

Dynamic detection and quantitative analysis of biological network structures from live images

Alessandro Abate, Angelo Cenedese, and Alberto Silletti

I. DESCRIPTION OF THE TECHNIQUE

We will consider the three phases of the procedure and approach the methodological issues of interest in order to compute the structure of interest in the image.

- 1) *Detection*. The first available image frame is analyzed searching for an underlying network structure, which is then extracted leveraging the use of a random walk model to navigate the network edges combined with a network agent to organize the retrieved information into a complex structure. The first frame is taken as a reference frame, also for the warping function w and the outcome of this procedure is a network graph (vertices, edges) with spatial information.
- 2) *Tracking*. The networked structure is then deformed in time starting from the information given by the reference frame and according to the visual data of the following frames. This phase resorts to the computation of optimal (or sub-optimal) warping maps w for motion and deformation that account for the underlying dynamics of the deformable system and at the same time accommodate the dataflow. A set of time stamped network structures is produced, each uniquely associated to the corresponding image frame.
- 3) *Registration*. Once the salient structures are detected and tracked in time, is of fundamental importance to establish a quantitative relationship among them by matching networks from consecutive frames.

A. Detection: random walk model and network model

The rationale behind the detection procedure is that of finding topologically continuous paths over the image \mathcal{I} exploiting the collective motion of a set of agents $\{\mathcal{A}_1, \dots, \mathcal{A}_{N_\tau}\}$ exploring the digital field of the frame, and tracking¹ their paths along (see Fig. 1(a)), similarly to what proposed in [1], [4]. The motion of these exploring agents is a random walk model driven by an external input that is related to the features of the already traveled path and some prediction on the following steps, in order to

ensure a level of continuity in the track. On the other hand, the number of the agents N_τ is not constant during the detection phase and determined by the complexity of the structures: At any bifurcation point one or more new agents are initialized with different heading directions in order to proceed towards the detection of new branches, and whenever a closed structure is completely detected the detecting agent ceases to exist.

More formally, each random walk agent \mathcal{A}_i is characterized at instant τ by a couple $\{\mathbf{x}_i, \theta_i\}(\tau)$, respectively the position on the image frame and the heading direction, and employs a motion law defined as:

$$\mathbf{x}_i(\tau + 1) = \mathbf{x}_i(\tau) + \alpha e^{j\theta_i(\tau)}, \quad (1)$$

where α is the step size of the advancement and j is the imaginary unit. The heading direction $\theta_i(\tau) \in \Theta = [0, 2\pi)$ is the polar angle seen by the walker agent obtained from the following solution set

$$\{\arg \max_{\theta \in \Theta_{ik}} (\mathcal{E}_I(\theta, \mathbf{x}_i(\tau)) | \Theta_{ik} \text{ s.t. } \mathcal{E}_I(\theta, \mathbf{x}_i(\tau)) > \bar{\mathcal{E}})\}, \quad (2)$$

where \mathcal{E}_I is a suitable energy function defined over the image in a \mathbf{x}_i neighborhood domain (see Fig. 1(b)), θ is the polar angle coordinate, and $\bar{\mathcal{E}}$ is a suitable threshold value. Basically, if there is only one subdomain of Θ where \mathcal{E}_I exceeds $\bar{\mathcal{E}}$, be it Θ_{i1} , there is a unique possible heading direction θ_i the agent keeps on traveling. Conversely, if the thresholding operation highlights more intervals $\{\Theta_{i1}, \dots, \Theta_{iN_\Theta}\}$, the extremization procedure of Eq. 2 suggests a set of heading directions $\{\theta_{i1} \in \Theta_{i1}, \dots, \theta_{iN_\Theta} \in \Theta_{iN_\Theta}\}$ as local extremal points in the intervals of Θ (see Fig. 1(c)). In such a case, N_Θ agent instances are initialized with the same position and different heading directions $\{\{\mathbf{x}_i, \theta_{i1}\}, \dots, \{\mathbf{x}_i, \theta_{iN_\Theta}\}\}$, and the procedure is able to accommodate path bifurcations.

Loosely speaking, the energy function \mathcal{E} gathers local information by exploring the surroundings of current position $\mathbf{x}_i(\tau)$ and the rôle of this energy term is to drive the agent towards the salient structure, being strictly related to the image intensity function and the visual data appearance. In the specific case, the reticular structure appears as a light network on a darkish cluttered and noisy background: Hence, the energy term \mathcal{E}_I related to the image intensity I is chosen with respect to a reference value I_0 as:

$$\mathcal{E}_I(\theta, \mathbf{x}_i(\tau)) = \frac{\int_{\Omega_i} (I(\mathbf{x}) - I_0(\mathbf{x}))^2 d\mathbf{x}}{\int_{\Omega_i} d\mathbf{x}}, \quad (3)$$

A. Abate is with the Delft Center for Systems and Control, TU Delft, The Netherlands; a.abate@tudelft.nl.

A. Cenedese is with the Department of Information Engineering, University of Padova, Italy; angelo.cenedese@unipd.it.

A. Silletti is with the Department of Information Engineering, University of Padova, Italy; silletti@dei.unipd.it.

The research leading to these results has received funding from ...

¹In this context, the temporal coordinate τ is set by the pseudo-time of the detection algorithm.

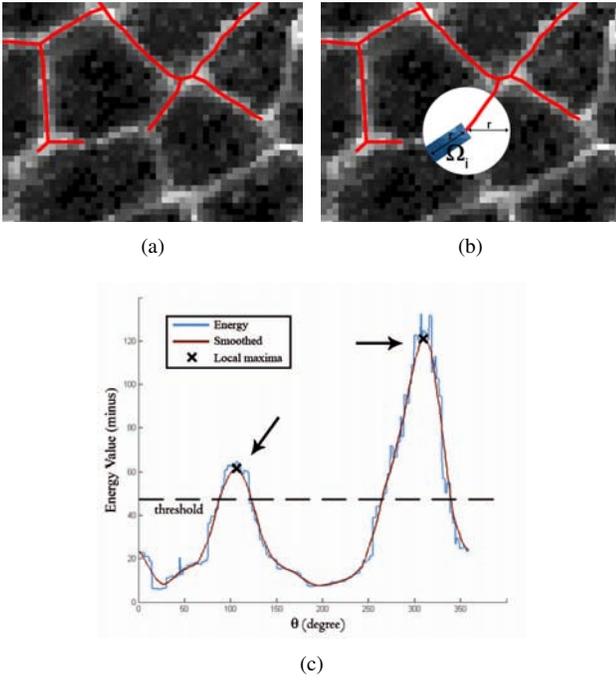


Fig. 1. **Walker agent approach.** (a) Several agents explore the frame. The union of the paths traveled by all agents provides the reticular structure detection. (b) The agent field of view, Ω , chosen as a rectangular shape, is employed as the domain to determine the heading direction. (c) Local maxima of the energy function \mathcal{E}_T correspond to possible heading directions for advancing a single agent or generating new agents. A smoothed version, filtering small random disturbances, is more reliable for noisy images.

where the integration domain depends on local position and orientation, $\Omega_i \triangleq \Omega(\theta, \mathbf{x}_i(\tau))$. Nicely, this domain represents the field of view of the agent in exploring its neighborhood.

The global description provided by the walker agents, obtained as the union of all states $\bigcup_{\tau} \bigcup_i \{\mathbf{x}_i, \theta_i\}(\tau)$ yields a refined representation of the networked structure and can be summarized through a graph model $\mathcal{G} = (\mathbf{N}, \mathbf{E})$ (nodes and edges, respectively), where each location reached by a random walk agent is a node in the graph and the edges keep track of the path traveled by the agents².

Being designed for representation aims, this model may be overabundant for modeling purposes, especially when the goal is to obtain compact-size models that are flexible and agile to be used both for simulation and for prediction of the behavior of the structure. In this sense, only the subset $\bar{\mathbf{N}} \subset \mathbf{N}$ of the nodes associated to path bifurcations (and the correspondent connecting edge set $\bar{\mathbf{E}}$) are interesting in the description. If border effects are neglected, the resulting graph $\bar{\mathcal{G}} = (\bar{\mathbf{N}}, \bar{\mathbf{E}})$ is a 3-connected graph, obtained by selecting the nodes with valency greater than 2, formally $\delta(\bar{\mathcal{G}}) = 3$, and creating connecting edges according to the path connections of \mathcal{G} (see Fig. 2(a)).

Such a procedure operates the abstraction of the retrieved visual data into a compact and versatile data structure,

²A loop closure procedure is also performed in order to detect closed structures, fusing contributions from different agents while avoiding at the same time closed loop artifacts in uniform image intensity areas.

that allows the definition of metrics for the quantitative evaluation of the salient objects detected in the image frame and the applications of methodologies and models from network and graph theory. Interestingly, in the case of cell clusters or similar agglomerates, the graph $\bar{\mathcal{G}}$ (obtained from the network of the objects boundaries) can be considered as a primary structure of nodes (all branching locations) and edges (virtual boundaries) whose dual graph $\bar{\mathcal{G}}' = (\bar{\mathbf{N}}', \bar{\mathbf{E}}')$, conversely, is composed by nodes as the object centers and edges as the connections with the neighbors, and is useful to understand collective behaviors during the evolution. An instance of such duality is shown in Fig. 2.

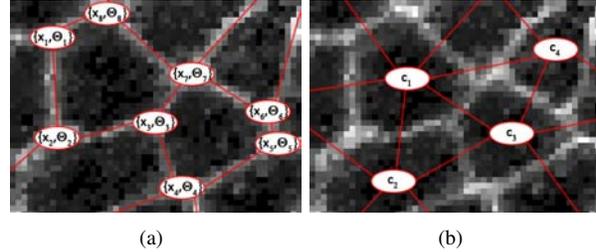


Fig. 2. **Primary and dual graphs.** (a) In the primary graph $\bar{\mathcal{G}} = (\bar{\mathbf{N}}, \bar{\mathbf{E}})$ nodes correspond to the bifurcation locations visited by the walker agents and the edges are directly related to the traveled paths. (b) Conversely, the dual graph $\bar{\mathcal{G}}' = (\bar{\mathbf{N}}', \bar{\mathbf{E}}')$ shows as nodes the centers of the cluster objects, and the connections with the neighbors assume the rôle of edges.

B. Tracking: motion and deformation maps

In the case a sequence of images is available in the discrete time domain $t = 1, \dots, T$ (t indicating the evolution time), represented by the image intensity set $\{I_1, \dots, I_T\}$, the temporal coherence, that is the property of an image frame I_t to be similar to its neighbor I_{t+1} , can be exploited to perform a dynamic detection of the structures or the tracking.

In general, to understand motion in an image sequence, computer vision literature focuses on two different concepts, those of motion field and optical flow: The former is a purely geometric and unique map $\mathcal{V} : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ that represents the projection of the real world 3D motion vectors on the 2D image plane, while the latter is the not uniquely determined map $\mathcal{F} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ that transforms one image into another [5]. Given these definitions, the purpose of motion estimation is to recover an optical flow that is as close as possible to the real motion field.

Considering two images I_t and I_{t+1} , the optical flow $\mathcal{F}(\mathbf{x})$ is a velocity vector \mathbf{v}_t that describes the displacement of pixel at \mathbf{x} of image I_t into I_{t+1} [2], [3]:

$$I_t(\mathbf{x}) = I_{t+1}(\mathbf{x} + \mathbf{v}_t). \quad (4)$$

The assumption is made of small displacement so that the linearization of (4) yields the image constraint equation

$$I_t(\mathbf{x}) = I_{t+\delta t}(\mathbf{x} + \delta \mathbf{x}) \approx I_t(\mathbf{x}) + \frac{\partial I}{\partial \mathbf{x}} \delta \mathbf{x} + \frac{\partial I}{\partial t} \delta t, \quad (5)$$

whence the optical flow constraint equation is derived, that can be written in a compact fashion as

$$\nabla I \cdot \mathcal{F} + \frac{\partial I}{\partial t} = 0. \quad (6)$$

To find the optical flow, another set of equations is needed, given by some additional constraint, which is often found in a smoothness constraint, $|\nabla \mathcal{F}| \approx 0$, assuming that motion field does not vary drastically between neighboring regions.

Accordingly, the graph structure \mathcal{G}_t deforms into \mathcal{G}_{t+1} and its *morphing* is the optical flow field restricted to the node set \mathbf{N}_t , where the pedices highlight the dynamic nature of these structures. The computation of such *morphing* is obtained as the solution to the following minimization problem:

Problem 1: For each point $\mathbf{x} \in \mathbf{N}_t$ a neighborhood $\mathcal{N}(\mathbf{x})$ in I_t is considered, and a corresponding point $(\mathbf{x} + \mathbf{v}_t) \in \mathbf{N}$ with neighborhood $\mathcal{N}(\mathbf{x} + \mathbf{v}_t)$ in I_{t+1} . The optimal value for the optical flow is obtained as

$$\mathcal{F}(\mathbf{x}) = \arg \min_{\mathbf{v}_t} \underbrace{\int_{\mathcal{N}} |I_t(\mathbf{x}) - I_{t+1}(\mathbf{x} + \mathbf{v}_t)| dx}_{J_{\mathbf{x}}}, \quad (7)$$

where the functional $J_{\mathbf{x}}(\mathbf{v}_t)$ as a function of \mathbf{v}_t , related to position \mathbf{x} , is highlighted.

While on images with high coherence, a unique minimum is found with the optimization procedure above, with low coherence images the solution to (7) may in fact yield a sub-optimal vector, due to the presence of several minima of the same magnitude in the functional $J_{\mathbf{x}}(\mathbf{v}_t)$. This issue is exemplified in Fig. 3 where two typical maps of $J_{\mathbf{x}}$, referred to as *J*-maps, are shown as a function of vector \mathbf{v}_t (the central point of the map corresponding to the null displacement, $\mathbf{v}_t = 0$): High coherence images (like that on the left) generate a precisely located and unique minimum, while low-coherence images (e.g. on the right) exhibit more than one good minima. In these

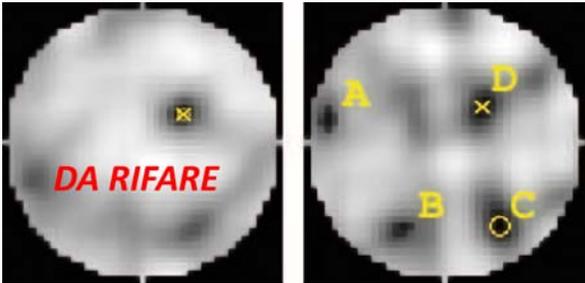


Fig. 3. *J*-maps. The yellow circle indicates the minimum of $J_{\mathbf{x}}(\mathbf{v}_t)$, while the yellow cross the ground-truth. A highly coherent image pair $\{I_t, I_{t+1}\}$ generates a *J*-map providing a well defined minimum (left). Instead, a low coherent image pair (right) generates a *J*-map with many local minima of similar value (A, B, C, D). In this case, it is more difficult to discriminate among these minima and, also, the global minimum (yellow cross, D) may not correspond to the ground truth (circle, C).

cases, to detect the position of the true minimum, the strategy is to consider instead of the single *J*-map, a set of neighboring *J*-maps (*J*-maps associated to neighbor points) that can mutually support the localization of the optimal displacement by averaging their values. This approach finds a justification in the smoothness constraint of the optical flow $\mathcal{F}(\mathbf{x})$: Neighbor *J*-maps will show similar patterns and approximately close minima

$$\mathbf{x}_1 \approx \mathbf{x}_2 \Rightarrow J_{\mathbf{x}_1}(\mathbf{v}_t) \approx J_{\mathbf{x}_2}(\mathbf{v}_t) \Rightarrow \mathcal{F}(\mathbf{x}_1) \approx \mathcal{F}(\mathbf{x}_2), \quad (8)$$

with straightforward meaning of the adopted notation.

Fig. 4 shows a set of neighbor *J*-maps, with a yellow cross on the global minima. Intuitively, it is possible to detect the error on the right and to (somehow) correct it.

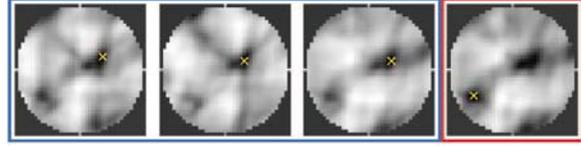


Fig. 4. **Low coherent J-maps.** *J*-maps from low coherent images show a number of local minima, and sometimes fail in detecting the ground truth. Nonetheless neighbor *J*-maps exhibits a similar appearance, and thus it's possible to cross check them and detect potent errors. For instance the image on the right shows a global minima not consistent with the global minima of the other *J*-maps - not consistent with the smoothness constraint

The *J*-maps correction is an iterative steps. For each image J_s we consider a set of neighbor images, and compute the global minima of each of them. The average value indicates respectively the optical flow average direction and the confidence of this measure.

Image 5 shows equation ?? applied to the example of image 4. The image on the right is \bar{J}_s , and the three remaining images on the left are $\{J_{s_1}, J_{s_2}, J_{s_3}\}$. After the

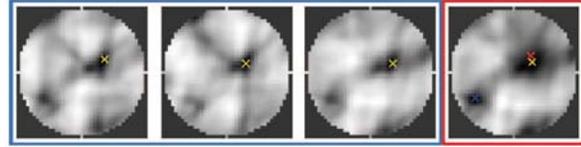


Fig. 5. **Correction of the J-Maps.** The *J*-map on the right, \bar{J}_s , is computed from J_s according to equation ?. The blu cross is the position of the old global minimum. The red cross is the center of the gaussian (\bar{v}_x, \bar{v}_y) . The yellow cross is the location of the new global minimum

correction step over all the images, the “new” \bar{J} overwrite the “old” J . The entire procedure may be repeated, until a steady state is reached.

Figure 6 compare the morphing of a reticular structure computed respectively with the raw optical flow algorithm and the improved one. Figure 7 compare the optical flows

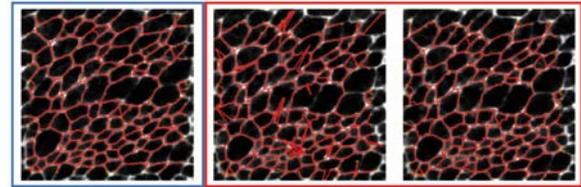


Fig. 6. **Morphing of a reticular structure.** The inexact computation of the optical flow can morph a shape (left, a reticular shape in red) into a wrong one (center). The improved optical flow algorithm guarantees instead optimal results (right). The images refer to a video sequence of the drosophila morphogenesis, captured by mean of a microscope. Red lines, over imposed on the original frames, show the cellular structure. The image on the left refers to frame 21, images on center and on the right refer to frame 22. See also images 7. Courtesy of Prof. Jeff Axelrod.

after and before refinement.

Evolving the Active Contour in a pseudo time removes false positives edges and refine the structure. Finally, we

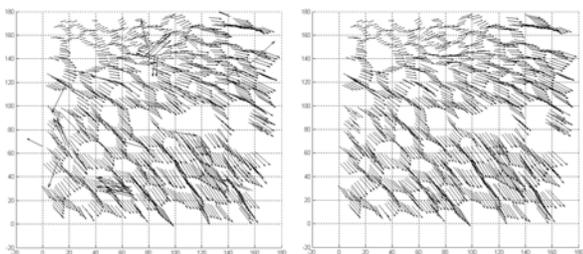


Fig. 7. **Optical flow field.** The two images refer to the optical flow vector field computed on 6. The raw computation (left) yields a wrong result. The smoothness constraint is not respected. The refined optical flow is instead a smooth vector field. See also images 6

generate a set of random walk agent from each node of the graph. This is necessary to fill holes and to “expand” the shape over the possibly appeared portion of the frame.

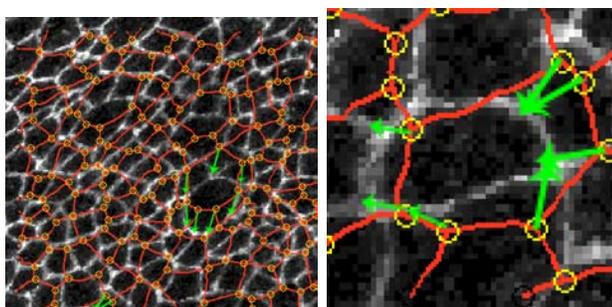


Fig. 8. **Morphing of the Drosophila reticular structure using J-Maps.** The image shows the reticular shape detected in frame 11 overimposed on frame 12. The green arrows schematically represents the morph obtained by means of J-maps, applied here to two cells

REFERENCES

- [1] K.W. Wojciechowski B. Smolka. Random walk approach to image enhancement. *Signal Processing* 81 (2001) 465-482, 2001.
- [2] S. S. Beauchemin and J. L. Barron. The computation of optical flow. *ACM Computing Surveys (CSUR)*, 27(2), 1995.
- [3] David J. Fleet and Yair Weiss. *Optical Flow Estimation - Handbook of Mathematical Models in Computer Vision*. Springer, 2006.
- [4] Leo Grady. Multilabel random walker image segmentation using prior models. *Proc. of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1(1), 2005.
- [5] Berthold K. P. Horn and Brian G. Schunk. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

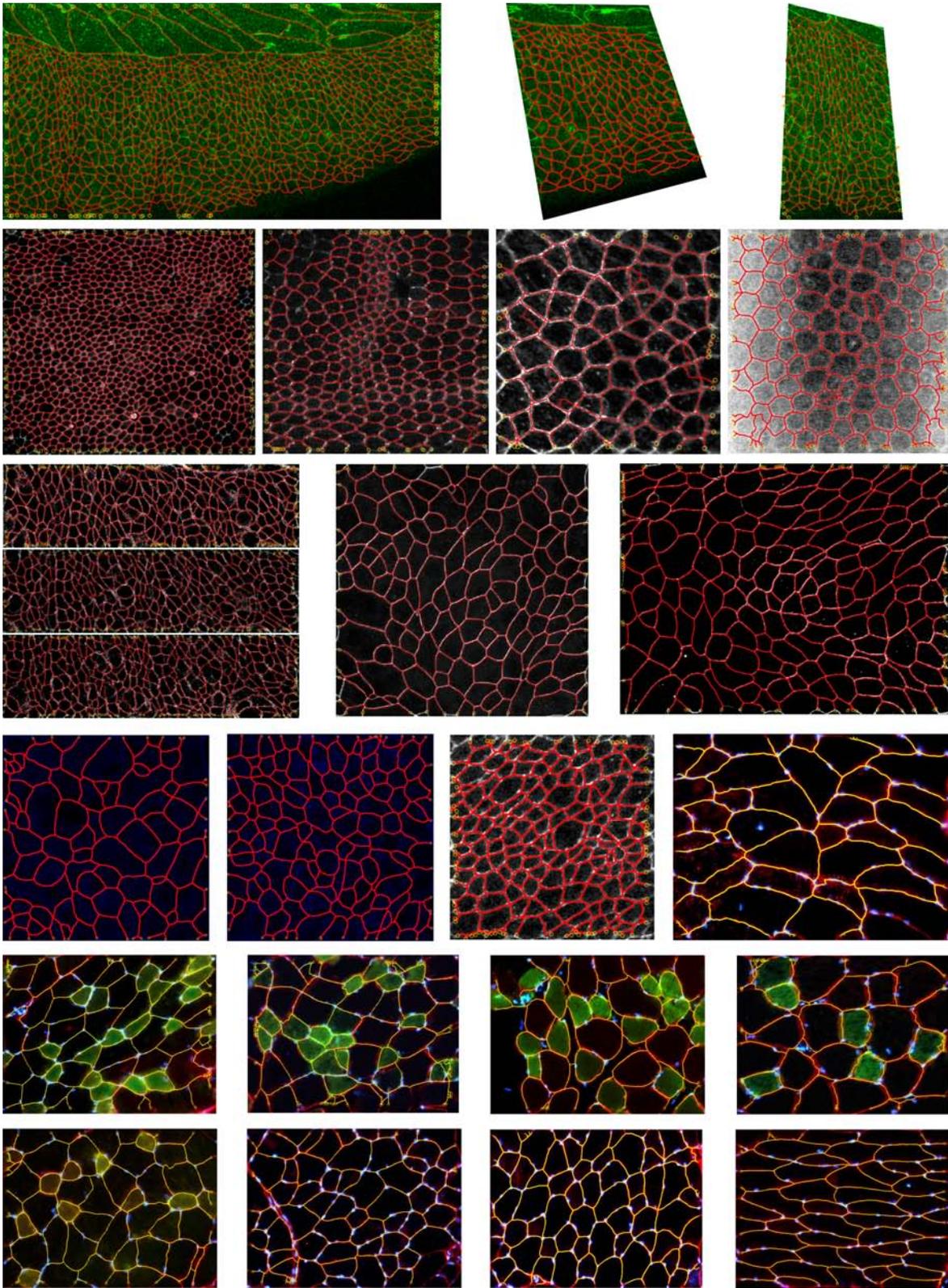


Fig. 9. Complete dataset IM1.