Cell lineage tree reconstruction from time series of 3D images of zebrafish embryogenesis

Robert Spir[1,2], Karol Mikula[1,2], Nadine Peyrieras[3]

[1]Department of Mathematics, Slovak University of Technology, Radlinskeho 11, 810 05 Bratislava, Slovakia
[2]Algoritmy:SK s.r.o., Sulekova 6, 811 06 Bratislava, Slovakia
[3]Institut de Neurobiologie Alfred Fessard, CNRS UPR 3294 Av. de la Terrasse, 91198 Gif-sur-Yvette, France

Abstract. The paper presents numerical algorithms, postprocessing and validation steps for an automated cell tracking and cell lineage tree reconstruction from large-scale 3D+time two-photon laser scanning microscopy images of early stages of zebrafish (Danio rerio) embryo development. The cell trajectories are extracted as centered paths inside segmented spatio-temporal tree structures representing cell movements and divisions. Such paths are found by using a suitably designed and computed constrained distance functions and by a backtracking in steepest descent direction of a potential field based on these distance functions combination. Since the calculations are performed on big data, parallelization is required to speed up the processing. By careful choice and tuning of algorithm parameters we can adapt the calculations to the microscope images of vertebrae species. Then we can compare the results with ground truth data obtained by manual checking of cell links by biologists and measure the accuracy of our algorithm. Using automatic validation process and visualisation tool that can display ground truth data and our result simultaneously, along with the original 3D data, we can easily verify the correctness of the tracking.

1 Introduction

The comprehensive image analysis for complex stages of embryogenesis is a difficult problem not yet solved satisfactory. Modern imaging technologies generate terabytes of image data, comprising up to tens of thousands of cells imaged for thousands of time points. However, existing manual or semi-automated approaches to reconstructing cell lineages do not scale to data sets of such complexity and size. Automated computational approaches have been developed to analyse such image data for small model organisms such as Caenorhabditis elegans [2] embryos and for early developmental stages of more complex organisms such as the early zebrafish blastula [20, 11] and the Drosophila blastoderm [23, 11]. However, a development of methods for accurate, automated cell lineaging in later stages of development is still a hot topic and open problem.

In [1], methods based on sequential Bayesian approach with Gaussian mixture models were developed. First, a partition of the 3D image volume recorded

at each time point into supervoxels is performed. A supervoxel is a connected set of voxels in space that all belong to a single nucleus, and each nucleus can be represented by multiple supervoxels. Second, an interconnection of supervoxels in space (segmentation) and time (tracking) is done to recover full cell lineages. The authors developed a sequential Bayesian approach with Gaussian mixture models (GMMs) to perform both tasks simultaneously using parametric contour evolution. The parametric model reduces the segmentation and tracking problem to finding ten parameters per nucleus: its 3D center, 3D covariance matrix (shape) and parent identity.

A recent work towards building the cell lineage tree for the complex stages of Zebrafish embryo development based on stochastic simulated annealing minimization of a heuristic energy functional has been presented in [9]. After the construction of the tree, the individual cells or cell populations are tracked producing lineage "forest" as an union of several disjoint trees representing cell lineage. First, the edges between detected cell centers at consecutive time steps are created using nearest-neighbour heuristic method. Next, the simulated annealing, variant of Metropolis algorithm, is used to progressively enforce a set of predefined constrains summarizing together a certain number of biological requirements, such as no cell should have more than two "daughters" and divisions should not occur too frequently.

The method presented in this paper is based on extraction of the cell trajectories as centered paths inside 4D spatio-temporal tree structures obtained by segmentation of 4D images. In addition to the approaches described by Mikula et al. in [18, 19], in the presented approach the 4D segmentation is obtained by creating spatio-temporal tubes using cell nuclei diameters obtained from real nuclei segmentations [6, 12, 21, 4, 17, 9] around the cell identifiers given as a result of suitable image filtering [8, 14, 4, 9] followed by a cell detection algorithm [10, 4, 9]. Then a computation of constrained distance functions inside 4D segmentation is performed by solving numerically a spatially 4D eikonal equation. By a suitable combination of computed distance functions we build a potential field which is backtracked in steepest descent direction in order to get the cell trajectories. In contrast to [18, 19], in this paper we significantly improve the tracking results by introducing a new parameter α in the construction of the potential field, weighting the distance functions influence. The cell lineage tree can be constructed by detecting merging trajectories when going backward in time indicating mitosis and thus a branching node of the cell lineage tree. In the paper we discuss the results of our improved method and perform its validation on real ground-truth data. This ground truth data contains 38797 manually checked, correct cell links. First we find the correspondence between trajectory points in our result and ground truth data and then we check for the exactly matching cell links. Using the presented method we are able to obtain 96.5% of correct cell links accuracy which is on the top of existing tracking methods.

The data we are dealing with are given by two-photon laser scanning microscopy and represent the first hours of zebrafish embryo development, approximately from the 4th until 10th-20th hour. The labeling of cell nuclei is obtained

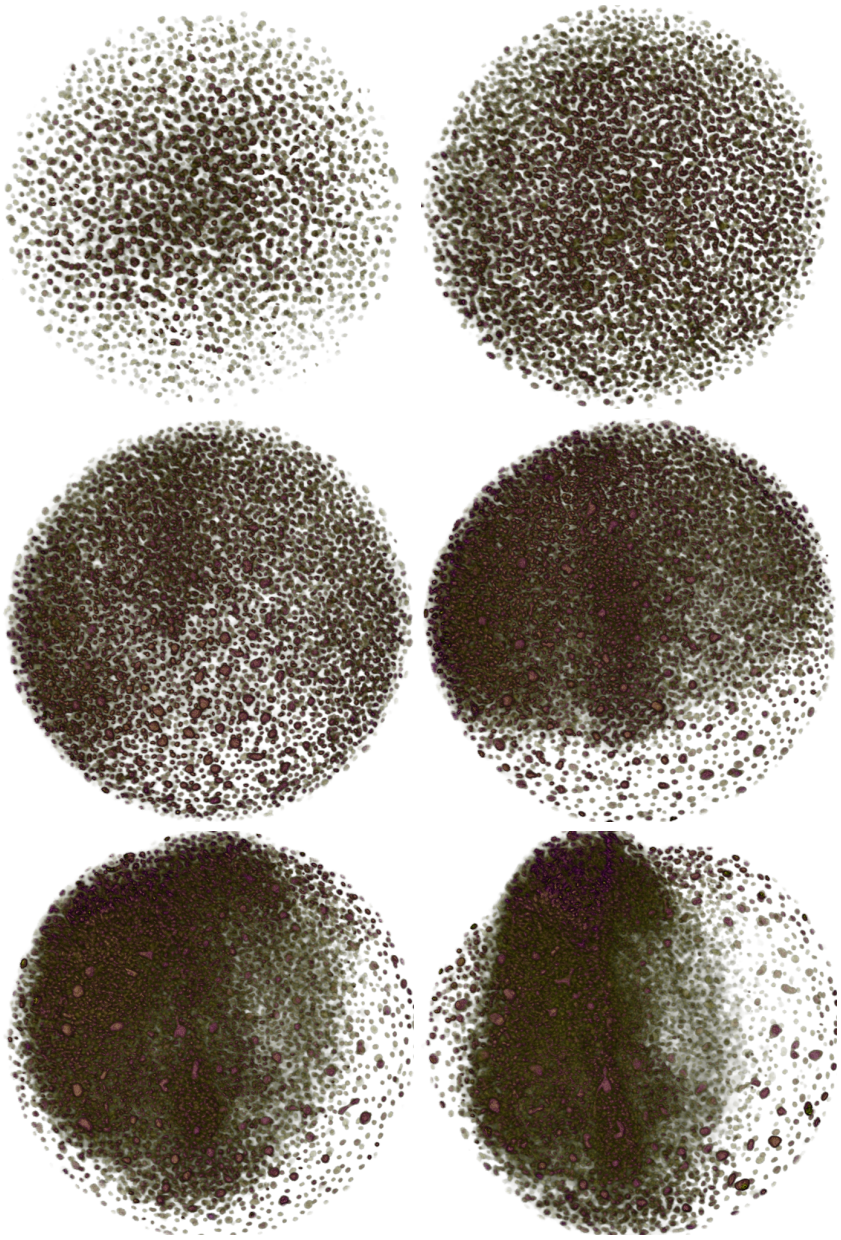

Fig. 1. Volume rendering of the cell nuclei data during the embryogenesis starting at 4 hours after fertilization until 13 hours after fertilization in time steps 1, 96, 192, 288, 384 and 480.

by expression of the fluorescence protein through its RNA injection performed at the one-cell stage. The 3D images are obtained by moving the focal plane from the top more deeply inside the embryo and their quality depends on the speed of scanning in one plane, we refer to a web page for various quality datasets (http://bioemergences.iscpif.fr/bioemergences/). The 3D image acquisition step ranges from 50 seconds to 5 minutes. A longer time step produces better image quality and such data is well suited for segmentation purposes e.g. for obtaining a shape of cells and other their characteristics during the embryogenesis [4, 17]. On the other hand, such data is not suitable for tracking since the cells move too far between single 3D images and consequently mother-daughters cell correspondences can be lost. In Fig. 1 we plot an example of embryo development from the beginning to the end of the imaging. We see visualization of 3D cell nuclei which are tracked (Fig. 4) by our method. One can clearly observe how the zebrafish embryo grows from an unorganized set of cells to complex stages of development containing presumptive organs of the future zebrafish adult.

The paper is organized as follows. In the next section we present our approach to cell trajectories and lineage tree extraction, discuss numerical approaches used in the tracking method and parameters that can be used to tune and improve tracking results. Then we discuss numerical experiments devoted to processing of real 3D+time image sequences of the early zebrafish embryogenesis. We also present comparison of our results with ground truth data containing correct cell links in time verified manually.

2 The cell lineage tree reconstruction algorithm

2.1 Filtering, cell nuclei detection and segmentation

These first three algorithmic steps are taken from [14, 10, 15] and for completeness are described below. First step in our approach is nonlinear diffusion filtering of the input data. The noise is intrinsically linked to the microscopy technique and its level increases with decreasing the time step $d\theta$ of the scanning. Here we use geodesic mean curvature flow (GMCF) [14, 9] which is based on discretization of the following nonlinear diffusion equation

$$u_t - |\nabla u| \nabla . \left(g \left(|\nabla G_\sigma * u| \right) \frac{\nabla u}{|\nabla u|} \right) = 0 \,, \tag{1}$$

where $u(t, x)$, $t > 0$, represents the filtered image intensity function. We start from the initial condition $u(0, x) = u^0(x)$, where u^0 is an original 3D image and we consider few discrete time steps of the discretized model. We consider the zero Neumann boundary conditions on the boundary $\partial\Omega$ of the 3D image domain Ω. In this model, the mean curvature motion of the level sets of function u is determined by the edge indicator function

$$g(s) = \frac{1}{1 + Ks^2}, \qquad K \geq 0 \tag{2}$$

where K is the edge detection sensibility parameter, that is applied to the image gradient presmoothed by the Gaussian kernel G_σ with a small variance σ. The essential property of this function is that its negative gradient points towards the edges in the image and overall nonlinear diffusion process given by (1) causes accumulation of the level sets of u along the boundaries of objects in the image and therefore the filtering is also edge preserving. The filtered image u^F is obtained as the solution of (1) at a time $t = T_F$. The optimal choice of the model and discretization parameters was studied in [14] on the basis of the mean Hausdorff distance of the level sets of the filtered image and a gold standard. For discretization of (1) in time we use the semi-implicit approach that guarantees unconditional stability of the numerical scheme, for details see [14] and for spatial discretization we use the finite volume method.

The second step of our approach is the detection of cell nuclei centers, we call them also cell identifiers. The cell center detection method is based on a fact that objects visible in the image can be seen as humps of relatively higher image intensity. Any such hump is represented by certain image intensity level sets. The diameter of these level sets allow us to distinguish between significant objects, e.g. cell nuclei, and spurious inner structures which still remain after GMCF filtering. For cell nuclei, the diameter d is relatively large, $0 << c_1 \leq d \leq c_2$, while the diameter of the spurious inner structures is much smaller, $0 < d << c_1$. If the level sets are moving (advected) at a constant speed in the direction of the inner normal, the encompassed volume is decreasing and finally the hump disappears. Our model is based on the fact that the level sets with small diameter corresponding to spurious structures disappear quickly while level sets representing cell nuclei are observable in a much longer time scale. Since the motion of every level set is given by the normal velocity $V = \delta + \mu k$ where δ and μ are constants (model parameters) and k is the mean curvature, we formulate our level set center detection (LSCD) method in the form of the following nonlinear advection-diffusion equation [10, 9]

$$u_t + \delta \frac{\nabla u}{|\nabla u|}.\nabla u - \mu |\nabla u| \nabla . \left(\frac{\nabla u}{|\nabla u|} \right) = 0 \tag{3}$$

which is applied to the initial condition given by u^F, the result of GMCF filtering. Again, we consider the zero Neumann boundary condition and the equation is solved in time interval $[0, T_C]$. Due to the shrinking and smoothing of all (real and spurious) structures in the evolutionary process represented by (3), we observe decrease of the number of local maxima M of the solution u as time proceeds. This decrease is fast in the beginning and much slower later. We stop this process when the slope of decrease is below a certain threshold, and then we use visual inspection in few 3D images of the whole sequence in order to choose an optimal evolutionary step of LSCD for finding cell identifiers. The time discretization of LSCD equation (3) is explicit in advective and semi-implicit in diffusion parts and it uses the finite volume method together with up-wind principle for space discretization [10].

The third step in our approach is segmentation of cell nuclei to obtain approximate cell diameters used in our new 4D segmentation approach. Here we use the generalized subjective surface method for 3D image segmentation [21, 15, 9]. Let $I^0 : \Omega \to R, \Omega \subset R^3$ represent the intensity function of an image, usually the image after GMCF filtering. If we want to segment an object, we need a segmentation seed - the starting point that determines the approximate position of the object in the image. Then we construct an initial segmentation function $u^0(x)$. For nuclei segmentation, all isosurfaces of the initial segmentation function were equal ellipsoids centered in the detected cell center. The radii were approximated from average nuclei radius given by biologists. The partial differential equation for generalized subjective surface (GSUBSURF) method is given by

$$u_t - w_a \nabla g \cdot \nabla u = w_c g \sqrt{\varepsilon^2 + |\nabla u|^2}\, \nabla \cdot \left(\frac{\nabla u}{\sqrt{\varepsilon^2 + |\nabla u|^2}} \right) \tag{4}$$

and comes from the level set formulation of the geodesic active contour model [6, 7, 12, 13, 21]. In the model (4), g is an edge detector function, for which we again use $g(s) = \frac{1}{1+Ks^2}$, where K is the edge detection sensibility parameter and $s = |\nabla I^0|$, where I^0 is the input image intensity function. Parameters w_a and w_c are weights for the advection and curvature terms of the model, respectively and ε is the regularization parameter, usually $\varepsilon << 1$. The same ε-regularization is used in eqs. (1) and (3). We choose zero Dirichlet boundary condition for the equation (4). In order to discretize (4) in time, we apply the semi-implicit approach that guarantees unconditional stability with respect to the diffusion term. In order to discretize (4) in space, we apply the so called flux-based level set finite volume method for advective part and in curvature part we use approach similar to the ones used in GMCF and LSCD discretizations [15, 17].

2.2 Tracking algorithm steps

Our final step, the method for cell trajectories extraction and cell lineage tree reconstruction is composed of the following steps:

- construction of a 4D segmentation yielding the 4D spatio-temporal tubular tree structure, chapter 2.3,
- computation of the first constrained distance function D giving distance of any point of 4D segmentation to the most far (backwards in time) cell identifier to which it is continuously connected, chapter 2.4,
- computation of the second constrained distance function D_B giving distance of any point of 4D segmentation to its boundary, chapter 2.4,
- building a potential field V for tracking by using a suitable combination of two computed distance functions, chapter 2.4,
- extraction of the steepest descent paths of the potential field inside all simply connected 4D segmentation regions, chapter 2.5,

– centering the extracted paths inside the 4D spatio-temporal trees in order to get unique cell trajectories, chapter 2.5,
– postprocessing and validation of the results, chapter 2.6

We note that after successful cell trajectories extraction the reconstruction of the cell lineage tree can be performed by detecting trajectories which merge together when going backward in time indicating mitosis and thus a branching node of the cell lineage tree.

2.3 Building the 4D segmentation

From mathematical point of view, the 3D+time image sequence is understood as a function $u(x_1, x_2, x_3, \theta)$, $u : \Lambda \to [0, 1]$, where Λ is a bounded spatio-temporal (rectangular) subdomain of R^4, (x_1, x_2, x_3) is a spatial point and θ represents a time.

The 4D segmentation is a spatio-temporal structure which approximates the space-time movement of cell nuclei. Due to [4, 15] the shape of cell nuclei during zebrafish embryogenesis is reasonably approximated by spheres or ellipsoids. Thus, in order to construct 4D segmentation we use cell identifiers detected in all time steps, $s_m^l, m = 1, \ldots, n_C^l, l = 1, \ldots, N_\theta$ (m denotes cell identifier index at time step l and N_θ is number of time steps) by method from [4, 10], and create 4D ellipsoids around all these points. To determine halfaxes of the ellipsoids we use real cell nuclei segmentations obtained using GSUBSURF method [4, 15] paired with cell coordinates from cell detection step. We calculate the volume of real segmented nucleus and compute the radius of a sphere with the same volume. This radius is then used as spatial halfaxes for constructed ellipsoids. Here, we also introduce a parameter S representing shrinking of the halfaxes (if $S < 0$) or expanding of them (if $S > 0$). A slight shrinking of real radius is used later in the tracking algorithm since it helps to have spatially non-overlapping tubular structure representing the cell movement. This parameter is tuned by comparison of tracking with ground truth data and its optimal choice improve the quality of tracking results. In temporal direction we are using halfaxis equal to $d\theta$ corresponding to the image acquisition interval. The nonzero temporal halfaxis is important due to the time overlap which we create and thus we improve connectivity of 4D spatio-temporal tree structures. Thanks to the time overlap we interconnect branches of the 4D spatio-temporal tree where a cell center was not detected in one frame but it was detected in two neighboring frames and thus we correct false negative errors of the cell center detection algorithm.

2.4 Building the potential field for tracking

As noticed above, for building the potential field V we compute two types of distance functions, D and D_B, inside the 4D spatio-temporal segmented tree structures. The distance functions are computed by OpenMP implementation of 4D Rouy-Tourin scheme [5] or they can be determined by fast-marching or

fast-sweeping methods [22, 24]. These distance functions are called constrained because all the calculations are constrained by the boundaries of the 4D segmentation. Due to that fact, the computed distances between doxels of the 3D+time image sequence are not a standard Euclidean distances in R^4 but they approximate a minimal Euclidean paths between the points inside the 4D segmentation.

We represent the 4D segmentation by a 4D piecewise constant function, with some BIG value (value that is bigger than the biggest distance in dataset) outside of the segmentation and with zero value inside it. The 4D distance function $D(x_1, x_2, x_3, \theta)$ is calculated gradually inside all simply connected regions, starting from cell centers in lowest possible time step θ. After the calculation is completed in all regions reachable from these cell centers, we fix the computed values and continue the calculation from centers in next time step, but only in regions where the values are not yet fixed. Using this approach we calculate the distance function D inside whole 4D segmentation. At the end all doxels inside the 4D segmentation contain the value of distance to the most far (backwardly in time) cell identifier to which it is continuosly connected. In Fig. 2 left we show that inside the regions encompassed by the BIG values, the value of D is growing from zero, in cell identifier where the simply connected component "begins", up to a locally maximal value, where the simply connected component "ends".

We could think about D as a potential field and traverse it in the steepest descent direction from the local maxima at every simply connected component to the zero value. The paths obtained in such way would represent good approximation of the space-time cell trajectories. And, if the 4D segmentation would contain only perfectly separated 4D spatio-temporal tree structures, we would obtain correctly all (also partial) cell trajectories which can be extracted from the data. Unfortunately, in the real 4D data it is not always possible and we must deal with imperfections given mainly by a cells overlapping. In order to overcome this difficulty we have to keep the extracted paths in a certain distance from the spatio-temporal cell boundaries or, in other words, they should be more centered inside the 4D segmentation.

This can be achieved by using the constrained distance function $D_B(x_1, x_2, x_3, \theta)$ [3, 16] values of which grow from boundaries to the center of the 4D spatio-temporal trees, see Fig. 2 right.

Finally, we build a potential field

$$V(x_1, x_2, x_3, \theta) = D(x_1, x_2, x_3, \theta) - \alpha D_B(x_1, x_2, x_3, \theta) \tag{5}$$

which is used for the extraction of cell trajectories. Parameter $\alpha > 0$ is introduced here to adjust the weight of D_B function. It is used to tune and improve tracking results.

2.5 Extraction of the cell trajectories

The cell trajectory will be represented by a series of points in space-time (discrete spatio-temporal curve) for which we prescribe the condition that there is exactly

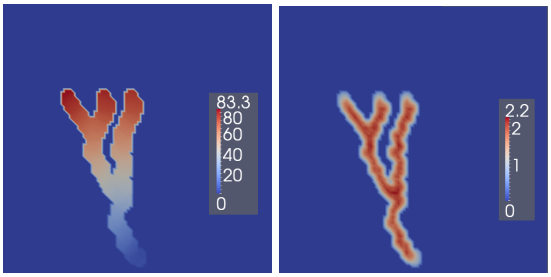

Fig. 2. On the left one can see the plot of constrained distance function D in one simply connected component of the 4D segmentation, on the right one can see the plot of the constrained distance function D_B in the same simply connected component.

one point in every time step $l = N_b, \ldots, N_e$, $1 \leq N_b < N_e \leq N_\theta$. The extraction of cell trajectories is realized in two steps

- first, we use backtracking in time by the steepest descent direction of the potential V built in (5) starting from all cell identifiers $s_m^l, m = 1, \ldots, n_C^l$ detected in all time steps $l = N_\theta, \ldots, 2$,
- then, we center all the extracted paths inside the 4D spatio-temporal trees by using constrained distance function D_B only, in order to eliminate duplicates and thus to obtain the unique cell trajectories.

The first step is realized as follows: Let s_m^l be one of the cell identifiers detected in the lth time step. Let us define a temporary point $P_T^l = s_m^l$. Then, we search recursively in the nearest vicinity of P_T^l, but only in the current time step l and previous time step $l-1$, for a doxel with the minimal value of the potential V which is also strictly less than the value of potential at the temporary point. The extracted path point P^l for the time step l is defined as the (last in search) doxel from which we move to a point in the previous time step $l-1$. The point where we moved becomes the temporary point P_T^{l-1} for time step $l-1$ and we continue the descent as above. We end the process when we cannot move from a time step N_b to a previous time step $N_b - 1$ by decreasing value of the potential V. Then the last point of the search in the time step N_b must be some detected center $s_m^{N_b}$ and it becomes the first point of the extracted path starting in the time step N_b and ending in the time step N_e where we started the descent. As an output of this first step, we get as many extracted paths as is the number of cell identifiers in all time steps except the first one, which means that we have $\sum_{l=2}^{N_\theta} n_C^l$ steepest descent paths.

After the first step of trajectories extraction there exist many duplicated paths (representing the same cell space-time movement but for a shorter time). To illustrate the above fact, let us consider a long cell trajectory going from the first to the last 3D volume of the image sequence. Since we start the descent from centers detected at every $l = N_\theta, \ldots, 2$, and they all lay inside the branch corresponding to that cell, we obtain $N_\theta - 1$ extracted paths laying inside the same branch of the 4D spatio-temporal tree. These paths can slightly differ

because the steepest descent search does not give necessarily the same set of points when starting at different time steps from different temporary points. Since all such paths lay in the same branch of the tree we center their points in 3D volumes by using the steepest growth direction of the constrained distance function D_B. After the centering step, the points of the shorter path become subset of points of the longer path and we can remove the shorter one just by comparing their points. The remaining (longest) path represents the cell trajectory. After this step we obtain a set of unique cell trajectories in the sense that a mother cell representative point is presented in as many trajectories as is the number of her descendant cells.

2.6 Lineage tree reconstruction

In this last step we build from unique cell trajectories the binary cell lineage tree by using biological coherence. First, we store the results in a tree-like data structure, where each point is stored only once with reference to mother cell and array of daughter cells. If the mother or daughters don't exist, we use null references.

Depending on the quality of input data, the resulting unique cell trajectories can be disjointed in a short time interval and sometimes more than two trajectories can merge going backward in time in single time step. To correct these situations we use the following postprocessing steps:

- We disconnect all trajectories in points where there more than two of them are merged. Then we allow only two trajectories having two nearest daughters to mother to merge, all others will remain disjointed.
- We allow reconnection of all disjointed trajectories. For each ending trajectory we are searching for nearest starting trajectory in the next two time steps and interconnect them by gradually increasing the search radius.

After such postprocessing steps we have built the cell lineage tree which is optimally stored and can be used for validation of the results and for visualization of cell dynamics.

3 Numerical experiment on real zebrafish embryogenesis data

We performed experiments on two representative real zebrafish embryogenesis datasets. First dataset has acquisition step $d\theta = 67$ seconds, $N_\theta = 480$ number of time steps and dimension of every 3D image is 512x512x104 voxels. The real voxel size is $dx_1 = dx_2 = dx_3 = 1.37$ micrometer in every spatial direction. In the last time step $N_\theta = 480$ biologist selected manually cells forming seven presumptive organs in the brain (hypoblast, presumptive hypothalamus, ventral telencephalon, right eye, right optic stalk, left eye, left optic stalk). Using developed approach we can track those cell populations backwards (and then also forward) in time and thus follow their dynamics and clonal history. For this

dataset we also have the ground truth data which contains about 39000 manually checked cell links, that can be used for validation of the tracking. We have to note that building such ground truth is extremely difficult and time consuming task for expert in biology. We highly appreciate this unique work, which can be used for tuning algorithm parameters and thus allowing to use the methods in practice when processing similar type zebrafish datasets. Second dataset has acquisiton step $d\theta = 154$ seconds, $N_\theta = 200$ and dimension of every 3D image is 512x512x120 voxels.

Before tracking, all 3D images of the processed data were filtered by 10 steps of geodesic mean curvature flow (GMCF) model [14, 4, 9] and the cell nuclei identifiers were detected by 15 steps of level set center detection (LSCD) algorithm [10, 4, 9] for the first dataset and by 4 steps of LSCD for the second dataset. The cell nuclei were segmented using generalized subjective surface (GSUBSURF) method [4, 9]. From several millions of cell identifiers we built the 4D segmentation and then the cell trajectories were extracted by the approach developed in section 2. The correctness of mother-daughter cell links for the first dataset were validated using ground truth data and the results are presented in Table 1 and Fig. 3.

To tune the tracking results we adjust two parameters mentioned in section 2, S and α. First, we can expand or shrink the nuclei radii used for building the 4D segmentation. By comparison with ground truth data we concluded that the best results for these two datasets were obtained when we shrink the radii by $S = -0.5$, cf. Fig. 3. The second parameter is α, used in the construction of the potential V. We tested tracking for $\alpha \in [0.5, 4]$ and obtained the best results around $\alpha = 3.2$, see Table 1.

Table 1. Comparison of the tracking result with ground truth data depending on α, with $S = -0.5$.

α	**Correct mother links**	**Correct daughter links**	**Wrong mother links**	**Wrong daughter links**
0.4	34856	34719	3935	4072
0.8	35510	35371	3280	3419
1.2	36064	35946	2726	2844
1.6	36551	36493	2240	2298
2.0	37125	37126	1668	1667
2.4	37288	37309	1506	1485
2.8	37349	37386	1448	1411
3.2	37402	37437	1395	1360
3.6	37357	37415	1440	1382
4.0	37377	37434	1420	1363

We present here also Figs. 4 and 5 showing results of the tracking procedure on our two datasets. For the cell trajectories visualization we built software,

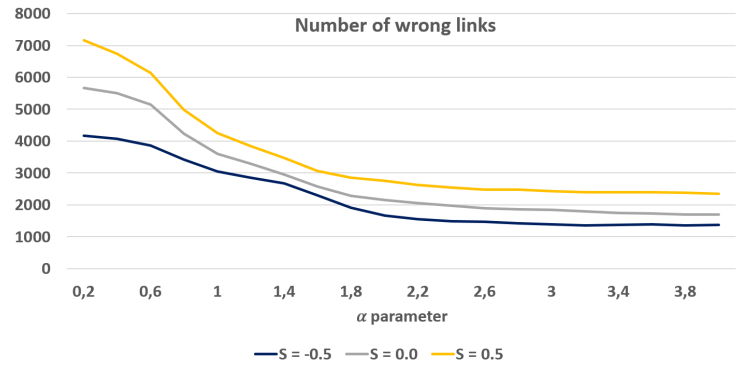

Fig. 3. Number of wrong links in tracking compared to ground truth data, depending on α and S. With α increasing towards 2, the number of wrong links is decreasing, then it stabilizes and the best result is obtained for $S = -0.5$, cf. also Table 1.

running on graphics card, where we can fluently zoom, rotate and animate in time a 3D scene even with very high number of trajectories. The trajectories are displayed as short lines (in various colors, according to speed, direction or manually defined color of cell population) connecting a few subsequent spatio-temporal points with a freely chosen starting time. In Fig. 4 we show trajectories of a second dataset with colors depending on the cell movement speed. Dark blue are the slowest, while red are the fastest moving cells. One can see the evolution of cells from the 1st still chaotic stage, through 96th, 192nd, 288th, 384th time steps where the cells are becoming more compactly localized up to 480th time step, cf. Fig. 1.

As noticed above, in the last time step of the first dataset, $N_\theta = 480$, biologist manually selected cells forming seven presumptive organs during the zebrafish brain early embryogenesis.

In Fig. 5 we show trajectories along with the slices of original data in time step $N_\theta = 260$. Here all trajectories have the same color and original data is displayed as black and white slices obtained from 3D volume. Only the cells that are near the slices are displayed along with the trajectories showing their movement in next and previous 20 time steps.

4 Conclusions

In this paper we presented algorithm for the cell tracking and lineage tree reconstruction from 3D+time microscopy data. We validate the tracking results by comparing them with manually verified ground truth data. The best results which we can achieve show more than 96% correctness of mother-daughter links. We applied the tracking method to complex stages of the zebrafish early embryogenesis images and visualize extracted cell trajectories and present mean velocity of selected cell populations.

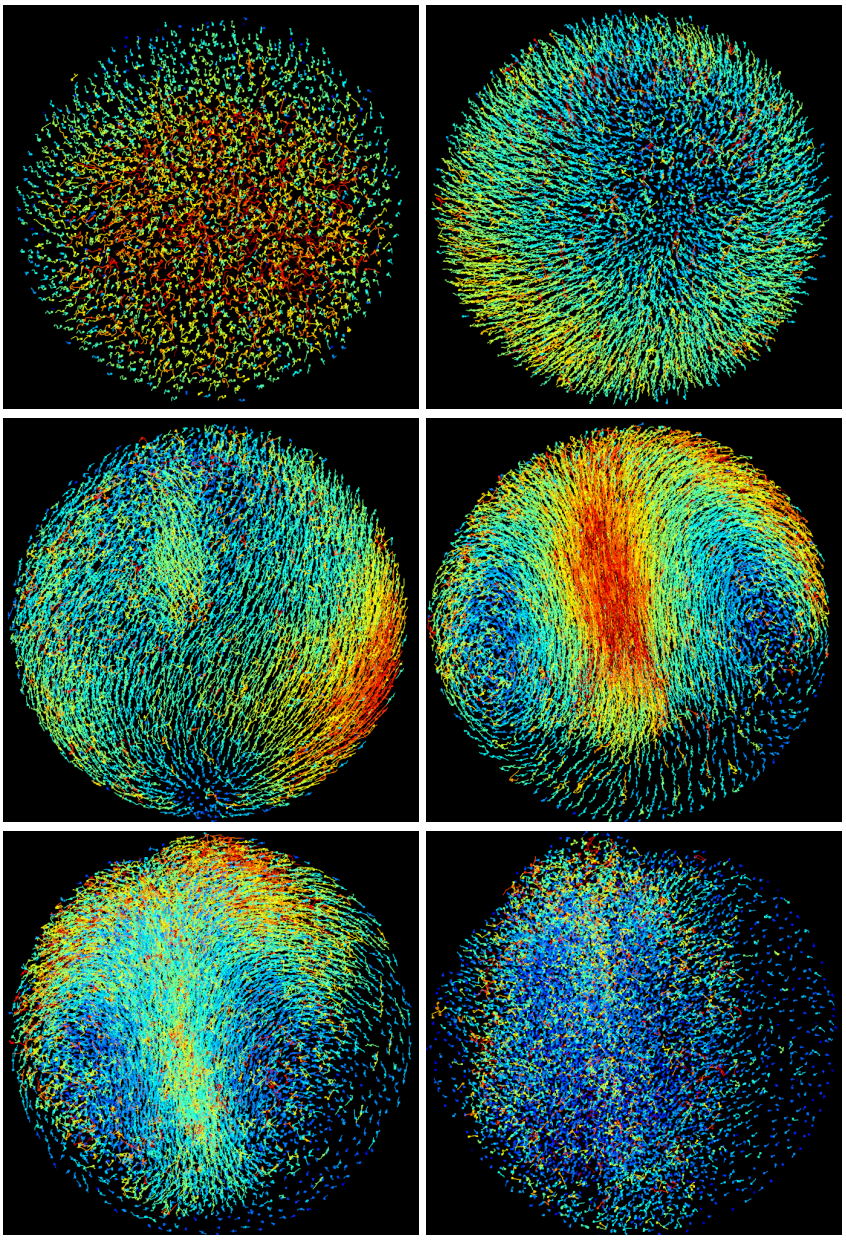

Fig. 4. Visualization of the cell movement speed during embryogenesis in time steps 1, 96, 192, 288, 384 and 480. Red trajectories are the fastest moving cells, while blue are the slowest.

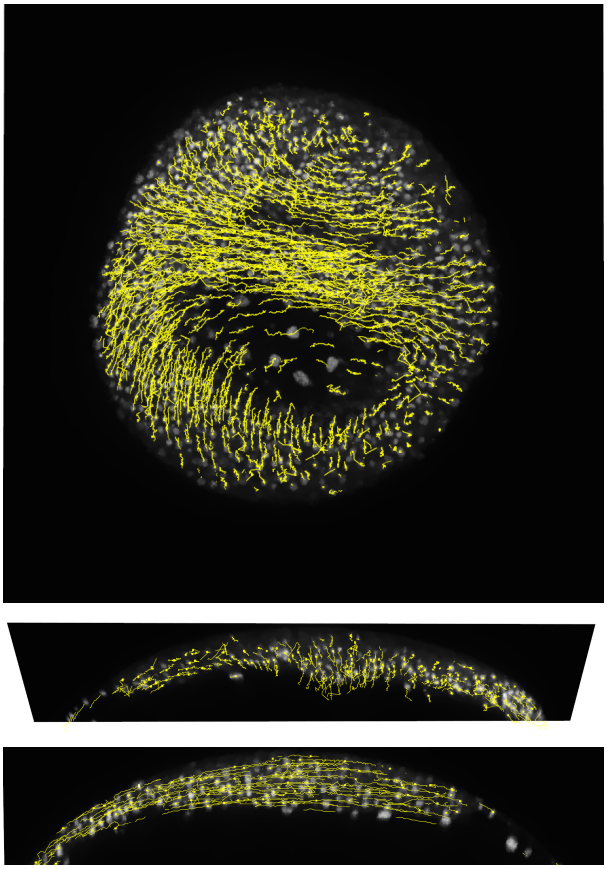

Fig. 5. Tracked cells with trajectories displayed along the slices of original data. Middle image is slightly tilted for better visibility since the trajectories are going outwards from the slice.

Acknowledgement. This work was supported by grants APVV-15-0522 and VEGA 1/0608/15.

References

1. Amat, F., Lemon, W., Mossing, D.P., McDole, K., Wan, Y., Branson, K., Myers, E.W., Keller, P.J.: Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data. Nature Methods **11** (2014) 951–958
2. Bao, Z., et al.: Automated cell lineage tracing in caenorhabditis elegans. Proc. Natl. Acad. Sci. USA **103** (2016) 2707—-2712
3. Bellaïche, Y., Bosveld, F., Graner, F., Mikula, K., Remešíková, M., Smíšek, M.: New robust algorithm for tracking cells in videos of drosophila morphogenesis based on finding an ideal path in segmented spatio-temporal cellular structures. Proceeding

of the 33rd Annual International IEEE EMBS Conference, Boston Marriott Copley Place, Boston, MA, USA, August 30 - September 3, 2011, IEEE Press (2011)
4. Bourgine, P., Čunderlík, R., Drblíková, O., Mikula, K., Peyriéras, N., Remešíková, M., Rizzi, B., Sarti, A.: 4d embryogenesis image analysis using pde methods of image processing. Kybernetika **46** (2010) 226–259
5. Bourgine, P., Frolkovič, P., Mikula, K., Peyriéras, N., Remešíková, M.: Extraction of the intercellular skeleton from 2d microscope images of early embryogenesis. Lecture Notes in Computer Science 5567 (Proceeding of the 2nd International Conference on Scale Space and Variational Methods in Computer Vision, Voss, Norway, June 1-5,2009) , Springer (2009) 38–49
6. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. International Journal of Computer Vision **22** (1997) 67–79
7. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. Proceedings International Conference on Computer Vision '95, Boston (1995), 694–699
8. Chen, Y., Vemuri, B.C., Wang, L.: Image denoising and segmentation via nonlinear diffusion. Comput. Math. Appl. **39** (2000) 131–149
9. Faure, E., Savy, T., Rizzi, B., Melani, C., Remešíková, M., Špir, R., Drblíková, O., Čunderlík, R., Recher, G., Lombardot, B., Hammons, M., Fabrèges, D., Duloquin, L., Colin, I., Kollár, J., Desnoulez, S., Affaticati, P., Maury, B., Boyreau, A., Nief, J.Y., Calvat, P., Vernier, P., Frain, M., Lutfalla, G., Kergosien, Y., Suret, P., Doursat, R., Sarti, A., Mikula, K., Peyriéras, N., Bourgine, P.: An algorithmic workflow for the automated processing of 3d+time microscopy images of developing organisms and the reconstruction of their cell lineage. Nature Communications **7, Article number: 8674** (2016)
10. Frolkovič, P., Mikula, K., Peyriéras, N., Sarti, A.: A counting number of cells and cell segmentation using advection-diffusion equations. Kybernetika **43** (2007) 817–829
11. Kausler, B.X., et al.: A discrete chain graph model for 3d+t cell trackingwith high misdetection robustness. ECCV (2012) 144–157
12. Kichenassamy, S., Kumar, A., Olver, P., Tannenbaum, A., Yezzi, A.: Conformal curvature flows: from phase transitions to active vision. Arch. Rational Mech. Anal. **134** (1996) 275–301
13. Kichenassamy, S., Kumar, A., Olver, P., Tannenbaum, A., Yezzi, A.: Gradient flows and geometric active contours model. Proceedings International Conference on Computer Vision '95, Boston (1995)
14. Krivá, Z., Mikula, K., Peyriéras, N., Rizzi, B., Sarti, A., Stašová, O.: 3d early embryogenesis image filtering by nonlinear partial differential equations. Medical Image Analysis **14** (2010) 510–526
15. Mikula, K., Peyriéras, N., Remešíková, M., Sarti, A.: 3d embryogenesis image segmentation by the generalized subjective surface method using the finite volume technique. Finite Volumes for Complex Applications V, Problems & Perspectives (Eds. R.Eymard, J.M.Herard), ISTE and WILEY, London (2008) 585–592
16. Mikula, K., Peyriéras, N., Remešíková, M., Smíšek, M.: 4d numerical schemes for cell image segmentation and tracking. Finite Volumes in Complex Applications VI, Problems & Perspectives, Eds. J. Fořt et al. (Proceedings of the Sixth International Conference on Finite Volumes in Complex Applications, Prague, June 6-10, 2011), Springer Verlag (2011) 693–702
17. Mikula, K., Peyriéras, N., Remešíková, M., Stašová, O.: Segmentation of 3d cell membrane images by pde methods and its applications. Computers in Biology and Medicine **41** (2011) 326–339

18. Mikula, K., Peyriéras, N., Špir, R.: Numerical algorithm for tracking cell dynamics in 4d biomedical images. Discrete and Continuous Dynamical Systems - Series S **8** (2015) 953–967
19. Mikula, K., Špir, R., Smíšek, M., Faure, E., Peyriéras, N.: Nonlinear pde based numerical methods for cell tracking in zebrafish embryogenesis. Aplied Numerical Mathematics **95** (2015) 250–266
20. Olivier, N., Luengo-Oroz, M.A., Duloquin, L., Faure, E., Savy, T., Veilleux, I., Solinas, X., Débarre, D., Bourgine, P., Santos, A., Peyriéras, N., Beaurepaire, E.: Cell lineage reconstruction of early zebrafish embryos using label-free nonlinear microscopy. Science **329** (2010) 967–971
21. Sarti, A., Malladi, R., Sethian, J.A.: Subjective surfaces: A method for completing missing boundaries. Proceedings of the National Academy of Sciences of the United States of America **97** (2000) 6258—-6263
22. Sethian, J.A.: A fast marching level set method for monotonically advancing fronts. Proceedings of the National Academy of Sciences **93** (1996) 1591–1595
23. Tomer, R., Khairy, K., Amat, F., Keller, P.J.: Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy. Nature Methods **9** (2012) 755—-763
24. Zhao, H.: A fast sweeping method for eikonal equations. Mathematics of Computation **74** (2005) 603–627