

Article

Automatic Stomatal Segmentation Based on Delaunay-Rayleigh Frequency Distance

Miguel Carrasco ^{1,*}, **Patricio A. Toledo** ¹, **Ramiro Velázquez** ² and **Odemir M. Bruno** ³¹ Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Av. Diagonal Las Torres, 2700 Santiago, Chile; patricio.toledo@uai.cl² Facultad de Ingeniería Josemaría Escrivá de Balaguer 101, Campus Aguascalientes, Universidad Panamericana, Aguascalientes 20290, Mexico; rvelazquez@ags.up.mx³ Scientific Computing Group, São Carlos Institute of Physics, University of São Paulo, P.O. Box 369, São Carlos, SP 13560-970, Brazil; bruno@ifsc.usp.br

* Correspondence: miguel.carrasco@uai.cl; Tel.: +562-2331-1269

Received: 3 November 2020; Accepted: 17 November 2020; Published: 20 November 2020



Abstract: The CO₂ and water vapor exchange between leaf and atmosphere are relevant for plant physiology. This process is done through the stomata. These structures are fundamental in the study of plants since their properties are linked to the evolutionary process of the plant, as well as its environmental and phytohormonal conditions. Stomatal detection is a complex task due to the noise and morphology of the microscopic images. Although in recent years segmentation algorithms have been developed that automate this process, they all use techniques that explore chromatic characteristics. This research explores a unique feature in plants, which corresponds to the stomatal spatial distribution within the leaf structure. Unlike segmentation techniques based on deep learning tools, we emphasize the search for an optimal threshold level, so that a high percentage of stomata can be detected, independent of the size and shape of the stomata. This last feature has not been reported in the literature, except for those results of geometric structure formation in the salt formation and other biological formations.

Keywords: stomatal segmentation; image segmentation; Delaunay-Rayleigh frequency

1. Introduction

Stomata are the gates through which gas exchange through the leaf takes place, involving carbon dioxide intake and water vapor loss (reviewed in Reference [1–3]). The carbon dioxide intake drives plant growth and productivity [4], while the water vapor loss through a process known as transpiration regulates leaf temperature, nutrient uptake, and root-to-shoot signaling [5,6]. The remaining of the epidermis is covered by an impervious cuticle, with restricted gas exchange properties [7,8]. Gas-exchange capacity depends on stomatal density (i.e., stomata number per unit area), size, and patterning (spacing) [9–11]. Methods for automatic segmentation of stomatal density and size have been earlier presented [12,13], whereas patterning (spacing) is currently performed manually [14].

From a computational point of view, stomatal analysis has been explored mainly on pixel or object analysis. Although localization and processing tools exists, these are limited to search for stomatal position through techniques based on chromatic and morphological combinatorial operations [15–17], texture analysis [16], fractal analysis [18], segmentation using object-oriented method in multiple resolutions [19], and, more recently, by Deep Convolutional Neural Networks [12,13,20–22]. Today, no technique allows full spectrum analysis of different stomatal types since they have a great variation according to species, shape, position, and, in general, noise and technique present in the microscopy

image [23–26]. At large, both the location and the analysis of the stomata require demanding work on the part of the biologists. This task is not simple, since there are great differences in the same species by the microscopy technique used [27]. This research proposes a new segmentation algorithm based on the spatial stomatal statistical distribution. We define a geometric model, revisited in recent years, implying a mechanism present in nature and allowing an ideal distribution for gas transfer [5,14,28].

The study and analysis of stomata is the first step towards phenotypic plasticity and the relationship with climate change, soil use, resource, drought, CO₂ concentration, and light level [4,8,29–33]. Consequently, segmentation process automation and stomatal morphology analysis is a relevant task, especially in high volume analysis of images needed as input for other studies [20]. The latter might not be viable by hand because of analysis complexity. Although we use a specific microscopy database, both for the design and for the analysis of the computational algorithm, our findings have potential use for any kind of stomatal images since it uses as key information stomata's geometric configuration between analyzed species. We generate a manual segmentation in different species of *monocot* and *dicot* families to test our hypothesis. The results section presents the main findings of our algorithm through a fully automatic system for the *Jatoba Hymenaea Courbaril* species. Subsequently, we manually analyze the algorithm proposed in other species in order to understand the existence of an underlying pattern between them. The Materials and Methods' section presents a description of the steps of the algorithm and, finally, the general conclusions of the method.

2. Results

Delaunay-Rayleigh Threshold Binarization (DRTB) algorithm has been evaluated over a set of 31 optic microscopy images from *Jatoba Hymenaea Courbaril* tree localized in Caribbean, Central and South American zones. The images were taken with an Olympus E-330 camera with optic microscope with 3136 × 2352 resolution pixels, at the Biosciences Institute at University of São Paulo, Brazil (USP). The images' set has 3087 regions hand-classified as stomata (see Supplementary Materials section) with its coordinates (*x,y*).

The abaxial surfaces of the epidermis were obtained from three or four leaves (one per individual) of each species. From each collected leaf, a sample of approximately 1 cm² was removed from the leaf's middle region, between the margin and midrib. Dissociation of the leaf epidermis was performed using a 1:1 solution of glacial acetic acid and hydrogen peroxide at 60 °C for 12 h, or the time required to completely decouple the epidermis [34].

The algorithm's performance was evaluated by two statistical indicators: precision (also known as positive predictive value, *PPV*) and recall (or true positive rate, *TPR*). The procedure consists of manually identifying stoma's center (*x,y*) coordinates. Then, the segmentation-algorithm identifies this position and proceeds with a comparison. If the distance between the manual coordinate is less than a threshold with respect to the coordinate of the segmented region, we consider that it has been classified correctly. This case is considered as True Positive (*TP*). In case the algorithm does not find the stoma, even when it is present in the image, we consider this case as False Negative (*FN*). Finally, when the algorithm classifies a region where there is no stoma, we consider this coordinate as False Positive (*FP*) (see Figure 1). Therefore, the evaluation metrics are:

$$PPV = \frac{TP}{TP + FP}, \quad (1)$$

$$TPR = \frac{TP}{TP + FN}. \quad (2)$$

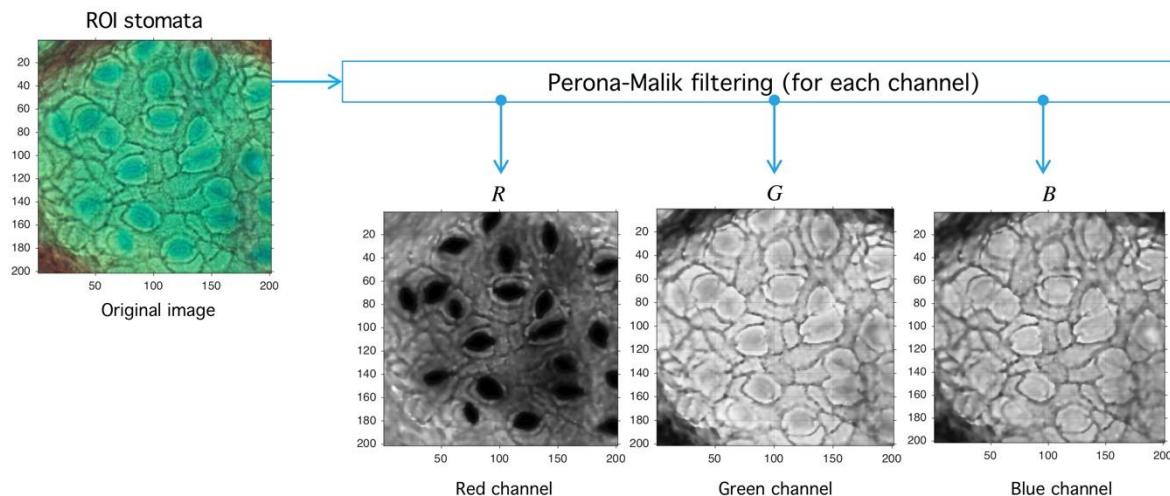


Figure 1. Illustration of precision and recall concepts. True positive (TP) occurs when the distance between manual-coordinate (yellow cross) is less than a threshold with respect to the coordinate of the segmented-region (ellipses). False negative (FN) occurs when the algorithm does not find the stoma, even when it is present in the image. Finally, when the algorithm classifies a region where there is no stoma, we consider this coordinate as False Positive (FP). The ratio $TP/(FP + TP)$ is known as precision (positive predictive value (PPV))), and the ratio $(TP/TP + FN)$ is called recall (true positive rate (TPR))).

From this set, the proposed algorithm is capable of 2752 detections (recall) of $89.14 \pm 8\%$. On the other hand, the number of false positive regions (precision) is $72.8 \pm 10\%$. As shown on Figure 2, the performance is variable with each sample, given structural conditions; however, a high classification rate is achieved as a result of multilevel threshold traveling design (see examples in Figures 3 and 4). In contrast, lower results occur at lower image quality where the stomata are located because they are out of focus. The combined performance of both measures reflects this fact, and it is shown in Figure 3 (specimen).

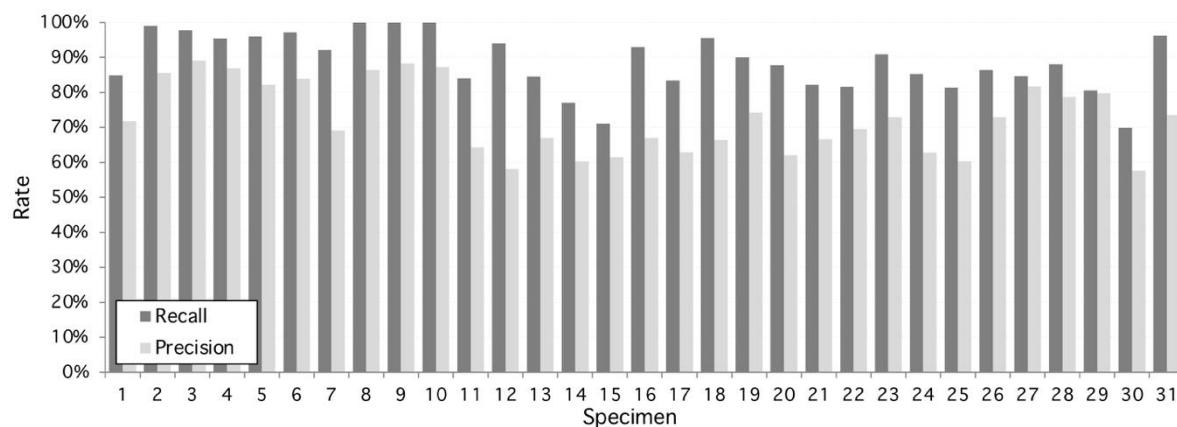


Figure 2. Performance of the proposed algorithm applied to 31 specimens of *Hymenaea Courbaril* (Jatoba Database). Overall recall and precision are 89.14% and 72.8% , respectively. Maximum recall performance is achieved at specimens #8, #9, and #10 with 100% and maximum precision performance is reached at 98% by specimen #3. Worst recall and precision performance is 70% and 58%, respectively, both at specimen #30.

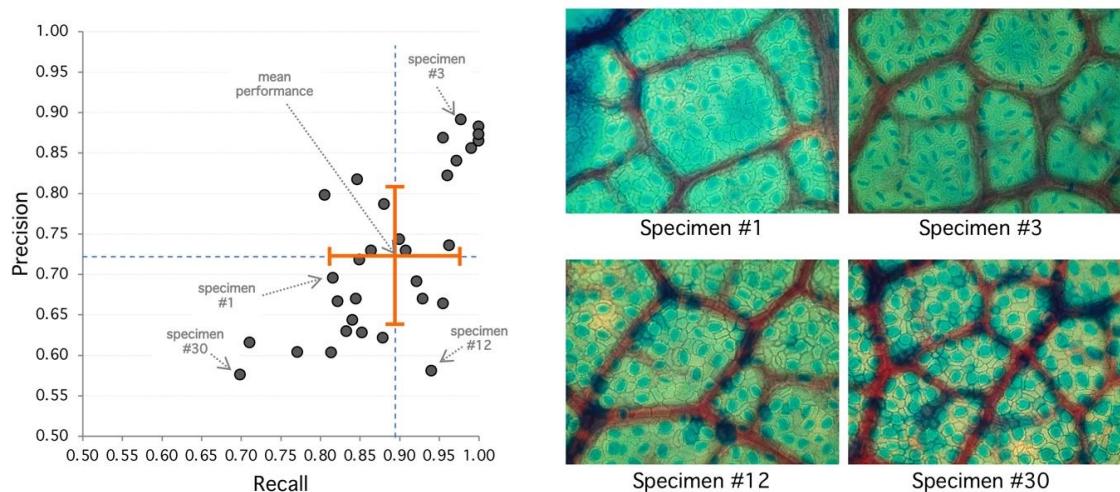


Figure 3. Analysis of selected specimens. (Left) Best performance is achieved at the upper-right area of the diagram with high recall and precision. Worst performance is achieved at the lower-left area. Orange cross indicates mean (recall/performance 83%/72%) and standard deviation (8%/10%). (Right) Specimen #1 represents an average performance (85%/72%) with clear regions but high diffusion. Specimen #3 has 98%/89% detection performance, and stomata shows sharp borders with clear regions leading to good detection metrics. Specimen #12 shows 94%/58% with high recall and lower precision, explained by diffuse borders and poor region definition. Specimen #30 has the poorest performance with 70%/58% with very diffuse borders and poor regions with high false positive rate.

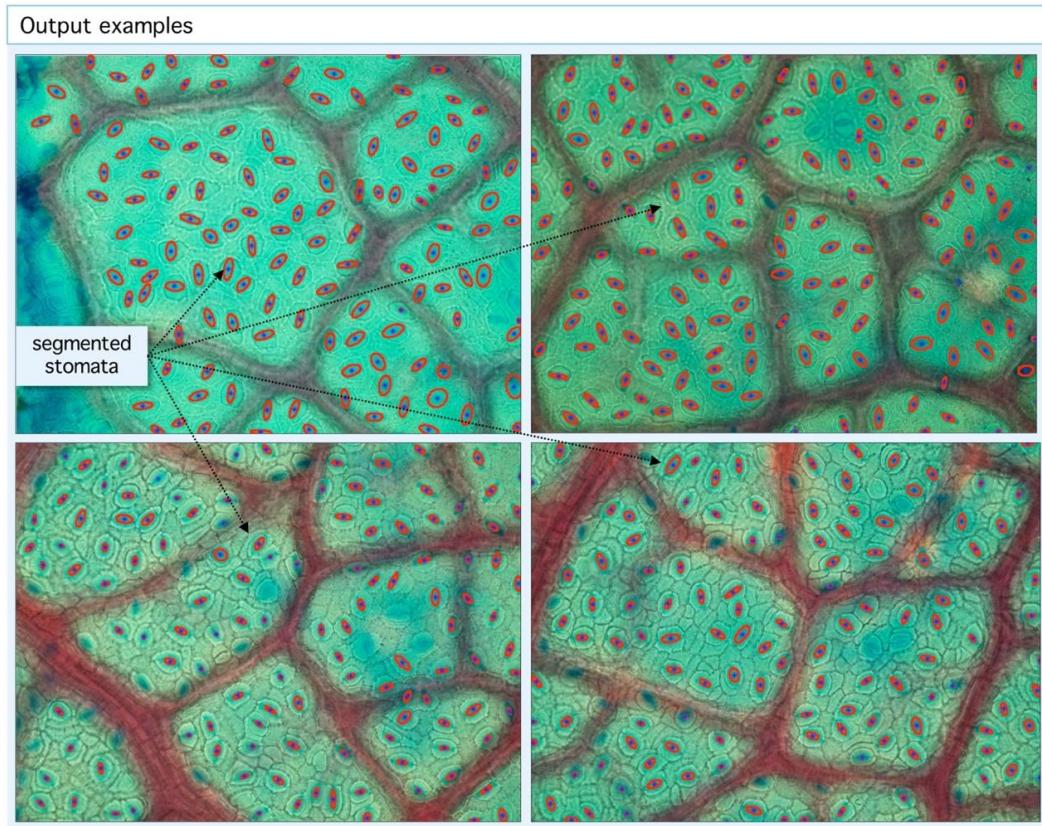


Figure 4. Example of segmentation output. Our algorithm is able to detect stomatal centroids (blue dots) and segmented areas (red ellipses). With the geometric information from centroid coordinates, statistics and tessellations are built.

3. Discussion

Previous results reveal a high precision rate in the detection of *Hymenaea Courbaril* stomata. However, is this distribution similar on other species? To answer this question, we analyzed seven species with different microscopy techniques. Experimental results show that not only *Hymenaea Courbaril* has a stable distribution, we applied a manual segmentation to 2842 stomata with 12 different configurations shown in Figure 5 and described in Table 1. Despite morphological and chromatic differences, the spatial distributions are similar, as it is shown in Figure 4. These results are consistent with the findings found by Croxdale [9], where the relationship between area and number is analyzed in the first instance. This result is relevant since it provides statistical evidence that would allow us to design an optimal algorithm for the relative distance between stomata, which is consistent with Peat & Fitter [35] developments. The only requirement is a preprocess with efficient acceptance/rejection range for stomatal segmentation. This last task could be performed by any segmentation algorithm, such as those mentioned above. It is known that environmental conditions during growth may also affect stomatal dispersion pattern among species [36,37]. Based on this background, stomatal spacing may be employed as a marker for studying plant adaptation to growth environment [38]. Next, we present a brief discussion of manual segmentation in different types of plants of the Monocot family, compared to the automatic evaluation of the Dicot family.

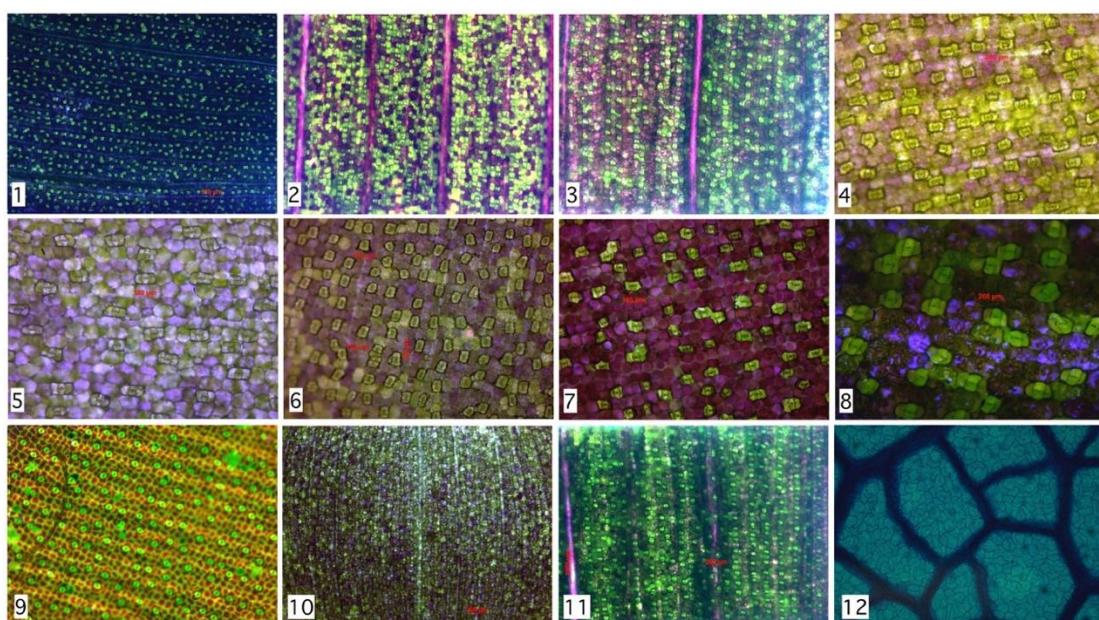


Figure 5. Hand segmentation to 2842 stomata within 12 species as described in Table 1. Despite morphological and chromatically differences, the spatial distributions are similar to the example shown in Figure 4 (*Hymenaea Courbaril*).

Table 1. Stability of Rayleigh parameter for various species. From 2842 regions analyzed, a stability range $3 < \sigma < 8$ might be proposed. Maximum Rayleigh parameter is 11.1 (*Tradescantia Pallida*) minimum value is 3.3 (*Tradescantia Zebrina*). Family name: C: Commelinaceae, M: Marantaceae, F: Fabaceae.

#	Species Analyzed	Stomata's Number	RMSD	Rayleigh Parameter	Group	Family
1	<i>Tradescantia Zebrina</i>	753	0.00832	3.3	Monocot	C
2	<i>Tradescantia Pallida</i> —under 24 h of light	394	0.00983	3.6	Monocot	C
3	<i>Tradescantia Pallida</i> —in natural condition	245	0.03327	7.1	Monocot	C
4	<i>Callisia repens</i>	65	0.05358	8.8	Monocot	C
5	<i>Callisia repens</i>	29	0.12969	9.7	Monocot	C
6	<i>Callisia repens</i>	104	0.02037	4.3	Monocot	C
7	<i>Tradescantia Zebrina</i>	69	0.05905	8.3	Monocot	C
8	<i>Tradescantia Pallid</i>	25	0.13832	11.1	Monocot	C
9	<i>Ctenanthe Oppenheimiana</i>	138	0.02531	7.1	Monocot	M
10	<i>Calisia repens</i> —using stereoscope 15×	586	0.01578	3.3	Monocot	C
11	<i>Tradescantia Pallida</i> using stereoscope 15×	295	0.01902	5.4	Monocot	C
12	<i>Hymenaea Courbaril</i>	139	0.02420	6.1	Dicot	F
Total Regions		2842	$\mu = 0.04473$			

3.1. Manual Segmentation

The automatic algorithm evaluation gives $TPR = 89.14 \pm 8\%$ and $PPV = 72.8 \pm 10\%$. Those results are relevant given image type. Our results are statistically valid, considering the total number of stomata studied. To assess if this relation is relevant for other species, we evaluate 12 different configurations on 7 different plant species. Despite interspecies variations, a relevant result is the Root-Mean-Square Deviation (RMSD) falling with segmented region number. In terms of distribution parameters, they are stable in the range $3 < \sigma < 8$ as long as region number is lower than 100. As future work, it remains to design an optimal search algorithm based on a distribution parameter in the appropriate range, depending on the type of species, the analysis of second (autocovariance) and higher order moments of geometric characteristics, and the cross covariances with other relevant properties captured by microscopes, like color.

According to the analyzed data, the segmented region number affects the algorithm performance. In general, the higher the segmented region number the lower RMSD between Rayleigh distribution and frequency histogram; which confirms the true nature of the spatial distribution of stomata (see Figure 6). Moreover, Rayleigh distribution ideal parameters (that minimize RMSD) have a narrow range between $3 < \sigma < 8$; see Figure 7. For each species, this parameter is different; however, results tend to cluster as the segmented region number reaches 100.

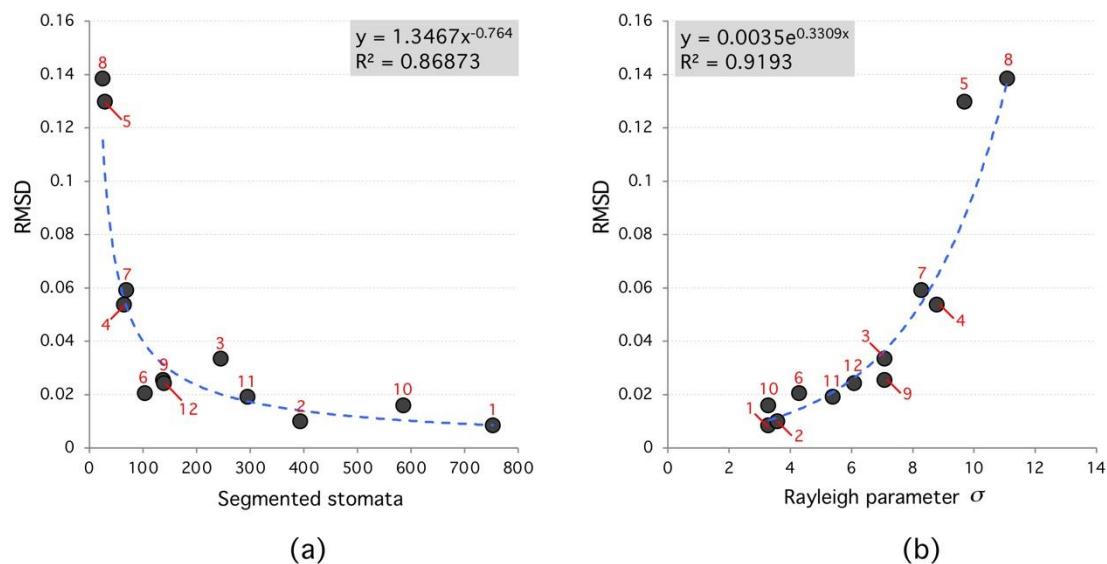


Figure 6. Root-Mean-Square Deviation (RMSD) versus segmented stomata number and Rayleigh parameter. (a), the RMSD decreases as a power law with number of segmented stomata. (b), RMSD increases exponentially with Rayleigh parameter. High RMSD species (red labels) tend to have lower number of segmented regions and higher Rayleigh parameter.

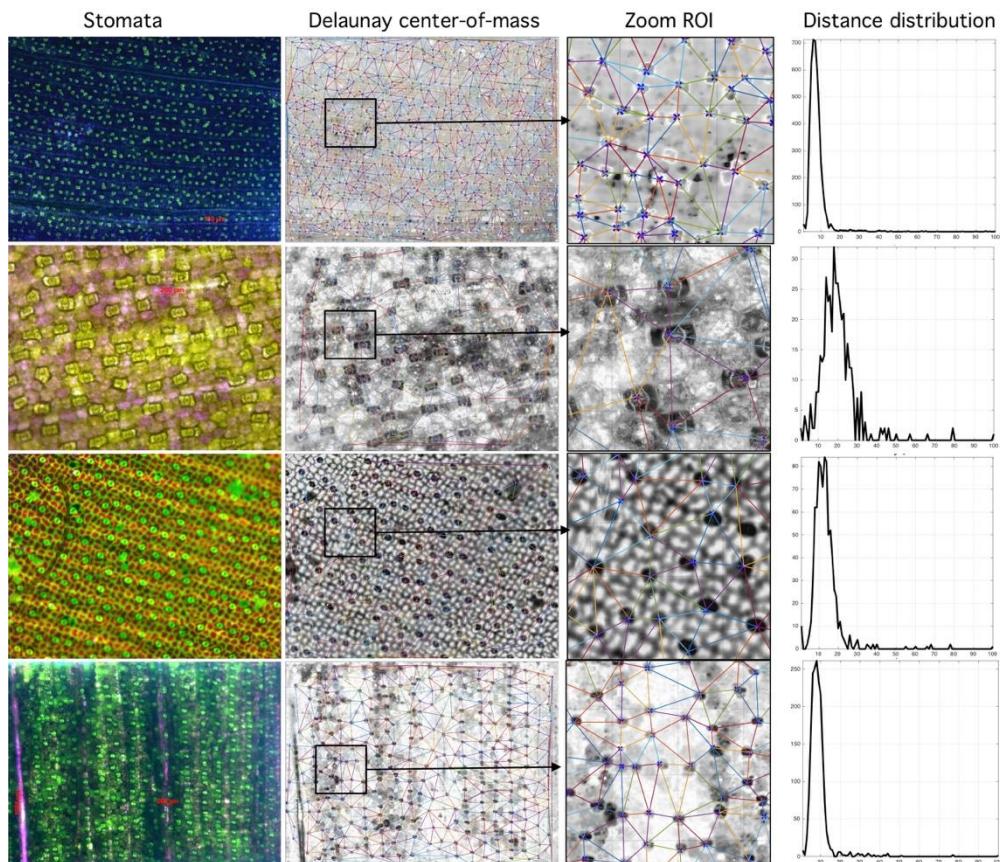


Figure 7. Relationship between segmented region and Rayleigh distribution histogram. First column (stomata) shows different species analyzed. Second column (Delaunay center-of-mass) shows mass centers and corresponding tessellations. Third column (Zoom Region-of-Interest, ROI) shows region of interest. Fourth column (Distance distribution) shows sensibility of histogram to spatial distribution (more at <https://github.com/mlacarrasco/drtb/tree/main/stomatasDB/output>).

3.2. Defocus Stomata in Automatic Mode

In the automatic segmentation mode, we observe relevant differences in the maximum and minimum performance of the proposed algorithm. When analyzing these differences in detail, we observe that they are due to the lack of focus of the stomata in some samples of the study (Figure 8).

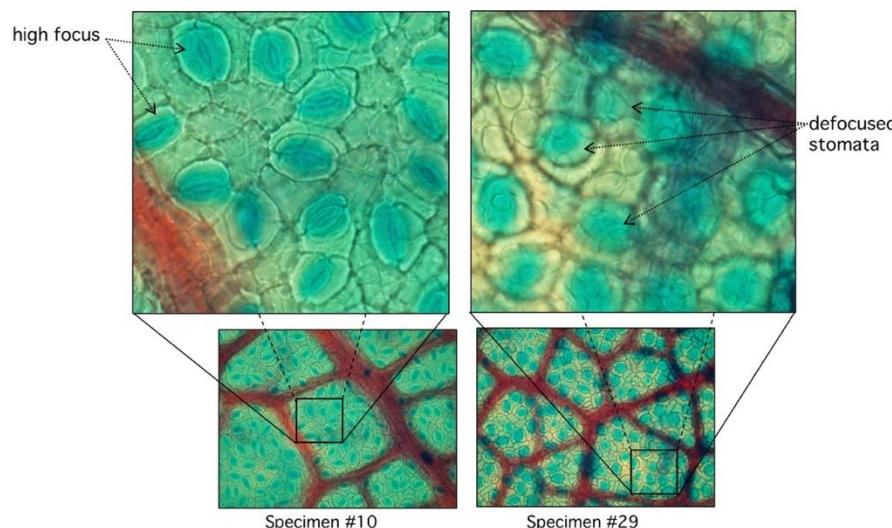


Figure 8. Comparison between high and low focused stomata. In specimen #10, stomata are clearly visible with sharp edges with very low RMSD (see Figure 7). Specimen #29 has diffused edges.

3.3. Main Advantages and Disadvantages

As a first advantage, we highlight that our method is based on the stomatal geometric properties, easily verifiable, such as centroids spatial distribution; which does not depend on the shape and size of the stomata. Second, the image-preprocessing method is a well-established procedure in the field of imaging, and it is inspired by the property of diffusion, which is one of the fundamental mechanisms of stomatal physiology. Third, stomatal geometry is described by a Delaunay tessellation, which is a technic known to captures diffusive-processes, such as those occurring at the cellular level. Fourth, the use of the median as the main statistic decreases outliers-weight, which increases robustness. Regarding the disadvantages, our proposal might be classified as standard as it was not implemented through a deep-learning scheme (although it would be possible to do so). Despite our efforts to capture a large number of samples, we recognize that the number of species is limited, although high variability between samples can be observed. Regarding the distribution's discrimination method, it is known that RMDS has some bias, which could be avoided with an Akaike scheme or similar. The preprocessing phase of our proposal is designed for *Hymenaea Courbaril*, as future work, and this process might be automated in such a way that it is possible to accept segmentation in other types of stomata.

4. Materials and Methods

The natural way to deal with structural complexity found in stomata images is noise analysis. The latter is associated, by large, to simplification and information-reduction processes, like anisotropic diffusion, wavelet transform techniques, and nonlinear, statistical, or adaptive filters [39–41]. Even when image segmentation is possible, for example, by morphological processing, the use of geometric properties is desirable. Therefore, we propose a novel segmentation method that seeks an optimal binarization threshold through Rayleigh-distribution distance-minimization. This idea was inspired by Staff et al. [42], where they analyzed stomatal pore center coordinates, but, in this case, only areas of the inner triangles and their angles were reported. In Reference [43], a statistical mechanics explanation of this phenomenon is given through particle interaction, in this case, the distance between stomata, as a means of regulation and state of the tissue in general. An algorithm summary is shown in Figure 9.

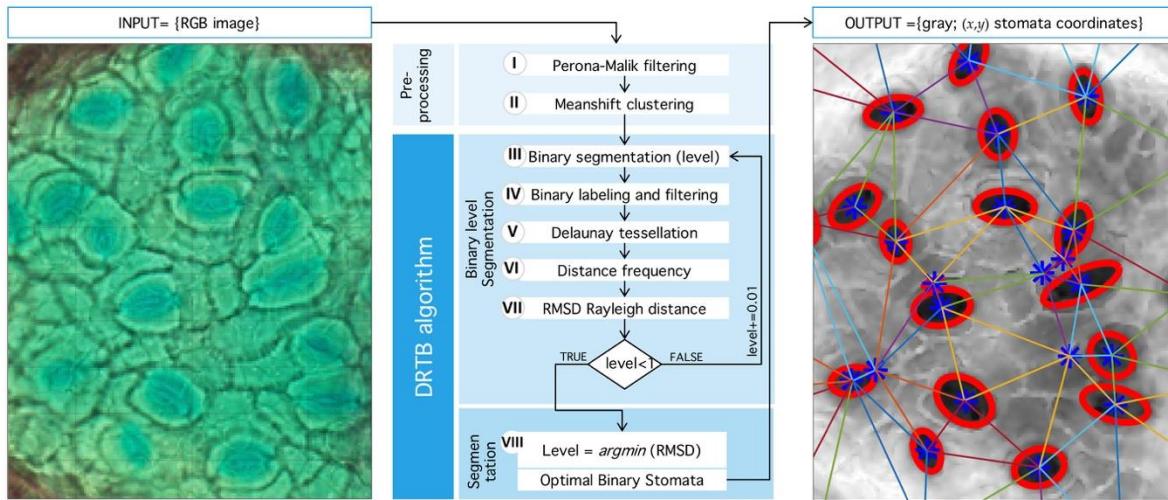


Figure 9. General process implemented in detection algorithm. (Left) Input RGB image from microscope. (Center) Sequence of processing: First, Preprocessing stage: Perona-Malik filtering followed by Meanshift clustering. Second, segmentation algorithm proposed (DRTB): binarization, labeling, tessellation, distance analysis, and segmentation stage plus optimal leveling. (Right) Final output image with segmented region and centroid based tessellation.

4.1. Preprocessing

Standard noise-reduction and boundary-preservation techniques are used given the common noise properties present in most stomata images. We propose a two-step methodology based on Perona-Malik (PM) diffusion and Meanshift-Hadamard intensity-reduction.

Step I. Perona-Malik: The first step is PM iterative filter application [44]. This algorithm reduces noise level preserving and, at the same time, boundary structure through diffusion adaptation [45], meaning a lower diffusion constant near boundaries, and higher otherwise. Being an interactive filter, level simplification of image structures is possible, which is a leading characteristic in stomata search. Filter design criteria defined by Perona and Malik [46] are causality, immediate localization, and piecewise smoothing. The last three properties are relevant in our problem given that stomata have well defined structures hard to segment, mainly by their surrounding structures. As for causality, PM filters all regions classified as noise. Immediate localization allows sharp boundaries, despite scale changes, and the most useful property here is piecewise smoothing since it allows smaller structures to collapse into larger ones sharing a visual similarity criterion.

Let P be an image defined over the space $M_{n,m}^3(\mathbb{N})$ of arrays with n rows and m columns with pixels in the natural numbers. Depending on the color-space, {Red, Green, Blue} (RGB) or {Hue, Saturation, Value} (HSV) superscripts are used to refer the components. The PM filter ∂_t defined from $M_{n,m}(\mathbb{N})$ into itself is a solution of the heat-equation parameterized by t representing the number of times the filter is applied (time in the original partial differential equation context) and a diffusion parameter controlling the noise level (see Reference [46] for more details) (see Figure 10). As a result of PM filter application over RGB space, the red channel increases separation levels with respect to background, and this result is relevant as the application of PM filter improves stomata profile detection, as it is exemplified in Figure 11.

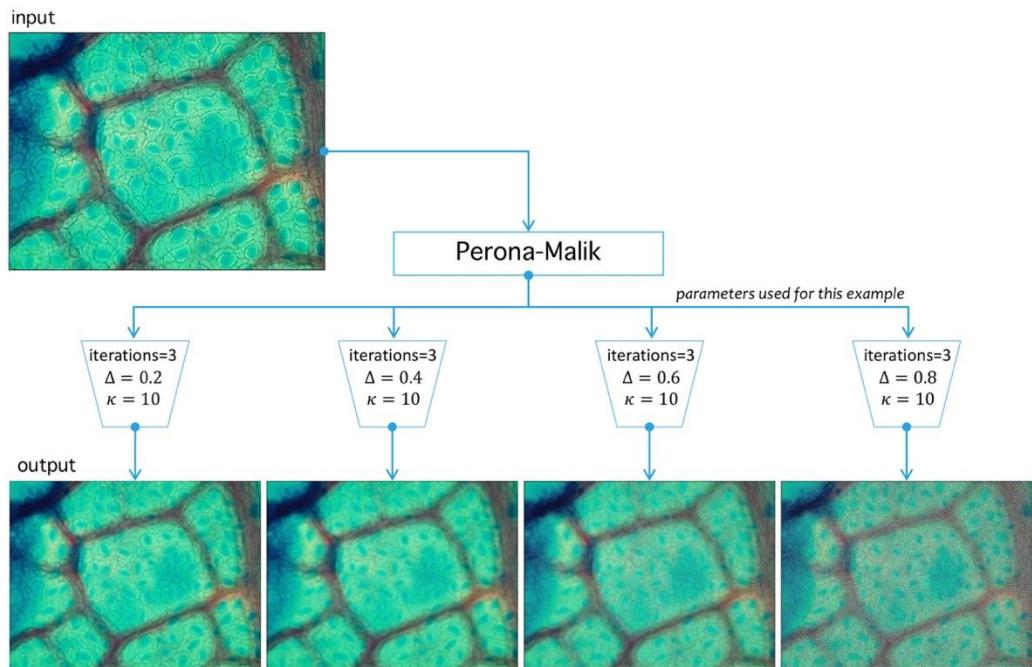


Figure 10. Image preprocessing with Perona-Malik (PM) filtering. PM has three main parameters: Δ , which represents the diffusion level; κ , which represents an advance-step (time in the original PDE framework); and the iteration number (advance-step times iteration number is the total time in the PDE framework). Four examples are shown to exemplify diffusion action over an image; higher Δ parameter means more diffuse image.

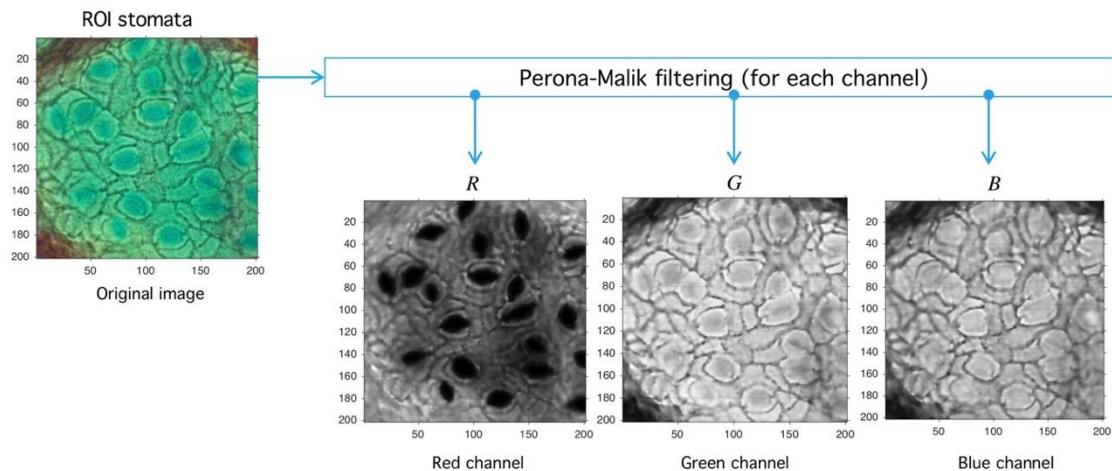


Figure 11. Image preprocessing with Perona-Malik (PM) filtering. After the PM usage (see Figure 10) the image is decomposed into red, green, and blue channels (RGB decomposition) with a standard routine (see shared code at GitHub). The red channel P^R is used later in the next process.

Step II. Meanshift-Hadamard: The second step is the application of Meanshift M operator over the red channel to obtain an image with a reduced intensity level but with sharp defined boundaries, which eases the segmentation process. This step uses an unsupervised clustering algorithm over the intensity-space through a minimization process between each energy cluster [47]. With the aim at improving signal-background splitting, a Hadamard division [48] operator between saturation (HSV space) and red (RGB space) channels is applied. This operation enhances signal splitting, as

it can be seen in Figure 12. The final operator is the consecutive application of PM, Meanshift and Hadamard division (3):

$$F = P^S \oslash M(\partial_t P^R), \quad (3)$$

where F is the resultant preprocessed image.

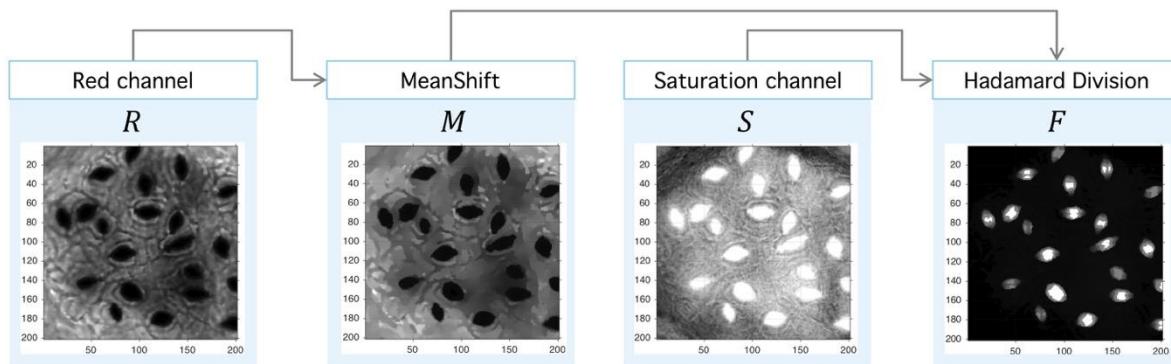


Figure 12. Image preprocessing with Meanshift-Hadamard division. After the PM usage (see Figure 10), the filtered red-channel $\partial_t P^R$ is used as input for Meanshift; the output $M(\partial_t P^R)$ is combined with the saturation channel of the original image P^S through the Hadamard division, and this later image is subject to clustering.

This two-step algorithm might be modified based upon stomata-image type. However, the proposed technique is useful as segmentation method as it helps with the post-labeling process, which is a main focus of the research.

4.2. Delaunay-Rayleigh Threshold Binarization (DRTB Algorithm)

Our main algorithm uses stomata spatial localization as a method to find an optimal binarization threshold. As input, it uses the grey-scale image F and outputs an optimal umbralization level such that error in Delaunay-distance frequencies be minimum with respect to an ideal Rayleigh distribution [28] (Figure 13). Next, we present detailed description of DRTB algorithm phases.

Step III. Threshold level binarization: First, phase is binary-image building given a threshold. Literature gives various binarization techniques; however, most of them use as input information the relationship between intensity levels of the image. Instead, our algorithm explores a geometric-analysis of image embedded structures as optimal binarization-level search method. Initially, all binarization level-space values are traversed by normalization of intensity-levels in $[0, 1]$ domain and then we explore such space. Binarization operator $\mathcal{H}_l(F)$ at level l is defined as:

$$\mathcal{H}_l(F) = \begin{cases} 0, & F \leq l, \\ 1, & F > l \end{cases} \quad (4)$$

which has the same size of F but with values in $[0, 1]$. All pixels with values higher than l are given the value 1 and the rest 0. From the image $\mathcal{H}_l(F)$, we can obtain a series of n_R regions with areas given by $r = [r_1, r_2, r_3, \dots, r_{n_R}]$.

Step IV. Binary labeling and filtering: Valid regions are those whose area be not an outlier. We propose [49] methodology for outlier detection. First, median absolute deviation (MAD) is determined as:

$$MAD = b \text{ MEDIAN}(r - \text{MEDIAN}(r)), \quad (5)$$

measured in pixels with value 1, and $b = 1.4826$ is a normality constant. Area outlier detection criterion is defined as:

$$\left| \frac{r_i - \text{MEDIAN}(r)}{\text{MAD}} \right| < 3, \quad i = 1, \dots, n_R. \quad (6)$$

This means that all regions not fulfilling the criterion will be discarded in the following analysis.

Step V. Delaunay tessellation: This phase consists in the generation of a triangular tessellation for all regions defined as segmented inliers. Mass centers for each i -th region are determined through central moment estimation [50].

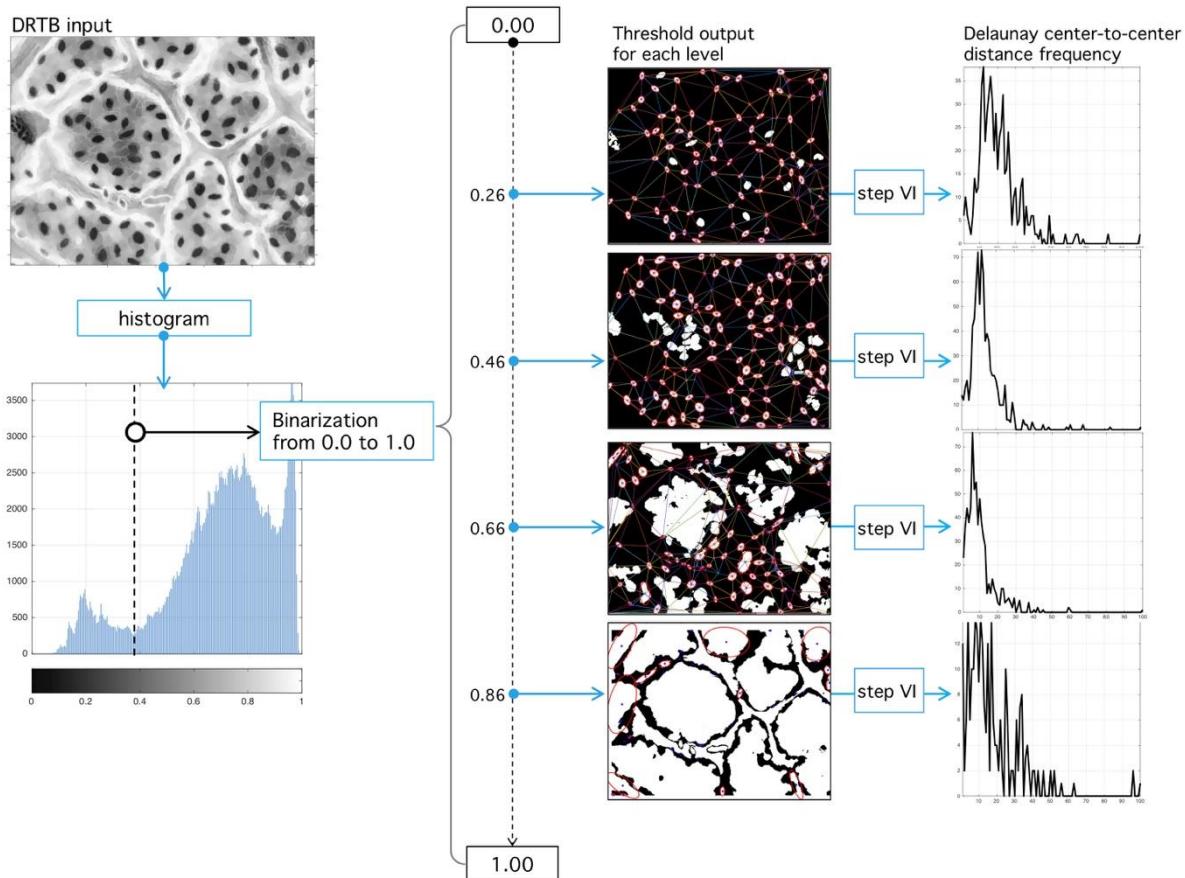


Figure 13. Binary segmentation and Delaunay tessellation. (Left) Grey-scale image (see Figure 12) is subject to the process of binarization at the l -level. (Right) Different binarization threshold results in different tessellations. The best tessellation is found through an optimization procedure applied over the RMSD, with respect to an ideal Rayleigh distribution and the empirical histogram.

Let m_i the mass center corresponding to i -th region (see Figure 14). A Delaunay algorithm uses an iterative process to generate a triangular tessellation considering a given region. The output is a set of 3 vertices for each triangle in terms of its coordinates $[m_i^1, m_i^2, m_i^3]$; given a two-dimensional set of points, it is easy to build a tessellation joining the points (vertices) with lines (edges) in a triangulation. For each triangle, a circumcircle might be drawn. A Delaunay tessellation is such a triangulation that no vertex from the tessellation is found inside the circles. Figure 14 shows a triangulation composed by 17 vertices. For each triangle, its circumcircle is drawn. A circle meets three vertices, and no vertex is inside. Delaunay tessellations tend to maximize the smallest inner angle possible. Diffusion operators complies with a maximum principle, which are relevant in existence and uniqueness of solutions under this operation; it is possible to extend these properties to numerical solution calculations obtained over Delaunay tessellations, mainly because of the geometric regularity

imposed over the angle distribution [51]. This kind of tessellation are used in partial differential equation analysis and morphogenesis studies, and we expect it might play a role given the importance of gas-diffusion in stomatal physiology.

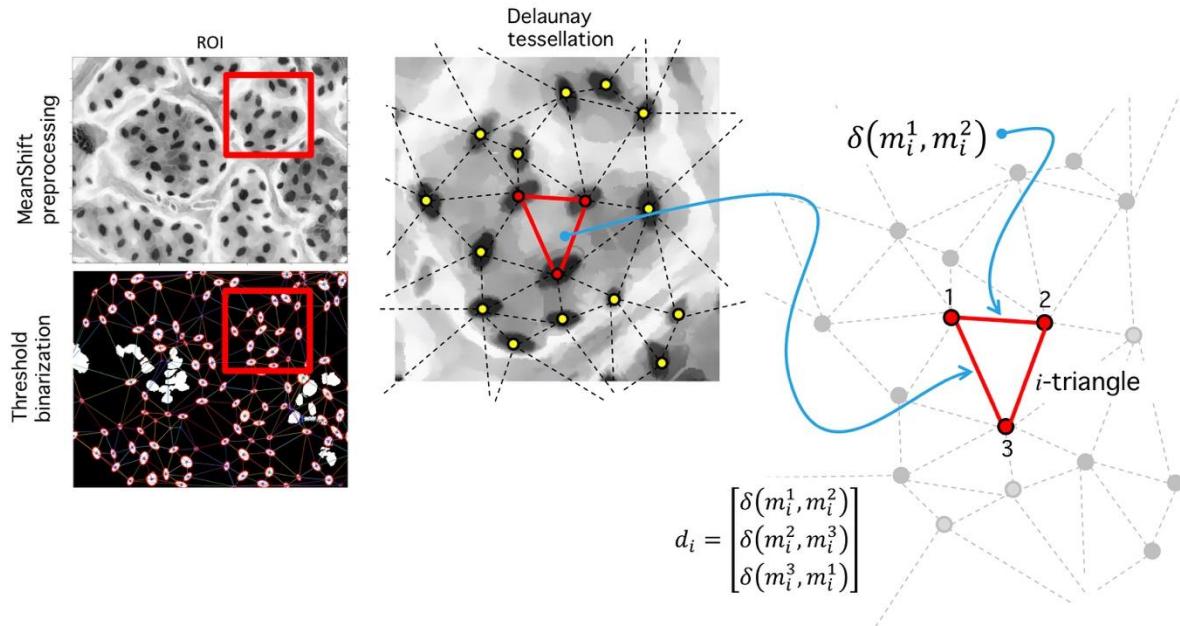


Figure 14. Delaunay tessellation over ROI (red square). (Left) Meanshift image and binarized image. (Center) Zoom over the ROI shows the segmented regions and its centroids (yellow dots). A Delaunay tessellation is built from centroids; note that centroid positions are dependent on the binarization level l . (Right) After the centroids are fixed and the tessellation is built, the set d_i of all found distances are used to calculate a histogram, which is used for RMSD analysis.

Step VI. Distance analysis: As result of the last process, a set of triangles in obtained (Figure 14), each vertex an inlier-region. We analyze the set of all distances:

$$d_i = [\delta(m_i^1, m_i^2), \delta(m_i^2, m_i^3), \delta(m_i^3, m_i^1)], i = 1, 2, 3, \dots, n, \quad (7)$$

where $\delta(\cdot, \cdot)$ is the Euclidian distance between vertices, and n is the number of tessellated triangles. Let $d = [d_1, d_2, \dots, d_i, \dots, d_{3n}]$, a vector composed of all edges, and let n_i the number of occurrences of distance d_i ; then:

$$p(d_i|l) = \frac{n_i}{\sum_{i=1}^{3n} n_i} \quad (8)$$

is the occurrence probability of d_i in the vector d for a given threshold l obtained from the binarization process.

Step VII. Error estimation: Consider the Rayleigh distribution of parameter σ :

$$\mathcal{R}(d_i|\sigma) = \frac{d_i}{\sigma^2} e^{-\frac{d_i^2}{2\sigma^2}}; \quad (9)$$

our experimental results point at its presence in all analyzed images (manual or automatic process); see Reference [28] for other applications of this probability in natural formations. Initial calibration points at $\sigma = 2$ and $t > 0$. By root mean-square (RMS) deviation (RMSD) minimization, we assess the proximity of $p(d_i|l)$ to $\mathcal{R}(d_i|\sigma)$ as a function of threshold l and σ -parameter:

$$RMSD(\sigma, l) = \sqrt{\frac{\sum_{i=1}^{3n} (\mathcal{R}(d_i|\sigma) - p(d_i|l))^2}{3n}}. \quad (10)$$

See the Discussion for comments on the significance of the parameter σ .

Step VIII. Optimal level selection: The ideal stomata distribution is associated with the minimum RMSD. The optimal parameter \hat{l} happens when such distance is minimum:

$$\hat{l} = \underset{0 < l < 100}{\operatorname{argmin}} (RMSD_{\sigma,l}). \quad (11)$$

Once the best level is found, the image is finally binarized. An example is shown in Figure 15.

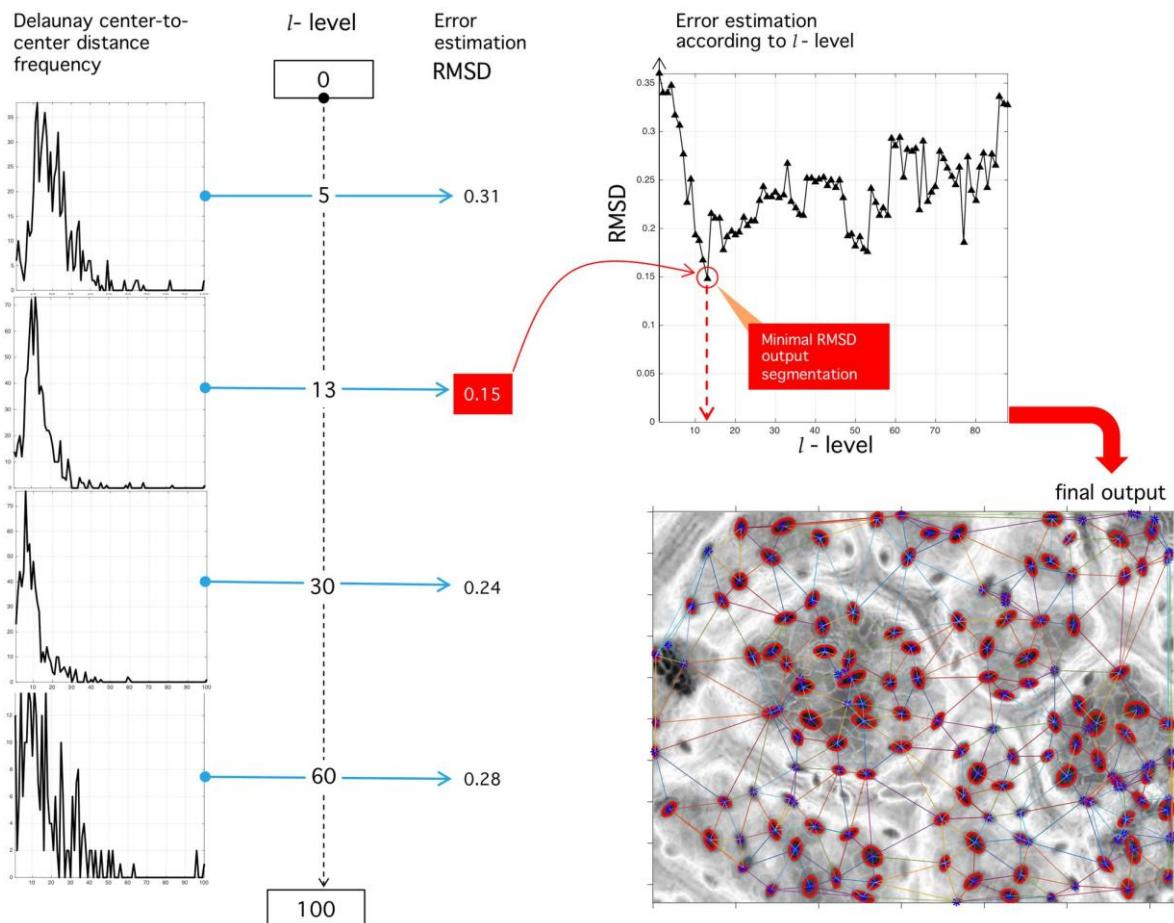


Figure 15. Sensibility analysis of binarization level and RMSD optimization procedure. (Left) Different histograms corresponding to different binarization levels. (Upper right) RMSD sensibility with respect to binarization level. The level of binarization minimizing the RMSD error is used for the final tessellation. (Lower right) Final output of the proposed algorithm with the positions of stomata fixed at tessellation nodes.

5. Conclusions

We showed an optimal segmentation algorithm based on stomata-position frequency-analysis through Delaunay tessellation and Rayleigh distribution. We tested the algorithm with *Hymenaea Courbaril* before a preprocess technique aimed at noise reduction. Our proposal is centered around a Delaunay-Rayleigh Threshold Binarization (DRTB algorithm) that allows an optimal binarization threshold with a posterior segmentation. The automatic segmentation shows the presence of a stable Rayleigh distribution, despite image differences. This result is relevant given that DRTB algorithm might be applied to other species, as was shown in the manual experimental phase with seven different species.

Supplementary Materials: Our solution can be accessed online at <https://github.com/mlacarrasco/drtb>, and images database are available online at <https://github.com/mlacarrasco/drtb/tree/main/database>.

Author Contributions: M.C. and P.A.T. performed manuscript conceptualization; R.V. performed formal analysis; P.A.T. analyzed the data; M.C., P.A.T., and R.V. wrote, edited the manuscript; O.M.B. performed project administration and funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been partially financed by STICAmSud project number 14STIC-12.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fanourakis, D.; Nikoloudakis, N.; Pappi, P.; Markakis, E.; Doupis, G.; Charova, S.; Delis, C.; Tsaniklidis, G. The Role of Proteases in Determining Stomatal Development and Tuning Pore Aperture: A Review. *Plants* **2020**, *9*, 340. [[CrossRef](#)] [[PubMed](#)]
2. Fanourakis, D.; Aliniaiefard, S.; Sellin, A.; Giday, H.; Körner, O.; Nejad, A.R.; Delis, C.; Bouranis, D.; Koubouris, G.; Kambourakis, E.; et al. Stomatal behavior following mid- or long-term exposure to high relative air humidity: A review. *Plant Physiol. Biochem.* **2020**, *153*, 92–105. [[CrossRef](#)] [[PubMed](#)]
3. Pillitteri, L.J.; Torii, K.U. Mechanisms of Stomatal Development. *Annu. Rev. Plant Biol.* **2012**, *63*, 591–614. [[CrossRef](#)] [[PubMed](#)]
4. Zhu, Z.; Piao, S.; Myneni, R.B.; Huang, M.; Zeng, Z.; Canadell, J.G.; Ciais, P.; Sitch, S.; Friedlingstein, P.; Arneth, A.; et al. Greening of the Earth and its drivers. *Nat. Clim. Chang.* **2016**, *6*, 791–795. [[CrossRef](#)]
5. Franks, P.J.; Drake, P.L.; Beerling, D.J. Plasticity in maximum stomatal conductance constrained by negative correlation between stomatal size and density: An analysis using *Eucalyptus globulus*. *Plant Cell Environ.* **2009**, *32*, 1737–1748. [[CrossRef](#)]
6. Zhang, L.; Liu, L.; Zhao, H.; Jiang, Z.; Cai, J. Differences in Near Isohydric and Anisohydric Behavior of Contrasting Poplar Hybrids (I-101 (*Populus alba* L.) × 84K (*Populus alba* L. × *Populus glandulosa* Uyeki)) under Drought-Rehydration Treatments. *Forests* **2020**, *11*, 402. [[CrossRef](#)]
7. Fanourakis, D.; Hyldgaard, B.; Giday, H.; Aulik, I.; Bouranis, D.; Körner, O.; Ottosen, C.-O. Stomatal anatomy and closing ability is affected by supplementary light intensity in rose (*Rosa hybrida* L.). *Hortic. Sci.* **2019**, *46*, 81–89. [[CrossRef](#)]
8. Niu, Y.; Ahammed, G.J.; Tang, C.; Guo, L.; Yu, J. Physiological and Transcriptome Responses to Combinations of Elevated CO₂ and Magnesium in *Arabidopsis thaliana*. *PLoS ONE* **2016**, *11*, e0149301. [[CrossRef](#)]
9. Croxdale, J.L. Stomatal patterning in angiosperms. *Am. J. Bot.* **2000**, *87*, 1069–1080. [[CrossRef](#)]
10. Franks, P.J.; Beerling, D.J. Maximum leaf conductance driven by CO₂ effects on stomatal size and density over geologic time. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 10343–10347. [[CrossRef](#)]
11. Dow, G.J.; Berry, J.A.; Bergmann, D.C. The physiological importance of developmental mechanisms that enforce proper stomatal spacing in *Arabidopsis thaliana*. *New Phytol.* **2014**, *201*, 1205–1217. [[CrossRef](#)]
12. Li, K.; Huang, J.; Song, W.; Wang, J.; Lv, S.; Wang, X. Automatic segmentation and measurement methods of living stomata of plants based on the CV model. *Plant Methods* **2019**, *15*, 1–12. [[CrossRef](#)] [[PubMed](#)]
13. Song, W.; Li, J.; Li, K.; Chen, J.; Huang, J. An Automatic Method for Stomatal Pore Detection and Measurement in Microscope Images of Plant Leaf Based on a Convolutional Neural Network Model. *Forests* **2020**, *11*, 954. [[CrossRef](#)]
14. Naulin, P.I.; Valenzuela, G.; Estay, S.A. Size matters: Point pattern analysis biases the estimation of spatial properties of stomata distribution. *New Phytol.* **2016**, *213*, 1956–1960. [[CrossRef](#)] [[PubMed](#)]
15. Bhaiswar, N.; Dixit, D.V.V.; Student, P.G. A Review: Methods of Automatic Stomata Detection and Counting Through Microscopic Images of a Leaf. *IJIRSET* **2007**, *5*, 6.
16. Da Silva, N.R.; Oliveira, M.W.D.S.; Filho, H.A.D.A.; Pinheiro, L.F.S.; Rossatto, D.R.; Kolb, R.M.; Bruno, O.M. Leaf epidermis images for robust identification of plants. *Sci. Rep.* **2016**, *6*, 25994. [[CrossRef](#)] [[PubMed](#)]
17. Jayakody, H.; Liu, S.; Whitty, M.; Petrie, P. Microscope image based fully automated stomata detection and pore measurement method for grapevines. *Plant Methods* **2017**, *13*, 1–12. [[CrossRef](#)] [[PubMed](#)]
18. Florindo, J.B.; Landini, G.; Filho, H.A.; Bruno, O.M. Analysis of Stomata Distribution Patterns for Quantification of the Foliar Plasticity of *Tradescantia Zebrina*. *J. Phys. Conf. Ser.* **2015**, *633*, 012113. [[CrossRef](#)]

19. Zhu, J.; Yu, Q.; Xu, C.; Li, J.; Qin, G. Rapid Estimation of Stomatal Density and Stomatal Area of Plant Leaves Based on Object-Oriented Classification and Its Ecological Trade-Off Strategy Analysis. *Forests* **2018**, *9*, 616. [[CrossRef](#)]
20. Fetter, K.C.; Eberhardt, S.; Barclay, R.S.; Wing, S.; Keller, S.R. StomataCounter: A neural network for automatic stomata identification and counting. *New Phytol.* **2019**, *223*, 1671–1681. [[CrossRef](#)]
21. Toda, Y.; Toh, S.; Bourdais, G.; Robatzek, S.; MacLean, D.; Kinoshita, T. DeepStomata: Facial Recognition Technology for Automated Stomatal Aperture Measurement. *Bioinformatics* **2018**, *36*, 5098. [[CrossRef](#)]
22. Casado-García, A.; Del-Canto, A.; Sanz-Saez, A.; Pérez-López, U.; Bilbao-Kareaga, A.; Fritsch, F.B.; Miranda-Apodaca, J.; Muñoz-Rueda, A.; Sillero-Martínez, A.; Yoldi-Achalandabaso, A.; et al. LabelStoma: A tool for stomata detection based on the YOLO algorithm. *Comput. Electron. Agric.* **2020**, *178*, 105751. [[CrossRef](#)]
23. García Álvarez, S.; Morla Juaristi, C.; Solana Gutiérrez, J.; García-Amorena, I. Taxonomic differences between *Pinus sylvestris* and *P. uncinata* revealed in the stomata and cuticle characters for use in the study of fossil material. *Rev. Palaeobot. Palynol.* **2009**, *155*, 61–68. [[CrossRef](#)]
24. Sweeney, C.A. A key for the identification of stomata of the native conifers of Scandinavia. *Rev. Palaeobot. Palynol.* **2004**, *128*, 281–290. [[CrossRef](#)]
25. Zhang, K.; Zhao, Y.; Guo, X. Conifer stomata analysis in paleoecological studies on the Loess Plateau: An example from Tianchi Lake, Liupan Mountains. *J. Arid. Environ.* **2011**, *75*, 1209–1213. [[CrossRef](#)]
26. Camargo, M.A.B.; Marenco, R.A. Density, size and distribution of stomata in 35 rainforest tree species in Central Amazonia. *Acta Amaz.* **2011**, *41*, 205–212. [[CrossRef](#)]
27. Eisele, J.F.; Fäßler, F.; Bürgel, P.F.; Chaban, C. A Rapid and Simple Method for Microscopy-Based Stomata Analyses. *PLoS ONE* **2016**, *11*, e0164576. [[CrossRef](#)]
28. Shapiro, B.E.; Jönsson, H.; Sahlin, P.; Heisler, M.; Roeder, A.; Burl, M.; Meyerowitz, E.M.; Mjolsness, E.D. Tessellations and Pattern Formation in Plant Growth and Development. In *Tessellations in the Sciences: Virtues, Techniques and Applications of Geometric Tilings*; van de Weijgaert, R., Vegter, G., Ritzerveld, J., Icke, V., Eds.; Springer: Berlin/Heidelberg, Germany, 2008.
29. Karabourniotis, G.; Tzobanoglou, D.; Nikolopoulos, D.; Liakopoulos, G. Epicuticular Phenolics Over Guard Cells: Exploitation for in situ Stomatal Counting by Fluorescence Microscopy and Combined Image Analysis. *Ann. Bot.* **2001**, *87*, 631–639. [[CrossRef](#)]
30. Cao, L.; Bala, G.; Caldeira, K.; Nemani, R.; Ban-Weiss, G. Importance of carbon dioxide physiological forcing to future climate change. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 9513–9518. [[CrossRef](#)]
31. Konrad, W.; Roth-Nebelsick, A.; Grein, M. Modelling of stomatal density response to atmospheric. *J. Theor. Biol.* **2008**, *253*, 638–658. [[CrossRef](#)] [[PubMed](#)]
32. Chater, C.; Peng, K.; Movahedi, M.; Dunn, J.A.; Walker, H.J.; Liang, Y.-K.; McLachlan, D.H.; Casson, S.A.; Isner, J.C.; Wilson, I.D.; et al. Elevated CO₂ -Induced Responses in Stomata Require ABA and ABA Signaling. *Curr. Biol.* **2015**, *25*, 2709–2716. [[CrossRef](#)] [[PubMed](#)]
33. Tomimatsu, H.; Tang, Y. Effects of high CO₂ levels on dynamic photosynthesis: Carbon gain, mechanisms, and environmental interactions. *J. Plant Res.* **2016**, *129*, 365–377. [[CrossRef](#)] [[PubMed](#)]
34. Franklin, G.L. Preparation of Thin Sections of Synthetic Resins and Wood-Resin Composites, and a New Macerating Method for Wood. *Nat. Cell Biol.* **1945**, *155*, 51. [[CrossRef](#)]
35. Peat, H.J.; Fitter, A.H. A comparative study of the distribution and density of stomata in the British flora. *Biol. J. Linn. Soc.* **1994**, *52*, 377–393. [[CrossRef](#)] [[PubMed](#)]
36. Zheng, Y.; Xu, M.; Hou, R.; Shen, R.; Qiu, S.; Ouyang, Z. Effects of experimental warming on stomatal traits in leaves of maize (*Zea mays* L.). *Ecol. Evol.* **2013**, *3*, 3095–3111. [[CrossRef](#)] [[PubMed](#)]
37. Fanourakis, D.; Heuvelink, E.; Carvalho, S.M.P. Spatial heterogeneity in stomatal features during leaf elongation: An analysis using Rosa hybrida. *Funct. Plant Biol.* **2015**, *42*, 737. [[CrossRef](#)] [[PubMed](#)]
38. Gan, Y.; Zhou, L.; Shen, Z.-J.; Shen, Z.-X.; Zhang, Y.-Q.; Wang, G.-X. Stomatal clustering, a new marker for environmental perception and adaptation in terrestrial plants. *Bot. Stud.* **2010**, *51*, 12.
39. Weickert, J. *Anisotropic Diffusion in Image Processing*; Teubner, B.G., Ed.; European Consortium for Mathematics in Industry: Stuttgart, Germany, 1998.
40. Buades, A.; Coll, B.; Morel, J.M. A Review of Image Denoising Algorithms, with a New One. *Multiscale Model. Simul.* **2005**, *4*, 490–530. [[CrossRef](#)]

41. Liu, K.; Tan, J.; Ai, L. Hybrid regularizers-based adaptive anisotropic diffusion for image denoising. *SpringerPlus* **2016**, *5*, 404. [[CrossRef](#)]
42. Staff, L.; Hurd, P.; Reale, L.; Seoighe, C.; Rockwood, A.; Gehring, C. The Hidden Geometries of the Arabidopsis thaliana Epidermis. *PLoS ONE* **2012**, *7*, e43546. [[CrossRef](#)]
43. Chen, D.; Aw, W.Y.; Devenport, D.; Torquato, S. Structural Characterization and Statistical-Mechanical Model of Epidermal Patterns. *Biophys. J.* **2016**, *111*, 2534–2545. [[CrossRef](#)] [[PubMed](#)]
44. Kamalaveni, V.; Rajalakshmi, R.A.; Narayanankutty, K. Image Denoising Using Variations of Perona-Malik Model with Different Edge Stopping Functions. *Procedia Comput. Sci.* **2015**, *58*, 673–682. [[CrossRef](#)]
45. Tschumperlé, D.; Deriche, R. *Anisotropic Diffusion Partial Differential Equations in MultiChannel Image Processing: Framework and Applications*; Advances in Imaging and Electron Physics (AIEP); Academic Press: Cambridge, MA, USA, 2007; pp. 145–209.
46. Perona, P.; Malik, J. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **1990**, *12*, 629–639. [[CrossRef](#)]
47. Fashing, M.; Tomasi, C. Mean shift is a bound optimization. *IEEE Trans. Pattern Anal. Mach. Intell.* **2005**, *27*, 471–474. [[CrossRef](#)] [[PubMed](#)]
48. Reams, R. Hadamard inverses, square roots and products of almost semidefinite matrices. *Linear Algebra Appl.* **1999**, *288*, 35–43. [[CrossRef](#)]
49. Leys, C.; Klein, O.; Bernard, P.; Licata, L. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *J. Exp. Soc. Psychol.* **2013**, *49*, 764–766. [[CrossRef](#)]
50. Grimmett, G.; Stirzaker, D. *Probability and Random Processes*; Oxford University Press: Oxford, UK, 2009.
51. Vanselow, R. About Delaunay Triangulations and Discrete Maximum Principles for the Linear Conforming FEM Applied to the Poisson Equation. *Appl. Math.* **2001**, *46*, 13–28. [[CrossRef](#)]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).