

L-systems from 3D-imaging of Phenotypes of Arborized Structures

Fons J. Verbeek*, Lu Cao

Imaging and Bioinformatics group

Leiden Institute of Advanced Computer Science (LIACS)

Leiden University, Leiden, The Netherlands

f.j.verbeek@liacs.leidenuniv.nl, l.cao@liacs.leidenuniv.nl

Abstract. Biology is 3D. Therefore, it is important to be able to analyze phenomena in a spatio-temporal manner. Different fields in computational sciences are useful for analysis in biology; i.e. image analysis, pattern recognition and machine learning. To fit an empirical model to a higher abstraction, however, theoretical computer science methods are probed. We explore the construction of empirical 3D graphical models and develop abstractions from these models in L-systems. These systems are provided with a profound formalization in a grammar allowing generalization and exploration of mathematical structures in topologies.

The connections between these computational approaches are illustrated by a case study of the development of the lactiferous duct in mice and the phenotypical effects from different environmental conditions we can observe on it. We have constructed a workflow to get 3D models from different experimental conditions and use these models to extract features. Our aim is to construct an abstraction of these 3D models to an L-system from features that we have measured. From our measurements we can make the productions for an L-system. In this manner we can formalize the arborization of the lactiferous duct under different environmental conditions and capture different observations.

All considered, this paper illustrates the joint of empirical with theoretical computational sciences and the augmentation of the interpretation of the results. At the same time, it shows a method to analyze complex 3D topologies and produces archetypes for developmental configurations.

Keywords: L-system, Phenotype analysis, Center-line topology, lactiferous duct, 3D model representation

* Address for correspondence: Imaging and Bioinformatics group, Leiden Institute of Advanced Computer Science (LIACS), Leiden University, Niels Bohrweg 1, 2333 CA, Leiden, The Netherlands

1. Introduction

In this paper we explore the use of L-systems to support the analysis of phenotypes in molecular developmental biology - that is the measurable characteristics of an organism from its interaction with the environment. We address a particular case study of the development of the lactiferous duct in newborn mice under normal circumstances and in conditions of environmental stress.

The murine lactiferous duct is a branching structure that is internal in the breast of the organism, from the outside only the nipple sheet is visible. Thus in order to visualize the structure of the lactiferous duct, newborn mice need to be prepared for imaging.

The imaging and the construction of a 3D model is the first step in a workflow for the analysis of the phenotype of the lactiferous duct as it is exposed to environmental stress. The 3D model will be used to extract features. Next, the features will be compared over different conditions requiring statistical tests and pattern recognition to take the different condition groups apart.

The lactiferous duct is developed and present in a rudimentary form at the foetal stage at birth. The development of the lactiferous duct is part of a maturation process that is parallel with the reproductive organs; for the female mouse this maturation primes the individual for reproduction. The developmental process is under hormonal control. The development can be distorted by over-exposure to certain hormones or compounds that are very similar to hormones and that mimic hormonal operation, so called endocrine disruptors. These chemicals are present in the environment and thus have a potential effect on the healthy maturation of the reproductive organs or secondary sexual organs. This effect of endocrine disruption can be studied by the phenotypical effect on the lactiferous duct. In a case study this is shown by a maternal exposure to endocrine disruptors.

The investigation of arborized structures requires a number of steps in imaging to come to a data set that can be further processed. Measurements can confirm that differences between conditions exist, but do not give us insight in where these differences occur nor how the initial layout of a structure changes under environmental pressure.

For our case study we start from a simple contour model that is further optimized to accommodate the analysis of the complex structure of the lactiferous duct in the best possible manner [1]. However, as accurate the 3D renditions of the lactiferous duct can be, they still remain rather complex. A further simplification can help in understanding the differences between models. Therefore, we investigate simplifications that preserve topology of the models. We assess the use of the skeleton and the center-line of the enhanced 3D model. From the enhanced 3D model and the center-line features can be extracted to obtain statistical evidence of the differences. Still that does not give an answer to the understanding on the innate model of the duct development and how environmental effects can alter the structure. Therefore we investigate a further abstraction and assess the use of L-systems in describing the core shape of the lactiferous duct and the effect of environmental factors. The idea is that the archetypal form as described by the L-system is represents the way the shape is programmed in the genetic markup, just like an L-system can represent the form of a plant. The intriguing question is whether or not the models from the imaging provide sufficient information to construct an L-system that represents the lactiferous duct well.

This paper is structured as follows. From the exposition of the problem given in the introduction we provide the background information of the components in the workflow in section 2. In the

following section we apply the workflow and illustrate the production of the models including some validation. In section 4 we extract the features, determine the prominent features and develop the abstraction to the L-system. Finally, in chapter 5 we present the conclusions and discuss the implications of the approach for the phenotypical analysis.

2. Background to the phenotype analysis workflow

In order to accomplish an analysis of our material, a number of components need to be explained in context and detail. Therefore, in this section these components, as depicted in Figure 1, are discussed.

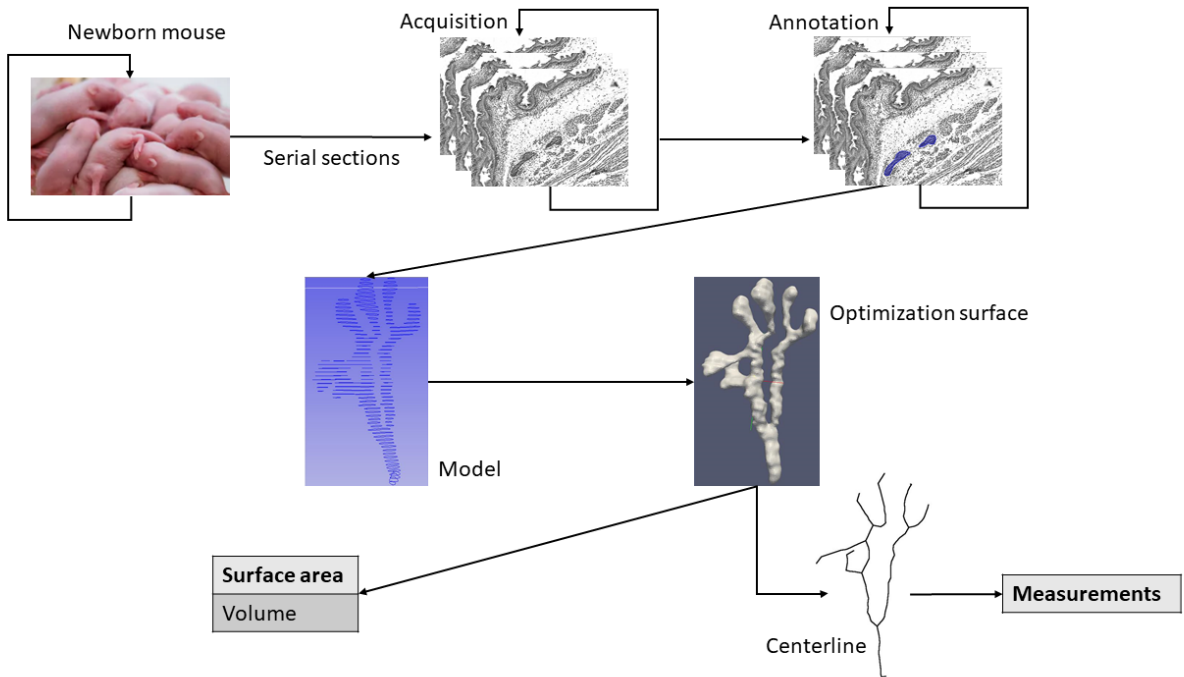


Figure 1. Overview of the workflow of the data processing for the phenotypical analysis of the lactiferous duct

2.1. Building a 3D models from images

In our case study we investigate the topology of the breast gland at birth, i.e. the topological status of fetal breast gland. In order to be able to produce a data set, fetal mice are sacrificed, fixed and prepared for serial sectioning. From physical serial sections we can obtain a 3D image from which we will be able to extract a model of the breast gland. To this end we make use of our dedicated setup for registration and realignment of serial sections [2].

In order to obtain contrast, sections are stained with a standard histological stain. Next, while observing the section under the microscope, in each section a part is selected that contains the lactiferous duct; the suggested continuity of the duct is obtained by starting at the section that contains the nipple

sheet. The lactiferous duct is then followed from there with all its branches. So, from acquisition each section is represented by a section image. Our microscope setup [2], guarantees that subsequent section images are aligned and over the extent of a lactiferous duct constitute a coherent 3D image.

Making a model from histological sections has the advantage that internal structure of an organism can be elucidated. This is particularly the case for larger sized objects for which other imaging is not an option and, moreover, where the histological staining is required to make the necessary contrast for the differentiation of the different constituents. This approach had, however, disadvantages. First, the sectioning will introduce some deformations and second, the sampling to images in the plane of sectioning is much higher than it is perpendicular to the direction of sectioning; commonly referred to as the sampling in the z -direction. This results in an undersampling of the model that we obtain. The sampling in the plane of sectioning is determined by the combination of the lens and the CCD camera that is mounted on the microscope. The sampling perpendicular the direction of sectioning is determined by the section thickness. This thickness is a compromise between higher sampling rate and more deformation, i.e. thinner sections result in more deformation.

From the section images a model is extracted using our 3Dbase software [3]. The coherent set of section images is used to find the structures of interest per section image. A structure of interest is extracted from the section image and a 3D model is constructed by repeating this over the stack of section images. Initially this 3D model is composed of contours of the selected structures. The next step is a conversion from the contour model to a volume or a surface model [3].

As indicated, in general, a model is constructed which is undersampled in the z -direction. We therefore need to improve the model through resampling between the initial contours, thereby enhancing the sampling and decreasing the space between contours. The subsampled contours do not have a relation to a section image. Interpolation of these contours is realized by the use of a distance transformation [4] and the projection of the contour in layer n to the distance transformation in layer $n+1$. The rationale is to have the spacing in the z -directions equidistant to that in the xy -plane. With a piecewise monotonous cubic spline [5] the contour points in the z -direction are connected. The equidistant planes are sampled over all points in equidistant steps [1]. In this manner contours are inserted to be able to generate a smoother 3D surface of the original 3D model of the duct. Thus, the resampled 3D contour model is the onset for a better sampled 3D surface model. To that end, from the contour model a point cloud is generated and from this 3D point cloud, using Poisson reconstruction [6], a 3D surface model is computed. The 3D surface model is represented by a triangular mesh.

We take this approach from our analysis that the Poisson reconstruction provides a good and accurate reconstruction of the surface that can henceforth be used in measurement of the surface area [37]. Surface area, and derived characteristics, are among the features we would like to include in the analysis of the phenotype. The dual of the surface model, the volume model, is used for the computation of the volume of the lactiferous duct.

2.2. The center-line as a topological descriptor for shape

In combination with the 3D model, we need to stress topological characteristics of the shape. Arborized structures can be complex and 3D phenotype analysis specifically requires topological features to detect morphological changes in different conditions. The feature extractions is part of the

workflow as depicted in Figure 1. Graph based measurements have the potential of encoding geometrical and topological shape properties in a more faithful and intuitive manner [7]. The skeleton is such typical descriptor that simplifies the shape while preserving its topology. It is defined as a thinned version of the shape equidistant to its boundaries. In 3D, the skeleton also contains surface patches. The center-line is a special case of the skeleton, it is a simplified 1D representation of the original 3D object, consisting only of curves [8]. The center-line is derived from the two outermost sections of an arborized structure which locally maximizes the distance from the boundary [9]. The center-line includes the topology information of arborized structures in a compact and efficient way. Different center-line extraction methods have been presented in bio-medical imaging. These can be separated into three categories: topological thinning, distance-based methods and polygon-based methods.

Topological thinning deals with a volumetric dataset. Boundary voxels are iteratively removed according to topological rules until the whole volume has been peeled and left with the center-line [10, 11]. Thinning procedures usually create a connected one voxel thick center-line. However, from the constraint of end-point preservation, undesirable small branches are also created on coarse object boundaries. An end-point means that the point on the one voxel graph has less than 2 neighbours. Therefore, the topological thinning method requires an additional efficient pruning method [12].

Distance-based methods compute a distance transform from the boundary and find the center-line as the local maximum or ridge. However, the output center-line can be discontinuous and is not guaranteed to be one voxel thick. Moreover, it is quite sensitive to noise on the boundary [13, 14]. The local maximum extraction might produce surface skeletons rather than center-lines. So, there might be a planar set of voxels in the middle of the object that comply to the local maximum criterion [15].

The polygon-based methods directly deal with the surface and the output center-line is suitable to topological analysis of arborized structures. It is a useful method to obtain a center-line of a complex arborized network. However, it is a heuristic method which lacks a theoretical framework and guaranteed solution [16].

The center-line preserves the minimal topology of an arborized surface, therefore, as phenotypical descriptors, the features from the center-line can be used to describe the arborization structure. Furthermore, several prominent features that are able to distinguish the various phenotypical conditions are helpful in the construction of a formalized abstraction. To that end we investigate the connection from an arbor-like center-line structures to L-systems [17, 18, 19] in order to reason over the phenotypical effects as a result of different conditions to which subjects are exposed.

2.3. The Lindenmayer system

The Lindenmayer-system, in short L-system, is typically known for its versatile ability to model plants. In 1968 the theoretical biologist Aristide Lindenmayer developed a mathematical notion to describe the development of algae and plants [17]. The typical characteristic of plants is the structure of arborization. From the onset of the description by Lindenmayer, the L-systems have been developed to the comprehensive modelling tool for plant topology as it is used today.

Underlying the modelling is a vigorous mathematics that sets out for a consistent and extensive toolset for the modelling of plant topology. That is, originally, it was set out as a tool for plant

development. However, focusing on the topological nature of the L-systems, we can understand that it can be generalized to arborized structures. Plants and trees are, to that end, a metaphor for the topology of structures that are able to branch out to a complex organization.

L-systems are generated from productions, or rewriting rules. These rules are written in their own grammar. So, an L-system is a system of rewriting. Rewriting is a technique for defining complex objects by successively replacing parts of a simple initial object using a set of rewriting rules or productions [20]. The most known rewriting rules operate on character strings. L-systems are one of them with distinctive characteristics. Specifically, L-systems consist of two trivial elements: an axiom, and a set of productions. The axiom is the starting point of the rewriting process. The set of productions are the derivative rules. In L-systems, the productions are applied on a parallel and simultaneous manner replacing all elements in a given string. A key element of the L-system is that it is extended with geometrical modelling in order to generate graphical models from the productions that are written. With the powerful graphics adaptors that computers are nowadays equipped with, this results in wonderful and realistic renditions of plant topologies generated from productions.

L-systems have already been applied successfully in other arborizations such as neural tissue [21, 22]. The computational construction of neurons using L-systems provide an efficient and reliable method to investigate the relationships between neuro-morphology and neurophysiology [23, 24].

From the perspective of modelling, the L-system provides a manner to elucidate the mathematical foundation of the topology of a particular organism or structure.

We will use the L-system approach as an inspiration to accomplish the description of a topology in a manner of reverse engineering. From real data we derive the specification of the productions, i.e. the rewriting rules, and use that to generate a generalization of a particular arborized topology. This generalization represents class of structures and portrays an archetype for its topology.

The mammary gland, i.e. the lactiferous duct, has a noticeable branching characteristic. Environmental effects seem to have an influence on the branching. We use L-systems to establish the innate archetype of the lactiferous duct, i.e. the blue print for its shape, and also characterize and simulate the phenotypes that have arisen from exposures to endocrine disruption.

2.4. Experimental setup to study disruption of the murine lactiferous duct

Our case study is on the development of the murine lactiferous duct and in order to study this we build 3D models of the lactiferous duct of newborn mice. Initially this is applied on the wild-type condition and this provides us with the archetypal shape of the lactiferous duct. In newborn mice the mammary gland is present in rudimentary form and defined to be in the stage of secondary sprout, i.e. lactiferous duct or epithelial cords. Unlike human, where the amount of lactiferous ducts could reach to 10-20 for each nipple, in the murine embryo only one epithelial cord grows from the mammary bud. Thus, only one lactiferous duct exists for each nipple. The early initial round of branching growth results in a primary duct with several initial branches [25]. This makes the murine lactiferous duct a suitable model for experimental studies.

The development of the lactiferous duct is under influence of hormones. Unbalanced or unexpected exposure to hormones or hormone-like substances can potentially disturb and unsettle the normal development of the duct. Hormone-like substances in the environment are referred to as

endocrine-disruptors and have known effects on the development of the reproductive organs and the secondary reproductive organs, i.e. the mammary gland.

An experiment is done in order to study the effects of endocrine-disruption on gland development. The exposure of endocrine-disruptors to newborn mice is accomplished through exposure of the mother by including compounds to the food. In this manner the foetus is exposed to the compound and a possible effect on the development of the lactiferous duct is investigated. This is realized by making a 3D model of various lactiferous ducts of the foetus. The distinct tree-like structure of the murine lactiferous duct is depicted in Figure 2.

For our case study we have 4 different conditions: a control group consisting of 7 samples, i.e. the unexposed wild-type (WT) condition ; a condition control group where an inert compound was added (OLIE) as shown in Figure 3(a),(b) consisting of 12 samples; a first treatment group exposed to diethylstilbestrol (DES) consisting of 8 samples; a second treatment group exposed to a cocktail of oestrogen and progesterone (EP) shown in Figure 3(c),(d) consisting of 8 samples. Each sample is a geometrical model that is derived from the imaging of a foetus.

So, we measure a maternal effect in the offspring of exposure to endocrine disruptors and model the lactiferous ducts. An overview of the workflow of the experiment is shown in Figure 1.

3. Methodology for analysis of arborization in 3D graphical models

In order to understand and analyze the structure of the arborization we first have to extract the topology features from the optimized 3D surface model. To that end we further reduce the model to a minimal representation of the topology of a model i.e. the center-line. We have assessed several methods for center-line extraction and have chosen the polygon-based center-line extraction as a vehicle for obtaining a useful center-line. The center-line provides the topology and further facilitates quantitative analysis of the structures of the arborization. We investigate topology related features that can be extracted from the center-line, i.e. number of branches, average length of branches, number of bifurcations, and so forth.

Arborization structure analysis is discussed in many research areas. Recently, methods have been proposed to trace neurons or reconstruct the neuronal structure by tracking arborized line structures in a 3D image volume [26, 27]. The quantitative analysis of the tree is substantially simplified by transforming a voxel-level tree object into a set of interconnected single-voxel center-lines representing individual tree branches [28]. The center-line is also used for the blood vessel quantification and meaningful parameters extraction from 3D vascular images [29].

3.1. Center-line extraction

The polygon-based center-line extraction method can create a connected one voxel thick center-line without undesired small branches. The center-line extraction is fully automatic and does not need any operator interference. This prevents subjective information to enter in the process and assures an objective subsequent phenotype measurement. For the identification of vessel bifurcations and arborized network organization, we incorporated the mesh-based method created by Piccinelli et al.

[9] in our workflow. First, the algorithm identifies the bifurcations by visiting the whole structure of the arborized model with a moving sphere. In this manner the method is able to decompose the structure into bifurcations and single branches. Since the complexity of the algorithm is linear with the number of surface mesh vertices, it is very efficient. There are, however, some limitations. Nested bifurcations with a very short distance between each other can potentially be grouped into a single bifurcation if the radius of the probing sphere on the junction part is greater than the distance between the bifurcations [9]. Center-lines may be positioned on the outer side of the model when the boundary has a particular non-convex shape [9]. Both these conditions can be solved by adding an extra heuristic.

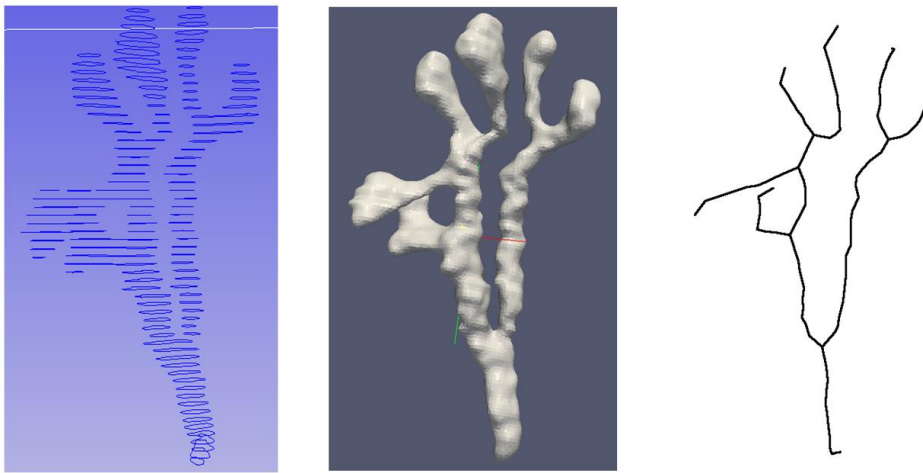


Figure 2. (a) original contour model from section images; (b) surface model; (c) center-line.

We have evaluated the application of center-line extraction method in our workflow to make sure branches and bifurcations are detected right. For a visual inspection of correctness of the center-line we have used ParaView [30] as shown in Figure 2. We compare each branch and bifurcation between original model and center-line so as to make sure they are correctly extracted. We are, in particular, aware of the limitations that are posed by the algorithm [9] and that are bound to occur when bifurcations are too close, in other words, the sampling is insufficient. In an initial evaluation, we focus on number of branches that are wrongly connected and low quality bifurcations. We have randomly selected 16 out of 42 mammary gland models and their automatically extracted center-lines and compute the number of branches and bifurcations per model as well as the wrongly connected branches and low quality bifurcations. We established from this analysis that 97.81% of the branches are detected with the correct topology of the corresponding structure of arborization in mammary gland and 97.65% of the bifurcations are properly positioned in the model. Errors occurred in models that were branched more and where branched from different parts were nearly touching each other. We prevented this to happen in the models and in this manner prevented the errors from occurring.

Our center-line extraction thus results in an accurate representation of the topology of the mammary glands in our experiment and we can use these representations to extract features to characterize the mammary glands.

3.2. Extraction of features from 3D models and center-lines

Each of the models represents a phenotype measurement. So from each 3D model we extract a number of features. The simple features are, the number of branches and bifurcations in a model. Next, for each global center-line we extract individual branches and for each branch we compute branch length, curvature, torsion, tortuosity, minimal radius, maximal radius, mean of the radius and median of the radius. Where radius relates to the radius of the duct.

The curvature of a branch is defined as the average of curvatures of each of the branch-points. The curvature of a point is the inverse of the radius of the local osculating circle. The radius is defined as the radius of maximum inscribed sphere for each point. So, curvature of one point on the center-line is defined as:

$$K = \frac{1}{R} \quad (1)$$

where K is the curvature at point "i" and R is the radius of the local osculating circle.

The torsion of the line $c(s)$ is the degree by which the osculating plane rotates along the line [31], defined as:

$$\tau(s) = \frac{[c'(s) \times c''(s)] \cdot c'''(s)}{|c'(s) \times c(s)|} \quad (2)$$

where *prime* denotes the derivative and *cross* (x) denotes the cross product. The tortuosity is the ratio between the branch length and the distance of the branch end-points, defined as:

$$\chi = \frac{L}{D} - 1 \quad (3)$$

where L is the length of the center-line and D is the Euclidean distance between its end-points.

So, after measurement extraction, for each model we have global features and features per branch that assemble in a feature vector per model. All of our models are labelled with the condition and therefore, we can establish what features are dominant for a treatment. The center-lines and its features are used for the construction of the L-system. This will provide us an abstracted archetype for each of the treatments.

3.3. L-system programming platform

In order to be able to incorporate our measurement results into an abstraction we need a versatile and programmable platform for L-systems. We use the L-Py platform to construct the L-system model for lactiferous duct. This is an L-system simulation framework initially used for modelling plant architecture development based on the Python language [32]. We use dynamic language to enhance the development of lactiferous duct growth models. To that end the turtle geometry is used - this is a tool for the geometric interpretation of a character string as a sequence of segments so that it can draw the information in the string as a geometrical representation in 3D space. During the string reading, the turtle moves on the line with varying segment lengths and angles. The turtle can move in the 3D space based on differential geometry and using quantities such as local tangent, curvature, and step size [32]. The L-Py platform provides seamless control of the differential turtle geometry.

Within the feature set we obtain from our experiments, we look for those features that are prominent in representing phenotypical changes of mammary gland under different environmental conditions. Subsequently, from these features we can generate an archetypal mammary gland as an L-system. This L-system then simulates a particular phenotype under normal circumstances. Next, from the experimental conditions, we will generate an L-system that is an expression of the environmental pressure that is imposed by the exposure to a compound. These conditions are elaborated in section 4.1, and the feature extraction are presented in section 4.2. The generation of the L-system abstraction is elaborated in section 4.3.

4. Analysis of arborization and L-system construction

Our case study is scaled such that we can do sufficient statistical analysis as well as develop the workflow from a 3D graphical model derived from section images to an L-system that represents the arborization of the lactiferous duct. The aim of this study is to come to a formalized abstraction of a phenomenon in biology. The scale of the study is, therefore, chosen such that we can accomplish this. In a larger study, the outcome of the process described here can be used to obtain further insights in the mechanical and molecular-genetic aspects of biology.

4.1. The different groups in the case study and measurement preparation

The section images of the lactiferous duct are acquired and a stack of aligned images is used as input for the 3D model [33, 34]. For each stack the lactiferous duct structures are delineated by a specialist resulting in initial 3D models. For all models, interpolation and surface reconstruction is applied. Subsequently, the center-lines of all branches are extracted. To assess the topological shape differences in the models the center-line of the lactiferous duct structure are used in the analysis. The distinct tree-like structure of the mouse lactiferous duct, and how it is obtained from the workflow, is depicted in Figure 1.

The collection of 3D models comprises 4 different experimental conditions, 3 of which are exposures to compounds that might have an effect of the development of the lactiferous duct. The different groups are (1) the WT control group, (2) OLIE control group, (3) the EP group and (4) the DES group. For this case study the same compound dose is administered for groups 2,3 and 4; i.e 10 microgram. In Figure 4a,b and example of a model in the OLIE group is depicted and in Figure 4b,c an example of a model in the EP group is depicted.

The exposures entail that a mouse embryo is exposed early in development to different environmental factors that can be recognized as hormone; a.k.a. endocrine disruptors. These compounds influence development of reproductive organs in primary and secondary stages by altering the structure of the duct.

Our hypothesis is that we expect to observe an effect in the EP and the DES group with respect to the development of the duct. Therefore the models from the group EP and the group DES should be significantly different from the WT group. The OLIE group is a control group that is exposed to an inert compound. Consequently, there should be no effect in the OLIE group and the results should be

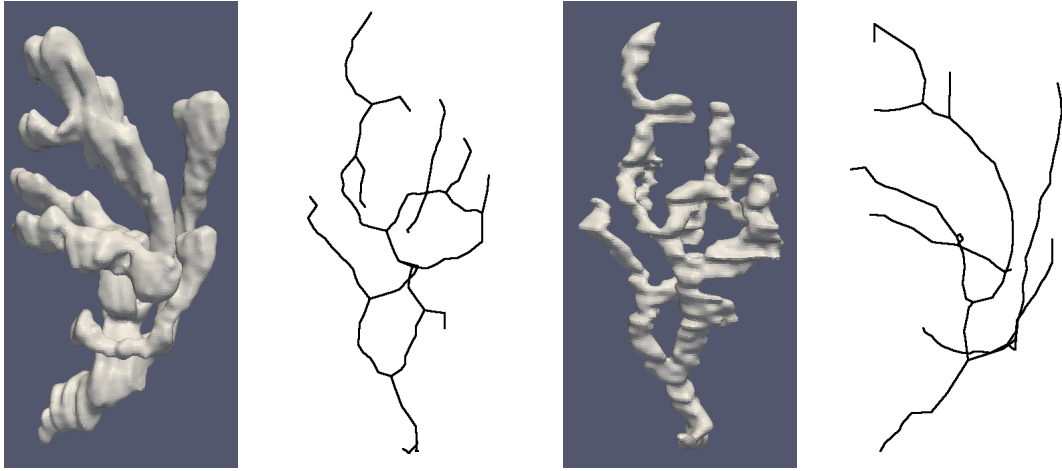


Figure 3. Different branch structures of lactiferous duct (a) surface model from group OLIE; (b) center-line of the model from group OLIE; (c) surface model from group EP; (d) center-line of the model from group EP.

similar to the WT group. The WT and OLIE group are considered as the ground truth group. The EP and the DES group should thus be significantly different from the ground truth groups.

4.2. Feature extraction and characteristics of the groups

The topological structure of the duct is derived from a center-line extraction method, cf. section 3.1 [9]. The features derived from center-lines are the basis for the L-system representation. Our primary question is whether we can accurately separate complex 3D structure on the basis of the selected morphology features. The data set is a collection of all branches of a group with their measurements. In order to check for significant differences between measurements and groups (cf. section 4.1) we use the KS-test. The KS-test is selected as we can not assume that the measurements are normally distributed - some of them are clearly not.

From the nature of the experimental setup we know that the WT and the OLIE group must be very similar, this will therefore be subject to a test. We do not know how similar the EP and the DES groups are; so this will be subject to a test as well. Finally, as given that WT and OLIE are similar and that EP and DES will resort in an effect in the development of the duct, we test these groups in pairs. That is, WT and OLIE (OW) are considered the ground truth and the DES with EP (DE) are the disruptor group. A KS-test is applied on the different measurements of these three selected combinations. If we set a null-hypothesis that there is no effect in the DE group with respect to the ground truth, we expect the null-hypothesis to be rejected.

The KS-test requires the significance value, i.e. p -value, to be set. For our case study the significance level was set to $p = 0.01$. The outcome of the KS-test is that if the tested value is significantly larger then value h is 1. Otherwise, if the tested value, the data set, is significantly smaller than the reference set, then h is -1. If there is no significant difference between the two datasets then value h is 0.

For the groups that are evaluated the results are shown in Table 1. We are able to conclude from these results that group OLIE and WT are exactly the same. Moreover, group DES and group EP are almost the same except for curvature. We can understand this result by inspecting the center-lines of the models respectively. We observe that the group EP clearly has more curved branches compared to group DES. Nevertheless, the combined group DES with EP and group OLIE with WT are significantly different in length, curvature, maximal radius, mean of the radius and median of the radius. These features are the prominent ones over torsion, tortuosity and minimal radius. The results support the hypothesis and also confirm the feasibility of our 3D analysis for this complex system.

Table 1. Result of the comparison of the data from the DES/EP samples, the OLIE/WT samples and the DES+EP/OLIE+WT samples.

		length	curvature	torsion	tortuosity	min-radius	max-radius	mean-radius	median-radius
DES-EP	h	0	-1	0	0	0	0	0	0
	p	0.2027	0.0098	0.5588	0.8170	0.7951	0.6364	0.8275	0.9412
OLIE-WT	h	0	0	0	0	0	0	0	0
	p	0.2564	0.8537	0.7724	0.9246	0.0416	0.2951	0.1581	0.3943
DE-OW	h	-1	1	0	0	0	-1	-1	-1
	p	0.0043	0.0031	0.2714	0.5649	0.1007	4.99E-5	0.0026	0.0040

Table 2. Input measurements for L-system of the lactiferous duct; for illustration, the WT and the DES samples are chosen.

Type	MBL	σ_m	MBC	σ_m	MBR	σ_m	MBN	σ_m
WT	303.1	34.19	0.0096	0.0004	28.45	1.37	6.17	0.75
DES	193.1	12.85	0.0101	0.0003	27.90	0.74	14.625	3.19

MBL: Mean of branch length in micrometer; MBC: Mean of branch curvature;

MBR: Mean of branch radius in micrometer; MBN: Mean of branches number;

σ_m : standard deviation of mean.

4.3. L-system modeling

The abstraction of the lactiferous duct to an L-system requires to have the right *rewriting rules*. We first focus on the realization of the elongation and the bifurcations of the duct. We thus integrate the available measurements in the system, i.e. branch length, curvature, radius and the number of branches. We consider these measurements more prominent as is shown from the statistical analysis of the variation in morphology of the experimental groups (cf. section 4.2).

The integration of branch length, curvature and the radius of maximal inscribed sphere is relatively straightforward. A change in the branching pattern in the L-system, i.e. the number of branches, is realized by changing the probability of bifurcation construction.

Algorithm 1: L-system generation from measured features

```

Initialize maximum number of branch order to 4
Initialize increment of radius through time to 0.0001
Initialize maximum duration of a branch relating to the branch order
Initialize branch angle with a random number from 15 to 25 weighted with the branch order
Initialize branch angle to a random number from 3 to 7
Initialize total number of iteration of the system
Set time to zero
Set branch order to zero
module A represents trunk apical meristem
module B represents apical meristem of lateral branches
module I Internode
Start recursion for production, using the native method for this in L-py.
production:
A(t,o):
if time t is shorter than maximum duration of a branch (o) then
|   Continue growing with a set branch angle and length: IA(t+1,o)
end
else
|   Create bifurcation with a set roll angle around previous branch and a set angle in between
|   the bifurcation: IB(0,o+1)
end
B(t,o) if time t is shorter than maximum duration of a branch (o) then
|   produce IB(t+1,o)
end
else
|   create bifurcation: B(0, o+1)
end
Decrement radius of the branch

```

This approach can be summarized in Algorithm 1, where we introduce the following variables for generating a model from our parameters. The branch order is retrieved from our observations in the models, we set it just 1 step higher than the maximum order that we have retrieved from the models, i.e. 4. The radius increment is set to a value that is hardly visible in the L-system model as, in the visual sense, we strive for an abstraction model. We allow maximum duration of the branch order so as to enable the model to bifurcate "fast". From the models angular bifurcation is measured between 10 and 30 degrees, therefore the value is randomized within that range and the result is weighted. The bending of the lateral branches is processed in similar fashion. The iteration time is used to constrain the complete development of the bifurcating model.

From the measurements we have established what measurements differentiate the condition from the ground truth, i.e. wild type. These measurements are used to construct an L-system. To this end we set the rewriting rules to model the L-system for each specific condition.

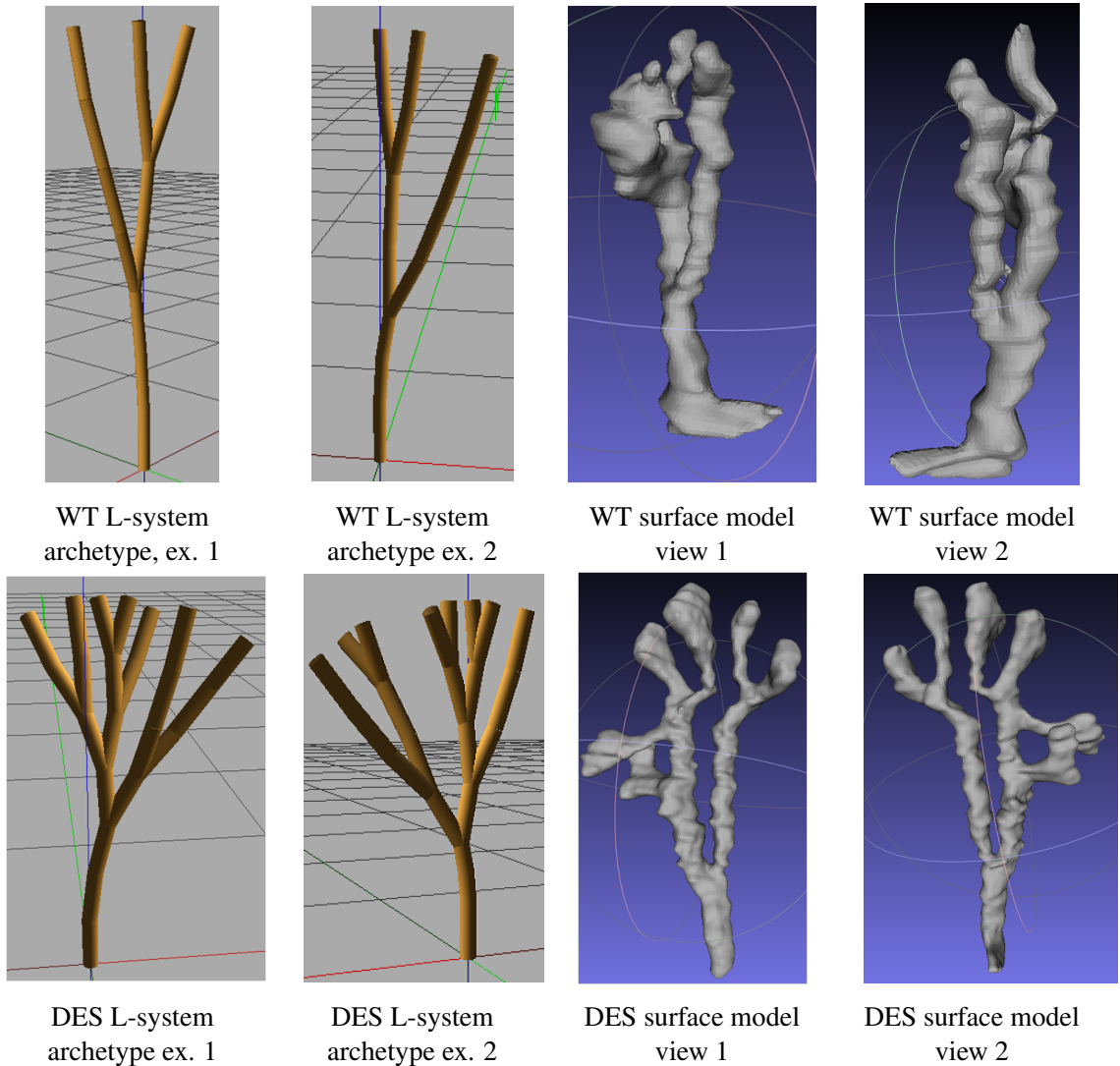


Figure 4. Two examples of L-sytem archetypes and real surface models that they were derived from. It can be appreciated that the WT L-system archetype is really similar for different 3D models. For the surface model we present two views of the same 3D model.

The L-system model mainly focuses on the abstraction of elongation and bifurcation of the center-line of the lactiferous duct. We illustrate the L-system abstraction for two different kinds of lactiferous ducts: i.e. the wild type and the DES exposed type. These can be considered extreme cases. In Table 2 the measurements used for these L-system abstractions are listed.

According to the hypothesis, compared to wild type (WT), the lactiferous duct with the DES condition would be significantly different. Here the effect is clearly seen as the DES-model has more branches with shorter branch length. In Figure 4, the results of the rewrite rules and L-system visualization in an abstracted model of the lactiferous duct are depicted.

WT III [II] [III] [IIII]

DES III [I] [III] [III]] [I [I [II] [III] [I [II] [II] [II]]]

The differences between the two example L-systems can be appreciated by visual inspection of Figure 4. However, one can also assess the production string that we can deduce from the production of the L-system. In Figure 5 a schematic rendition of the production string for the two cases of Figure 4 is depicted. The increasing complexity and branching can be read from this string; the color coding is added to interpret the iterations that have passed in the production of the system.

The formal abstraction to the L-system enables to reason over these topologies. The L-system of the normal development, i.e. the wildtype, embodies a blue print of the innate program of the forma-

tion of the lactiferous duct, we can add more time points to further understand this particular program. At the same time, the phenotype measurement is formalized in an L-system and the productions from the innate program to a specific phenotype can be inferred through the L-system. The phenotype from some environmental condition can be considered as a sequence of alterations from the innate program of duct formation. This augments the toolbox for analysis a great deal and will further contribute to the understanding of branching patterns under exposure of compounds in the environment. In this manner it has great potential in the contribution of research to the development of the lactiferous duct system as well as studies on effects of potential pollutants.

The work presented in this paper takes a phenomenological approach to the analysis of the murine lactiferous duct. It constructs the abstractions, i.e. L-system, just from measurements taken from samples exposed to different conditions. The condition, i.e. the disruptor, will alter the molecular genetics of the animal and thus influence the formation of the lactiferous duct. A lot of Gene Networks, involving different growth factors, participate in the formation of the lactiferous duct [35, 36]. The alterations as we observe them, are engaged by the interference to these Gene Networks. In fact, these networks are instrumental in making the alterations from the blue print shape as captured the the L-system of the lactiferous duct.

The construction of the abstraction in the L-system from measurement is retrospective. The rewriting rules, however, can be used in *what-if* scenarios in a prospective as well as predictive manner. The complex ontogeny [35] of the lactiferous duct is now abstracted in an L-system and this can bring parameters from molecular genetics and tissue mechanics into play so as to further elaborate on these models. This will enable us to really make inferences from the model with molecular parameters. One can make observations from dose effects as well as relate the gene expression to changes in the duct topology.

This comprehensive analysis of induced phenotypes, from model to abstraction helps us in bringing new inspiration to the manipulation of the L-system which together with the features provide an important quantifiable twist to this system. Our primary L-system model illustrates the potential for the simulation of mammary gland formation. The L-system can provide a ground truth model from real data. In the future we will direct our research further to this delicate interplay of nature inspired systems and nature driven models and combine empirical computing with fundamental abstractions in computing.

6. Acknowledgements

The authors wish to express their gratitude to Dr Paul van der Saag, in dear memory, who inspired the research on the endocrine disruption and to Dr Fienke Lemmen for her contribution of material for the case study. With the first author, both were affiliated with the Hubrecht Institute in the Netherlands.

References

- [1] Cao L, Verbeek FJ. Nature Inspired Phenotype Analysis with 3D Model Representation Optimization. In: Abraham A, Krömer P, Snášel V (eds.), *Innovations in Bio-inspired Computing and Applications*. Springer International Publishing, Cham. 2014 pp. 165–174. ISBN:978-3-319-01781-5.

- [2] Verbeek F, Boon P. High resolution 3D-reconstruction from serial sections - Microscope instrumentation, software design and its implementations. *Proceedings of SPIE - Three Dimensional and Multi-Dimensional Microscopy. Image Acquisition and Processing IX.*, 2002. **4621**:65–76. doi:10.1117/12.467833.
- [3] Verbeek FJ, de Groot MM, Huijsmans DP, Lamers WH, Young IT. 3D base: A geometrical data base system for the analysis and visualisation of 3D-shapes obtained from parallel serial sections including three different geometrical representations. *Computerized Medical Imaging and Graphics*, 1993. **17**(3):151–163. Confocal Microscopy.
- [4] Borgefors G. Distance transformations in digital images. *Computer Vision, Graphics, and Image Processing*, 1986. **34**(3):344 – 371. doi:[https://doi.org/10.1016/S0734-189X\(86\)80047-0](https://doi.org/10.1016/S0734-189X(86)80047-0).
- [5] Fritsch FN, Carlson RE. Monotone Piecewise Cubic Interpolation. *SIAM Journal on Numerical Analysis*, 1980. **17**(2):238–246. URL <http://scitation.aip.org/getabs/servlet/GetabsServlet?prog=normal&id=SJNAAM000017000002000238000001&idtype=cvips&gifs=yes>.
- [6] Kazhdan M, Bolitho M, Hoppe H. Poisson surface reconstruction. In: *Proceedings of the fourth Eurographics symposium on Geometry processing, SGP '06*. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland. ISBN:3-905673-36-3, 2006 pp. 61–70. URL <http://portal.acm.org/citation.cfm?id=1281957.1281965>.
- [7] Akgül CB, Sankur B, Yemez Y, Schmitt F. 3D Model Retrieval Using Probability Density-Based Shape Descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2009. **31**(6):1117–1133. doi:10.1109/TPAMI.2009.25.
- [8] Dey TK, Sun J. Defining and Computing Curve-skeletons with Medial Geodesic Function. In: *Proceedings of the Fourth Eurographics Symposium on Geometry Processing, SGP '06*. Eurographics Association, Aire-la-Ville, Switzerland, Switzerland. ISBN:3-905673-36-3, 2006 pp. 143–152. URL <http://dl.acm.org/citation.cfm?id=1281957.1281975>.
- [9] Piccinelli M, Veneziani A, Steinman DA, Remuzzi A, Antiga L. A Framework for Geometric Analysis of Vascular Structures: Application to Cerebral Aneurysms. *IEEE Trans. Med. Imaging*, 2009. **28**(8):1141–1155. doi:<http://dx.doi.org/10.1109/TMI.2009.2021652>.
- [10] Lee TC, Kashyap RL, Chu CN. Building Skeleton Models via 3-D Medial Surface Axis Thinning Algorithms. *CVGIP: Graphical Models and Image Processing*, 1994. **56**(6):462–478. doi:10.1006/cgip.1994.1042.
- [11] Xie WJ, Thompson RP, Perucchio R. A topology-preserving parallel 3D thinning algorithm for extracting the curve skeleton. *Pattern Recognition*, 2003. **36**(7):1529–1544. doi:10.1016/S0031-3203(02)00348-5.
- [12] Nmeth G, Kardos P, Palgyi K. Topology Preserving 3D Thinning Algorithms Using Four and Eight Subfields. 2010. **6111**:316–325. doi:10.1007/978-3-642-13772-3_32.
- [13] Cardenes R, Bogunovic H, Frangi AF. Fast 3D centerline computation for tubular structures by front collapsing and fast marching. In: *Proceedings of the International Conference on Image Processing, ICIP 2010*, September 26-29, Hong Kong, China. IEEE. ISBN: 978-1-4244-7994-8, 2010 pp. 4109–4112. doi:<http://dx.doi.org/10.1109/ICIP.2010.5651609>.
- [14] Van Uitert R, Bitter I. Subvoxel precise skeletons of volumetric data based on fast marching methods. *Medical Physics*, 2007. **34**(2):627–638. doi:10.1118/1.2409238.
- [15] Telea A, Vilanova A. A robust level-set algorithm for centerline extraction. In: *Joint EUROGRAPHICS-IEEE TCVG Symposium on visualization*. 2003.

- [16] Verroust A, Lazarus F. Extracting skeletal curves from 3D scattered data. *The Visual Computer*, 2000. **16**(1):15–25. doi:10.1007/pl00007210.
- [17] Lindenmayer A. Mathematical models for cellular interactions in development I. Filaments with one-sided inputs. *Journal of Theoretical Biology*, 1968. **18**(3):280 – 299.
- [18] Rozenberg G, Salomaa A. Mathematical Theory of L Systems. Academic Press, Inc., Orlando, FL, USA, 1980. ISBN:0125971400.
- [19] Rozenberg G, Salomaa A (eds.). The book of L. Springer-Verlag New York, Inc., New York, NY, USA, 1986. ISBN:0-387-16022-1.
- [20] Prusinkiewicz P, Lindenmayer A. The Algorithmic Beauty of Plants. Springer-Verlag New York, Inc., New York, NY, USA, 1990. ISBN:0-387-97297-8.
- [21] Jelinek H, Karperien A, Cornforth D, Cesar Junior RM, G LJJ. MicroMod - an L-systems approach to neuron modelling. In: National University. 2002 pp. 156–163.
- [22] Ascoli GA, Krichmar JL. L-neuron: A modeling tool for the efficient generation and parsimonious description of dendritic morphology. *Neurocomputing*, 2000. **3233**(0):1003–1011. doi:10.1016/S0925-2312(00)00272-1.
- [23] Ascoli GA. Progress and perspectives in computational neuroanatomy. *The Anatomical Record*, 1999. **257**(6):195–207. doi:10.1002/(SICI)1097-0185(19991215)257:6<195::AID-AR53.0.CO;2-H.
- [24] Jelinek HF, Cesar RM, Leandro JJG, Spence I. Automated morphometric analysis of the cat retinal alpha/Y, beta/X and delta ganglion cells using wavelet statistical moment and clustering algorithms. *J Integr Neurosci*, 2004. **3**(4):415–32. URL <http://www.biomedsearch.com/nih/Automated-morphometric-analysis-cat-retinal/15657977.html>.
- [25] Cowin P, Wysolmerski J. Molecular mechanisms guiding embryonic mammary gland development. *Cold Spring Harb Perspect Biol*, 2010. **2**(6):a003251. URL <http://www.biomedsearch.com/nih/Molecular-Mechanisms-Guiding-Embryonic-Mammary/20484386.html>.
- [26] Lee PC, Chuang CC, Chiang AS, Ching YT. High-throughput Computer Method for 3D Neuronal Structure Reconstruction from the Image Stack of the Drosophila Brain and Its Applications. *PLoS Computational Biology*, 2012. **8**(9). doi:10.1371/journal.pcbi.1002658.
- [27] Zhao T, Xie J, Amat F, Clack N, Ahammad P, Peng HC, Long FH, Myers E. Automated Reconstruction of Neuronal Morphology Based on Local Geometrical and Global Structural Models. *Neuroinformatics*, 2011. **9**(2-3):247–261. doi:10.1007/s12021-011-9120-3.
- [28] Kalman P, Juerg T, Eric AH, Milan S. Quantitative analysis of pulmonary airway tree structures. *Computers in Biology and Medicine*, 2006. **36**(9):974–996. doi:10.1016/j.compbiomed.2005.05.004.
- [29] G KD, C SD, B RJ. Three-Dimensional Blood Vessel Quantification via Centerline Deformation. *Medical Imaging, IEEE Transactions on*, 2009. **28**(3):405–414. doi:10.1109/TMI.2008.2004651.
- [30] Henderson A. The {ParaView} Guide: A Parallel Visualization Application. Kitware, 2004. ISBN-10:1930934149, 13:978-1930934146.
- [31] Pressley A. Elementary differential geometry. Springer, London; New York, 2010. ISBN:9781848828902 184882890X.
- [32] Boudon F, Pradal C, Cokelaer T, Prusinkiewicz P, Godin C. L-py: an L-system simulation framework for modeling plant architecture development based on a dynamic language. *Front Plant Sci*, 2012. **3**:76. URL <http://www.biomedsearch.com/nih/L-py-system-simulation-framework/22670147.html>.

- [33] Verbeek FJ, Huijsmans DP. A Graphical Database for 3D Reconstruction Supporting (4) Different Geometrical Representations. 1998. **465**:117–144. doi:10.1007/978-1-4615-5553-7_5.
- [34] Verbeek FJ, Huijsmans DP, Baeten RJAM, Schoutsen NJC, Lamers WH. Design and implementation of a database and program for 3D reconstruction from serial sections: A data-driven approach. *Microscopy Research and Technique*, 1995. **30**(6):496–512. doi:10.1002/jemt.1070300607.
- [35] Propper AY, Howard BA, Veltmaat JM. Prenatal morphogenesis of mammary glands in mouse and rabbit. *J Mammary Gland Biol Neoplasia*, 2013. **18**(2):93–104. doi:10.1007/s10911-013-9298-0.
- [36] Rappoport JZ, Simon SM. Endocytic trafficking of activated EGFR is AP-2 dependent and occurs through preformed clathrin spots. *Journal of Cell Science*, 2009. **122**(9):1301–305.
- [37] Cao L, Verbeek FJ. Analytical evaluation of algorithms for point cloud surface reconstruction using shape features. *Journal of Electronic Imaging*, 2013. **22**(4):043008–043008. doi:10.1117/1.JEI.22.4.043008.