On the value of graph-based segmentation for the analysis of structural networks in life sciences

Denis BUJOREANU[1],PEJMAN RASTI[2], DAVID ROUSSEAU[1,2]

(1) Université de Lyon, CREATIS, CNRS UMR5220, INSERM U1206, Université Lyon 1, INSA-Lyon, 69621 Villeurbanne, France
(2) INRA, UMR 1345 Institut de Recherche en Horticulture et Semences (IRHS), Université d'Angers, F-49071 Beaucouzé, France
Email: davis.rousseau@univ-lyon1.fr

***Abstract*—We propose, under the form of a short overview, to stress the interest of graph to encode the "topological" structure of networks hidden in images especially when applied in life sciences. We point toward existing computer science tools to extract such structural graph from images. We then illustrate different applications, such as segmentation, denoising, and simulation on practical examples of various bioimaging domains including vascular networks observed with fluorescent microscopy in 2D imaging, macroscopic root systems observed in 2D optical intensity imaging, and 3D porosity networks of seed observed in absorption X-ray microtomography.**

I. INTRODUCTION

Graph-based image processing is a growing field of information sciences [7]. In a large part of this field the graph are mainly associated to concepts of high level such as energy minimization, partial differential equation, mathematical morphology, ... By contrast, in this work, we consider situations where the graph directly corresponds to an intuitive representation of a structure. In this communication, we consider the broad domain of bioimaging which produces images constituted by underlying structural networks from which one intend to extract topological informations. The extraction of information from such networks requires segmentation methods specifically designed to preserve the topological structure of the network hidden in the image. A first attempt in this direction with an original graph-based representation approach has recently been introduced [3]. We propose three extensions of this work to illustrate its generical value.

II. ROBUST GRAPH REPRESENTATION OF MUSCLE FIBERS IN FLUORESCENT MICROSCOPY

We shortly recall the image processing algorithm recently introduced in [3] for graph-based segmentation. In [3] this algorithm was illustrated on a specific biomedical problem, mice colon's characterization and classification. We demonstrate in this section that the approach of [3] is generic and can be successfully applied to other classes of bioimaging problems.

The pipeline introduced in [3] takes as input images that can be seen as noisy images of a structural network over a background. To produce a representation of this network that allows to extract its topology, a first step is a segmentation block assumed controlled by the parameter θ. A second step is a skeletonization block followed by the third step which corresponds to the creation of a graph. Therefrom, the last step consists in analyzing the topology of the graph with various graph metrics. To ensure that the topology of the graph extracted from our pipeline fits with the topology of the real structural graph imaged, the segmentation of the image is realized jointly with the topological analysis of the graph. The pipeline is applied for a set of the segmentation parameters θ. The selected segmentation parameters are considered optimal if they produce a graph with topology satisfying a criteria of stability thus producing a robust graph representation.

The algorithm of Fig. 1 is applied to fluorescent microscopy images of muscles where the informational task of biological interest is the measure of the muscle fibers. The muscle used comes from mice. The Tibialis Anterior of these mice has been extracted and frozen. Transversal section of 7 μm have been realized and received a dye which highlights the tissue surrounding each fiber. Image acquisition is done with a X20 lens on a Zeiss Axio observer Z1. As visible in Fig. 2, the images are rather well contrasted. The difficulty comes from the possible blur or decrease of the fluorescence on some fibers. We applied the algorithm of Fig. 1 on the images of Fig. 2 A and B. We use the same segmentation algorithm as in [3] with a simple adaptive local thresholding method [11] based on a single segmentation parameter θ corresponding to the size of the local patch on which the threshold is computed. The evolution of the graph metrics plotted as a function of the segmentation parameter θ are given in Fig. 3. Here we plotted the average degree which is the average number of neighbors in the graph for each node, the average distance which is the average distance between two nodes in the graph and the clustering coefficient [12] which corresponds to the average value of the ratio between the existing and possible edges in a node's neighborhood. The typical evolution for these three graph metrics, is the same as the one described in [3] for another problem, i.e. a non monotonic evolution with a rapid variation peak followed by a stationary "plateau" reached when the segmentation parameter overpasses the typical size of the muscle fiber (10 μm). The skeleton extracted with the graph are not perfect especially on very tiny structures as visible on Fig. 2C. Other segmentation

ISBN 978-0-9928626-7-1

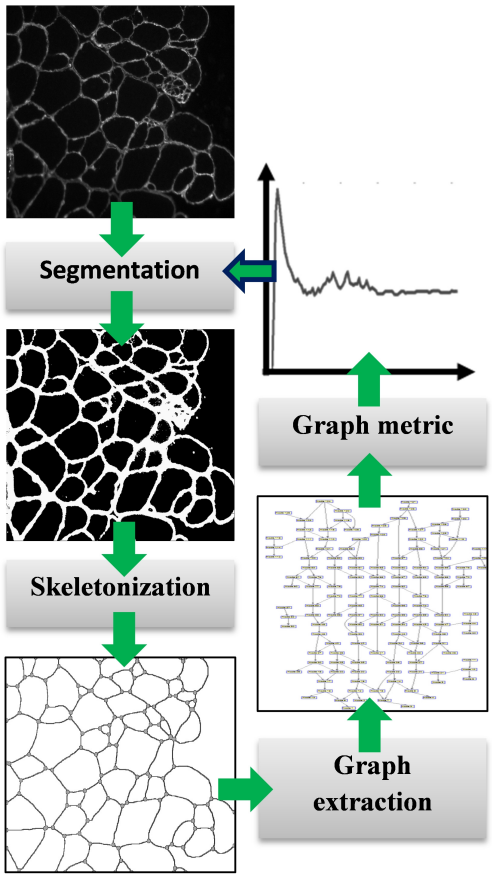

Fig. 1. General description of the image processing pipeline proposed for the graph representation of images with underlying networks. The feedback loop expresses that the segmentation parameters are tuned to satisfy a stability criterion of the graph metric.

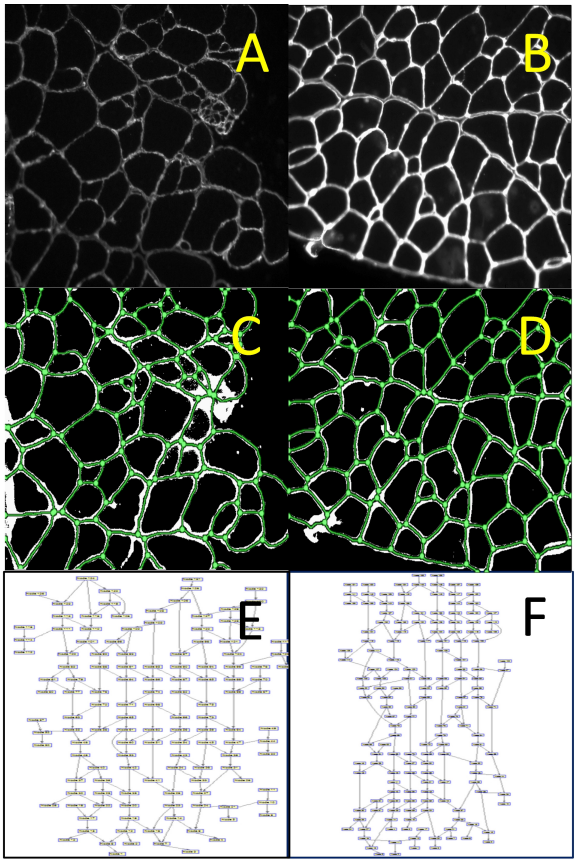

Fig. 2. Example of graph extracted from the pipeline of Fig. 1. On panel A, and B raw fluorescence microscopy images of muscle fibers. Image B is well contrasted and Image A shows typical blur and spatial inhomogeneity of fluorescence. On panel C, D superposition of the segmented images and the topological graph with in green dots the position of the nodes in the graph and in green lines the edges. Panels E, F show the graph corresponding to the images A and B.

techniques have recently been proposed in the literature for the muscle fiber segmentation [10], [6] of similar images. They are based on different approaches (active contour for [10] and erosion-dilatation morphomathematics for [6]) and it would be interesting to compare the results. Instead, we choose in the following sections to further demonstrate the interest of the method [3] for other bioimaging problems.

III. Graph-based denoising of skeletonized root-systems

In this section, we demonstrate that the robust graph-based representation recalled in the previous section can be used to other tasks than the segmentation. To this end, we consider another bioimaging problem, the analysis of root systems in reflectance intensity imaging. Roots continuously grow hidden in the soil, and form complex 3D structures. Monitoring this biological process by automated computer vision is one of the most challenging tasks in imaging and image analysis for the plant sciences. The development of various imaging systems to address this challenge with rhizotron in projected 2D or X-ray tomography in 3D has recently triggered many developments in image analysis [8] to extract phenotypic traits

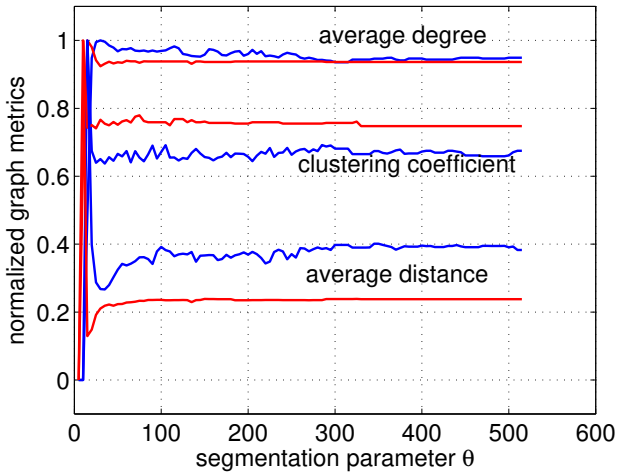

Fig. 3. Normalized graph metrics as a function of the segmentation parameter for the two images of Fig. 2. The red line is for the image of Fig. 2A and the blue line for the image of Fig. 2B.

ISBN 978-0-9928626-7-1 © EURASIP 2017

of root systems or in informatics to define universal description formats for root systems based on graphs [9]. However, despite these important efforts to promote reproducible science practices, specific issues, brought by the diversity of imaging systems and the richness of fundamental biological questions, continue to pop up and remain to be addressed. We focus here on the problem of single root systems composed of a primary root and secondary roots imaged with a rhizotron at a stage of development where secondary roots touch. The problem of root crossings has recently been addressed [2], [1] with inpainting methods based on anisotropic diffusion with partial differential equations. This technique gives good performances to separate roots when they cross with rather large angle (typically larger than 3 °). However, the situation of almost parallel roots is very common. This may cause artifacts in the root-system analysis and thus remains, to the best of our knowledge, to be addressed. For illustration, we process the image of Fig. 4A where a vertical primary root is connected to horizontal secondary roots which touch themselves. A current approach to extract the architecture of the root system is to perform after several preprocesses a morphological skeleton of the root system. As visible in Fig. 4B, such an approach fails to separate the secondary roots which are connected in the skeleton (D1 in Fig. 4B). Also, skeletonization artifacts (D2, in Fig 4B) may cause false detections of secondary roots. A possible way to overcome this would be to use more sophisticated image preprocessing to enhance the separation between secondary roots and smooth the spurious features in the image. Instead, we propose to use the topology of the root system as a prior to denoise the skeleton.

In a first stage, the algorithm of Fig. 1 is applied to produce a graph and all the useful data (length, position, name, etc.) about crossings, or nodes in graph nomenclature, and connections, or edges in graph nomenclature, are stored in a node/edge structure. This structure allows us, in the second stage of our approach, to filter skeletonization artifacts (D2) by erasing all the "short" edges that are only connected to one node. The results can be seen as a graph representation of the adjacency matrix in Fig. 4D or as a detailed representation of the node/edge structure in Fig. 4C. In a second stage, the adjacency matrix is analyzed in order to detect the primary root and to direct the matrix. Considering that the primary root contains the most crossings, the detection of primary root is trivial, and can be done by searching the longest path in the adjacency matrix via the Dijkstra's algorithm [5]. Then all the edges are oriented one by one to flee from the primary root and the new oriented graph, shown in Fig. 4E, is created. In the last stage, the secondary roots are detected. Considering that for each secondary root there is a node (crossing) on the primary root and each secondary root ends with a leaf-node in the oriented adjacency matrix, the chain of nodes and edges of each secondary root is calculated using a shortest path algorithm. Our process produces the oriented graph of Fig. 4G which correctly restitutes the topology of the root system and encodes the number, position, and length of the primary and secondary roots. Artifacts in the skeletonized root-system of Fig. 4B have been denoised with the graph instead of a more conventional approach that would have focussed on denoising the skeleton image.

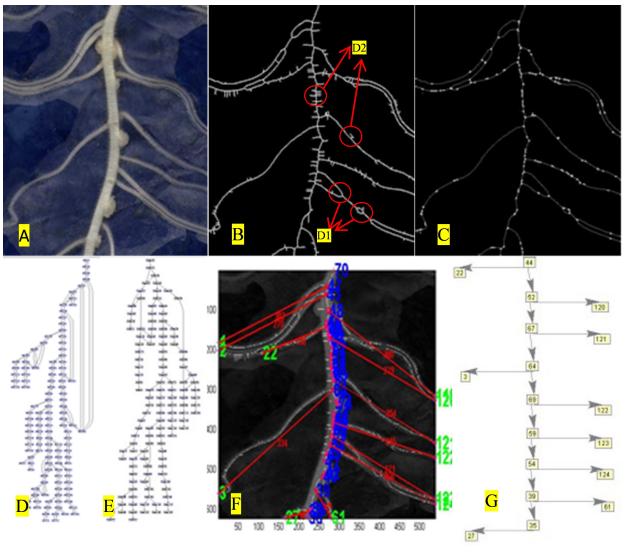

Fig. 4. (A) Rhizotron image of a taproot system and its skeleton (B). Node/edge structure is given in (C) and the non-oriented graph in (D). (E) is the oriented graph. Panel (F) shows the nodes of the primary root (blue) and the calculated endpoints of each secondary root (green). The red lines represent the correspondence between the starting node and end node of each secondary root, the red numbers represent the estimated length of each secondary root. Panel (G) gives the final graph of the taproot system.

IV. Graph-based simulation in 3D porous media

The graph-based representation discussed in the previous section is not limited to the encoding of 2D structures. We also applied it to the 3D images of another bioimaging problem as demonstrated in this section. To this end, we considered X-ray microtomography images of sugar beet seeds. As visible in Fig.5, in these images a strong contrast is observed between the air network and the tissue in the seed. These air networks are known to be specially useful during the imbibition process when an entry of water triggers the transition from an inert dry seed into a living plant. An imbibition process is expected to be faster when the air networks interconnect all the parts of the seed with each other [4]. The inspection of this connectivity can be done directly from the images. To this end, a segmentation of the air network has first to be realized. Then a numerical imbibition corresponding to a simple region growing process with a zero tolerance can be launched. As illustrated in Fig. 6 panels A to D, two points are connected if after convergence the region growing process includes these two points. This iterative region growing process is very time-consuming. That is why we consider the possibility to instead use the graph representation of the segmented air network. The graph-based encoding of the air networks enables, as shown in Fig. 6 panel E and F, by direct computation along the adjacency matrix to determine whether a pore is connected to another one. The computation cost information concerning the comparison of the region growing and graph-based approach

ISBN 978-0-9928626-7-1 © EURASIP 2017

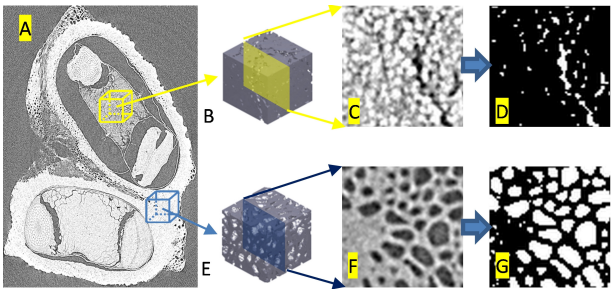

Fig. 5. Panel A: X-ray microtomography of a dry sugar beet seed. Solid blue and yellow boxes highlight the contrast between air network in dark and the seed tissue in hypersignal with panels B, C located in the starch and panels E, F in the cork. The contrast is sufficient in panel C and F so that it is easy by simple thresholding to segment as shown in panels D and G the respective air network in white from the surrounding tissue in black.

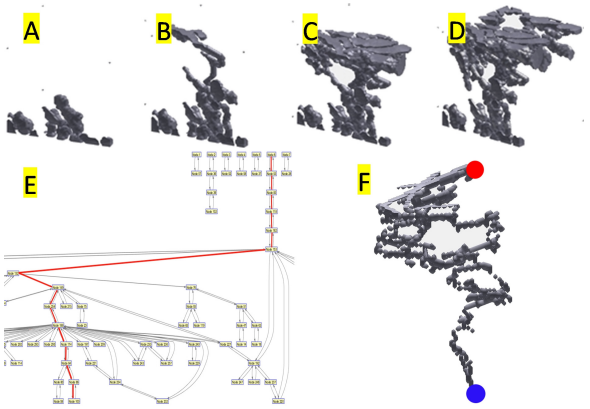

Fig. 6. Panels A to D, are snapshots of the imbibition process simulated as an iterative region growing process from one point in the segmented binary data included in the blue cube in Fig. 5E. Panel E represents the graph corresponding to the same cube of data. The red line highlights the possibility of a path from one point to another in the cube from the graph. Panel F corresponds to the path, inside the binary data, between the two points built using the red path from panel E and the node/edge structure.

is given in Fig. 5 where the size of the data and the number of floating-point operations (FLOPs) that each process required on our computer. With our 2 processors Intel Xeon E5620@ 2.4Ghz with 4 cores we found that simulating the imbibition along the imaging with region growing was 2700 times slower than doing it along a graph and that the manipulated data was 10 times bigger when simulating region growing than with the graph approach. This demonstrates, on an example distinct from the previous sections of this communication, the added value of a graph-based representation for the analysis of network in bioimaging.

V. Conclusion

We have briefly illustrated the genericity of the algorithm recently introduced in [3] for the analysis of networks in life sciences with various bioimaging problems and for various informational tasks (segmentation, denoising, simulation). Other tasks may benefit from this technically simple approach (registration for instance) where some prior on the structural topology of the underlying network helps in the image

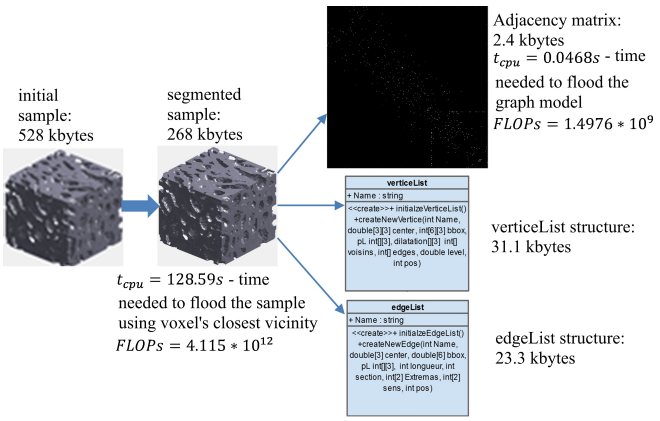

Fig. 7. Overall view of the size of the structures generated on a 3D X-ray microtomography samples of sugar beet seed with $64 \times 64 \times 64$ pixels. The size of the data structure attached to the graph is (56.8 k octets) including the adjacency matrix, the list of vertices $verticeList$ and the list of edges $edgeList$. This is approximately 10 times smaller than the size of the segmented image. In terms of computation time to reach the results, the computation of the connectivity of the air network takes approximately 2700 times longer to simulate it with numerical imbibition, i.e. region growing, than by flying along the extracted graph.

processing step. Each bioimaging problem presented would also, beyond the proof of feasibility given here, deserve more attention by benchmark with alternative methods (based on graph or not) on larger datasets.

Acknowledgment

Authors thank Frédéric COINTAULT for image on root systems and Bénédicte CHAZAUD for image on muscles. This work was supported by the LABEX PRIMES (ANR-11-LABX- 0063) of Université de Lyon, within the program Investissements dAvenir (ANR-11-IDEX-0007) and in the framework of the program "Investissements d'Avenir" under reference ANR-11-BTBR-0007 (AKER program) both operated by the French National Research Agency (ANR).

References

[1] L. Benoit, D. Rousseau, É. Belin, D. Demilly, and F. Chapeau-Blondeau, "Simulation of image acquisition in machine vision dedicated to seedling elongation to validate image processing root segmentation algorithms," *Computers and Electronics in Agriculture*, vol. 104, pp. 84–92, 2014.

[2] L. Benoit, D. Rousseau, E. Belin, D. Demilly, S. Ducournau, F. Chapeau-Blondeau, and C. Dürr, "Locally oriented anisotropic image diffusion: application to phenotyping of seedlings," in *8th international conference on Computer Vision Theory and Applications (VISAPP 2013), Barcelona*, 2013, pp. 21–24.

[3] D. Bujoreanu, H. Dorez, W. Boutegrabet, D. Moussata, R. Sablong, and D. Rousseau, "Robust graph representation of images with underlying structural networks. application to the classification of vascular networks of mices colon," *Pattern Recognition Letters*, vol. In Press, no. Special Issue on Graph, pp. 1–11.

[4] P. Cloetens, R. Mache, M. Schlenker, and S. Lerbs-Mache, "Quantitative phase tomography of arabidopsis seeds reveals intercellular void network," *Proceedings of the National Academy of Sciences*, vol. 103, no. 39, pp. 14 626–14 630, 2006.

[5] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische mathematik*, vol. 1, no. 1, pp. 269–271, 1959.

[6] A. C. Keefe, J. A. Lawson, S. D. Flygare, Z. D. Fox, M. P. Colasanto, S. J. Mathew, M. Yandell, and G. Kardon, "Muscle stem cells contribute to myofibres in sedentary adult mice," *Nature communications*, vol. 6, 2015.

[7] O. Lezoray and L. Grady, "Image processing and analysis with graphes: Theory and practice." CRC Press, 2012.
[8] G. Lobet, X. Draye, and C. Périlleux, "An online database for plant image analysis software tools," *Plant methods*, vol. 9, no. 1, p. 38, 2013.
[9] G. Lobet, M. P. Pound, J. Diener, C. Pradal, X. Draye, C. Godin, M. Javaux, D. Leitner, F. Meunier, P. Nacry *et al.*, "Root system markup language: toward a unified root architecture description language," *Plant Physiology*, vol. 167, no. 3, pp. 617–627, 2015.
[10] J. Mula, J. D. Lee, F. Liu, L. Yang, and C. A. Peterson, "Automated image analysis of skeletal muscle fiber cross-sectional area," *Journal of Applied Physiology*, vol. 114, no. 1, pp. 148–155, 2013.
[11] J. Sauvola and M. Pietikäinen, "Adaptive document image binarization," *Pattern recognition*, vol. 33, no. 2, pp. 225–236, 2000.
[12] D. J. Watts and S. H. Strogatz, "Collective dynamics of small-world networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.

ISBN 978-0-9928626-7-1