CELL PAIRINGS FOR ASCIDIAN EMBRYO REGISTRATION

Gaël Michelin[⋆1] *Léo Guignard*[†2] *Ulla-Maj Fiuza*[†]
Patrick Lemaire[†] *Christophe Godin*[‡2] *Grégoire Malandain*[⋆1]

[⋆]Inria, 06900 Sophia Antipolis, France
[†]CRBM, UMR 5237, CNRS, Univ. Montpellier 1 & 2, 34293 Montpellier, France
[‡]Inria, 34095 Montpellier, France

ABSTRACT

Recent microscopy techniques allow imaging temporal 3D stacks of developing organs or embryos with a cellular level of resolution and with a sufficient acquisition frequency to accurately track cell lineages. Imaging multiple organs or embryos in different experimental conditions may help decipher the impact of genetic backgrounds and environmental inputs on the developmental program. For this, we need to precisely compare distinct individuals and to compute population statistics. The first step of this procedure is to develop methods to register individuals.

From a previous work of cell segmentation from microscopy images, we here demonstrate how to extract the symmetry plane of embryos at early stages, and how to use this information as a geometrical constraint to both register these embryos and obtain a cell-to-cell mapping.

***Index Terms*—** Fluorescence microscopy, symmetry detection, embryogenesis, cell-to-cell mapping

1. INTRODUCTION

A central aim in developmental biology is to better understand how each tissue of an embryo progressively acquires its functional shape, a process called morphogenesis. Image-based studies therefore represent a method of choice. Current live microscopy techniques allow the acquisition of temporal sequences of 3D images with a spatio-temporal resolution high enough to follow embryo development at sub-cellular scale [1]. An automatic framework to register individual cells from distinct developing embryos would allow quantifying the variability in embryo development at the cellular level, which is a major issue in morphogenesis studies.

The present work describes a complete framework to register two embryos at similar developmental stages and to provide cell-to-cell mapping. The proposed framework takes advantage of a cell segmentation method based on differential and structural information (section 2) to derive a novel symmetry plane extractor (section 3), and cell-to-cell mapping registration method (section 4). Its efficiency is demonstrated on embryo microscopy images of the simple marine invertebrate chordate *Phallusia mammillata*.

Contact: {gael.michelin,gregoire.malandain}@inria.fr
[1]GM and GM are with Inria team Morpheme.
[2]LG and CG are with Inria team Virtual Plants, UMR AGAP, Montpellier, France.

2. CELL SEGMENTATION FRAMEWORK

Embryo image segmentation is performed with the method proposed in [2]. We recall here its main steps (Fig. 1):

1. Planar response filter using 1st and 2nd order derivatives, see Fig. 1(b),
2. Extrema extraction and binarization, see Fig. 1(c),
3. Gap filling by Tensor Voting framework, see Fig. 1(d), and
4. Cell detection and segmentation by seeded watershed approach, see Fig. 1(e)-(f).

The planar response filter, inspired from [3], uses the property that a membrane has homogeneous intensity locally in the plane tangential to the membrane whereas the intensity varies strongly in its normal direction. Therefore, the Hessian matrix eigenvector, denoted $\mathbf{n}$ in the next section, associated to the largest eigenvalue in absolute value indicates the normal direction to the membrane, while the two other eigenvectors are tangential to the membrane. Using 1st order derivatives of the image, an edge response at a constant distance to the voxel is then integrated on either side in the normal direction to the membrane. This yields the planar response filter.

Then, the maxima of the planar response filter with respect to the normal direction to the membrane, are extracted, so that most of the irrelevant information is removed. The resulting image is finally thresholded in order to segment membrane voxels. This last image will be denoted B in next section.

The unavoidable gaps are filled by the application of the tensor voting framework [4]. The binarised voxels are encoded as 2nd order tensors that describe local shape of the data, using the voxel orientation information obtained from Hessian matrix. The voting process consists, for each voxel, in propagating its local shape in its neighborhood. The accumulation of votes results in a new tensor map in which each tensor encodes structural information. A scalar surfaceness map is then extracted from this tensor map.

The surfaceness map is used as an input of a dedicated process for cell detection and segmentation by seeded watershed algorithm [5].

In the following sections, we will use the result of membranes binarisation (Fig. 1(c)) with associated membrane normal orientations to extract the embryo symmetry plane. The individualized cells (Fig. 1(f)) will be used for embryo co-registration.

3. SYMMETRY PLANE EXTRACTION

The embryo morphology holds a left-right (L-R) symmetry across early stages until the neurula rotation that takes place during the late neurula stage [6]. The interface between cells of both sides of the L-R symmetry forms a plane-like structure (Fig. 2).

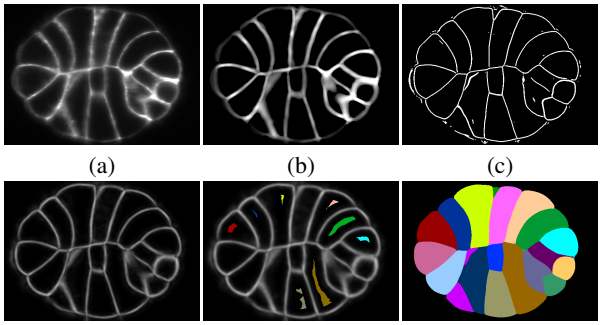


Fig. 1. Cell segmentation framework. (a) A section through a 3D embryo image. (b) Planar response filter. (c) Thresholded extrema. (d) Surfaceness map after tensor voting. (e) Seeds detection. (f) Cell segmentation by seeded watershed.

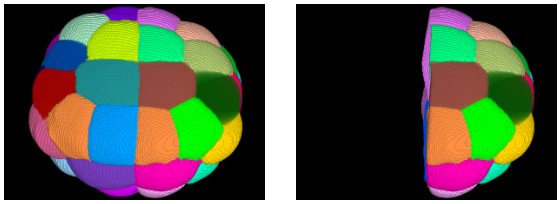


Fig. 2. 3D view of embryo cell segmentation. (a) Full embryo. (b) Embryo with cells of its left half hidden. This illustrates the plane-like structure of Left-Right cell interface.

The principle of the symmetry plane extraction is threefold:

1. estimation of a set of plane normals through the study of the membrane orientation distribution;
2. for each normal, the optimal plane is computed with an associated figure of merit: only the one with the best figure of merit is retained; and
3. the plane equation is refined using a least squares minimization.

B denotes the thresholded extrema of the segmentation method (see section 2), while $\mathbf{n}(b)$ is the normal orientation (i.e. the unit eigenvector associated to the Hessian matrix largest eigenvalue) of point $b \in B$. Since B may be noisy, we mask B with the image M of dilated frontiers of the segmented cells. $B \cap M$ is an image in which most of the false membrane detections are removed.

3.1. Symmetry plane normal estimation

Since many cell membranes participate or are parallel to the symmetry plane, it is hypothesized that a large number of membrane normal directions are aligned with the symmetry plane normal, thus that the distribution of membrane normal directions will exhibit a maximum for the symmetry plane direction.

To compute this distribution, we discretize the unit sphere into the set of vectors $\mathbf{N} = \{\mathbf{n}_i\}$ such that $-\mathbf{n}_i \in \mathbf{N}, \forall \mathbf{n}_i \in \mathbf{N}$ (typically around 2000 vectors), and compute a kernel density estimate of the distribution of the membrane normal directions onto the $\{\mathbf{n}_i\}$ (Fig. 3). The distribution value at $\mathbf{n}_i$ is given by

$$D(\mathbf{n}_i) = \frac{1}{|B \cap M|} \times \sum_{b \in B \cap M} \frac{1}{N(b)} w(\mathbf{n}_i, \mathbf{n}(b)), \qquad (1)$$

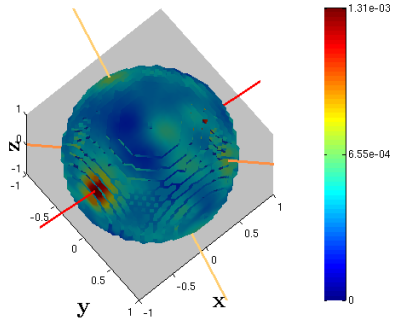


Fig. 3. Kernel density estimate of the orientation of binarised voxels. Red axis corresponds to the highest density. The other axes are local maxima greater than half the value of the global maximum.

where $|B \cap M|$ denotes the cardinal of $B \cap M$, $N(b)$ is a normalization constant, and the kernel $w(\cdot, \cdot)$ is defined by

$$w(\mathbf{n}_i, \mathbf{n}(b)) = \exp \frac{-\arccos(|\mathbf{n}_i.\mathbf{n}(b)|)^2}{2\sigma^2} \qquad (2)$$

with σ fixed to 2° since it ensures a maxima extraction fine enough to ensure satisfying candidate axes extraction. Let $\{\bar{\mathbf{n}}_k\}$ be the set of local maxima of the distribution D, $\{\bar{\mathbf{n}}_k\} \subset \{\mathbf{n}_i\}$. By construction, D is symmetrical, and since the search plane orientation is unoriented, we picked maxima in the half unit sphere. Moreover, let D_{max} be the maximal value of the distribution D, i.e. $D_{max} = \max_{\mathbf{n}_i} D(\mathbf{n}_i)$, we restrict the maxima to those that verify $D(\bar{\mathbf{n}}_k) \geq D_{max}/2$.

3.2. Symmetry plane equation global estimation

A plane P is defined by its normal $\mathbf{n} = (n_x, n_y, n_z)^T$ and a scalar d that sets the plane position along its normal axis. The plane equation is

$$P_{\mathbf{n},d}(x, y, z) = n_x x + n_y y + n_z z + d = 0. \qquad (3)$$

We hypothesize that the symmetry plane of normal $\bar{\mathbf{n}}_k$ should be at *the middle* of the embryo, more precisely that the voxels $b \in B \cap M$ that contribute to the local maximum $\bar{\mathbf{n}}_k$ are equally parted by the symmetry plane , which can be formalized by

$$d_k = \arg\min_d \left| \sum_{b \in B^+(\bar{\mathbf{n}}_k, d)} w(\bar{\mathbf{n}}_k, \mathbf{n}(b)) - \sum_{b \in B^-(\bar{\mathbf{n}}_k, d)} w(\bar{\mathbf{n}}_k, \mathbf{n}(b)) \right| \qquad (4)$$

where $B^-(\bar{\mathbf{n}}_k, d) = \{b \in B \cap M | P_{\bar{\mathbf{n}}_k, d}(b) < 0\}$ and $B^+(\bar{\mathbf{n}}_k, d) = (B \cap M) \setminus B^-(\bar{\mathbf{n}}_k, d)$.

3.3. Symmetry plane figure of merit

From the local maxima $\{\bar{\mathbf{n}}_k\}$, we have defined a set of planes $\{(\bar{\mathbf{n}}_k, d_k)\}$. Each plane can now be assessed with respect to the original data. Let S be the embryo segmentation, i.e. the union of the individualized cells, and $S \circ P$ the symmetrical of S with respect to plane P. If P is a symmetry plane, then S and $S \circ P$ are identical. We use the Dice coefficient to compare S and $S \circ P$ and finally retain as the symmetry plane the one that maximize the Dice coefficient,

$$(\hat{\mathbf{n}}, \hat{d}) = \arg \max_{(\bar{\mathbf{n}}_k, d_k)} \text{Dice}(S, S \circ P(\bar{\mathbf{n}}_k, d_k)). \qquad (5)$$

3.4. Symmetry plane final estimation

Experiments demonstrated that $P(\hat{\mathbf{n}}, \hat{d})$ is a quite good symmetry plane, but may be slightly different from the L-R cell interface. Indeed all points $b \in B$ contribute to the distribution D and the retained symmetry plane $P(\hat{\mathbf{n}}, \hat{d})$ may be impaired by points b that did not belong to the L-R cell interface.

For a more accurate estimation of the symmetry plane, we perform a last iterative least-squares estimation until plane convergence where points b far away from the symmetry plane or with an orientation different from the plane normal are penalized:

$$(\mathbf{n}, d)^{(i+1)} = \arg\min_{(\mathbf{n},d)} \sum_{b \in B} \rho_{(\mathbf{n},d)^{(i)}}(b) \|P_{\mathbf{n},d}(b)\|^2, \qquad (6)$$

with

$$\rho_{(\mathbf{n},d)^{(i)}}(b) = \exp \frac{-\left(P_{(\mathbf{n},d)^{(i)}}(b)\right)^2}{2\sigma_d^2} \exp \frac{-\arccos(|\mathbf{n}^{(i)}.\mathbf{n}(b)|)^2}{2\sigma_a^2}$$

$(\mathbf{n}, d)^{(0)}$ being initialized by $(\hat{\mathbf{n}}, \hat{d})$. The parameters σ_d and σ_a are the standard deviations weighting respectively the distance to the plane and the angular difference between $\mathbf{n}_b$ and the plane normal. We fix σ_d at half the approximate diameter of an embryo cell (it depends on the embryo stage) and arbitrarily set the parameter σ_a to $5°$.

3.5. Experiments

We worked on *Phallusia mammillata* embryo images coming from 5 different samples.

For the first embryo, the membranes were marked by a lipophilic dye (FM4-64). The other embryo membranes were genetically marked using PH-GFP [7]. Imaging of all samples was done at each time point from 4 different angles separated by 90 degrees with a MuViSPIM microscope [8]. The 4 raw images are fused into one in order to have a time point image less sensitive to imaging defects and higher and isotropic resolution.

We tested the proposed method on 115 images: 99 images of the first embryo (ranging from 32 cells to 172 cells stages), 9 images of the second embryo (64 cells and 112 cells stage), 5 images of the third embryo (64 to 70 cells stage), one image of the fourth embryo (114 cells stage) and one image of the fifth embryo (mid neurula stage) that counts around 450 cells.

Visual inspection demonstrates that the method worked perfectly on all of the tested images (see Fig. 4). The last least squares estimation yields often similar result to the global estimation $P(\hat{\mathbf{n}}, \hat{d})$, but clear differences appear on some images (typically Fig. 4(c)-(f)).

With our implementation, the kernel density estimate building is computationally the most expensive step and is processed in about 20 minutes for an image of size $400 \times 400 \times 400$, while the plane initialization and refinement steps take less than 10 seconds.

4. EMBRYOS REGISTRATION

At early stages, ascidian development is stereotyped and invariant, i.e., there exists a cell-to-cell mapping between different embryos at the same stage. Such mapping offers a powerful means for population studies with single cell resolution but is challenging to establish. Indeed, at early stages, the embryo resembles a sphere tiled with cells, and pairing cells requires experimented researchers. Even after gastrulation, where a concavity (the blastopore) appears that is a natural landmark, recognizing pairs is made difficult by the increasing number of cells. We propose a method to register embryos (here with an affine transformation), which enables to build a further cell-to-cell mapping.

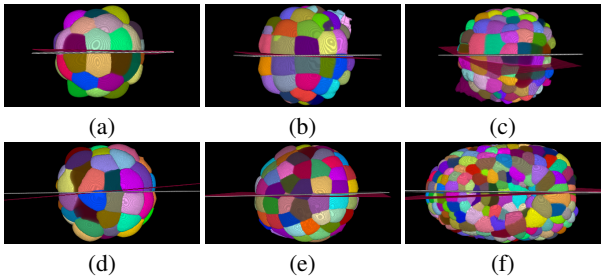


Fig. 4. L-R symmetry plane initialization ($P(\hat{\mathbf{n}}, \hat{d})$, in red) and after least squares estimation (in white). (a-c) First embryo, 32-, 76-cells and early gastrula stages. (d-e) Second embryo, 64- and 112-cells stages. (f) Fifth embryo, mid neurula stage.

In this section, E and F are two embryo images at the same developing stage, with respectively the following notations:

- (P) and (Q) are their computed symmetry planes,
- $\mathbf{n}_P$, $\mathbf{n}_Q$ are normal vectors of (P), (Q),
- G, H are the centers of mass of the entire embryo cell segmentation,
- g, h are the projections of G and H on (P) and (Q),
- $C = \{c_i\}_{i \in [1,n]}$, $D = \{d_i\}_{i \in [1,m]}$ are the centers of mass of the segmented cells. Please note that it is not required to have an error-free segmentation.

4.1. Embryo registration

To register two embryos, the transformation that optimizes a cell-to-cell mapping is computed. Since the true mapping is unknown, several of them are compared. Obviously, testing all cell-to-cell pairings between two embryos is computationally intractable. However, it is reasonable to assume that registering embryos implies that their symmetry planes superimpose. Using this geometrical constraint allows to reduce greatly the number of cell pairings to be tested.

Let $\mathbf{I}^{(+)}$ and $\mathbf{I}^{(-)}$ be the rigid transformations that align h on g and $\mathbf{n}_Q$ on either $\mathbf{n}_P$ or $-\mathbf{n}_P$. This provides a reasonable alignment of both embryos up to one last degree of freedom, namely the angle θ of a rotation $\mathbf{R}(\theta)$ of center h and axis $\mathbf{n}_Q$. Let $\mathbf{T}^{(+)}(\theta) = I^{(+)} \circ \mathbf{R}(\theta)$, $\mathbf{T}^{(-)}(\theta) = I^{(-)} \circ \mathbf{R}(\theta)$. $[0, 2\pi]$ is discretized into the set of angles $\{\theta_i\}$ (we took 128 angles), yielding a set of transformations $\{\mathbf{T}^{(+)}(\theta_i)\} \cup \{\mathbf{T}^{(-)}(\theta_i)\}$.

For each transformation $\mathbf{T}$ of this set, we built a set of pairs of cell barycenters $\{(c_j, d_j)\}_{\mathbf{T}}$ such that they are the closest to each other, meaning that

$$\|c_j \mathbf{T}(d_j)\| = \min_{d_i \in D} \|c_j \mathbf{T}(d_i)\| = \min_{c_i \in C} \|c_i \mathbf{T}(d_j)\|. \qquad (7)$$

Some cell barycenters c_i or d_i may remain unpaired. From the pairings $\{(c_j, d_j)\}_{\mathbf{T}}$, we estimated a transformation (we chose affine transformations) that minimizes the average of the square residuals:

$$\hat{\mathbf{T}}(\mathbf{T}) = \arg\min_{\mathbf{t}} r_{\mathbf{T}}(\mathbf{t}) \text{ with } r_{\mathbf{T}}(\mathbf{t}) = \frac{1}{|\{(c_j, d_j)\}_{\mathbf{T}}|} \sum_j \|c_j \mathbf{t}(d_j)\|^2.$$

Instead of a least squares estimation, and to discard erroneous pairings (e.g. due to segmentation errors), we preferred a least trimmed squares estimation [9], so that the pairings with the largest residuals are excluded from the transformation calculation. We chose to discard 20% of the pairings, i.e. the $k = \lfloor 0.2 \times |\{(c_j, d_j)\}_{\mathbf{T}}| \rfloor$ worst pairings, since it has been estimated that the segmentations have less than 10% of errors.

Last, the transformation $\mathbf{T}_{opt}$ that best registers the two embryos is the one that has the minimum average of the square residuals:

$$\mathbf{T}_{opt} = \arg \min_{\hat{\mathbf{T}}(\mathbf{T}), \mathbf{T} \in \{\mathbf{T}^{(+)}(\theta_i)\} \cup \{\mathbf{T}^{(-)}(\theta_i)\}} r_{\mathbf{T}}(\hat{\mathbf{T}}(\mathbf{T})). \quad (8)$$

4.2. Cell-to-cell mapping

The transformation $\mathbf{T}_{opt}$ allowed to build the final cell-to-cell mapping $\{(c_j, d_j)\}_{\mathbf{T}_{opt}}$ between the two embryos, using the same symmetrical constraint as Eq. 7

$$\|c_j \mathbf{T}_{opt}(d_j)\| = \min_{d_i \in D} \|c_j \mathbf{T}_{opt}(d_i)\| = \min_{c_i \in C} \|c_i \mathbf{T}_{opt}(d_j)\|. \quad (9)$$

4.3. Experiments

We used images from first to fourth embryos (described in section 3.5) for our experiments. We extracted from the set of images of the first, second and third embryos an image corresponding to the 64 cells stage and from the first, second and fourth embryos another image corresponding to the 112 cells stage, so that for both 64 cells stage and 112 cells stage, we could experiment the registration method on three image pairs coming from three distinct embryos.

Since we are interested in automated processing of high throughput acquisitions, we do not perform any correction of the cell segmentation provided by the framework of section 2. This test enables to evaluate the method's robustness facing segmentation errors. At the 64 cells stage, the first embryo has 71 segmented regions (instead of 64) due to 7 over-segmentation errors while the second one has 66 regions due to 2 over-segmentation errors and the third one has 69 regions due to 5 over-segmentation errors. At the 112 cells stage, the first embryo has 133 regions (instead of 112) due to 21 over-segmentation errors while the second one has 113 regions due to 1 under-segmentation error and 2 over-segmentation errors and the fourth one has 114 regions correctly segmented.

- For the 64 cells stage, the final cell-to-cell mapping built 64, 63 and 60 pairs between respectively the first and the second, the first and the third, and the second and the third embryo.
- For the 112 cells stage, the final cell-to-cell mapping built 105, 102 and 100 pairs between respectively the first and the second, the first and the fourth, and the second and the fourth embryo.

As demonstrated by Fig. 5, and confirmed by careful visual inspection, almost all the pairings are correct. Mapping errors may be due to segmentation errors and/or limitations of affine transformation that did not allow a perfect superimposition of both embryos. Given that no other information has been incorporated (as adjacency between cells, etc.), there is room for further improvements that are however beyond the scope of this paper.

5. CONCLUSION AND FUTURE WORK

We proposed a novel method for early embryo Left-Right symmetry detection and showed that it enables to register embryos at a single cell level. The ability of robustly identifying corresponding cells in different embryos is a breakthrough for embryology, since it offers the means to first, conduct population statistical analysis at a cell level, and second, to register an atlas (i.e. a perfectly segmented template) onto any embryo which will thus allow to correct unavoidable segmentation errors.

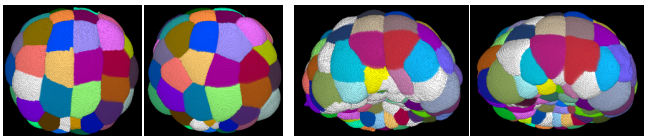


Fig. 5. Final region correspondence maps between the first and the second embryos. Left: 64 cells stage embryos. Right: 112 cells stage embryos. The white regions are unpaired ones.

Future research directions consist in improving the cell-to-cell mapping by incorporating additional information (lineages, cell adjacencies, etc.) and in building an average developing embryo from a population.

Acknowledgment: G. Michelin is funded by the Inria Project-Lab Morphogenetics. U.M. Fiuza was initially funded by the 2008 GeneShape ANR contract (Sysoomm program), then by Fondation pour la Recherche Médicale. P. Lemaire is a member of CNRS.

6. REFERENCES

[1] PJ Keller, "Imaging Morphogenesis: Technological Advances and Biological Insights," *Science*, vol. 340, no. 6137, pp. 1234168+, June 2013.

[2] G Michelin, L Guignard, U Fiuza, and G Malandain, "Embryo Cell Membranes Reconstruction by Tensor Voting," in *ISBI*, Beijing, China, Apr. 2014, IEEE.

[3] K Krissian, G Malandain, N Ayache, R Vaillant, and Y Trousset, "Model-Based Detection of Tubular Structures in 3D Images," *Comput Vis Image Underst*, vol. 80, no. 2, pp. 130–171, 2000.

[4] G Medioni, MS Lee, and CK Tang, *Computational Framework for Segmentation and Grouping*, Elsevier Science Inc., New York, NY, USA, 2000.

[5] R Fernandez, P Das, V Mirabet, E Moscardi, J Traas, JL Verdeil, G Malandain, and C Godin, "Imaging plant growth in 4-D: robust tissue reconstruction and lineaging at cell resolution," *Nat Meth*, vol. 7, pp. 547–553, 2010.

[6] K Nishide, M Mugitani, G Kumano, and H Nishida, "Neurula rotation determines left-right asymmetry in ascidian tadpole larvae," *Development*, vol. 139, no. 8, pp. 1467–75, April 2012.

[7] M Carroll, M Levasseur, C Wood, M Whitaker, KT Jones, and A McDougall, "Exploring the mechanism of action of the sperm-triggered calcium-wave pacemaker in ascidian zygotes," *Journal of Cell Science*, vol. 116, no. 24, pp. 4997–5004, 2003.

[8] U Krzic, S Gunther, TE Saunders, SJ Streichan, and L Hufnagel, "Multiview light-sheet microscope for rapid in toto imaging," *Nat Methods*, vol. 9, no. 7, pp. 730–3, 2012.

[9] PJ Rousseeuw and AM Leroy, *Robust Regression and Outlier Detection*, John Wiley & Sons, New York, NY, USA, 1987.