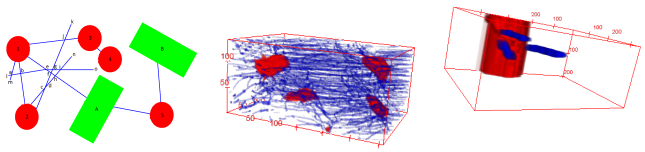
Graph encoding of multiscale structural networks from binary images with application to bio imaging.

Nicolas Parisse[1]
nicolas.parisse@gmail.com
Aurelien Gourrier[2]
aurelien.gourrier@univ-grenoble-alpes.fr
Rachel Genthial[2]
rachel.genthial@gmail.com
Delphine Debarre[2]
delphine.debarre@univ-grenoble-alpes.fr
Andrea Bassi[3]
andreabassi@polimi.it
David Rousseau[1]
david.rousseau@univ-angers.fr

[1] LARIS, UMR INRA IRHS, Université d'Angers, 62 Avenue de Notre Dame du Lac, 49000 Angers France
[2] Laboratoire Interdisciplinaire de Physique, UMR 5588 CNRS, Université Joseph Fourier Grenoble, Saint Martin d'Hères, France.
[3] Politecnico di Milano, Dipartimento di Fisica, Piazza L. da Vinci 32, 20133 Milano, Italy.

1 Introduction

Graph-based image processing is a growing field of information sciences [9]. In a large part of this field, graphs are mainly associated to concepts of high level representations such as energy minimization, partial differential equation, mathematical morphology... By contrast, here, we consider situations where the graph corresponds directly to an intuitive representation of a structure that can be visually distinguished in the image. Such situations occur in the broad domain of bioimaging which often generates images including underlying structural networks [1, 2] from which one intends to extract topological information. The novelty of this study resides in the multiscale graph extraction scheme which we propose to address the specific case of multiply-connected structures at different sizes. Synthetic and real world examples of images of this type of networks are shown in Fig. 1 in which the different scales are colorized. These images consist of an empty white background and of a foreground composed of thicker and more compact objects of various sizes connected by thinner, elongated branches. A parallel can be drawn between the foreground structures in Fig. 1 and graphs by associating thicker objects to graph nodes and the thin connections to graph edges. A pre-requisite of our analysis is, therefore, that the initial raw measured images can be segmented to produce binary volumes. This procedure is generally specific to a given imaging modality and sample nature and may require complex optimization procedures as recently illustrated on a set of 2D and 3D bioimaging problems with underlying structural networks [1, 2]. In this work, we focus on the more generic aspect associated to graph encoding of the binary data. A classical approach to extract graphs from binary images consists in analyzing its morphological skeleton. The skeleton makes graph extraction easier since the nodes will be associated to skeleton pixels with a connectivity equal to one (for leaf nodes) or higher than two. This approach is powerful to extract robust graphs for pattern recognition, where nodes do not necessarily correspond to structural junctions of a real network [8, 12]. To encode structural networks, skeletonization is only adapted when the networks are composed of a single scale association of elongated objects sparsely touching each others. For networks with multiple scales a brute force skeletonization on images such as Fig. 1 will very likely fail to produce a faithful graph due to the intrinsincly multiscale nature of the structure, even after extensive pruning of the skeleton. Indeed true small branches are present in the elongated part of the network while artifact branches of similar lengths may appear within the volume of the larger scale objects upon skeletonization. While new skeletonization schemes continue to be developed [13], a simple workaround to extract structural graphs from binary images was proposed by Iwanowski et al. [7] that detect objects and connections using morphological image processing, in particular morphological openings for node detection and anchored homotopic skeletonization for connection detections. Apart from nodes and connections, their joints are also detected and labeled. In the next step, based on the labeling of these three items (objects, connections and joints), the graph adjacency matrix is computed. In this work, we push forward the graph encoding scheme proposed in [7] to the case where multiple scales in the network are to be analyzed.



(a) (b) (c)

Figure 1: Model and real world multiscale network images. Panel (a) is a synthetic image used to test the encoding scheme with three distinct types of interconnected objects of very different size (scales) represented by a specific color code. Panel (b) is a 3D representation of a stack of confocal fluorescence microscopy images (182(X) x 180 (Y) x 113 (Z) voxels of 189.3 (X) x 189.3 (Y) x 350 (Z) μm^3) of a bovine femur [4] including two scales: lacunae, represented in red and caniculi in blue. Panel (c) is a 3D representation of a stack of light sheet fluorescence microscopy images with same characteristic as in Panel (b) of an arabidopsis root system [3] including two scales: radicle represented in red and hair cells in blue. The root system is monitored during elongation of hair cells.

2 Graph encoding scheme

The proposed graph encoding scheme is described in Fig. 2. The algorithm takes a binary image as input and we assume that all the objects corresponding to the different scales in the image can be extracted in separate binary images. This preprocessing can typically be performed with multiscale analysis such as wavelets [10] in the original raw images or with morphomathematic filtering, provided that some prior information on the value of these scales (size) can be assumed. This is a reasonable assumption in bioimaging when anatomical models exist.

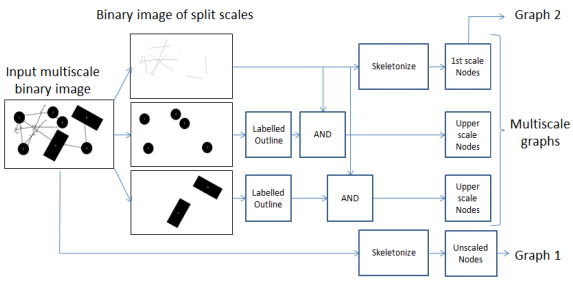


Figure 2: Pipeline presenting three methods to extract the graph of the image on Fig. 1. Graph 1 is unscaled, i.e. obtained after skeletonization of the full raw binary image. In contrast, Graph 2 corresponds to the graph of the first scale level only, i.e. without taking the higher scales into account. The proposed multiscale method, separates the objects at different scales, as in the graph 2 method, but then identifies the elements which connect each upper scale levels to the first one.

The smallest scale corresponding to thinner and elongated objects is first skeletonized. The nodes of this scale correspond to the points with a connectivity higher than 2 or equal to 1. Then, all the over objects that

have intersections with this first scale are coded as nodes with a label corresponding to the scale. Multiple adjacency matrices can then be produced from such a labelled data structure, matrix including all scales as produced by [7], but also the adjacency matrices corresponding to each individual scale. To better understand this multiscale approach let us apply the pipeline presented in Fig. 2 to the image in Fig. 1a. The resulting graphs, corresponding to each of the 3 scales of objects contained in the image, are presented in Fig. 3. The result for the unscaled graph, graph 1 in Fig. 2, are presented in Fig. 3a. Since this graph is extracted directly from the morphological skeleton of the image, one can observe that it contains some artifacts. Specifically, there are 4 nodes with label A and two nodes with label B while in the image there is only one object A and one object B. This result is directly linked to the fact that the skeleton of the image, without pruning, does not perform well on large objects: multiple quench lines are generated that in the graph will generate interconnected nodes with the same label (i. e. A-A or B-B). The graph extracted at the first scale, graph 2 in Fig. 2, is shown in Fig. 3b. Here again, some artifacts appear around the bigger objects A and B. Finally the multiscale graph approach proposed here is presented in Fig. 3c where one can observe that, only in this case, a faithful representation of the topology of the structure shown in Fig. 1a is obtained.

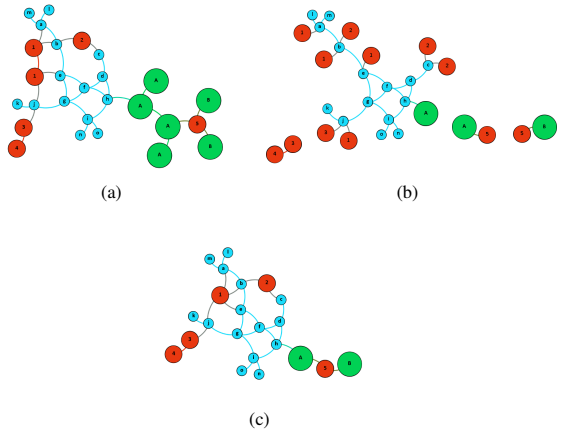
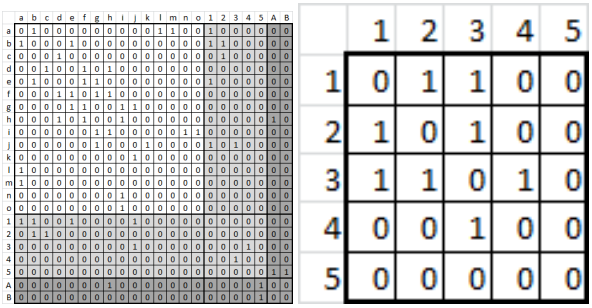


Figure 3: 2D visualization of graphs from image of Fig. 1. Panel (a) shows the graph of the unscaled method. This graph is obtained by skeletonizing the image in Fig. 1 without removing the upper-scale objects. Panel (b) displays the graph of the first scale only. Finally, panel (c) shows the graph obtained with our multiscale method: the image in Fig. 1 is skeletonized by removing the large objects, then the multiscale graph is computed by finding the nodes which belong to each of those large objects.

The associated multiple adjacency matrices corresponding to the connectivity of each labelled scale is presented in Fig. 4. The adjacency matrices of the larger scales are of particular interest since they would not be accessible from a coarse grained version of the image only preserving the large objects and loosing their tiny connections. It is clearly visible from Fig. 3 and Fig. 4 that the multiscale approach matches exactly the structural connectivity of the synthetic image of Fig. 1 while respecting the different classes (scales) of the objects. With real networks imaged in 3D, such a visual inspection would be too time consuming and a comparison of the different graph encoding approaches is usually done from average graph metrics. A huge literature of such graph metrics already exists [9]. As an example we list in Table 1 the number of nodes, the edge density (i.e. (the number of edges) / (the number of nodes)), the average clustering coefficient (i.e. average percentage of how close neighbor points are to a complete graph) and the average degree of connectivity (i.e. average number of connected neighbours). It clearly appears at this statistical level also that the different graph encoding approaches are distinct. The reason for this difference becomes evident when plotting the histogram of the nodes connectivity. The multiscale approach fuses all the nodes connected to a given object at one scale into a single node. This increases the relative weights of high connectivity nodes. In contrast, the unscaled graph produces a higher number of single connectivity nodes (leaf nodes) due to artifacts caused by the brute force skeletonization applied to non elongated objects. In addition to providing a more accurate graph representation and associated metrics, our multiscale approach also allows focusing on specific scales and testing different connectivity hypothesis.



	1	2	3	4	5
1	0	1	1	0	1
2	1	0	1	0	1
3	1	1	0	1	1
4	0	0	1	0	0
5	1	1	1	0	0

(c)

Figure 4: Adjacency matrices obtained from the multiscale graph encoding method. Matrix (a) is the adjacency matrix of the whole graph sorted by increasing scale level from left to right indicated by darker grayscale levels. The matrices below correspond to the adjacency matrices of the second level structures assuming connections with only the first level (b) or all levels (c) .

Kind of measurements	unscale	first scale	multiscale
Edge density	0.0985	0.0667	0.1212
Average clustering coefficient	0.0705	0.0389	0.9850
Average degree of connectivity	2.4615	1.9333	2.5455
Nodes number	26	30	22

Table 1: Graph metrics computed on the three graph encoding methods applied to the model scheme of Fig. 1.

For example, in our simplified scheme the structures of the second scale are not fully connected if we only consider the first two scales (Fig. 3b) since element 5 is only connected to the third scale. Hence, if we take this third scale into account, the second level becomes fully connected (Fig. 3c). In biological systems, where hierarchical levels are associated with specific functions, such that a detailed graph analysis conducted at different levels could provide valuable functional informations.

3 Application to bioimaging

Following the test analysis of our model network structure (Fig. 1a), we will illustrate on the poster results on real world examples of Fig. 1b and Fig. 1c where two connected structures of two different size are presented in red and blue. These could represent many structures in bioimaging such as in plant science voids in soil, or in dry seeds[2], vascular systems in maize ears [11]. However, such multiscale structure also occur in general in life science for instance in neural networks or vascular systems. To stress the generic value of our proposal we illustrate with an application in the biomedical domain and in the plant science domain and provide codes to explore with your own data [5, 6].

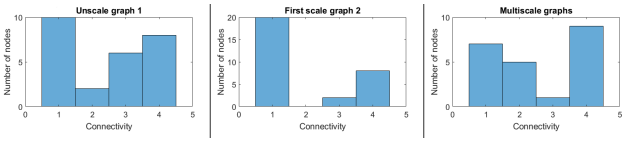


Figure 5: Histogram of the number of nodes as a function of their connectivity given by the three graph encoding methods applied to the model scheme of Fig. 1.

.

[1] Denis Bujoreanu, Hugo Dorez, Warda Boutegrabet, Driffa Moussata, Raphaël Sablong, and David Rousseau. Robust graph representation of images with underlying structural networks. application to the classification of vascular networks of mice's colon. *Pattern Recognition Letters*, 87:29–37, 2017.

[2] Denis Bujoreanu, Pejman Rasti, and David Rousseau. On the value of graph-based segmentation for the analysis of structural networks in life sciences. In *Signal Processing Conference (EUSIPCO), 2017 25th European*, pages 2664–2668. IEEE, 2017.

[3] Alessia Candeo, Fabrizio G Doccula, Gianluca Valentini, Andrea Bassi, and Alex Costa. Light sheet fluorescence microscopy quantifies calcium oscillations in root hairs of arabidopsis thaliana. *Plant and Cell Physiology*, 58(7):1161–1172, 2017.

[4] Rachel Genthial, Emmanuel Beaurepaire, Marie-Claire Schanne-Klein, Françoise Peyrin, Delphine Farlay, Cécile Olivier, Yohann Bala, Georges Boivin, Jean-Claude Vial, Delphine Débarre, et al. Label-free imaging of bone multiscale porosity and interfaces using third-harmonic generation microscopy. *Scientific Reports*, 7:3419, 2017.

[5] https://uabox.univ angers.fr/index.php/s/BPzkOjOe7n92bKK.

[6] https://uabox.univ angers.fr/index.php/s/Djt4UH7hjW7OWVf.

[7] Marcin Iwanowski. Automatic extraction of graph-like structures from binary images. In *Computer Vision and Graphics, International Conference, ICCVG 2008, Warsaw, Poland, November 10-12, 2008, Revised Papers*, pages 461–468, 2008. doi: 10.1007/978-3-642-02345-3_45. URL https://doi.org/10.1007/978-3-642-02345-3_45.

[8] Aurélie Leborgne, Julien Mille, and Laure Tougne. Noise-resistant digital euclidean connected skeleton for graph-based shape matching. *Journal of Visual Communication and Image Representation*, 31:165–176, 2015.

[9] Olivier Lézoray and Leo Grady. *Image processing and analysis with graphs: theory and practice*. CRC Press, 2012.

[10] Stéphane Mallat. *A wavelet tour of signal processing*. Academic press, 1999.

[11] David Rousseau, Thomas Widiez, Sylvaine Tommaso, Hugo Rositi, Jerome Adrien, Eric Maire, Max Langer, Cécile Olivier, Francoise Peyrin, and Peter Rogowsky. Fast virtual histology using x-ray inline phase tomography: application to the 3d anatomy of maize developing seeds. *Plant Methods*, 11(1):55, 2015.

[12] Khalid Saeed, Marek Tabędzki, Mariusz Rybnik, and Marcin Adamski. K3m: A universal algorithm for image skeletonization and a review of thinning techniques. *International Journal of Applied Mathematics and Computer Science*, 20(2):317–335, 2010.

[13] Punam K Saha, Gunilla Borgefors, and Gabriella Sanniti di Baja. A survey on skeletonization algorithms and their applications. *Pattern Recognition Letters*, 76:3–12, 2016.