Reconstruction and Visualization of Coordinated 3D Cell Migration Based on Optical Flow

Christopher P. Kappe, *Student Member, IEEE*, Lucas Schütz, Stefan Gunther,
Lars Hufnagel, Steffen Lemke, and Heike Leitte, *Member, IEEE*

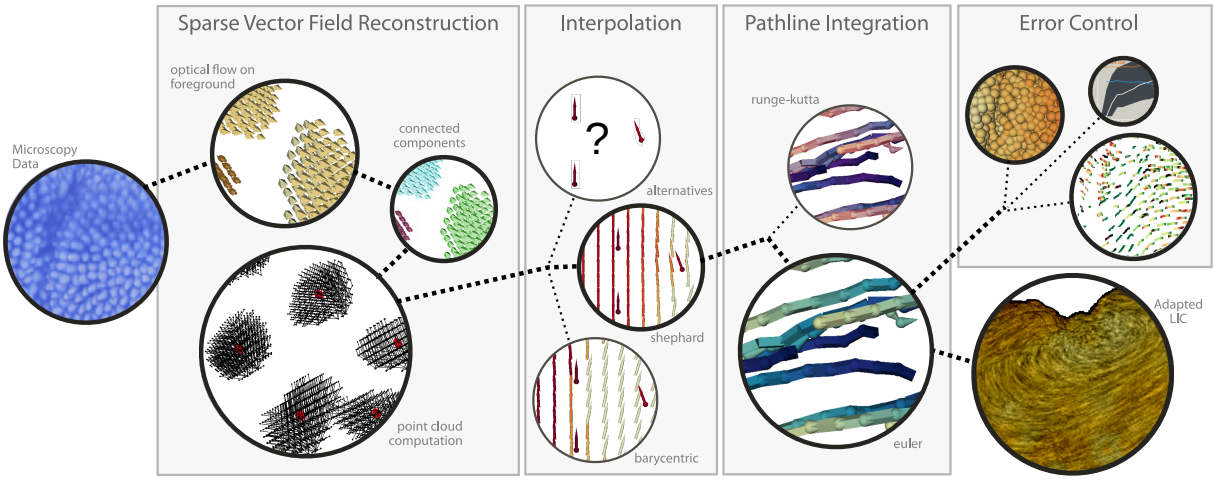


Fig. 1: Analysis pipeline: In a first step, we reconstruct local flow information from the 3D+t microscopy data using a combination of image processing and optical flow computation. Next, the global vector field is reconstructed using scattered point interpolation. This serves as input to the pathline integration which results in a reconstruction of cell trajectories. The pipeline finishes with a set of visual and numerical methods for error control and a new LIC variety for better pattern analysis.

Abstract—Animal development is marked by the repeated reorganization of cells and cell populations, which ultimately determine form and shape of the growing organism. One of the central questions in developmental biology is to understand precisely how cells reorganize, as well as how and to what extent this reorganization is coordinated. While modern microscopes can record video data for every cell during animal development in 3D+t, analyzing these videos remains a major challenge: reconstruction of comprehensive cell tracks turned out to be very demanding especially with decreasing data quality and increasing cell densities. In this paper, we present an analysis pipeline for coordinated cellular motions in developing embryos based on the optical flow of a series of 3D images. We use numerical integration to reconstruct cellular long-term motions in the optical flow of the video, we take care of data validation, and we derive a LIC-based, dense flow visualization for the resulting pathlines. This approach allows us to handle low video quality such as noisy data or poorly separated cells, and it allows the biologists to get a comprehensive understanding of their data by capturing dynamic growth processes in stills. We validate our methods using three videos of growing fruit fly embryos.

Index Terms—Cell migration, vector field, 3D, time-dependent, LIC, tracking, validation

1 Introduction

Light-sheet microscopy has been considered one of the major breakthroughs in biology in recent years. The novel microscopes can record high resolution 4D videos of growing organisms that capture intricate structure formation in 3D at low photoxicity, i.e. without damaging the imaged organism. Expectations towards this novel technique are high as all relevant information for a holistic analysis of morphogenetic processes during animal development are captured in this data [18, 23, 25]. The data, however, proves to be hard to analyze. Major hurdles are described by biologists as the lack of dedicated analysis software, the increasing scale of the data (up to 4 terabyte of raw data per data set), and the inherent noise in the data [9, 44]. This, in essence, translates into known visualization problems [51], namely, (i) depiction of nested objects and application of "biologically plausible" colors, (ii) simultaneous visualization of experimental data and derived structures and properties, (iii) display of multidimensional data into biological context, (iv) generation of abstract representation and illustrative renderings, (v) concurrent analysis of data from multiple experiments, and (vi) interaction with raw and processed data.

In this paper, we concentrate on the reliable reconstruction of coordinated cell migration and its intuitive depiction in developing organisms to address challenges (ii), (iv), and (vi). We particularly consider these visualization challenges in the context of animal gastrulation. Animal gastrulation comprises the first visible, morphological modification of the growing organism at the beginning of its development (fig. 2), and it describes the transition from a sphere of cells to an embryo with three layers of cells, with each layer giving rise to different parts of the adult animal. In biology, the process of gastrulation serves

- *Christopher P. Kappe and Heike Leitte are with IWR, Heidelberg University. E-mail: {christopher.kappe,heike.leitte}@iwr.uni-heidelberg.de.*
- *Stefan Gunther and Lars Hufnagel are with EMBL, Heidelberg, Germany. E-mail: {guenther,hufnagel}@embl.de*
- *Lucas Schütz and Steffen Lemke are with COS, Heidelberg University. E-mail: {lucas.schuetz,steffen.lemke}@cos.uni-heidelberg.de.*

Manuscript received 31 Mar. 2015; accepted 1 Aug. 2015; date of publication xx Aug. 2015; date of current version 25 Oct. 2015. For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.

as widely adapted model to study how organisms generate morphological order and structure and how this is achieved in an interplay of genetic regulation and mechanophysical mechanisms, a process often referred to as morphogenesis [24, 50].

A critical step in the analysis of morphogenetic processes like gastrulation is the understanding of cell migration and cell reorganization during development. Direct animations of the data are only a first simple step in this direction. The more crucial and challenging task is the reliable extraction of motion patterns from the given video data. The extraction of motion patterns allows the biologist to identify areas of coordinated cell migration, which is a prerequisite to link morphogenesis with the underlying activity of developmental genetic networks. From the perspective of data visualization, the extraction of motion patterns requires the solution of a spatio-temporal correspondence problem. The video captures the individual cells in 3D at each time step and the temporal resolution is commonly high enough to ensure spatial overlap between cells in consecutive time steps. This renders two conceptually different approaches possible for the reconstruction of cell trajectories: (a) The images are segmented into individual nuclei and subsequently a correspondence problem is solved that matches respective nuclei between time steps [21, 42, 2]. (b) A flow field is reconstructed based on the 3D optical flow between subsequent images, and pathlines are integrated in the reconstructed vector field. While the first method results in an exact reconstruction of the cell lineages which is of high importance for many biological questions, it also requires a very good nuclei segmentation and a robust tracking procedure to achieve satisfactory results. The second method is better suited for more global patterns of cell migrations and has lesser requirements to data quality.

While most available approaches concentrate on techniques as in (a), we opt for the approach using the optical flow of video data to reconstruct the underlying vector field (b). The reconstructed velocity field now opens up the wealth of existing flow visualization techniques. We found texture-based approaches particularly suited and adapt time-dependent volumetric line integral convolution (LIC) to our data.

We make the following contributions:

- We present a pipeline to reconstruct the flow field from underlying cellular motions in growing embryos.
- We integrate pathlines in these vector fields and validate their correctness with respect to the imagery data.
- We adapt LIC to operate on this special type of time-dependent volumetric data.
- We analyze the performance of the pipeline using videos from developing fruit fly embryos and detail how these visualizations can be used for advanced analysis.

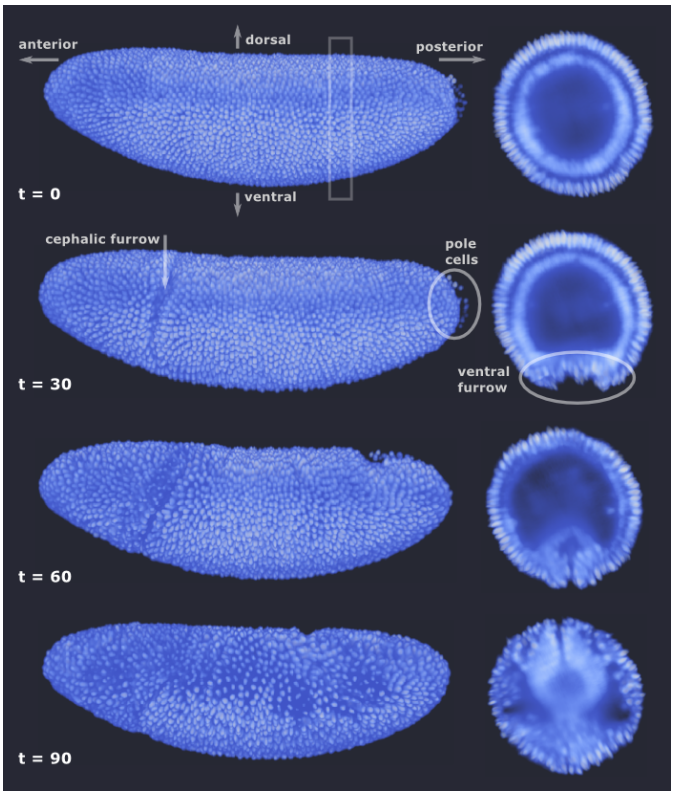


Fig. 2: Gastrulation in Drosophila: (left) Volume rendering of the microscopy data shows the morphogenetic changes in Drosophila over time. (right) Volume rendering of cross-sections (white box in top image) visualize processes in the interior and show surface shape changes.

2 Related Work

In this paper, we develop a method to reconstruct a time-dependent vector field from 3D videos and propose an adapted LIC version for improved visualization.

Flow Reconstruction: Flow field reconstruction from image data is a common problem in experimental fluid dynamics with a large volume of sophisticated solutions. Particle image velocimetry (PIV) is a well-established method in this area using images of tracer particles in the flow for the reconstruction [40]. PIV reconstructs the vector field based on the solution of the correspondence problem between two successive particle images [1] and is in spirit very close to the nuclei segmentation and tracking approach. A complementary approach pursued in recent years is the reconstruction of fluid flow by means of optical flow [17]. Liu and Shen [34] carefully derive the mathematical connection between fluid flow and optical flow. Unlike PIV, the optical flow based approach is very much suited for continuous patterns in imagery data which are predominant in many biological application settings.

Unsteady Flow Visualization: Since flow visualization is a very established topic in the visualization community, numerous state-of-the art reports have been written over time. In 2004, Laramee et al. have surveyed a broad hierarchy of research that covers 2D, 2.5D, 3D, steady and unsteady flow visualization problems and explicitly mark the listed publications as such [31]. Post et al. give an overview over techniques to compute various meta-information in flow data and track it over time. There is also a survey particularly examining surface-based flow visualization [13].

Kuhn et al. propose an approach based on pathlines enhancing the existing techniques in two main points: Firstly, the seeding of the pathlines is arranged in such a way that visual clutter appears as little as possible – which is always a problem with these 3D visualizations; and secondly, a carefully implemented new rendering technique allows a quick execution. Aiming to improve particle-based flow visualization, Günther, Kuhn et al. [16] explore a way of incorporating particle mass into the typical integral curve visualizations. Krishnan et al. also introduce a visualization tool for unsteady 3D flow [29]. It relies on the extraction of apt surfaces in the volume to visualize the data. Time surfaces show the temporal development of the flow; and streak surfaces represent the spatial properties of it. Ferstl et al. have further explored the field of topology-driven streak surface construction [14].

Line Integral Convolution (LIC): The LIC method [7] has proven to be an outstanding technique to produce flow visualizations. The original algorithm has been enhanced and extended a lot in the course of time, advancing from steady 2D to time-dependent 3D flow and incorporating more and more additional features or increasing the performance [49, 28, 56, 5]. Pioneering work that has enabled LIC visualizations of time-dependent data includes UFLIC [46] by Shen and Kao and its accelerated variant AUFLIC [35] by Liu and Moorhead II. They evolve from stream- to pathlines and use a successive feed-forward scheme after the generation of one frame to realize a temporal coherence between the output images for the subsequent time steps. Several papers such as [47] by Shen and Kao and [33] by Li et al. propose techniques that allow the visualization of unsteady flow on surfaces. Oth-

ers focus also on efficiency, performing the LIC computations only in image space and utilizing the computational power of modern graphics hardware where suitable [32, 52]. By 2005 the AUFLIC algorithm was extended to handle time-dependent volumetric data. In this approach, called VAUFLIC [36], pathline integral convolution occurs in a 3D input texture that is given by a noise function. Volume rendering techniques are used for the final representation. For example, the volume can be sliced according to the region of interest defined by the user; and areas of low velocity magnitude are made transparent to reduce occlusion. Weiskopf et al. proposed a hardware-accelerated approach to visualize time-dependent 2D and 3D vector fields [55].

Optical Flow: Optical flow [19] is a well-established technique to reconstruct the correspondence problem in 2D videos. It approximates the movement from one time step to the next and outputs the resulting flow field. Most current algorithms tackle 2D videos to address problems such as segmentation and object tracking. The recently proposed 3D optical flow by Amat et al. [3] is dedicated to fluorescence videos of growing embryos which suffer from general fuzziness and feature complex motions. The algorithm first segments the data into fore- and background based on simple intensity thresholding. On the foreground super-voxels a Markov Random Field is computed that incorporates motion smoothness constraints. The algorithm usually achieves very precise results. Known issues arise in areas of complex motion such as with cell divisions.

3 Biological Background

While the conceptual pipeline we are going to propose is very generic, we refer during the algorithmic description to a particular type of data originating from light sheet-based microscopy. Hence, we work with series of 3D images. Movement per voxel is detected using an optical flow algorithm especially designed for the given 3D+t data. The particular organisms we have worked with are the fruit fly species *Drosophila melanogaster* and *Drosophila pseudoobscura*.

3.1 Light-sheet Microscopy

For translucent organisms the necessary 3D+t video data can be acquired using optical microscopy such as the recently developed light sheet-based fluorescence microscopy (LSFM) [20, 22, 26]. The general idea is to label cell nuclei with a fluorescent marker that is stimulated during recording and shows up as high intensity values in the resulting data. A strength of LSFM is the very low invasiveness of the method that allows recording for extended periods of time (hours to days) without damage to the specimen. The 3D videos feature a high spatio-temporal resolution which allows for automatic segmentation of small individual structures such as cells and cell nuclei [38, 45].

In this paper, we concentrate on morphogenetic processes at the beginning of gastrulation in Drosophila. Images were recorded every 30 seconds on a MuVi-SPIM [30] using four distinct views of the embryo, which are created using two opposing cameras and rotating the specimen by 90 degrees. Each view has a spatial lateral resolution of 0.26μm/pixel and an axial resolution of 2μm/pixel. All views are registered with respect to each other into a common coordinate system and fused into a single 3D image using a sigmoidal tissue scattering model. The final image has an isotropic spatial resolution of 0.52μm/pixel.

We focused on a period of 50 minutes, corresponding to 100 time steps at size $1008 \times 351 \times 335$ voxels.

3.2 Embryonic Morphogenesis in Drosophila

Embryonic morphogenesis in Drosophila has been well characterized in the past 30 years, leading to a comprehensive understanding of the core aspects of gastrulation, both at the morphological and genetic level [8]. Just prior to gastrulation, the Drosophila embryo consists of a single epithelial sheet, the cellular blastoderm, consisting of about 6000 tall, columnar cells enclosing the central yolk, and a group of about 30 germ line cells lying on the surface of this sheet at the posterior pole of the embryo (the pole cells) [8] (fig. 2). At this stage, the shape of the embryo along the head-to-tail (anterior-to-posterior) axis resembles an irregular ellipsoid with a slimmer belly line at the head than the tail. All relevant cells reside in the periphery of this irregular ellipsoid, permitting to unroll the surface and study initial deformations in 2D.

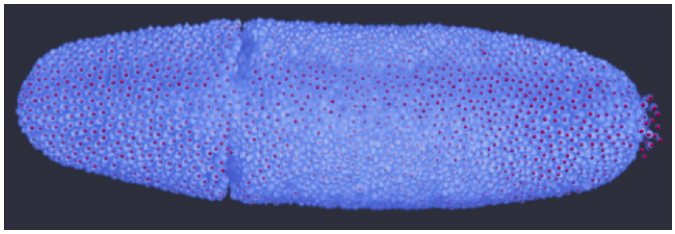


Fig. 3: Overlay of raw data and scattered point data: The translucent volume rendering depicts the microscopy data. The red spheres indicate positions of extracted consistent connected components in the optical flow.

With the onset of gastrulation, the cellular blastoderm first folds a longitudinal patch on the bottom (ventral) side of the blastoderm embryo. At either end of this ventral fold, but in processes developmentally and genetically distinct from the ventral fold, further areas of cells are gradually internalized. Additionally, a transient fold forms in the anterior third of the embryo and separates the future head from the remaining embryo (cephalic furrow), and two dorsal transversal folds appear behind the cephalic furrow. Superimposed on invagination and folding of the trunk of the embryo, the epithelial structures in the bottom tail domain of the embryo (the so-called germband) start to extend towards the tail. This extension is physically limited through the space defined by the egg shell, which is why the tissue at the posterior pole starts to fold back onto itself. During this extension, the blastoderm cells on the top (dorsal) side of the embryo flatten into a single thin sheet, which then folds up between extending germband and head. Overall, gastrulation in Drosophila takes about 2.5 to 3 hours [8, 12].

Several of these core aspects of Drosophila gastrulation have been studied by classical mutant studies. Typically in such studies a certain gene function is switched off, which presumably leads to a change in cell properties: the biologist then analyzes the embryo and decides, whether the morphogenetic result is the same or different from a non-affected embryo. Often these analyses have been carried out in fixed (i.e. dead) animals, excluding most of the inherent dynamic from the analysis. While numerous core aspects of gastrulation could be analyzed and genetically explained, many aspects of Drosophila epithelial morphogenesis during gastrulation are still poorly understood at the mechanistic level and have started to get investigated only in recent years and with the help of image analysis and visualization tools [41, 54, 39, 6, 27]. The advent of in-toto live imaging at high spatio-temporal resolution in LSFM carries the promise that more such advances can be made, provided that the raw data can be effectively and meaningfully visualized for analysis.

3.3 Biological Data Analysis Roadmap

The short-term goal regarding the presented data is to derive a quantitative description of in toto morphogenetic processes, ideally aiming for a delineation of domains (surface areas) with similar or identical cell behavior. To this end, coordinated motion patches have to be reconstructed from the video data. This task comprises two stages: First, cellular motions have to be reconstructed from the video data. Second, feature-based analysis is required to identify coherent motion patches in a 4D spatio-temporal setting. The resulting motion maps allow for the definition of areas with coordinated morphogenetic activity, with areas of high correlation indicating a strong and robust underlying genetic regulation of morphogenesis. Using existing maps of gene activity, the biologist can correlate areas of coordinated morphogenetic activity with patterns of gene expression and thereby identify previously unknown regulators of cell behavior, epithelial morphogenesis, and organismal development. In this paper, we address the first problem in this pipeline. We work towards the reliable reconstruction and depiction of cellular migrations in 3D videos.

4 Analysis Pipeline

To support biologists in the analysis of coordinated cell migrations in 3D+t, we have developed the following five-step pipeline (fig. 1):

Sparse vector field reconstruction: In the first step, we aim at the robust flow reconstruction between consecutive time steps. The input data are 3D gray-scale image series as detailed in sec. 3.1. Using the segmentation software Ilastik [48], we obtain a foreground mask that identifies cell nuclei in the image data. Within the boundaries of the foreground mask, the optical flow is computed using the algorithm proposed by Amat et al. [3]. This results in a robust vector field reconstruction with large undefined areas between cell nuclei. (Due to the imaging modality a more detailed reconstruction is not possible.). To reduce storage size and facilitate standard robust interpolation schemes, the field is further reduced to singular representatives per cell nuclei. The output of the first step is scattered point vector data. ($\rightarrow$ sec. 5.1)

Interpolation: In the second step, we employ scattered point data interpolation to obtain continuous vector information in the domain of the embryo. ($\rightarrow$ sec. 5.2)

Pathline Integration: In the third step, we use the reconstructed vector field to integrate pathlines in the unsteady vector field. These lines retrace the motion of the embryonic cells. ($\rightarrow$ sec. 5.3)

Error Control: To validate the correctness of our results, we propose in the fourth step a set of tools for error control. We integrate methods for both visual and numerical analysis. Using the visual methods, the user can compare the reconstructed data to the input data and compare arbitrary intermediate steps in the pipeline. The numerical approaches employ the distances of the pathlines to the nearest cell nuclei in respective time steps. Combining both approaches, we obtain expressive visualizations that guide the user to potentially erroneous parts of the data. ($\rightarrow$ sec. 7.3)

Adapted LIC: The last step in the pipeline is an adapted LIC variety. Our application partners found texture-based approaches particularly useful for pattern detection in 3D+t. Hence, we propose a novel LIC-variety in section 6 that takes the special hollow body structure of our data into account.

5 Flow Reconstruction

As detailed before, we are given a vector field based on the optical flow of a video (section 3). Building upon the work of Liu and Shen [34] and Corpetti et al. [11], we use the optical flow data to reconstruct the fluid flow in the observational domain Ω. We assume that the underlying flow field $\mathbf{f}$ is continuous and want to approximate this field in the entire domain Ω using the optical flow as reference data.

This approach holds the following challenges: (i) The optical flow can only be computed in areas of the video where enough textural information is present to estimate changes. Hence, no flow estimates are given in the areas between the nuclei, though these parts belong to the respective cells and move as well. (ii) Like all automatic analysis techniques optical flow results come with uncertainty. Especially in areas that are challenging to track the flow may be erroneous. (iii) The data that we are working with is very large. One time lapse consists of around 100 3D images, each of which taking up approximately 850 MiB of memory. We target at a visualization on a desktop machine to allow for ready to use, interactive data analysis.

The goal of this section is to estimate the fluid flow of the cells based on the optical flow. In the following, we will refer to the interpolated vector field as a function

$$\mathbf{f}: \Omega \rightarrow V,\ \Omega \subset \mathbb{R}^3, V \subset \mathbb{R}^3.$$

5.1 Robust Sparse Flow Representation

The first problem that we address is making the measurements more robust. The embryo videos often suffer from noise and varying intensities as can be seen in fig. 2. Especially in the later time steps, the flow reconstruction becomes progressively difficult for all currently available techniques as contrast decreases and structures become less distinct. As the optical flow already misses information in the embryo domain in the first place (the non-fluorescent cell parts around the nuclei) which makes some reconstruction necessary anyway, we decided to further reduce the flow data to individual robust representations per cell. As cells move cohesively and are not subject to strong deformations in our data, this is a valid reduction. In fact, for components where small irregularities exist (like a few mismatched voxels), this approach conveniently smoothes them away whereas one would have trouble working on the fine grid data later in the pipeline.

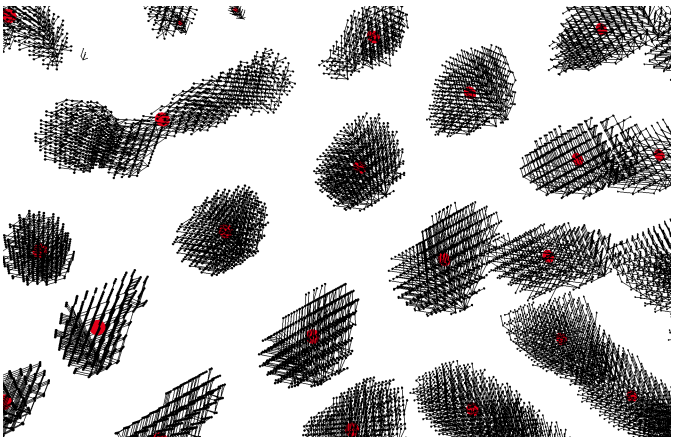


Fig. 4: Connected components (black edges) and their respective center of mass (red spheres).

5.1.1 Regular Connected Components

To identify consistent motion patches, we extract connected components in the grid-aligned optical flow. We only consider areas with non-zero velocity which already provides a fairly good segmentation of individual nuclei (see fig. 4). Connected components are computed based on a $3^3 - 1$-neighborhood. The resulting structures coincide in 95% of the cases with individual nuclei.

If the variance within a connected component C is low (see below for details), we summarize the contributing voxel information in a single representative c by averaging its positions and flow vectors:

$$\mathbf{pos}(c) := \frac{1}{|C|} \sum_{u \in C} \mathbf{pos}(u) \quad \mathbf{vel}(c) := \frac{1}{|C|} \sum_{u \in C} \mathbf{vel}(u)$$

We found that the center of mass is a good cell representative and coincides well with the imagery data, as can be seen from fig. 3.

To further improve the reconstruction quality, we take two measures: First, we filter connected components with a high variance in the corresponding velocity vectors, and second, we filter connected components that feature a very unusual flow pattern.

5.1.2 High-variance Connected Components

In many cases, the connected components are a good first approximation of individual nuclei. For densely packed cell compounds or dividing cells, however, they are not able to separate the structures. If the motion of all cells is very consistent, the joint representation is not critical as our primary goal is to reconstruct the underlying flow field. But in case of diverging motion, as occurs for example in the case of dividing cells, this reduction would lead to an erroneous oversimplification of the flow. To identify such high-variance connected components, we investigate the standard deviation of the contributing velocity vectors. This gives a precise decision criterion: Either it is very close to zero or it is clearly greater. To have some concrete threshold, for our implementation, we decided that if it is larger than 0.15 times the mean velocity magnitude of the respective component, we partition the connected component into coherent substructures.

Therefore, we chose an iterative approach based on k-means clustering [37]. For a good initialization of the algorithm, we compute the flow direction of highest variance using principal components analysis

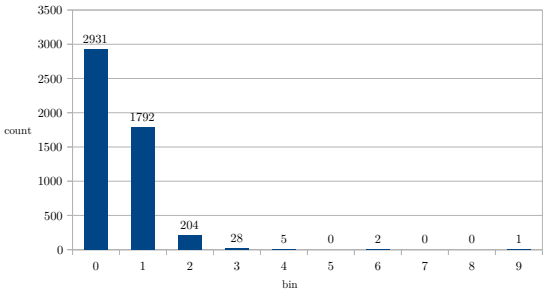


Fig. 5: Typical velocity magnitude histogram. The velocity magnitude range was divided in ten equidistant bins. 8 points out of 4963 have been classified as outliers.

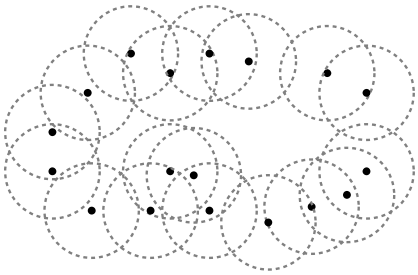


Fig. 6: Structure of the observational domain: Points of measurements are indicated as black dots. Vector values are now reconstructed in the vicinity of these points resulting in a bumpy volume.

on the vector data. Let $\boldsymbol{a}$ be the vector representing the axis with the greatest variance (the eigenvector with the greatest norm). With the mean velocity $\boldsymbol{\mu}$, we then set $\boldsymbol{c}_1 = \boldsymbol{\mu} - \boldsymbol{a}$ and $\boldsymbol{c}_2 = \boldsymbol{\mu} + \boldsymbol{a}$ as the initial cluster centers. Using these starting positions the algorithm with $k = 2$ quickly converges to high-quality directions. In rare cases more than two cells moving in different directions may be merged. Hence, we continue the splitting until the variance is below the threshold.

5.1.3 Outliers

In our study, we found that we have to expect a few heavy outliers in our data in terms of the velocity magnitude and direction. These originate from errors in the optical flow and hence, we exclude them from subsequent analysis. We base outlier classification on deviation from the mean velocity magnitude μ and keep a point with velocity magnitude v if

$$|v - \mu| \leq \sigma \cdot c,$$

where σ is the standard deviation of all velocity magnitudes in the data and c is a user-defined value to control the classification as outlier. By default, we set a relatively high value of $c = 5.15$ for the embryo data, to make sure we do not eliminate any actual data that might be part of an interesting phenomenon. Figure 5 shows a typical histogram of velocity magnitudes for one time step.

5.1.4 Data Compression

The condensed representation has the additional benefit that it significantly reduces the memory footprint of the data. While the optical flow data commonly amounts to approx. 80 GiB, the scattered point data uses only 50 MiB – almost a factor 2000 which significantly increases interaction performance and data loading.

5.2 Vector Field Interpolation

To reconstruct cell trajectories from the optical flow, we need to be able to estimate flow information at arbitrary positions in the observational domain Ω. This is commonly achieved using interpolation. We considered local methods only, to ensure acceptable computational costs [15]. The general form of these local interpolants for scattered point data is

$$\mathbf{f}(\boldsymbol{x}) = \frac{1}{W} \cdot \sum_{i=1}^{n} w_i \cdot \mathbf{f}(\boldsymbol{q}_i) \tag{1}$$

where $\boldsymbol{q}_i \in Q$ are the local reference points and w_i their respective weights. W is a normalization factor.

5.2.1 Weighting Methods for Scattered Points

The weighting in eq. (1) defines how strongly each data point influences the interpolants in its vicinity. Two widely used weighting methods are Barycentric coordinates and inverse distance weighting. A comprehensive survey of alternative strategies is given in [15].

Barycentric coordinate systems represent points as combinations of polytope vertices and were classically defined with respect to triangles and general simplices [43]. Barycentric coordinates are well suited for scattered point data and a neighborhood graph that forms a convex polytope. This notion has been extended to general polytopes [53] and point clouds [10]. We employ the following generalized Barycentric coordinates for weighting:

$$w_i = \sum_{j=1}^{n} d_j - d_i \qquad W = (n-1) \cdot \sum_{i=1}^{n} d_i$$

where d_i represents the Euclidean distance between $\boldsymbol{x}$ and the i-th neighbor.

Inverse distance weighting directly models the decreasing influence of measurement points with increasing distance. Shepard's interpolation uses the reciprocal distance for the weighting:

$$w_i = \frac{1}{1 + d_i^{\alpha}} \qquad W = \sum_{i=1}^{n} w_i$$

where $d_i = \|\boldsymbol{x} - \boldsymbol{q}_i\|$ represents the Euclidean distance between $\boldsymbol{x}$ and the i-th neighbor. Commonly the distance contains a power parameter α to directly manipulate the spatial decay.

5.2.2 Local Neighborhood

To evaluate the interpolation, we have to define a neighborhood for each position $\boldsymbol{x}$ in the observational domain Ω. Cells in tissue have a clearly structured layout and are densely packed. The motion patterns are commonly very smooth and we assume that cells only influence their direct neighbors. Hence, we restrict the interpolation schemes to the nearest neighbors. We construct the surrounding polytope Q of the query position $\boldsymbol{x}$ using the four closest points within a maximum distance r. If the points are roughly coplanar, this amounts to a quadrilateral in 3D; if they are not, we approximate interpolation in a tetrahedron. We achieved best interpolation results by setting r to the 66%-quantile of the 3-nearest neighbor distances of all the data points. We explored alternative settings and found the value using this heuristic most reliable across time steps and data sets.

For a fast computation of the k-nearest neighbor distances, we build a k-d tree for each time step [4].

5.3 Pathline Integration

With the reconstructed vector field in the entire domain Ω, we can now retrace the paths of cell nuclei using numerical integration. In accordance with [46] we define them as follows. Let the time-dependent input vector field be

$$\mathbf{f} : \Omega \times T \to V, \;\; \Omega \subset \mathbb{R}^3, T \subset \mathbb{R}, V \subset \mathbb{R}^3.$$

Then a pathline $\mathbf{p}$ in Ω, parametrized by a time $t \in T$, can be defined in terms of an integration procedure:

$$\mathbf{p}(t + \Delta t) = \mathbf{p}(t) + \int_{t}^{t+\Delta t} \mathbf{f}(\mathbf{p}(t), t) \mathrm{d}t \tag{2}$$

where $\mathbf{p}(t)$ is the position of an imagined particle at time t and $\mathbf{p}(t + \Delta t)$ is its position after time Δt has passed; the integral part describes the particle following the flow for the said time span.

For the given data we implemented Euler integration and a fourth-order Runge-Kutta scheme with adaptive step-size.

6 Hollow Body LIC

The visualizations presented so far are still not ideal for motion analysis in the embryo data. Animated volume rendering cannot be analyzed in stills which makes the analysis of motion patterns rather difficult. Geometric depictions of the pathlines are already quite accessible but strongly suffer from occlusion especially if the entire embryo is rendered (we commonly only depict one half to avoid visual clutter). To equip biologists with a visualization that can be easily used in presentations and papers, we designed a modified LIC version adapted to the given data. LIC is a well-established technique that has seen many improvements over the years. In this section we summarize the settings and alterations we made to fit time-dependent volume LIC [56, 46, 47, 35, 36] to the given data.

Opacity: As we aim at the visualization of the flow inside a hollow body, we can employ 3D LIC and obtain already good results by omitting the zero velocity parts. Care has to be taken not to introduce holes in the surface as there are time spans in the development when the local cell motion is very low in parts of the embryo. A second challenge is the surface of the data. Due to the interpolation scheme and the shape of the cells, the observational domain is a bumpy hollow body as depicted in fig. 6.

The opacity of a voxel is determined using a combination of different concepts in our algorithm. Firstly, there are two cases when a voxel, associated with a grid position $\boldsymbol{x}$, is made completely transparent. These are

1. The voxel has not been hit by a sufficient number of pathlines.
2. The velocity magnitude at the voxel position is too low.

Secondly, if a voxel does not fall into one of the categories above, the velocity magnitude v at $\boldsymbol{x}$ is linearly mapped to an opacity value $\alpha \in [0,1]$:

$$\alpha = \max\left(\alpha_{\min}, \frac{v - v_{\min}}{v_{\max} - v_{\min}}\right)$$

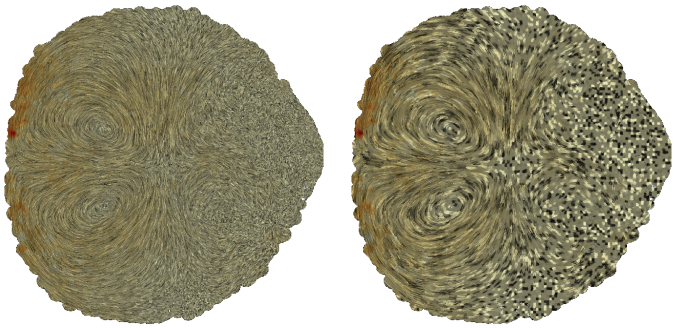
Here, $v_{\min}$ and $v_{\max}$ are the minimum and maximum velocity magnitudes occurring in the full data set (regarding all time steps). $\alpha_{\min}$ is an adjustable parameter that guarantees a minimum opacity. This helps avoiding undesired holes in the topology of the displayed objects and achieves a smooth blending at the boundaries.

Input noise: For the initialization the LIC algorithm requires a noise texture, in our case a 3D texture. An important parameter in this context is the noise frequency φ [28]. Since the noise values are saved in discrete voxels, the frequency describes how likely it is for neighboring voxels to have the same value. The default is to have the frequency match the voxel edge length l with $\varphi = \frac{1}{l}$. A comparison with a lower frequency is depicted in fig. 7. We see that, with decreasing frequency, the patterns become coarser. Thus a slightly lower frequency can make the visualization clearer.

Seeding strategy: The next step is the pathline seeding. Here we follow a two step approach. The first set of lines are started at the nuclei positions stored in the point-based flow field. At these positions we have the most accurate representation of the underlying field and want to use the respective lines as much as possible. In a second round, pathlines are started in the voxels that have not yet been hit by the LIC algorithm and feature non-zero velocity. For good visual results we employ the seeding strategy proposed by Stalling and Hege [49].

Convolution: Along the computed pathlines the texture is folded. We closely adopt the method presented by Shen and Kao [46] with a maximum hit value of 1 and a sampling size of $\Delta s = 0.5 \cdot l$. For smoother pathlines we use a cubic spline interpolation on the integration results.

Color-coding: By construction, the direct results of LIC algorithms are gray-scale images. Additional color-coding can be used to encode relevant flow properties such as pressure, vorticity or the density in the point cloud. In our examples, we always display the velocity magnitude since it is not directly discernible in the raw LIC image and is an important characteristic of the flow.



(a) $\varphi = 1$. (b) $\varphi = \frac{1}{8}$.

Fig. 7: Influence of the input noise frequency on the final image for different frequencies φ normalized by the output pixel edge length. The vortices in the left part of the images are depicted by increasingly rough lines with dropping noise frequency.

7 Evaluation

As we work with data that suffers from much uncertainty in several of the processing steps, validation is a critical issue to ensure trustworthy results. This means primarily that the computed pathlines – if started at the position of a cell nucleus – should follow this point over time. Validation against ground truth data, however, is in general hardly applicable as the manual tracking of cells in 3D is very cumbersome and time-consuming. So we did this only for synthetic 2D data and performed a more user-friendly qualitative and a fully automatic quantitative validation on the real data. The individual steps we took are described in detail in the following sections.

7.1 Analysis of Robust Flow Representation

We manually checked the quality of the compressed flow representatives. To this end, we concurrently displayed volume renderings of the microscopy data and the extracted cell nuclei (fig. 3). In general, we observed a very good agreement of the two modalities. As expected, the largest errors occur in areas of poor video quality. The most common problem is closely located cells that could not be separated. But as detailed before (see section 5.1), we found that such cells often feature a similar motion direction and their unified representation in the reconstructed vector field did not introduce relevant errors.

7.2 Comparison of Interpolation Methods

To find the best suited interpolation settings for our data, we investigated the quality of the local interpolation as well as the quality of the resulting pathlines. We designed synthetic 2D data sets that modeled common flow structures during embryonic morphogenesis. Figure 8 (left) gives a 5 point example of such a flow field. On the left and the right side we have a pair of consistent vectors respectively, while the central vector was slightly distorted and represents a contradicting measurement. Using such examples we explore which technique gives most reasonable results in the presence of noisy data.

The data in fig. 8 (left) was interpolated using the two presented methods and a neighborhood of size 4 ($k = 4$) and quadratic distance ($\alpha = 2$). The example illustrates the fundamental properties of the two interpolation methods: While Barycentric coordinates result in a very smooth vector field that reduces the influence of individual data points, Shepard's interpolation more closely follows the data. With Barycentric coordinates, the vectors rapidly start to bend towards the right side; this behavior is far less pronounced in the case of Shepard's method. Similar observations can be made for the distribution of norms of velocity encoded in color. While large velocity magnitudes are present only in a small area on the left-hand side for Barycentric coordinates, the high-velocity area is far larger in Shepard's method and better reflects the underlying data. Overlaying the interpolation with the microscopy data, we could confirm the better match of Shepard's interpolation.

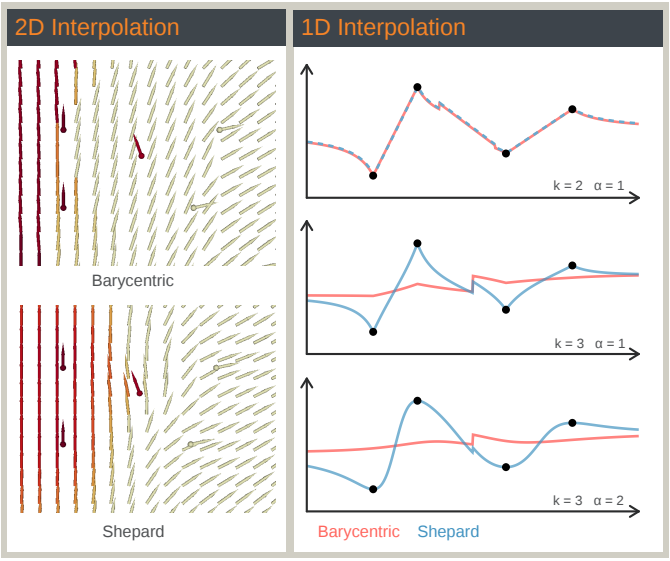
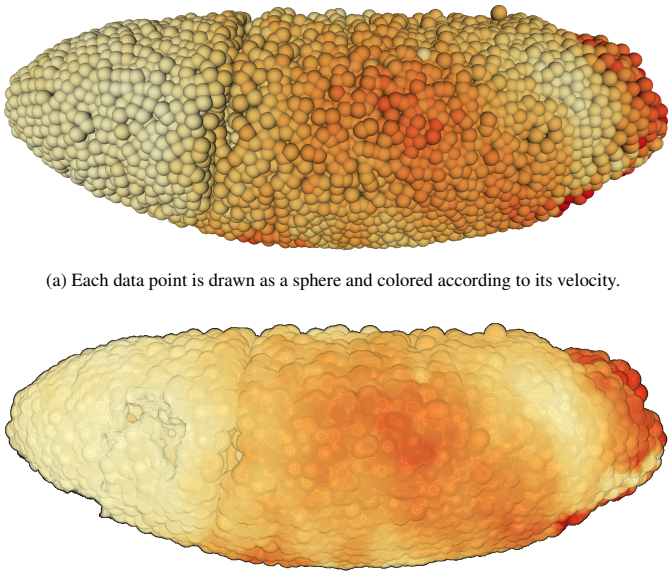


Fig. 8: Comparison of interpolation methods: The choice of interpolation method has a strong influence on the resulting field as is demonstrated for 2D (left) and 1D (right).



(a) Each data point is drawn as a sphere and colored according to its velocity.

(b) The data is rendered as a 3D texture where each voxel is colored according to the velocity interpolated at its center (very small values lead to transparency).

Fig. 9: Interpolation example: Comparison between a depiction of the raw data (a) and a volume rendering of interpolated values (b).

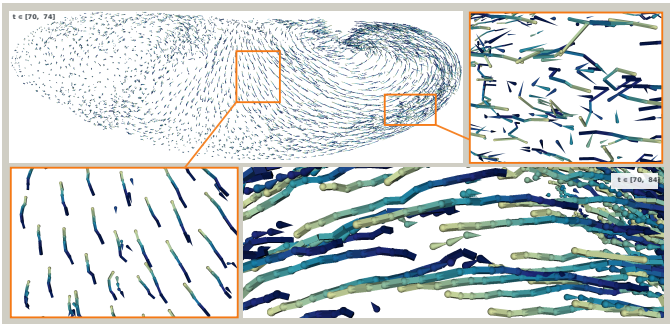


Fig. 10: Visual pathline validation: Pathlines were tracked over five time steps. The images show a concurrent depiction of extracted cell positions and pathline trajectories. The closeup on the right shows a difficult area with poor results. The optical flow in the center is very good as can be seen in the central closeup. In this area, pathlines can be computed over long time frames reliably (bottom right).

Figure 9 gives a comparison between a raw point cloud visualization and an image based on interpolation. For the raw image we rendered a color-coded sphere at each measurement point. The same color-coding was used for the interpolated image where we derived interpolated values for each voxel and used volume rendering to represent the data.

The influence of the parameters k and α on the interpolation can be seen more easily in the 1D example (fig. 8 (right)). For a 2-neighborhood, Barycentric and Shepard's interpolation are very similar. The reconstructed function goes through the measurement points in both cases and results in a roughly linear interpolation between data points. Deviations from linear interpolation, as can be observed at the point of discontinuity between the second and third point, result from the fixed local support and the strong variations in function values which we chose to clearly demonstrate the characteristics of the two methods. As the flow field of the cells is much smoother, such artifacts are not to be expected in real data. As we increase the neighborhood size, we can clearly see the (in this case unwanted) balancing effect of the Barycentric interpolation. For $k = 3$ the interpolated function no longer passes through the extremal points but is a smoothed version. Increasing the power of the distance measure α increases the local influence of the measurement points. $k = 2$ is the commonly recommended choice and results in smoother transitions between the values.

Based on these observations, we chose Shepard's interpolation for our data with $k = 4$ and $\alpha = 2$.

7.3 Visual Validation

Our first validation approach aims at the visual integration of the data processing stages. For each stage we calibrated the visualizations such that they can be combined with each other and allow for concurrent investigation:

Microscopy Data: To depict the microscopy data, we use basic volume rendering to provide data context. The transfer function contains a single color and sets the background pixels to fully transparent. The transparency for the nuclei can be modified depending on whether the data is combined with other modalities or not. Time-series are depicted by means of animation (fig. 3).

Scattered Point Data: The scattered point data is represented by color-coded arrow glyphs that are a combination of a sphere centered at the point location and a cone pointing in the direction of the local flow vector. We use color-coding to encode time. A fixed color-coding for the given time frame enables the user to identify missing cells and infer time points of erroneous cells in the data (fig. 10).

Pathlines: To depict the quality of the integration, we encode the pathlines as cuboid strips. We do not employ smoothing schemes such as splines to correctly represent the computed data at this stage. The diameter of the cuboids is slightly smaller than those of the sphere of the point data to allow for concurrent inspection. We use the same color coding for both pathlines and point data to support the user in the temporal orientation (fig. 10).

Using this setup we inspected the reconstructed cell trajectories in the three data sets. Figure 10 shows examples for the *D. pseudoobscura* data set where cells were traced from time step 70 through 74. We can clearly see the different flow rates resulting in pathlines of variable length. While there is barely any motion in the anterior part of the fly, strong dorso-ventral motions towards the posterior pole can be observed. The pathlines in the central part are very reliable and closely follow the extracted nuclei positions as can be seen in the closeup. In this area even longer pathlines over 15 steps can be computed that still give high-quality results.

Towards the posterior pole, the results get progressively worse due to decreasing quality of the optical flow. This is already clearly visible

Table 1: Pathline Integration Error by Step Size

Δt	2	1	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$
error	23.82	16.88	16.96	17.16	17.37

in the depictions of the scattered point data where the arrow glyphs feature a very chaotic direction profile. No consistent motion can be identified in this area. The pathlines consequently are of poor quality.

7.4 Quantitative Validation

For a quantitative evaluation of the pathline quality, we trace pathlines along the imaged nuclei. In the ideal case, the pathlines would move exactly along the nuclei positions. As mentioned before, perfect ground truth is not available for the given data sets. The best approximation that is accessible with acceptable costs are the approximated cell positions from the optical flow. Hence, we rely on this data for validation.

In each time step we calculate the distance between the pathline position and the closest nuclei position derived from the optical flow. This data can be used to colorcode the pathlines and provide the user with an enhanced uncertainty visualization. Figure 11 gives examples for the *D. pseudoobscura* and the *D. melanogaster* data sets. While the pathlines of the *D. pseudoobsura* data set closely follow the extracted nuclei positions (mostly green color), the pathlines for the *D. melanogaster* data set are colored in red in several areas indicating a large error. Already by visual inspection of the pathlines biologists can easily identify the implausible anterior-posterior motion of cells, which should feature a dorso-ventral profile similar to the *D. pseudoobscura* data set. Backtracking the error in the pipeline, we found that the errors stem from the optical flow computation. This example illustrates the importance of a good error control especially in such multi-step analysis pipelines.

To explore the error distribution over longer time-frames we added a summary chart that provides the user with key numbers for the integration. Figure 12 shows the error chart for pathlines that were integrated over 40 time steps in the *D. pseudoobscura* data set. The white line gives the median error, which increases from zero (the pathlines perfectly agree with the nuclei positions in time step 60) to roughly 6 units (voxel lengths) over the course of the tracking. The surrounding gray area depicts the range between the first and third quartile. To provide reference, we also depict the half average distance between cells. If this boundary is crossed, the cells are very likely no longer close to their "real" cell, but closer to a neighboring one. The blue line gives the average cell motion, which decreases from 4 to 3 units. Here we see that the difference between cell motion and cell distance is not very close, which underlines the request for an accurate tracking scheme.

7.5 Analysis of Integration Step Size

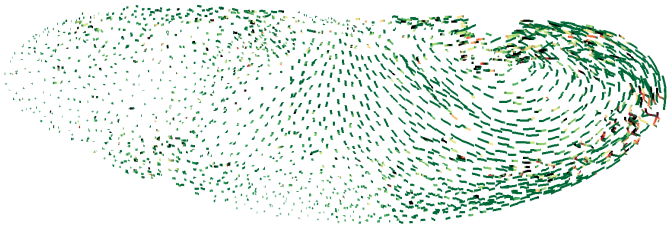
Using the quantitative error analysis of pathlines we can also automatically investigate the influence of different integrators and step sizes on the integration. Due to the coarse nature of the optical flow, we found no significant differences between the Euler integrator and more advanced schemes such as Runge-Kutta-integration.

We also explored the influence of step sizes in the integration process. Table 1 gives the errors for different settings. We see that the error first decreases with decreasing step size. After reaching the optimum at $\Delta t = 1$, the error slightly increases again. Again we account this fact to the discretization initiated by the image resolution and the respective optical flow which also can only achieve pixel accuracy.

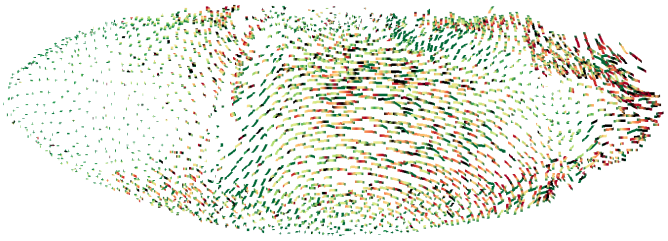
Due to these observations, we currently use Euler integration with fixed step size 1 and no interpolation between time steps.

7.6 LIC Results

Using the method described in section 7.3, we reconstructed the vector fields for three embryonic movie data sets. As detailed before, we found in the data quality analysis that the reconstruction of the optical



(a) *Drosophila pseudoobscura*, time steps 70–74. Only front part.



(b) *Drosophila melanogaster*, time steps 100–104. Only front part.

(c) Uncertainty-based color map from 0 to to "avg. inter-pt. dist.".

Fig. 11: Quantitative analysis of pathlines: Pathlines are colored according to their quality. Green color indicates good quality, red highlights likely errors in the data.

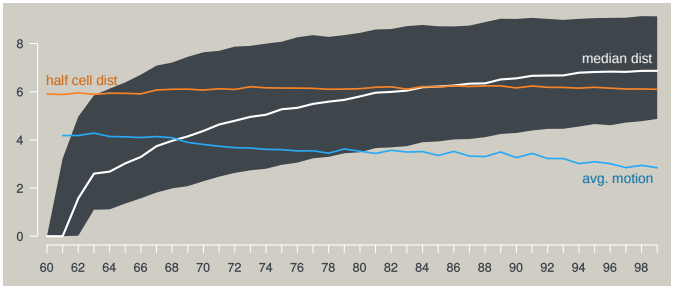
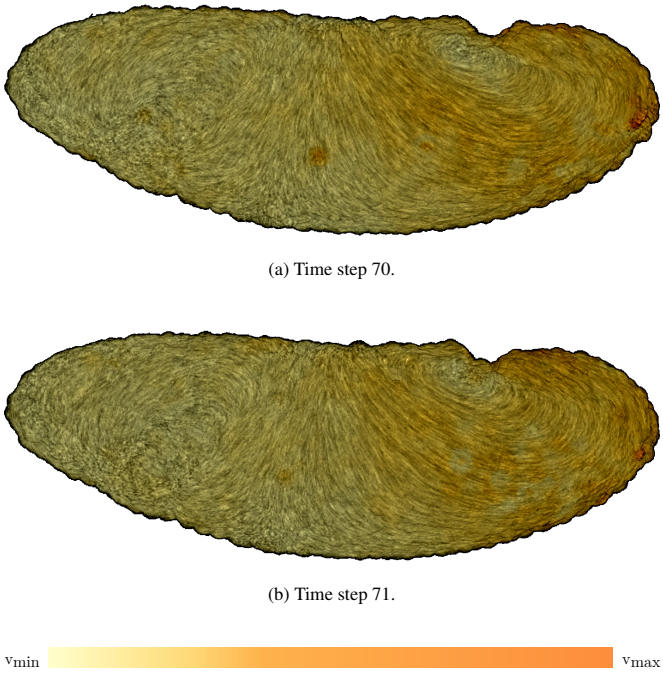


Fig. 12: Quantitative validation: Particles were started at $t = 60$ at nuclei positions and tracked over 40 time steps. The chart depicts the three distance quartiles to the nearest measurement. Lines for the half cell distance (orange) and average cell motion (blue) serve as reference.

flow field for one of the data sets is of poor quality and results in implausible motions. The other two data sets passed the quality checks.

In joint work with the biologists we adapted the settings to give most meaningful results. Here we controlled the length of the pathlines and decided on camera settings for the final videos. For each of these data sets we computed LIC videos from multiple directions. One example is given in the supplementary material (`lic_pseudo1_t040-t087.mp4`) displaying the development of the *D. pseudoobscura* embryo from time step 40 to 87 using the LIC technique.

Figure 13 shows two snapshots of time steps 70 and 71. From the color profile, we can directly observe that there is a strong velocity gradient from anterior to posterior (left to right). While the cells in the head area are barely moving, strong motions from the dorsal to the ventral side (top to bottom) are visible in the central part. The motions extend around the posterior pole and indicate germ band extension. Two prominent low-velocity areas are visible in the flow field, the center of rotation (top right) and a saddle region on the dorsal side. Looking at the pathlines, we readily recognize the surrounding motion



(a) Time step 70.

(b) Time step 71.

(c) Velocity magnitude-based color map.

Fig. 13: *Drosophila pseudoobscura*, time steps 70 and 71. LIC visualization snapshots.

for the 15 time steps that contributed to the image. The direct observation and spatial localization of these patterns was highly appreciated by the biologists as these structures were known but so far hard to visualize.

A second aspect, the biologists are very much interested in, is the temporal development of the motion patterns. For this task we rendered LIC movies from multiple sides. In these videos, it can be observed that the symmetric vortex cores move as the germ band extends. The visualizations also illustrate the difficulties that feature extraction methods are going to face as topological structures are not singular points, but consist of regions with little flow.

Overall the biologists were very impressed with the results. They particularly liked the spatio-temporal consistency in the visualization that makes structural analysis much easier. Another well received aspect of the new representation is the fact that emerging patterns can be seen easily.

8 User Integration and Feedback

The presented pipeline was developed in close collaboration with domain experts in biology (Lemke and Hufnagel lab from COS and EMBL, Heidelberg). The collaboration started in 2010 with a general request to be able to analyze video-microscopy data of morphogenetic processes. After having tested several standard approaches based on segmentation and tracking, the team shifted to the novel direction (integration-based reconstruction). Over the course of the collaboration, regular monthly meetings were supplemented by meetings after significant progress regarding data analysis.

Progress in data visualization capabilities repeatedly led to refined user questions by the biologists. Examples of this positive feedback loop are: (i) The initial plan to visualize fly development was to render long-term trajectories using geometric approaches. This approach proved to be little informative because global cell movements were changing too much. As a consequence, biologists defined more narrow time-windows that give meaningful results. (ii) We proposed texture-based visualization that had not been used previously and which helped the biologists rapidly pick up flow features in their data and allowed them to postulate novel hypothesis that they are currently testing. (iii) With increasing visualization capabilities it became apparent that robustness of data acquisition and analysis needed to be controlled for in order to account for e.g. drift of the embryo during imaging or artifacts introduced during large scale automated segmentation.

9 Conclusion

In this paper, we presented a pipeline for the analysis of 3D+t embryonic videos. The major challenges were the large data size (up to terabytes per data set), the unconventional data structure, and high error rates and uncertainty in the input data. We reconstructed the motion field of cells using the optical flow in a 3D video. We investigated robust measurement extraction, spatio-temporal interpolation, and pathline integration. We particularly concentrated on good data evaluation methods to assess the quality of the resulting data. We further advanced current LIC methods to be applicable in the given 3D+t setting with a hollow body object.

Several improvements are possible for the different steps in the pipeline. Interpolation of scattered point data is a well researched topic with a long history. The application of more advanced schemes would help to guarantee vector field properties such as higher-order continuity. A crucial aspect is still the reconstruction of robust measurement points. With our current scheme we already achieve very robust results which might be further improved by a more dedicated analysis of the super-voxels resulting from the optical flow.

On the visualization side, we want to further improve the visual quality of the LIC as proposed in the related work. A very helpful but challenging project is the automatic extraction of flow features, such as vortex cores and saddle points, in this uncertain data.

Acknowledgments

This research was supported by a stipend of the Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (HGS MathComp). And also by a fellowship from the Hartmut-Hoffman-Berling International Graduate School (L.S.), DFG grant LE 2787/1-1 (S.L.), and NVIDIA Corporation (donation of a Tesla K40 GPU).

References

[1] R. J. Adrian. *Particle Image Velocimetry*. Cambridge University Press, 2011.

[2] F. Amat, W. Lemon, D. P. Mossing, K. McDole, Y. Wan, K. Branson, E. W. Myers, and P. J. Keller. Fast, accurate reconstruction of cell lineages from large-scale fluorescence microscopy data. *Nature methods*, 2014.

[3] F. Amat, E. W. Myers, and P. J. Keller. Fast and robust optical flow for time-lapse microscopy using super-voxels. *Bioinformatics*, 29(3):373–380, 2013.

[4] J. L. Bentley. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18(9):509–517, 1975.

[5] S. Berger and E. Gröller. Color-table animation of fast oriented line integral convolution for vector field visualization. In *WSCG*, 2000.

[6] G. B. Blanchard, A. J. Kabla, N. L. Schultz, L. C. Butler, B. Sanson, N. Gorfinkiel, L. Mahadevan, and R. J. Adams. Tissue tectonics: morphogenetic strain rates, cell shape change and intercalation. *Nature Methods*, 6(6):458–464, 2009.

[7] B. Cabral and L. C. Leedom. Imaging vector fields using line integral convolution. In *Proceedings of the 20th annual conference on Computer graphics and interactive techniques*, pages 263–270. ACM, 1993.

[8] J. A. Campos-Ortega and V. Hartenstein. *The embryonic development of Drosophila melanogaster*. Springer Science & Business Media, 2013.

[9] A. Cardona and P. Tomancak. Current challenges in open-source bioimage informatics. *Nature Methods*, 9:661–665, 2012.

[10] N. Christ, R. Friedberg, and T. Lee. Weights of links and plaquettes in a random lattice. *Nuclear Physics B*, 210(3):337 – 346, 1982.

[11] T. Corpetti, D. Heitz, G. Arroyo, E. Memin, and A. Santa-Cruz. Fluid experimental flow estimation based on an optical-flow scheme. *Experiments in fluids*, 40(1):80–97, 2006.

[12] M. Costa, D. Sweeton, and E. Wieschaus. Gastrulation in drosophila: cellular mechanisms of morphogenetic movements. in the development of

drosophila melanogaster. pages 425–465. Cold Spring Harbor Laboratory Press (New York), 1993.

[13] M. Edmunds, R. S. Laramee, G. Chen, N. Max, E. Zhang, and C. Ware. Surface-based flow visualization. *Computers & Graphics*, 36(8):974–990, 2012.

[14] F. Ferstl, K. Burger, H. Theisel, and R. Westermann. Interactive separating streak surfaces. *Visualization and Computer Graphics, IEEE Transactions on*, 16(6):1569–1577, 2010.

[15] R. Franke and G. M. Nielson. Scattered data interpolation and applications: A tutorial and survey. In *Geometric Modeling*, pages 131–160. Springer, 1991.

[16] T. Günther, A. Kuhn, B. Kutz, and H. Theisel. Mass-dependent integral curves in unsteady vector fields. In *Computer Graphics Forum*, volume 32, pages 211–220. Wiley Online Library, 2013.

[17] D. Heitz, E. Mémin, and C. Schnörr. Variational fluid flow measurements from image sequences: synopsis and perspectives. *Experiments in fluids*, 48(3):369–393, 2010.

[18] B. Höckendorf, T. Thumberger, and J. Wittbrodt. Quantitative analysis of embryogenesis: A perspective for light sheet microscopy. *Developmental Cell*, 23(6):1111 – 1120, 2012.

[19] B. K. Horn and B. G. Schunck. Determining optical flow. In *1981 Technical Symposium East*, pages 319–331. International Society for Optics and Photonics, 1981.

[20] J. Huisken, J. Swoger, D. F, J. Wittbrodt, and E. Stelzer. Optical sectioning deep inside live embryos by selective plane illumination microscopy. *Science*, 305(5686):1007–1009, 2004.

[21] B. X. Kausler, M. Schiegg, B. Andres, M. Lindner, U. Koethe, H. Leitte, J. Wittbrodt, L. Hufnagel, and F. A. Hamprecht. A discrete chain graph model for 3d+ t cell tracking with high misdetection robustness. In *Computer Vision–ECCV 2012*, pages 144–157. Springer, 2012.

[22] P. J. Keller, A. D. Schmidt, J. Wittbrodt, and E. H. Stelzer. Reconstruction of zebrafish early embryonic development by scanned light sheet microscopy. *Science*, 322(5904), 2008.

[23] P. J. Keller and E. H. Stelzer. Quantitative in vivo imaging of entire embryos with digital scanned laser light sheet fluorescence microscopy. *Current Opinion in Neurobiology*, 18(6):624 – 632, 2008.

[24] R. Keller. Physical biology returns to morphogenesis. *Science*, 338(6104):201–203, 2012.

[25] K. Khairy and P. J. Keller. Reconstructing embryonic development. *genesis*, 49(7):488–513, 2011.

[26] K. Khairy and P. J. Keller. Reconstructing embryonic development. *genesis*, 2011.

[27] Z. Khan, Y.-C. Wang, E. F. Wieschaus, and M. Kaschube. Quantitative 4d analyses of epithelial folding during drosophila gastrulation. *Development*, 141(14):2895–2900, 2014.

[28] M.-H. Kiu and D. C. Banks. Multi-frequency noise for lic. In *Proceedings of the 7th Conference on Visualization'96*, pages 121–126. IEEE Computer Society Press, 1996.

[29] H. Krishnan, C. Garth, and K. I. Joy. Time and streak surfaces for flow visualization in large time-varying data sets. *Visualization and Computer Graphics, IEEE Transactions on*, 15(6):1267–1274, 2009.

[30] U. Krzic, S. Gunther, T. E. Saunders, S. J. Streichan, and L. Hufnagel. Multiview light-sheet microscope for rapid in toto imaging. *Nat Meth*, 9(7):730–733, July 2012.

[31] R. S. Laramee, H. Hauser, H. Doleisch, B. Vrolijk, F. H. Post, and D. Weiskopf. The state of the art in flow visualization: Dense and texture-based techniques. In *Computer Graphics Forum*, volume 23, pages 203–221. Wiley Online Library, 2004.

[32] R. S. Laramee, B. Jobard, and H. Hauser. Image space based visualization of unsteady flow on surfaces. In *Proceedings of the 14th IEEE Visualization 2003 (VIS'03)*, page 18. IEEE Computer Society, 2003.

[33] G.-S. Li, X. Tricoche, D. Weiskopf, and C. Hansen. Flow charts: Visualization of vector fields on arbitrary surfaces. *Visualization and Computer Graphics, IEEE Transactions on*, 14(5):1067–1080, 2008.

[34] T. Liu and L. Shen. Fluid flow and optical flow. *Journal of Fluid Mechanics*, 614:253–291, 2008.

[35] Z. Liu and R. J. Moorhead II. Auflic: An accelerated algorithm for unsteady flow line integral convolution. In *Proceedings of the symposium on Data Visualisation 2002*, pages 43–ff. Eurographics Association, 2002.

[36] Z. Liu and R. J. Moorhead II. A texture-based hardware-independent technique for time-varying volume flow visualization. *Journal of visualization*, 8(3):235–244, 2005.

[37] S. Lloyd. Least squares quantization in pcm. *Information Theory, IEEE Transactions on*, 28(2):129–137, 1982.

[38] X. Lou, F. O. Kaster, M. S. Lindner, B. X. Kausler, U. Köthe, B. Höckendorf, J. Wittbrodt, H. Jänicke, and F. A. Hamprecht. Deltr: Digital embryo lineage tree reconstructor. In *ISBI 2011: IEEE International Symposium on Biomedical Imaging*, 2011.

[39] C. M. Lye and B. Sanson. 5 tension and epithelial morphogenesis in drosophila early embryos. *Current topics in developmental biology*, 95:145, 2011.

[40] N. Malik, T. Dracos, and D. Papantoniou. Particle tracking velocimetry in three-dimensional flows. *Experiments in Fluids*, 15(4-5):279–294, 1993.

[41] A. McMahon, W. Supatto, S. E. Fraser, and A. Stathopoulos. Dynamic analyses of drosophila gastrulation provide insights into collective cell migration. *Science*, 322(5907):1546–1550, 2008.

[42] R. Mikut, T. Dickmeis, W. Driever, P. Geurts, F. A. Hamprecht, B. X. Kausler, M. J. Ledesma-Carbayo, R. Marée, K. Mikula, P. Pantazis, et al. Automated processing of zebrafish imaging data: a survey. *Zebrafish*, 10(3):401–421, 2013.

[43] A. F. Möbius. *Der Barycentrische Calcul*. 1827.

[44] G. Myers. Why bioimage informatics matters. *Nature Methods*, 9:659–660, 2012.

[45] N. Olivier, M. A. Luengo-Oroz, L. Duloquin, E. Faure, T. Savy, I. Veilleux, X. Solinas, D. Débarre, P. Bourgine, A. Santos, N. Peyriéras, and E. Beaurepaire. Cell lineage reconstruction of early zebrafish embryos using label-free nonlinear microscopy. *Science*, 329(5994):967–971, 2010.

[46] H.-W. Shen and D. L. Kao. Uflic: a line integral convolution algorithm for visualizing unsteady flows. In *Proceedings of the 8th conference on Visualization'97*, pages 317–ff. IEEE Computer Society Press, 1997.

[47] H.-W. Shen and D. L. Kao. A new line integral convolution algorithm for visualizing time-varying flow fields. *Visualization and Computer Graphics, IEEE Transactions on*, 4(2):98–108, 1998.

[48] C. Sommer, C. Straehle, U. Koethe, F. Hamprecht, et al. ilastik: Interactive learning and segmentation toolkit. In *Eighth IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 230–233. IEEE, 2011.

[49] D. Stalling and H.-C. Hege. Fast and resolution independent line integral convolution. In *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques*, pages 249–256. ACM, 1995.

[50] C. D. Stern. *Gastrulation: From Cells to Embryo*. Cold Spring Harbor Laboratory Press; 1st edition, 2004.

[51] T. V. Truong and W. Supatto. Toward high-content/high-throughput imaging and analysis of embryonic morphogenesis. *genesis*, 49(7):555–569, 2011.

[52] J. J. van Wijk. Image based flow visualization for curved surfaces. In *Visualization, 2003. VIS 2003. IEEE*, pages 123–130. IEEE, 2003.

[53] E. L. Wachpress. *A Rational Finite Element Basis*. Academic Press, New York, 1975.

[54] Y.-C. Wang, Z. Khan, M. Kaschube, and E. F. Wieschaus. Differential positioning of adherens junctions is associated with initiation of epithelial folding. *Nature*, 2012.

[55] D. Weiskopf, M. Hopf, and T. Ertl. Hardware-accelerated visualization of time-varying 2d and 3d vector fields by texture advection via programmable per-pixel operations. In *In Vision, Modeling, and Visualization VMV01 Conference Proceedings*. Citeseer, 2001.

[56] M. Zöckler, D. Stalling, and H.-C. Hege. Parallel line integral convolution. *Parallel Computing*, 23(7):975–989, 1997.