



Universiteit
Leiden
The Netherlands

Biological model representation and analysis

Cao, L.

Citation

Cao, L. (2014, November 20). *Biological model representation and analysis*. Retrieved from <https://hdl.handle.net/1887/29754>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/29754>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/29754> holds various files of this Leiden University dissertation

Author: Cao, Lu

Title: Biological model representation and analysis

Issue Date: 2014-11-20

Chapter 5

Optimizing 3D model representations for 3D phenotype analysis

Partially based on:

L. Cao, F.J. Verbeek. Nature inspired phenotype analysis with 3D model representation optimization. The 4th International Conference on Innovations in Bio-Inspired Computing and Applications, Ostrava, 2013

Abstract: In support of research in biology 3D models are constructed. These are made for visualization and analysis. In micro-anatomy these models often suffer from undersampling in the plane perpendicular to the sampling plane. In this paper we present a pipeline for optimization of 3D models obtained from plan-parallel images so that the 3D model is more suitable for shape analysis. In particular, we intend to create an optimized 3D model for analyzing phenotypical differences originating from different conditions and we extract shape features from the 3D models. The optimization is required for obtaining shape features that properly represent the object. The 3D model contains structures represented as contours that were extracted from the image stack. Starting from a point cloud based reconstruction method, i.e. Poisson reconstruction, we devised a method to convert a stack of contours into a uniformly distributed point cloud. The entire point cloud is integrated in the surface construction resulting in a surface that accurately represents the shape. The feasibility of our method has been confirmed by a representative case study in zebrafish development. The method can be successfully used for datasets from different types of imaging modalities.

5.1 Introduction

In recent studies [Long et al., 2012; Ng et al., 2012; Rubel et al., 2010] three-dimensional(3D) morphological information is used to find the various phenotypes in a sample population. Here a sample population consists of biological objects from images captured with a microscope at a certain magnification. The phenotype differences in the population are found in the micro-anatomy and characterized by shape features. It is clear that 3D models derived from microscope images potentially offer deeper insights for analysis. Therefore, 3D images and 3D models are frequently used for biological object visualization and analysis. With 3D microscopy techniques such as bright field [Tadrous, 2012; Willis et al., 1993] and confocal microscopy [Gouaillard et al., 2009; Natalie et al., 2004] it is possible to obtain 3D information from a plan-parallel stack of 2D slices. For bright field microscopy, the sampling is often realized by invasive physical sections that are acquired to capture images and then reassembled to a 3D image stack. Another option is to acquire images on the confocal laser scanning microscope (CLSM) and process the non-invasive optical sections to 3D models. The physical sectioning technique is very suitable for modelling histological information and larger structures such as tissues or organisms. The confocal technique is more geared towards imaging specific small structures such as cells and small multi-cellular structures such as embryos.

From the 3D stack of images a graphical model can be derived by segmentation or manual delineation of structures of interest. Manual delineation is used when specific structural knowledge cannot directly be derived from the image; a specialist then selects the specific information through graphical annotation [Verbeek et al., 1999a], aka delineation. A set of contours as extracted from the stack subsequently represents the 3D model. The general observation is that the output stack of 2D contours is (nearly) always under-sampled perpendicular to the direction of sampling. In order to improve the model some kind of interpolation can be applied. The classical way of performing such interpolation between section images (slices) is an interpolation of the gray values in the slices so as to estimate the gray values in the missing slices [Herman et al., 1992]. A linear interpolation for estimating the missing slices, however, may lead to artifacts. A

more advanced manner is a shape-based interpolation [Herman et al., 1992] which is applied directly to the contours of the model.

Two dominant methods of 3D shape reconstruction from consecutive contours are contour stitching and volumetric methods. Contour stitching methods directly connect the vertices of adjacent contours and produce a mesh that passes through all contours such as the methods provided by Keppel [Keppel, 1975] and Boissonnat [Boissonnat, 1988]. Keppel's method intends to maximize a function based on the volume of the triangulated surface model. However, this early technique is not handling special cases such as branching. Boissonnat utilizes Delaunay tetrahedralization to successfully cope with branching structures in the model. Contour stitching methods construct the surface by consecutively building up the triangulated patches slice by slice. It disregards the whole picture of the object which could result in noise remaining in between the slices and creating incorrect topologies in the structure of the model. The volumetric methods treat the stack of images as a whole by first interpolating intermediate gray-values and extracting the isosurfaces from a volumetric field. Representative methods are described by Levin [Levin, 1987] and Barrett et al. [Barrett et al., 1994]. Volumetric methods derive the isosurfaces directly from the interpolated gray-values. The smoothness of the surface model depends on the interpolation scheme. We tackle the problem of surface reconstruction using a point cloud based surface reconstruction method. The point cloud reconstruction method treats each point as a feature and inputs features into a mathematical model for an isosurface construction. The merit of a point cloud based reconstruction method is that as abstraction from a specific case to a general problem it sheds light on the critical aspects of the problem [Hoppe et al., 1992]. The process of mathematical model construction facilitates unwanted noise suppression and prominent feature preserving of the surface. In our analytical evaluation of point cloud based reconstruction methods [Cao and Verbeek, 2012] a review of recent methods is given. We aim at reconstructing the surface and suppressing the noise with the information from three directions (xyz). As a result, we develop a pipeline to produce a precise and smooth surface from a set of plane-parallel contours and extract important surface features to analyze the model that is derived from sampling under different conditions.

The remainder of this paper is structured as follows. In section 5.2 we introduce our methodology. In section 5.3 we validate our method with two case studies: zebrafish embryo and the mouse mammary gland. In section 5.4 we present our conclusions and discuss the results.

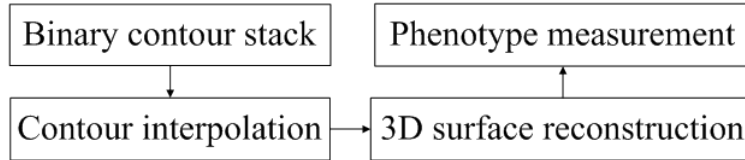


Figure 5.1. Pipeline of our system.

5.2 Methodology for model optimization

The input of our pipeline for model optimization is a stack of nicely aligned binary images with annotation information e.g. as obtained from an annotation software (TDR). Our pipeline is divided into several steps; the individual steps are as schematically shown in Figure 5.1: (1) contour interpolation, (2) 3D surface reconstruction and (3) phenotype analysis. Following to acquisition, the data are organized in a database and, structures apparent in the images are delineated by a contours. This is done either, for each slice through a manual delineation using our dedicated annotation software (TDR [Verbeek et al., 1993]) or automated segmentation; for manual delineation a WACOM digitizer tablet (WACOM, Cintiq LCD-tablet) is used. The contour stacks that are the basis of our models are well aligned by procedures prior to the modeling [Boon et al., 2000] and in case of the confocal images the alignment is an intrinsic quality of the microscope. In our pipeline we use the stack of images with only the contours as input.

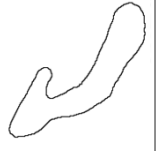
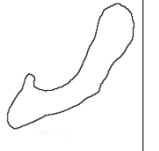
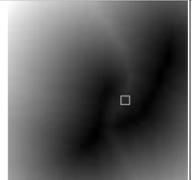
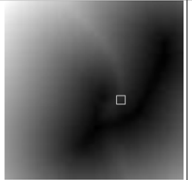
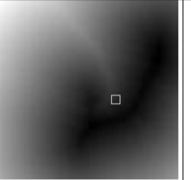
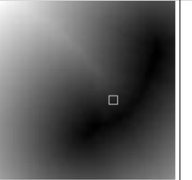
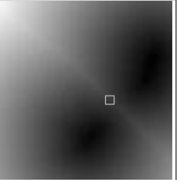
Slice	Slice 1	Slice 2	Slice 3	Slice 4	Slice 5
Contour					
Mask					
DT					
Local distance	-0.5 0.5 1.5	-0.5 0.5 1.5	-0.5 -0.5 -0.5	-0.5 -0.5 -0.5	5.5 5.5 5.5
	-0.5 0.5 1.5	-0.5 0.5 0.5	-1.5 -1.5 -1.5	-1.5 -1.5 -0.9	4.5 4.5 4.5
	-0.5 0.5 0.9	-0.9 -0.5 -0.5	-2.5 -2.5 -2.5	-2.5 -2.3 -1.7	3.5 3.5 3.5

Figure 5.2. Contour interpolation procedure from 5 example slices. Mask= Binary mask of the contour; DT= Distance transformation represented by Gray-value.

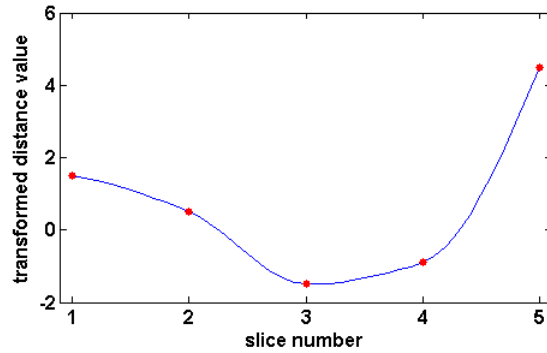


Figure 5.3. Piecewise cubic Hermite interpolation.

5.2.1 Contour interpolation

The sampling of an image stack derived from physical sectioning specifically is, in general, insufficient in the direction of sectioning (z direction). The insufficient

5. 3D MODEL PHENOTYPE ANALYSIS

information in the Z-direction would prevent us from making a good surface description. Therefore, we introduce a method of subsampling through interpolation and develop the method to produce a truthful interpolation. In this process, we make use of a shape-based contour interpolation method [Verbeek, 1992; Verbeek et al., 1995]. To that end a distance transform is applied in the contour as shown in figure 5.2; each point in the resulting image represents the shortest distance to the contour. The points on the contour are zero. Now, the distances inside the contour set positive whereas the distances outside the contour are multiplied with -1 and thus these have a negative value. Setting the negative value is to distinguish between the pixels outside of the object and the pixels belonging to the object. For each pixel-column in the z-direction, a 1D monotone piecewise cubic spline [Fritsch and Carlson, 1980] is constructed to interpolate distance values in z-direction [Braude et al., 2006]. A monotone cubic interpolation is based on the cubic Hermite spline with a coefficient matrix that looks as follows:

$$P(t) = \begin{bmatrix} t^3 & t^2 & t & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 & 1 & 1 \\ -3 & 3 & -2 & -1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{p}_0 \\ \mathbf{p}_1 \\ \mathbf{v}_0 \\ \mathbf{v}_1 \end{bmatrix} \quad (5.1)$$

where $\mathbf{p}_0$ and $\mathbf{p}_1$ are two endpoints; $\mathbf{v}_0$ and $\mathbf{v}_1$ are two related tangents. If a cubic Hermite spline is used for interpolation of a monotonic data set, the interpolated function will not necessarily be monotonic, but monotonicity can be preserved by adjusting the tangents. In Figure 5.3 an example for the construction of a 1D-interpolation in z-direction is shown. A monotone piecewise cubic spline is used to obtain a smooth function from interpolation. The output spline preserves the shape of the data and at the same time respects monotonicity. In this manner derived overshooting artifacts are eliminated by the method. So as to say, it will not introduce extra or artificial surface information to the biological model. We project the biological model in a bounding box before we do the interpolation so as to neglect unnecessary interpolation calculation. Once the spline is constructed from each vertical column within the bounding box, the intensity of intermediate missing slices at the same column can be evaluated by providing different position values in the z-direction. Finally, the interpolated contour is extracted by setting

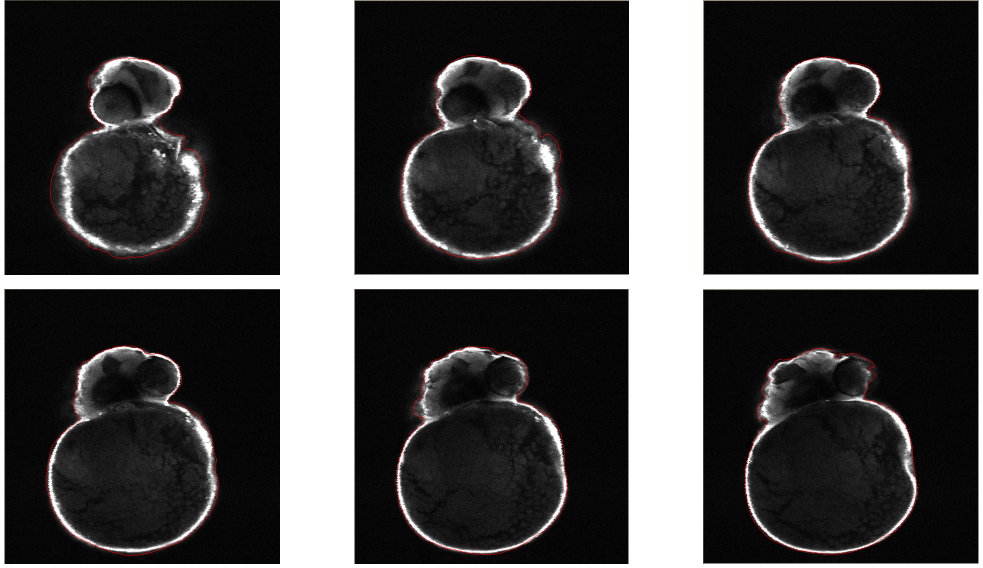


Figure 5.4. Example of original images of zebrafish embryo from confocal microscope.

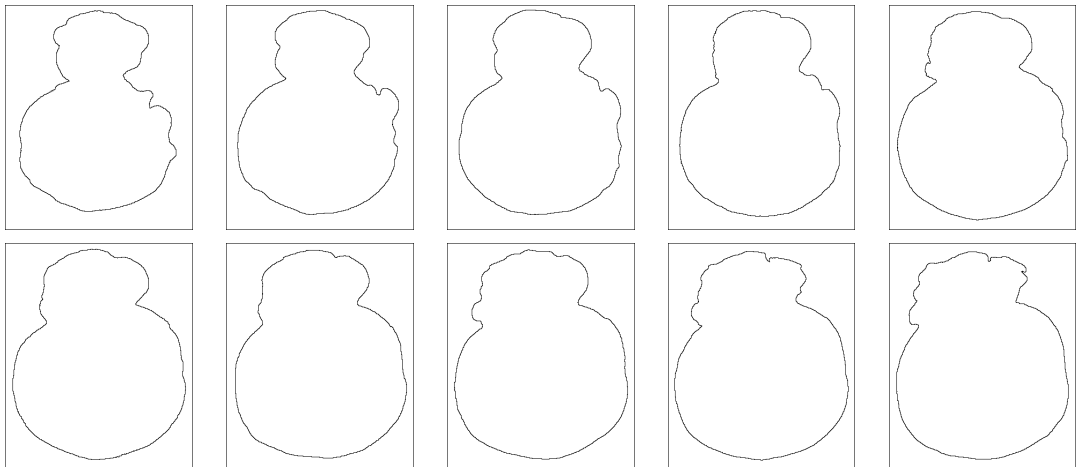


Figure 5.5. Example of contour interpolation of zebrafish embryo.

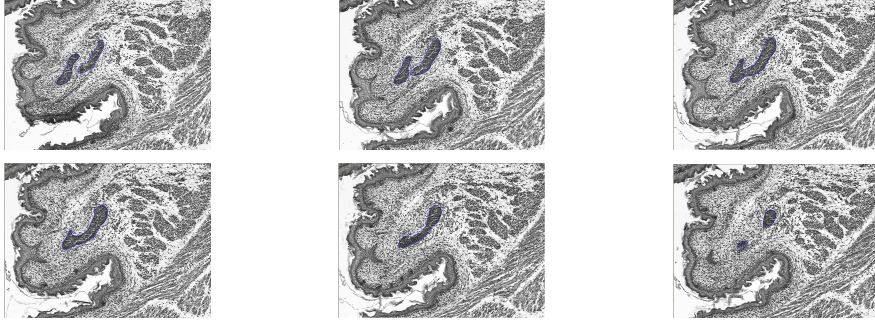


Figure 5.6. Example of original images of mammary gland from bright field microscope.

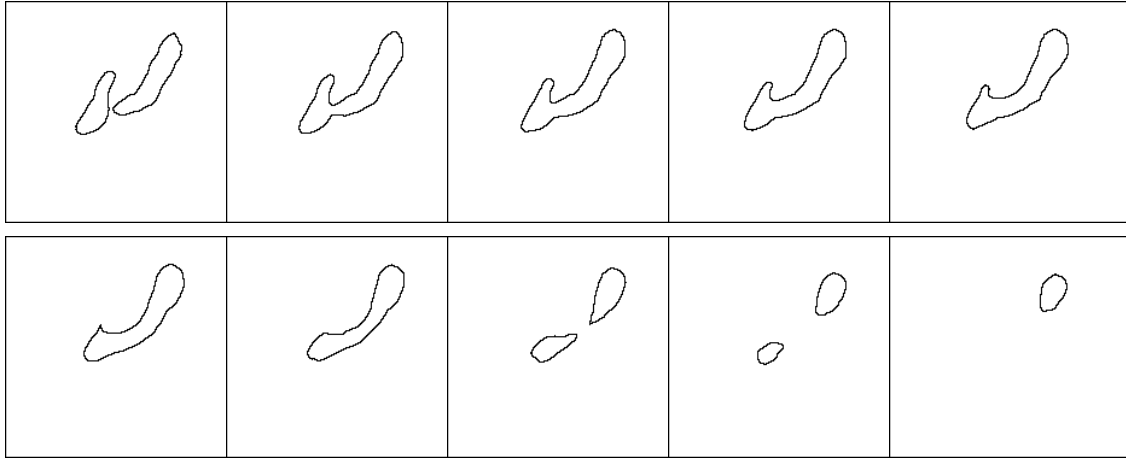


Figure 5.7. Example of contour interpolation of mammary gland.

the threshold of gray-value to zero for each slice. This approach results in an interpolated and equidistant sampled boundary for the model. The number of extra slices we have added for our model is based on the amount of difference in resolution at x,y directions and at z direction. For example, we capture a zebrafish embryo with a confocal microscope using a 10x objective (N.A. 0.5). The average volume of the zebrafish embryo is 0.1136 cubic millimeter. The pixel size in x,y direction is 0.821 μm and the optical cutting distance in z direction is 5 μm . Therefore, we added 6 extra slices in between two consecutive slices to obtain an equidistant sampling.

5.2.2 3D surface reconstruction

In order to be able to use the point cloud based reconstruction method, the stack of binary contour images needs to be converted into a point cloud in 3D space. Each pixel is converted to a point in 3D space taking the corresponding z-position of the slice into consideration. In this manner the point cloud is created without losing any detail and at the same time it is indisputably oversampled in both x,y and z directions. From a previously performed evaluation of point cloud based reconstruction methods [Cao and Verbeek, 2013], we concluded that the Poisson reconstruction method [Kazhdan et al., 2006b] performs the best for taking shape preservation and noise suppression into account. The Poisson reconstruction method, however, requires an oriented point cloud as an input. This means the method does not only need the location but also the normal of each point in the point cloud data. Therefore the normal for each point in the point cloud data is calculated using Hoppe’s algorithm [Hoppe et al., 1992]. The normal of a point X_i is determined by gathering together the k nearest points to X_i ; expressed as a nearest neighborhood function ($Nbhd$). The tangent plane of X_i is constructed as the least squares best fitting plane to $Nbhd(X_i)$. The normal of point X_i is calculated using a principle component analysis. The covariance matrix of $Nbhd(X_i)$ is formulated from:

$$Cov = \sum_{y \in Nbhd(X_i)} (y - o_i) \otimes (y - o_i) \quad (5.2)$$

where $\otimes$ denotes the operation of outer product vector. y is one point in the nearest neighborhood set. o_i is the centroid of $Nbhd(X_i)$. The normal of point X_i is either positive or negative of the smallest eigenvalues of Cov . The positive or negative selection depends on the consistency of orientation of nearby tangent planes.

After having obtained on oriented point cloud data set, the Poisson reconstruction method is applied to create a precise and smooth surface for the model. If necessary, the resolution of the resulting surface model can be tuned by changing the scale parameter which is part of the Poisson reconstruction method. We evaluated the effect of the scale parameter on the reconstruction step by chang-

ing the scale value and compare the error between the ideal and real output with the analytical models. The scale parameter we have used here is based on this evaluation and sufficient for the biological model reconstruction (scale=2). The scale parameter still retains sufficient details on the surface for phenotype measurement.

5.2.3 Phenotype measurement

With the reconstructed surface model, we need to find a way for 3D shape description so as to distinguish a difference in phenotype with various treatments. The type of measurements extracted from the 3D model strongly depends on the model at hand and the hypothesis posed to resolve phenotypical differences. We extract a range of different features as shown in Figure 5.8, some are derived directly from surface including global shape, i.e. volume and surface area, and local features such as surface curvature per point.

More features can be extracted from a graph based representation of the shape, i.e. skeleton or centerline. These features encode geometrical and topological shape properties in a faithful and intuitive manner [Akgül et al., 2009]. The centerline is useful to describe the topology information of tubular structures such as blood vessels. Additional features can be extracted from the centerline such as number of branches and nerves, average branch-length, number of bifurcations, and so on. These basically follow from the graph structure.

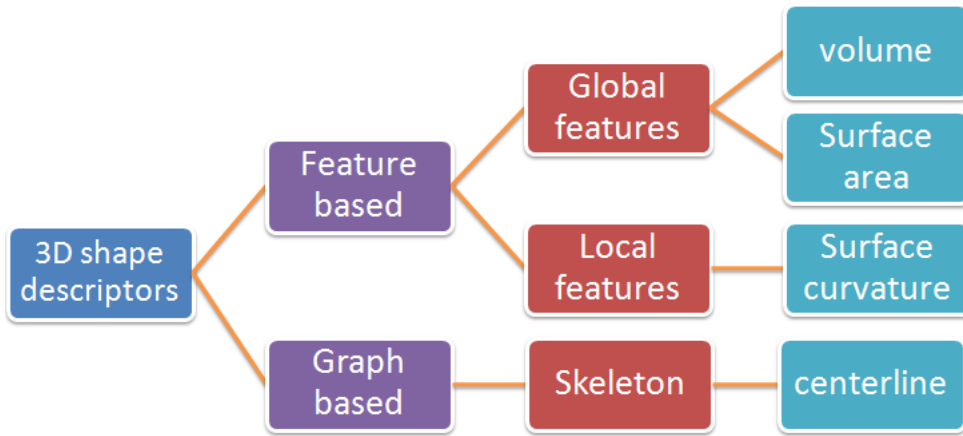


Figure 5.8. 3D shape descriptors.

5.3 Evaluation of the methodology

After the illustration of the methodology of model optimization, we continue with the evaluation of several important steps. We have three main steps as shown in Figure 5.1. Apart from the 3D surface reconstruction step, it is a parametric free pipeline. The step of 3D surface reconstruction is already evaluated in our former work [Cao and Verbeek, 2013]. In this section, we will look in detail on the contour interpolation, because this method might introduce interpolation error when the model reaches to a higher level of complexity. We want to check how well our solution is in dealing with different kinds of models. The second evaluation is given a general view on the level of improvement of the surface reconstruction compared to methods used earlier.

5.3.1 Evaluation of interpolation method

The interpolation by the 1D monotone piecewise cubic spline can introduce inaccuracies for topologically complex 3D models that are common in biology. Thus, in order to check the topological correctness of this important processing step applied to the 3D model, we also use a more complex 3D model for our evaluation. To this end the mammary gland is very suitable because it has a higher level of complexity in the shape structure, i.e. branches.

We design following evaluation method. The interpolation part is a crucial step in the construction of the relationship between slices and at the same time the correct topology of the model needs to be preserved.

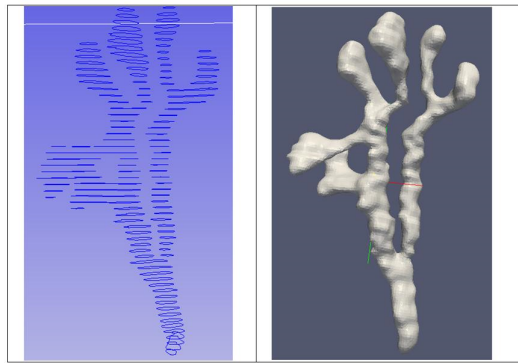


Figure 5.9. (a) original contour sections; (b) surface model.

5. 3D MODEL PHENOTYPE ANALYSIS

Table 5.1. Evaluation result of interpolation part

Model name	#Contours	#Incorrect contours	Percentage
DES	1126/1271	0/0	100%/100%
EP	904/508	4/9	99.56%/99.01%
OLIE	852/1071	0/0	100%/100%
WT	746/924	0/0	100%/100%
Total	7402	13	99.82%

DES: a group exposed to a range of concentration of diethylstilbestrol; EP: a group exposed to a cocktail of estrogen and progesterone; OLIE: a condition control group exposed to an inert component; WT: a control group which is not exposed.

Note: Only in EP there is a problem.

- Therefore, first of all, we labeled the branches in original model with different IDs as shown in figure 5.9(c).
- Next, we manually annotate each contour derived from interpolation method with the same ID system relating to a specific branch by checking the correlated location on the output surfaces.
- Subsequently, we validate each contour ID in one section from the interpolation method with the corresponding contour ID in two closest slides from original model.

In this manner, we know how many contours from interpolation method mismatch with each other. By treating each contour as an individual object, we finally calculate the percentage of correctness.

From a selection of 8 mammary gland models that are obtained from species that have been exposed to different conditions (cf. Chapter 6) the evaluation result is shown in Table 5.1. These models are constructed through physical sectioning with a standard histology staining and acquired on a section by section basis with a bright field microscope. Therefore, it is undersampled in Z direction and also deformed. The incorrect contours we found were mostly because two contours are too close to each other, the interpolation method cannot nicely separate them. However, the percentage of correctness is still high which shows that the interpolation method can properly preserve the complex topology such

as in our mammary gland models. Thus, we could confidently apply this method in our process.

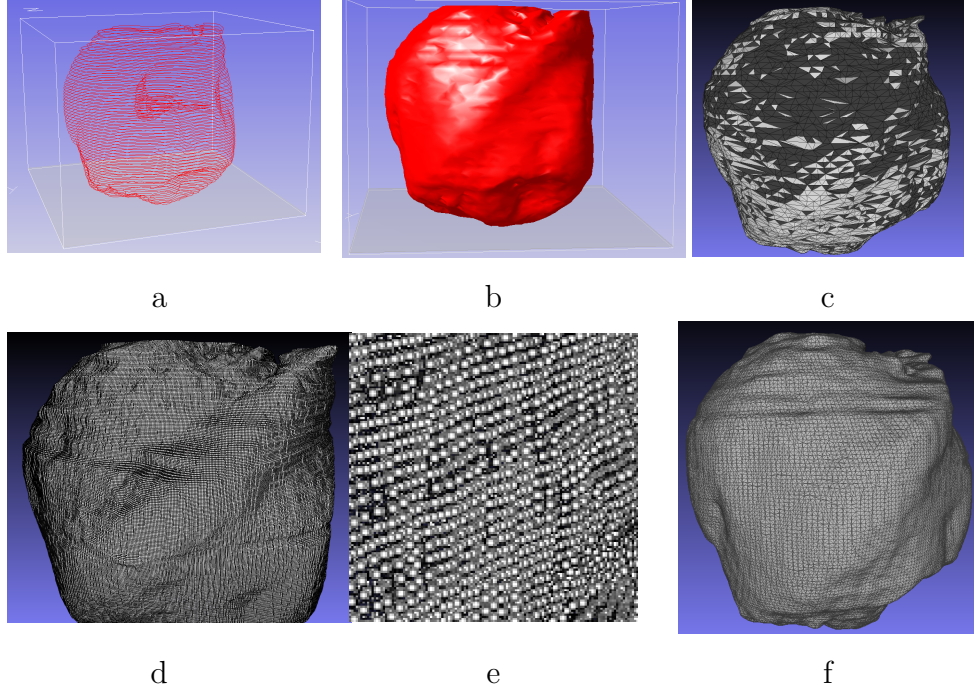


Figure 5.10. Comparison between two reconstruction methods. a. stack of contours; b. surface model from Boissonnat method in TDR; c. surface model from Boissonnat method in Meshlab; d. interpolated point cloud; e. close view of the interpolated point cloud; f. surface model from Poisson reconstruction method.

5.3.2 Evaluation of surface reconstruction method

The requirements of the surface quality for phenotype measurement are two fold. First, the surface should have a consistent surface orientation. Second, the surface should be a closed 2-manifold surface without holes on it.

A method that we successfully used for contour-based reconstruction is designed by Boissonnat and Geiger [Boissonnat and Geiger, 1993; Verbeek et al., 1995]. Their method is based on tetrahedralization of a volume. The method uses the medial axis to correct a nearest neighbor connection between adjacent slices. It is good at reconstructing complex contours such as complicated branching patterns and topologies with holes. However, extra postprocessing is required so as to

make the surface model ideal for phenotype measurements.

However, we need extra processing to adjust the output surface for the phenotype measurement. As we can see from Figure 5.10c, the reconstructed surface model does not guarantee a consistent orientation of the faces. We need post processing to reorient the faces. Furthermore, a closed surface is a compact connected 2-manifold. But the surface model has some not 2-manifold faces which makes the re-orientation of the surface even harder, since we need to find the location of the problematic faces, delete it and try to fill the holes on the surface. These problems can be addressed by the Poisson reconstruction method as shown in Figure 5.10f. It is an implicit reconstruction method which consider all points as a whole. The output surface is closed and consistently faced. Additionally, the surface is much smoother with more characteristics preserved.

5.4 Results

The dataset resulting from the surface reconstruction method is used for the shape analysis. In this chapter, we will illustrate this with models obtained through confocal imaging i.e. the point cloud of zebrafish embryos.

5.4.1 Zebrafish embryo, measurement verification

We want to illustrate that the surface reconstruction process can be very well applied to a stack of images. As a test data set, an experiment with confocal image stacks of zebrafish embryos are used. The images are part of an experiment that consists of two treatments and therefore two different groups; i.e., a control group that develops under normal oxygen levels and a treatment group that develops under a condition of hypoxia; i.e. oxygen deprivation. For each stack, a 3D model of the embryo is constructed by extracting embryo contour per slice. In total the set consists of 21 embryo models (14 with normal treatment, 7 with hypoxia treatment). In Figure 5.11(a) and (b) an example of embryo models with different treatments are shown.

For each model we calculated the surface area and the volume. The result from the point-cloud reconstruction to shape measurements is a triangulated surface.

The surface area is computed by integration over all the triangle patches on the surface. The volume is calculated directly from the grid using divergence theorem introduced in [Zhang and Chen, 2001]. For an objective assessment of the differences between two treatments, we use the sphericity [Wadell, 1935] as shape descriptor, described as follows:

$$\Psi = \frac{\pi^{\frac{1}{3}}(6V)^{\frac{2}{3}}}{A} \quad (5.3)$$

where V represents the volume of the object and A represents the surface area of the object. The sphericity is a shape descriptor and for a sphere it equals 1. The higher the sphericity of the object the more it resembles a perfect sphere.

Our dataset is not guaranteed to adhere to a normal distribution. To test for differences, the Kolmogorov-Smirnov Test (KS-test) is used [Massey, 1951b]. The KS-test is a better choice than Student t-test when the dataset is not guaranteed to adhere to a normal distribution. Sphericity is computed for all our models and thus mean and standard deviation of the set are available. We can classify two groups on the basis of the measurement of sphericity. The results are depicted in Figure 5.11(c) and Table 5.2. On the basis of the measurements we could classify into two classes. The test results indicate that sphericity of the control group is significantly smaller than that of hypoxia group. This is consistent with the biology; the hypothesis states that the embryo is spherical at the very beginning and starts to unfold during the development. Hypoxia slows the development of the zebrafish embryo. Under normal levels of oxygen the development is faster. The embryo develops uneven in different directions and does not resemble the spherical shape anymore that it started from. The development of embryo with hypoxia, however, is restrained by the lack of oxygen as a result the shape is closer to spherical state that it started from.

Table 5.2. Measurements of Zebrafish embryo

Type	$A(mm^2)$	σ	$V(mm^3)$	σ	S	σ
Normal	0.3397	0.0209	0.0127	0.0007	0.7759	0.037
Hypoxia	0.3568	0.02695	0.0154	0.0016	0.8388	0.0136

A=Area; V=Volume; S=Sphericity; σ =Standard deviation.

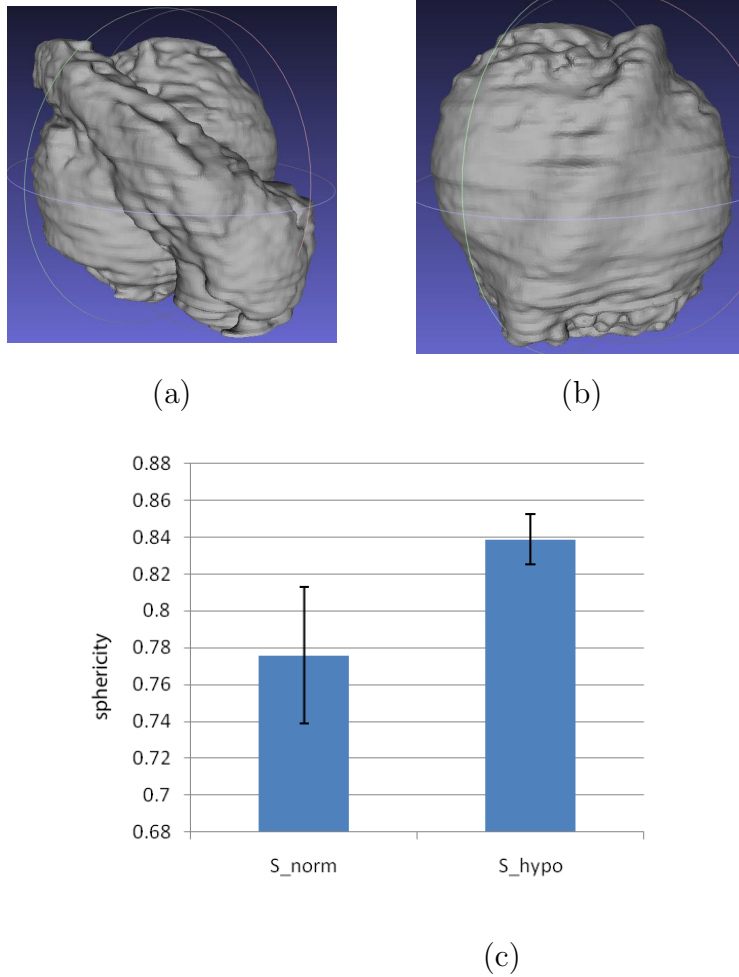


Figure 5.11. Results of Zebrafish embryo (a) Normal condition (b) Hypoxia condition (c) Bar chart of sphericity feature indicating the constrained embryo development under the condition of hypoxia.

5.5 Discussion & conclusion

In this chapter, we have computed a system for 3D representation and analysis from stack of images through a 3D model representation that is derived from that stack. We intend to use the representation for phenotype analysis. The input data for our system is a stack of images whose annotation and alignment are completed. However, the 3D model starts undersampled which is fine for visualization but not for quantification and therefore optimization is required. To that end a

point cloud based reconstruction method is used. In this manner the stack of contours in the model is efficiently converted into a uniformly distributed point cloud in 3D space. As such it changes a specific contour based reconstruction problem to a much more generalized point cloud based reconstruction method. Since we take all the 3D points at once to construct the surface representation it makes full use of the boundary point cloud relationship in 3D space. In this way we overcome the restriction imposed by the stack of contours which does not efficiently use the relationship in the z-direction. From former results [Cao and Verbeek, 2012, 2013] we have learned that the Poisson reconstruction method performs well in both shape preserving and noise suppression. This is confirmed in our current study.

Our purpose, i.e. representation and analysis of 3D biological models, the analysis pipeline presented here combines all best techniques and helps us observe the real situation. The verification of the pipeline which includes evaluation of important steps and zebrafish embryo surface analysis confirms the potential of our system in dealing with shape related studies in biology. The pipeline is applicable to different kinds of datasets that originate from different microscopes and sampling conditions. The number of additional slices interpolated in between two consecutive slices are defined by the differences in resolution at x,y direction and z direction. The complexity that one can find in micro-anatomy can be well covered by the pipeline. The pipeline needs further tuning though the type of models that we have now worked with illustrate the soundness of the pipeline. Henceforward, we will continue to develop this pipeline and include more techniques to be able to deal with the large variation of experimental settings for 3D phenotype analysis in biomedical research.