic cytopathology*. 2018 Apr 19.
 11. Platiša L, Van Brantegem L, Kumcu A, Ducatelle R, Philips W. Influence of study design on digital pathology image quality evaluation: the need to define a clinical task. *Journal of Medical Imaging*. 2017 Jun;4(2):021108
 12. Bueno G, Déniz O, Fernández-Carrobles MD, Váñez N, Salido J. An automated system for whole microscopic image acquisition and analysis. *Microscopy research and technique*. 2014 Sep 1;77(9):697-713.
 13. Chan CC, Huang SK, Chen HH. Enhancement of phase detection for autofocus. In *Image Processing (ICIP)*, 2017 IEEE International Conference on 2017 Sep 17 (pp. 41-45). IEEE.
 14. Wang Y, Feng H, Xu Z, Li Q, Chen Y, Cen M. Fast auto-focus scheme based on optical defocus fitting model. *Journal of Modern Optics*. 2018 Apr 16;65(7):858-68.
 15. Fu G, Cao Y, Lu M. A fast auto-focusing method of microscopic imaging based on an improved MCS algorithm. *Journal of Innovative Optical Health Sciences*. 2015 Sep;8(05):1550020.
 16. Zhang C, Hou G, Zhang Z, Sun Z, Tan T. Efficient auto-refocusing for light field camera. *Pattern Recognition*. 2018 Sep 1;81:176-89.
 17. Wang Y, Feng H, Xu Z, Li Q, Chen Y, Cen M. Fast auto-focus scheme based on optical defocus fitting model. *Journal of Modern Optics*. 2018 Apr 16;65(7):858-68.
 18. Cabazos-Marín AR, Álvarez-Borrego J. Automatic focus and fusion image algorithm using nonlinear correlation: Image quality evaluation. *Optik*. 2018 Jul 1;164:224-42.
 19. Nadeau JL, Bedrossian M, Lindensmith CA. Imaging technologies and strategies for detection of extant extraterrestrial microorganisms. *Advances in Physics: X*. 2018 Jan 1;3(1):1424032.
 20. Choi S, Min SW. Depth extraction using depth of field imaging with tilted retroreflective structure. In *Consumer Electronics (ICCE)*, 2018 IEEE International Conference on 2018 Jan 12 (pp. 1-3). IEEE.
 21. Tian Y, Cui H, Pan Z, Liu J, Yang S, Liu L, Wang W, Li L. Improved three-dimensional reconstruction algorithm from a multifocus microscopic image sequence based on a nonsubsampling wavelet transform. *Applied Optics*. 2018 May 10;57(14):3864-72.
 22. Li Q, Zhong J, Zhong S, Chen Z, Qi S, Zhang H. An self-focusing imaging method for leukocyte recognition. In *Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 2017 10th International Congress on 2017 Oct 14 (pp. 1-5). IEEE.
 23. Ishihara, T., Kitani, K. M., Asakawa, C., & Hirose, M. (2018, March). Deep Radio-Visual Localization. In *Applications of Computer Vision (WACV)*, 2018 IEEE Winter Conference on (pp. 596-605). IEEE.