Shape analysis for phenotype characterisation from high-throughput imaging

Guo, Y.; Guo Y.

Citation

Guo, Y. (2017, October 17). *Shape analysis for phenotype characterisation from high-throughput imaging*. *SIKS Dissertation Series*. Retrieved from https://hdl.handle.net/1887/56254

Author: Guo Yuanhao
Title: Shape analysis for phenotype characterisation from high-throughput imaging
Date: 2017-10-17

Shape Analysis for Phenotype Characterisation from High-throughput Imaging

Yuanhao Guo

ISBN 978-94-6299-734-9
Printed by RIDDERPRINT

Cover painting and design: Yang Lili, Zhou Jinglong
Cover painting: Human beings and nature are in harmony. In humility, we are grateful to make use of the materials (zebrafish, butterfly, orchid, wood, etc.) that nature offers for our research. Indeed, the output of this research improves our living conditions.

Universiteit Leiden

Shape Analysis for Phenotype Characterisation from High-throughput Imaging

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 17 oktober 2017
klokke 15.00 uur

door

Yuanhao Guo

geboren te Jinan, China
in 1986

Promotiecommissie

Promoteres:	prof. dr. ir. F.J. Verbeek	
	prof. dr. H.J. van den Herik	
Overige leden:	prof. dr. T. Bäck	
	prof. dr. T. Gevers	(University of Amsterdam)
	prof. dr. H.P. Spaink	
	prof. dr. H. Trautmann	(University of Münster)

Yuanhao Guo was financially supported through the China Scholarship Council (CSC) to participate in the PhD programme of Leiden University. Grant number 201206220108.

The research reported in this thesis was partially funded by the Netherlands Organisation for Scientific Research (NWO), under grant number 834.14.001.

SIKS Dissertation Series No. 2017-36.
The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

Contents

Chapter 1

Introduction

Computational approaches are important to characterise phenotypes in modern life-science research, in developmental biology, (patho)physiology, toxicology, pharmacology, etc.. According to the definition, the phenotype is recognised as an observable trait in the whole appearance of an organism from molecular to organism scale [1]. The study of gene expression can be used to visualise the phenotype. The process of gene expression operates on the level of synthesis and structuring of proteins which subsequently contribute to the organism's appearance in any form. In order to obtain an understanding of a phenotype, imaging is used to explore the phenotype characteristics. Thus, a readout of the phenotype by means of images is accomplished. Consequently, the phenotype characteristics need to be extracted from images and this can be realised through image analysis. Therefore, computational approaches for image analysis have to be developed. Recently, some approaches have been developed for this purpose (see [2, 3, 4]). In this thesis, we aim at developing dedicated computational approaches which may achieve an *efficient* and *robust* performance in generalising the whole description for phenotype characterisation.

1.1 Importance of shape

In this thesis, a *shape* in an image is formally defined as the quality of an object that depends on the relative position of all points on its surface (adapted from [5]). With respect to the whole phenotype of an organism, shape appears to be directly relevant to the physical development of that organism (or the organ) under investigation. For example, some shapes can directly reflect apparent variations of an object. The shape variations are often caused by the exposure of the samples in a compound for screening or by the timing control for wild type individuals.

However, in some circumstances, the shape variation is subtle, so that the variation is difficult to observe and analyse. In light microscopy, for example, it is difficult to compare the whole-mount of a sample with another from the same model organism in the same developmental stage. Yet, subtle shape variations can play an important role in toxicology, since the size of the organism reflects its response to certain drugs. Indeed, the shape variations in some experimental settings can be observed, but empirical interventions may introduce and propagate subjective errors. For instance, interactive annotation of key points in images is often used to analyse delicate structures such as the skeleton of an organism [6], but different annotators may have their own assessments for the objectivity.

As a result, accurate shape analysis of an organism will lead to reliable characterisation of phenotypes. To this end, we first need a good shape representation to describe the shape of an object precisely. In order to validate the performance of such shape representation, the scalar primitives for a shape, e.g., (a) perimeter and area in 2D shape representation, and (b) volume and surface area in 3D shape representation, can be used. With a validated shape representation, the geometrical primitives for the shape can be further involved for more sophisticated shape analysis. In practice, the simple scalar primitives for the shape are very important as they provide us with intuitive, stable, and accurate 2D/3D measurements, for a delicate exploration of subtle shape variations in the phenotype characterisation. Therefore, we are motivated, in this thesis, to develop new computational approaches based on images (1) to promote precise shape description and (2) to make reliable and accurate 2D/3D measurements possible.

1.2 High-throughput imaging

We should be aware that an unbiased shape analysis is available in a population of the model. In practice, this is conditioned by the choice of the sampling size for the population. As a result, a sufficiently large sampling size may reflect the general properties of the samples, which will result in an accurate statistical assessment. In order to obtain adequate sampling size, high-throughput (HT) screening was initiated for the applications of cytomics and toxicology, and has been applied on the application on organism scale such as zebrafish [7]. The HT screening facilitated the fast development of high-throughput imaging (HTI). A feasible HTI architecture can easily acquire a sufficiently large volume of data represented by images of the subject under study. This HTI architecture has the following advantages: (1) we can use bright-field microscopy producing the images, representing an overall shape for a specimen; (2) fluorescence microscopy can be used to produce the images presenting the fluorescently marked components such as detailed inner structures of a specimen; (3) synchronization of bright and fluorescence microscopy results in the so-called multi-modal microscopy producing fused multi-modal images which can be used to represent and evaluate the comparison between different modes; (4) we can potentially obtain shape analysis at a very high-resolution with the help of a better quality objective lens, in which structures such as tissues and cell type can be reflected from micro-scale texture in images.

1.3 Model organism

In modern life-science, a feasible and convenient model organism also plays an important role since many human diseases can be cultivated in a model organism. A good understanding of the model organism can obtain insights into the disease or treatment, and the obtained knowledge can be transferred to research on humans. In practice, we have many options for the model system. Invertebrate models such as fruit fly [8] and c. elegans [9] are intensively used in molecular genetics. As a comparison, vertebrate models like mouse are suitable and commonly used in the research of human diseases [10]. However, the growth of mouse is slow and it is difficult to get access to a large sampling size. Alternatively, in the last decade, zebrafish has been increasingly used for human disease studies as they present many remarkable characteristics [11], among which the most significant one is its 70% genome equivalence to human [12]. The development of zebrafish is pretty fast as its organs develop within 36 hours post fertilisation (hdf) [13]. The zebrafish are fertile, and one adult couple can easily produce 300 eggs per week [14]. In early larval stages, zebrafish are quite small ($< 1mm$) and optically transparent thus the whole body of a zebrafish as well as partial inner organs are observable using microscopy [15]. In particular, with the availability of many transgenic lines, the zebrafish is genetically modified with fluorescent markers like green fluorescent protein (GFP) [14, 16]. Fluorescence microscopy can be employed to visualise the fluorescently marked structures within a specimen such as organs and infectious diseases. Taking all the properties into consideration, zebrafish is very suitable to be used in high-throughput applications. Therefore, in this thesis, we will consistently use the zebrafish larvae as our model system to illustrate and validate our approaches.

1.4 Problem statement and research questions

In section 1.1, we have discussed the importance of shape analysis for phenotype characterisation. The HTI architecture will serve as the basis for the production

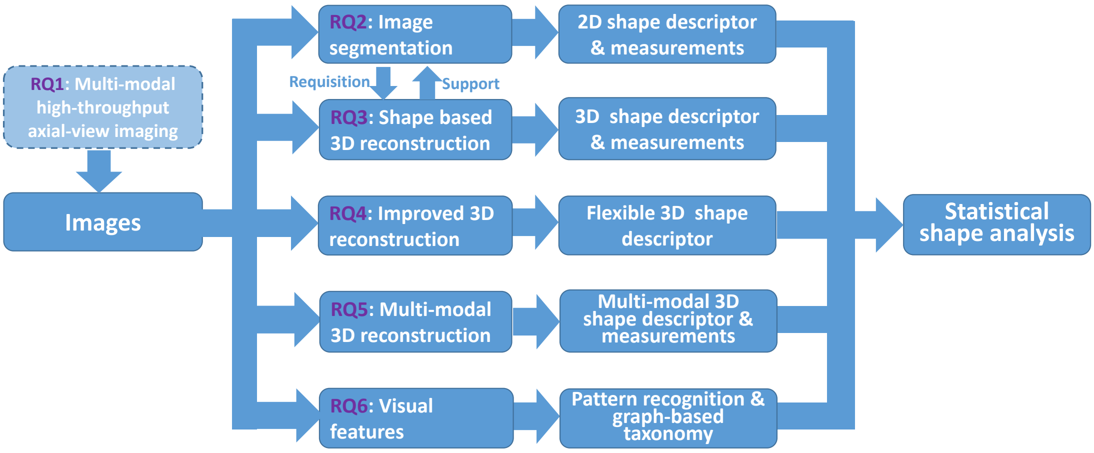


Figure 1.1: A schematic representation of a unified system for shape analysis in support of phenotype characterisation using an MM-HTAI architecture. The six RQs formulated in the text are indicated in the boxes.

of adequate data. We expect to obtain an *efficient* and *robust* shape analysis to characterise the phenotype from HTI. This idea has inspired the formulation of our problem statement (PS).

> **PS**: *To what extent can we develop a stable HTI architecture and produce a robust and accurate shape analysis for the phenotype characterisation from the HTI architecture?*

To address the PS, we are motivated to design a unified system integrating multiple functional modules. They should correspond to new computational approaches for shape analysis in support of the phenotype characterisation from HTI. We start remarking that the system needs to deal with the problem of a precise shape representation. Then a validation of the shape representation is required. This will imply a delicate shape analysis using geometrical primitives. The validation can be realised through scalar primitives for the shape in the form of 2D/3D measurements. To reach our goal, we specify and investigate six research questions (RQs). In Fig. 1.1, we provide a schematic representation corresponding to the six RQs.

As stated above, we need to develop a feasible HTI architecture for image acquisition of the zebrafish larvae. In practice, the zebrafish larvae are always positioned along their longitudinal axis when they are in an imaging modality. In this manner they are easy to manipulate. Moreover, most of its features are then observable and can be visualised. This gives rise to the high-throughput axial-view imaging (HTAI) architecture. Since the HTI can adopt multi-modal microscopy, we can obtain a multi-modal high-throughput axial-view imaging (MM-HTAI) architecture. Our first RQ is thus formulated as follows.

> **RQ 1**: *To what extent is it possible to develop an MM-HTAI architecture for the zebrafish larvae*?

From the axial-view images of the zebrafish acquired by the MM-HTAI architecture, one should first obtain a well-defined 2D shape descriptor and then derive accurate 2D measurements such as the perimeter and the area of an object. With respect to zebrafish, we can see the following properties. On the one hand, an accurate description (both for a whole-shape and its organs on microscopic level) will ensure a reliable quantitative assessment in the applications, such as drug targeting. On the other hand, we have to realise that the part of zebrafish tail is particularly relevant to some diseases since the hemopoietic stem cells in the zebrafish are found predominantly in the tail. An detailed description for every section of the zebrafish such as its tail will ensure accurate localisation, tracking and evaluation of some infectious diseases, such as the spread and development of cancer cells [17, 18]. However, in applications using the zebrafish as model system, the partial transparency across the whole specimen and the weakly defined boundaries which are mainly distributed around the tail in early larval stages, are commonly existing and observable. Hence, we formulate our second RQ as follows.

> **RQ 2**: *To what extent is it possible to obtain an accurate 2D shape description for the zebrafish from the MM-HTAI architecture*?

With respect to shape analysis, a 3D description is more reliable due to the 3D nature of an object's shape than a 2D descriptor. The confocal laser scanning microscope (CLSM) can obtain 3D imaging for a fluorescently labelled structure in an organism. 3D reconstruction has been produced from the images acquired by CLSM using the TDR-3Dbase software [19, 20]. However, there are two obstacles. (1) It is difficult to depict an overall shape of the whole organism for the CLSM, and (2) the CLSM has a low efficiency of image acquisition. Recently, an attempt for 3D imaging has been reported as optical projection tomography (OPT) [21, 22], whilst rather dense scanning is required and extra processing like 3D image segmentation should be employed for further image analysis. In OPT, the sample preparation is also rather time-consuming. To the best of our knowledge, there are few systematic assessments of 3D measurements, e.g., the volume and surface area, in real metrics for the zebrafish in phenotype research. Actually, these 3D measurements are essential in many applications. For example, statistical representations of the 3D measurements for the whole-mount of zebrafish in various developmental stages will give insights into the accuracy of shape analysis which enables HT compound screening. In fact, the axial-view images acquired by the MM-HTAI provide sufficient information for a good 3D description which is a prerequisite for reliable estimation of 3D measurements. This observation results in our third RQ.

> **RQ 3**: *To what extent is it possible to obtain a precise 3D shape description and derive accurate 3D measurements that are statistically relevant for the zebrafish from the MM-HTAI architecture*?

In our study, translucency and transparency often occur in light microscopy. Admittedly, in some cases, the boundaries of an organism are weakly defined. Yet, the qualities will still present a good 3D description for the whole-mount of the zebrafish, even without accurate 2D shape descriptions. Hence we formulate our fourth RQ as follows.

> **RQ 4**: *How can we efficiently deal with the translucency and transparency of specimen in light microscopy and still obtain a good 3D shape description from the MM-HTAI architecture*?

In life-science research such as toxicology, quantitative endpoints like organ size or growth retardation play significant roles. This requires an accurate 3D shape description and rather precise measurements on organ scale, such as the evaluation of organ susceptibility of toxicology in the zebrafish larvae [23]. The MM-HTAI is capable of producing multi-modal images including (1) bright-field images presenting the overall shape of the zebrafish and (2) fluorescence images presenting the detailed inner structure like zebrafish liver. So, we formulate our fifth RQ as follows.

> **RQ 5**: *How can we obtain a multi-modal 3D description and the corresponding measurements for the zebrafish from the MM-HTAI architecture*?

If we go to a higher resolution scale, i.e., on the cellular or tissue level, textures such as detailed fibrous structures in the specimens can contribute to even better shape analysis. In this case we will represent the shapes as well-defined features according to geometrical and textural information extracted from an image [24] or even representative features such as the convolutional neural networks (CNN) [25]. Hence we formulate RQ 6 as follows.

> **RQ 6**: *To what extent is it possible that the classification models (or regression models) are able to validate the performance of the image features to characterise the phenotypes in support of shape analysis*?

1.5 Research methodology

The research methodology in this thesis consists of (1) literature study and analysis, (2) development and implementation of new computational approaches, (3) performance validation and evaluation for the approaches. The literature study is realised by reading and investigation; the analysis by a comparison with the development of state-of-the-art. The new computational approaches are inspired by ideas from other well-developed research fields, such as computer vision and machine learning. The performance validation is achieved by applying the methodologies on datasets and comparing the results with state-of-the-art. We elaborate the methodologies in the analysis of the six RQs. We do so as follows.

In RQ 1, we employ the Vertebrate Automated Screening Technology (VAST BioImager) [26] and light microscopy to develop the MM-HTAI architecture. The VAST BioImager is used to manipulate the input. The positioning module of the VAST BioImager consists of a delicate capillary which is held by a pair of stepper motors. The stepper motors can manipulate the positioning module to revolve for 360 degrees. A pumping action system is loading a zebrafish larva into the positioning module; a mounted camera which we refer to as VAST camera is used to detect and localise the object and then manage the system to position the specimen in the view of an observer from an arbitrary axial-view. The observer can be either the VAST camera or a microscope camera. In this manner, the MM-HTAI can be accomplished and a sequence of axial-view images including bright-field and fluorescence of the specimen can be acquired.

In RQ 2, we take inspiration from the field of computer vision [27]. We consider an advanced method as image segmentation or edge detection for the acquisition of the 2D shape description of an organism represented in the images obtained from the MM-HTAI architecture. The methodology focusses on an *efficient* and *robust* segmentation method that incorporates conventional segmentation methods, such as variational based methods [28] and unsupervised learning based methods [29].

In RQ 3, a 3D shape representation for a scene or an object can be obtained from a range of multi-view images using multi-view stereo [30]. Our axial-view image sequence is a particular case for the multi-view, which is commonly referred to as turn-table data [31]. We can resort to a shape-based 3D reconstruction method to solve the problem of 3D shape description from the MM-HTAI architecture. Subsequently, 3D measurements can be directly derived from the results of the shape-based 3D reconstruction method. Importantly, we show that we can implement this 3D shape acquisition in an efficient manner.

In RQ 4, we need to solve the problem of 3D shape description for an organism that is partial translucent/transparent. It might result in weakly-defined boundaries. Therefore, we require an improved 3D reconstruction method which does not require the most accurate initial 2D shape descriptions. The method can be considered as an extension of the shape-based 3D reconstruction. We incorporate texture information from the axial-view images to infer a more flexible volumetric representation. We use probabilistic models and further validate accurate 3D measurements.

In RQ 5, the multi-modal images acquired by the MM-HTAI are used to obtain a multi-modal 3D shape description by the fusion of the 3D shapes both on organism and organ scale for the zebrafish larvae. This is supported by the shape-based and improved 3D reconstruction methods. The 3D shape description on organism scale presents a shape reference for the normalisation of the 3D description on organ scale. It requires an alignment of the multiple 3D shape descriptions resulting in a natural visualisation and a high quality 3D image fusion for the organism and its organs.

In RQ 6, we first extract features in the images from annotated datasets, and then apply classification (or regression) models to validate the performance of the features [32, 33]. Currently our research is hampered by the availability of sufficiently large annotated datasets for the zebrafish. Therefore, we choose to study the behaviour of the features and classification models in a collection of datasets concerning phenotypes/gene expression, e.g., humans, animals, and plants. We do so in order to obtain a balanced understanding of the methodologies. We believe that it is possible to transfer the knowledge in these datasets towards the application area of the zebrafish for the phenotype characterisation. The reason is that we are able to provide sufficient evidence for the generality and accuracy of the system in similar domains. Therefore, we will investigate various texture features and develop a graph-based local-global strategy for taxonomy prediction in several datasets. One of the particular cases would be kinship recognition in humans by facial analysis; other cases will refer to a diverse collection of datasets including butterflies, orchids, and wood species.

1.6 Thesis structure

The structure of this thesis is as follows. We address at least one research question in a chapter by presenting a new approach. Careful discussions and thorough inferences will be given. Chapter 1 provides the PS, the six RQs and the research methodologies.

Chapter 2 aims to answer RQ 2 (For RQ 1, see Chapter 3). To this end, we present an efficient and robust hybrid method for zebrafish image segmentation for bright-field microscopy of the MM-HTAI architecture. We integrate the merits of conventional segmentation methods, i.e., the variational based segmentation method and the unsupervised learning based segmentation method. Then we propose a sequential refinement on the hybrid segmentation, resulting in a better 2D shape description. The results present an overview for the zebrafish larvae.

Chapter 3 addresses RQ 1 and RQ 3. We first specify the MM-HTAI architecture based on the VAST BioImager and the light microscopy. From the acquired images, we address the problem of 3D shape acquisition through a shape-based 3D reconstruction method. The method uses the 2D shapes obtained in Chapter 2. An accurate camera motion estimation is the basis for this method. We solve the problem by presenting a novel method as the voxel residual volume (VRV) maximisation algorithm. We validate our method through particles of known size. In this chapter we also report a 3D shape reference using statistical distributions from 3D measurements of the zebrafish for three commonly used larval stages, i.e., 3, 4, 5 days post fertilisation (dpf). According to the best of our knowledge, this is the first validated and justified report on the topic in this field; the results have already been successfully used in pharmacokinetics and toxicology [34, 35].

Chapter 4 answers the RQ 4. We improve the 3D reconstruction by the incorporation of texture information from the original axial-view images, since in some cases, a 2D shape is difficult to obtain due to partial transparency. So, we take the texture distribution sampled from the images into consideration to estimate a more flexible 3D volumetric representation with a confidence score as entry. We demonstrate the successful application of the method in the MM-HTAI architecture.

Chapter 5 presents a solution for RQ 5. We propose the methodology of multimodal 3D reconstruction for the zebrafish larvae on both organism and organ scale. We use the feature of our MM-HTAI architecture to produce images for both the whole organism in bright-field and detailed organ structures in fluorescence. We take the zebrafish larvae and its liver as examples to explain our method. The shape-based 3D reconstruction method is applied to obtain the multiple 3D shape description; an alignment and a fusion of the multiple 3D shapes are integrated to obtain a good visualisation of the results.

Chapter 6 concerns RQ 6. We apply a hand-crafted feature, the Local Binary Patterns (LBP), on human facial appearance. Then we propose a graphical model to predict the taxonomy (kinship) for genetic related family members. We also apply a CNN architecture to acquire representative features. Subsequently, we design a multi-output layer to enable taxonomy prediction for a set of datasets of biological specimens, i.e., butterflies, orchids and wood species. As a result, we have successfully applied our methods in the applications mentioned earlier. The experiment provides suitable knowledge and understanding for the behaviour of the method when transferring the knowledgeable items from current applications to the phenotype characterisation using the zebrafish.

Chapter 7 summaries the answers to the six RQs and answers the PS. We list a few limitations of the whole work and propose possible solutions. Finally, we offer six recommendations for further research.

Table 1.1: The structure of the thesis

Chapter	RQ 1	RQ 2	RQ 3	RQ 4	RQ 5	RQ 6	PS
1	✓	✓	✓	✓	✓	✓	✓
2		✓					
3	✓		✓				
4				✓			
5					✓		
6						✓	
7	✓	✓	✓	✓	✓	✓	✓

Chapter 2

A Hybrid Segmentation Method for 2D Shape Description

Based on:

- Y. Guo, Z. Xiong & F.J. Verbeek, “An efficient and robust hybrid method for segmentation of zebrafish objects from bright-field microscope images,” submitted to Machine Vision and Applications, 2017.

This chapter addresses RQ 2.

> **RQ 2**: *To what extent is it possible to obtain an accurate 2D shape description for the zebrafish from the MM-HTAI architecture?*

Abstract – Accurate segmentation of zebrafish from bright-field microscope images is crucial to many applications in the life-sciences. Early zebrafish stages are used and in these stages the zebrafish is partially transparent. This transparency leads to edge ambiguity as is typically seen in the larval stages. Therefore, segmentation of zebrafish objects from images is a challenging task in computational bio-imaging (see Section 2.1). Popular computational models fail to segment the relevant edges which subsequently results in inaccurate measurements and evaluations (see Section 2.2). Here we present a hybrid method to accomplish accurate and efficient segmentation of zebrafish specimens from bright-field microscope images (see Section 2.3). (A) We employ the mean shift algorithm to augment the color representation in the images (see Section 2.3.1). This (1) improves the discrimination of the specimen to the background and (2) provides a segmentation candidate retaining an overall shape of the zebrafish. (B) A distance regularised level set function is initialised from this segmentation candidate and fed to an improved level set method, such that we can obtain another segmentation candidate which preserves the explicit contour of the object (see Section 2.3.2). (C) The two candidates are fused using heuristics and the hybrid result is refined to represent the contour of the zebrafish specimen (see Section 2.3.3). We have applied the proposed method on two typical datasets (see Section 2.4.1). From experiments, we conclude that the proposed hybrid method improves both efficiency and accuracy of the segmentation of images with zebrafish specimen (see Section 2.4.2 & 2.4.3). The results are going to be used for high-throughput applications with zebrafish (see Section 2.5).

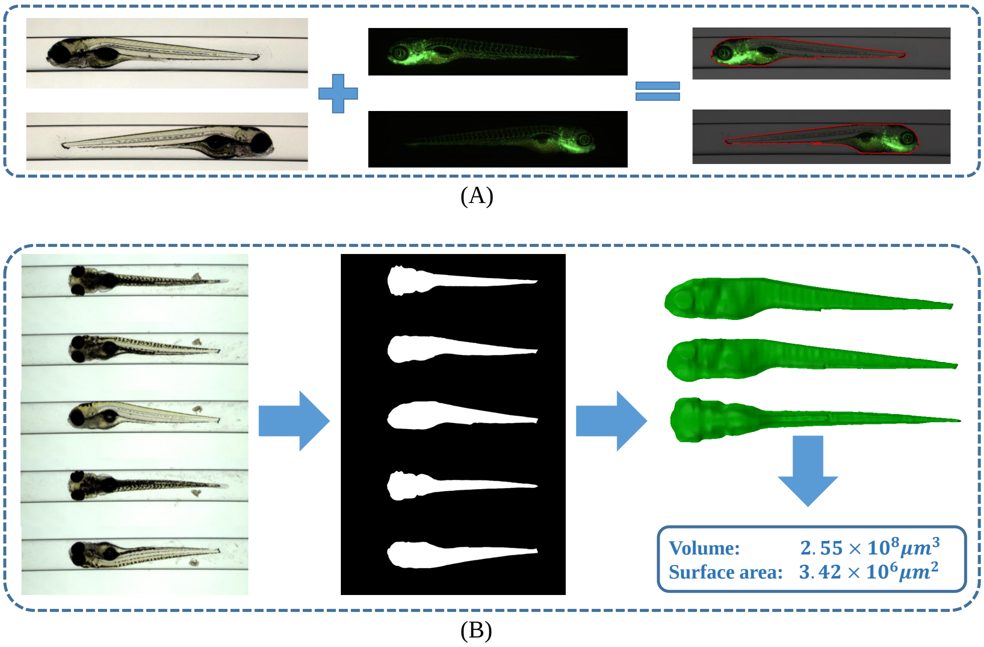


Figure 2.1: Typical applications of zebrafish segmentation. (A) Fluorescence images visualization and evaluation. Bright-field zebrafish images offer reference for the shape of the specimen (column one). Fluorescent images present informative signals, e.g. the blood vessels in green (column two). Accurate segmentation of the bright-field image provides a good shape reference to evaluate the fluorescent signals, for example, the development and concentration of specific cells (column three). (B) 3D zebrafish reconstruction from axial-views. Axial-view zebrafish images (column one) are segmented to obtain 2D binary shapes (column two), from which the axial-view based 3D reconstruction produces 3D models as well as 3D measurements (column three).

2.1 2D shape description

High-throughput imaging applications pose a challenge to the image acquisition in that in some cases the quality of the imaging is compromised at the cost of the speed of the imaging. Often this compromise is well-studied and the loss of quality is relatively mild. We have studied high-throughput applications for zebrafish; the zebrafish is a popular model system in bio-medical research. At present, high-throughput applications for zebrafish can be found, among others, in the fields of toxicology, cytology and oncology [36, 37].

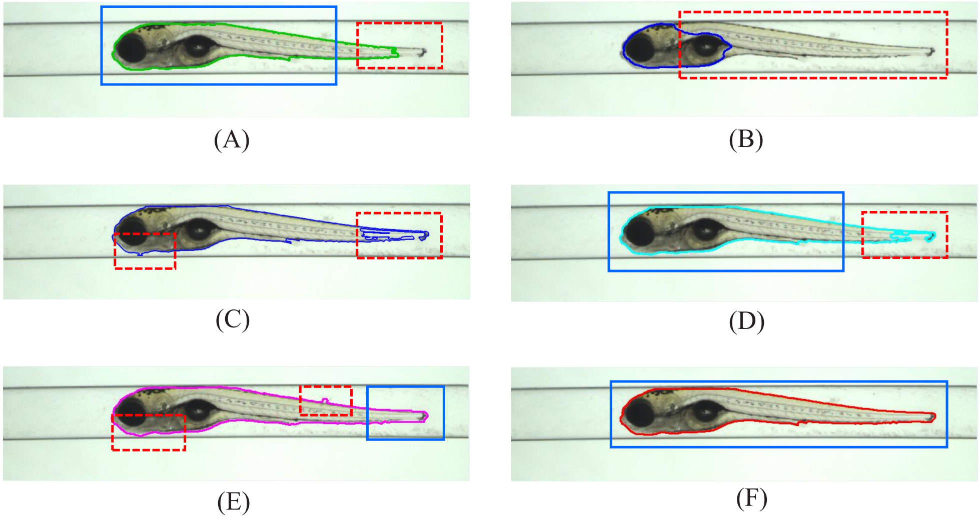


Figure 2.2: Segmentations by different methods for a zebrafish specimen in lateral position. Blue bounding box indicates the expected segmentations and red bounding box indicates inaccurate segmentations. (A) Segmentation by the geodesic active contours (GAC) model. Due to the edge sensitivity, the GAC model fails to detect the tail of the specimen. (B) Segmentation by Chan-Vese (CV) model. The partial transparency of the specimen makes it difficult for a region based method to discriminate the object from the background. (C) Segmentation by a local region based level set (LRLS) model. Similar problem occurs that the tail of the specimen is incorrectly segmented. (D) Segmentation by an improved level set (ILS) method. (E) Segmentation by mean shift (MS) algorithm. Better results are obtained though, edge sensitivity becomes worse. (F) Segmentation by the proposed hybrid (HY) method. The accurate segmentation presents a natural and compact shape description for the zebrafish specimen.

The development of zebrafish high-throughput imaging [7] has resulted in massive amounts of data, i.e. images, becoming available. This requires an *efficient* and *robust* analysis for the images, so that phenotype descriptions of the zebrafish can be generated. Genetically engineered zebrafish can be labelled with fluorescent markers. Images from fluorescence present good properties of visibility and measurability for cancer cells and organs. In order to evaluate the features which are usually represented as color intensity and concentration from the fluorescence, accurate segmentation of the zebrafish in bright-field images is quite essential to offer a shape reference for the measurements [38]. So, feature evaluations from control and experimental groups become comparable. In Fig. 2.1(A), an example of this application is depicted.

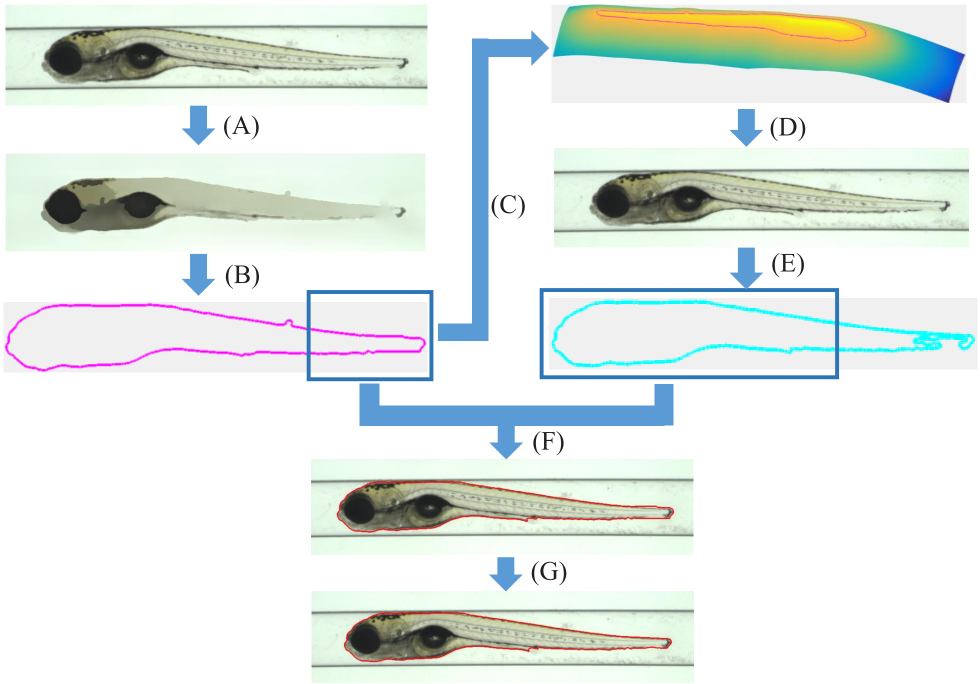


Figure 2.3: A pipeline schematic of the hybrid method. (A) MS algorithm is applied to improve the visibility of the transparent regions and weak edges. (B) An enclosed contour is extracted from the segmentation candidate in (A). (C) A distance regularized level set function (LSF) is initialized from the zebrafish contour in (B). (D) The ILS method is activated and applied on the original image. (E) Another segmentation candidate is generated. (F) An initial hybrid segmentation of the zebrafish is obtained by stitching the remarkable segments from the two candidates according to pre-defined protocols. (G) A refinement is followed to fine-tune the segmentation which can accurately represent the shape of the zebrafish.

Moreover, we can observe more informative features, e.g. volume, surface area and 3D shape variation, in 3D zebrafish imaging [39]. To this end, we need accurate 2D zebrafish segmentation to obtain sufficient shape priors for the axial-view based 3D zebrafish reconstruction [40]. In Fig. 2.1(B), we show this application.

In a particular case, according to the observation that the hemopoietic stem cells in zebrafish predominantly distribute in the tail, an accurate description of the overall shape of the zebrafish will ensure the evaluation of particular diseases by detecting and localising the tail region [17, 18]. Thus, an accurate segmentation of zebrafish objects in bright field microscopy is very significant for a large range of biomedical applications.

Computational methods from the field of computer vision can, in principle, help to accomplish the image segmentation task in zebrafish imaging. However, when popular image segmentation methods are applied, for example, the geodesic active contours (GAC) model [28] and the Chan-Vese (CV) model [41], the inhomogeneity of the intensity distribution caused by partial transparency and edge discontinuity of zebrafish larvae usually results in an inaccurate segmentation. To illustrate these effects, in Fig. 2.2(A) and Fig. 2.2(B), the segmentation results from, respectively, the GAC model and the CV model are shown. These segmentations show that the CV model converges at the most observable region but fails to retain the whole shape of the object; the GAC model obtains a poor shape description for the zebrafish tail. As shown in Fig. 2.2(C) and Fig. 2.2(D), other improved algorithms, such as the local region based level set (IRLS) model [42] and the improved level set (ILS) method [43] also do not result in an accurate segmentation of the zebrafish.

In fact, the edge based methods including the GAC model and the ILS method are able to accurately discriminate the visible edges but suffer from the problem of edge leakage. In contrast, as depicted in Fig. 2.2(E), unsupervised learning methods such as the mean shift (MS) algorithm [29] can obtain an overview shape description for the object, whilst the explicit edge will be, to a certain extent, contaminated due to region fusion effects.

For this particular research project, we aim at an *efficient* and *robust* solution for accurate zebrafish segmentation from bright-field microscope images. We, therefore, have developed the hybrid (HY) method to combine the advantages of various models. The objective of the HY method is to largely preserve the prominent contour of the object and discriminate the transparent regions and weak edges. In Fig. 2.2(F), we show the segmentation result. A schema of the HY method is depicted in Fig. 2.3, and below we elaborate the method.

In (A), we apply the MS algorithm on the original image to improve the color representation from the transparent object with respect to the background and obtain a segmentation candidate. This initial segmentation retains and approximates an overall shape of the zebrafish. In (B), we extract an enclosed contour for the object from the results obtained in (A). In (C), a distance regularised level set function is initialised from the result obtained in (B). In (D), with the initialised level set function, the ILS method is applied on the original image to obtain another segmentation candidate. It is important that this manner of initialisation significantly accelerates the curve convergence of the level set method and improves the segmentation accuracy. Because the initialisation already approaches to the edge potentials, the local minimum problem is solved to a certain extent. In (F), according to pre-defined protocols, we heuristically fuse the two segmentation candidates. In (G), a cascaded refinement module aims to fine-tune the segmentation result, which drives the contour to describe the shape of the zebrafish in a compact and accurate form.

A similar initialisation idea to step (C) is proposed in [44]. However, the employment of the MS algorithm in this work is not only to accelerate and stabilize the curve evolution, but also to obtain an overall view of the shape of the zebrafish which is beneficial for the following hybrid result. In other words, compared to the problem presented in [44], our zebrafish segmentation problem presents a more challenging task; the segmentors with just the improved initialisation is insufficient to achieve the best performance.

The remainder of this chapter is structured as follows. In Section 2.2, we review the related work and derive the level set method. We elaborate the HY method in Section 2.3. In Section 2.4, we present two datasets of zebrafish objects from bright-field microscope imaging. The experimental setup is depicted and the experimental results to evaluate the performance of the proposed method are presented. In Section 2.5, we summarise the research and indicate future developments.

2.2 Related work and background

In medical imaging, the functional based segmentation methods have been successfully developed and obtained good performance. These methods seem to be suitable for bright-field microscope imaging where complex scenes and noise are common. These methods aim at optimising an energy functional to estimate the optimal enclosed contour attaching the object boundary.

An early version of this technique is proposed as the classic active contours (snakes) model [45], from which the more advanced algorithms have been derived. The snakes model detects the object boundary by parameterising it as an enclosed curve $\mathcal{C}(p) \in \mathbf{R}^2,\ p \in [0,1]$. The curve will topographically evolve to minimise an energy functional formulated as $E(\mathcal{C})$ which incorporates an internal force considering the total length and the smoothness of the curve, and an external force derived from the image to encourage the curve to approach the object boundary. However, the snakes model cannot deal with changes in topology, in other words, it cannot detect all the boundaries in an image with multiple objects. Moreover, this method is rather sensitive to blurred edges.

The level set method is developed to handle the problems of topological merging and breaking [46]. The idea is to formulate the object boundary as the zero level set contour implicitly embedded in a three dimensional function which is known as the level set function ((LSF)) $\phi(\mathbf{x}, t) : \Omega \rightarrow R$, where the t is an artificial time variable presenting the time evolution procedure and the Ω is the image domain. The ϕ is usually assigned with positive and negative values in and out of the zero level set contour. The energy functional is transformed to $\widehat{E}(\phi)$ from $E(\mathcal{C})$.

Subsequently, a region based level set (CV) model is proposed [41]. With the introduction of the Heaviside function

$$H(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases} \tag{2.1}$$

the energy functional is defined as

$$\begin{aligned} \widehat{E}(\phi) \;=\; & \underbrace{\mu \int_{\Omega} |\nabla H(\phi)| d\mathbf{x}}_{\text{Length term}} \\ & \underbrace{+ \, \upsilon \int_{\Omega} \left(|I - u_{in}|^2 H(\phi) d\mathbf{x} \; + \; |I - u_{out}|^2 (1 - H(\phi)) \right) d\mathbf{x},}_{\text{External forece}} \end{aligned} \tag{2.2}$$

where, u_{in} and u_{out} represent the mean intensity of the image inside and outside of the curve, μ and υ are constants which can be tuned to balance different forces. The CV model can deal with the edge blurred images without employing edge terms. Based on the Euler-Lagrange equation, the gradient descent can solve the curve evolution problem. The gradient flow is computed as follows:

$$\frac{\partial \phi}{\partial t} = -\frac{\partial \widehat{E}}{\partial \phi}. \tag{2.3}$$

However, as shown in Fig. 2.1(B), the CV model fails to segment the zebrafish because of severe intensity inhomogeneity in the images. A local region based level set (LRLS) method is proposed to model the intensity variation as a bias term for each of the local region generated from intensity clustering [42].

Differently, the geodesic active contours (GAC) model [28, 47] which originates from the snakes model has its advantage of edge preserving, of which the energy functional is proposed as

$$\begin{aligned}\widehat{E}(\phi) &= \underbrace{\mu\int_{\Omega} g(|\nabla I|)|\nabla H(\phi)|d\mathbf{x}}_{\text{Length term}} \;+\; \underbrace{v\int_{\Omega} g(|\nabla I|)H(\phi)d\mathbf{x}}_{\text{Area term}} \\ &= \mu\int_{\Omega} g(|\nabla I|)\delta(\phi)|\nabla\phi|d\mathbf{x} \;+\; v\int_{\Omega} g(|\nabla I|)H(\phi)d\mathbf{x},\end{aligned} \tag{2.4}$$

where the g is known as the edge indicator which is formulated as

$$g(|\nabla I|) = \frac{1}{1 - c|\nabla I|^2}. \tag{2.5}$$

The values of g are close to zero at the region of object edges and one at the region of non-edges. This definition encourages the curve to converge at the object boundary when the energy functional is minimised. To derive the level set based GAC model, the gradient flow can be computed according to Eq. (2.3) as:

$$\begin{aligned}\frac{\partial\phi}{\partial t} &= \mu\delta(\phi)div\left(g(|\nabla I|)\frac{\nabla\phi}{|\nabla\phi|}\right) \;+\; vg(|\nabla I|)\}\delta(\phi) \\ &= \mu\delta(\phi)[g(|\nabla I|)div\left(\frac{\nabla\phi}{|\nabla\phi|}\right) + \;\nabla g(|\nabla I|)|\nabla\phi|] \;+\; vg(|\nabla I|)\delta(\phi).\end{aligned} \tag{2.6}$$

Finally, the curve evolution problem is transformed as a level set surface evolution problem

$$\phi_{i+1} = \phi_i + \Delta t\frac{\partial\phi}{\partial t}, \tag{2.7}$$

where the step size controller of Δt is tunable during solution search. This search is a standard gradient descent approach which can quickly locate the minimum of the functional.

From the observations of our bright field images, the contour of the zebrafish is more discriminative than the color. So, the edge based level set method should be suitable for our problem. However, from Figs. 2.1(A), 2.1(C) and 2.1(D), the boundary defects of zebrafish result in the problem of edge leakage for the aforementioned methods. To solve this problem, the shape prior based level set method are proposed [48, 49, 50]. This type of methods uses pre-defined shape templates to constrain the curve evolution. The employment of the shape constraint enforces the curve to approach the linear transformed template. However, the methods can only deal with the problems with limited shape deformations. Moreover, the methods including curvature constraint try to minimise the total curvature of the curve in order to control curve smoothness [51, 52]. However, these methods are difficult to implement with numerical solutions.

Besides, the performance of the GAC model also depends on the initialisation of LSF. A bad initialiser may lead the curve to converge at a local minimum, for example, the boundaries of the capillary as present in the images of the zebrafish. Cohen and Chen [53, 54] propose to find the global minimum of the geodesic energy by solving the Eikonal equation, but those methods require initial and end points from user input. In zebrafish high-throughput imaging, we prefer an automated manner.

Unsupervised learning based methods, e.g. k-means clustering [55, 56], superpixels [57, 58] and mean shift algorithm [29, 59], represent also a broad category of image segmentation techniques. Those methods can cope with complicated images by merging similar local regions and offer reasonable pre-segmentations.

Supervised learning based models [60, 61, 62] have drawn a lot of attention. Recently, the Convolutional Neural Networks (CNN)approach is becoming very popular and being successfully applied in many computer vision applications [25]. For the problem of image segmentation, some architectures are proposed and achieve great performance [63, 64, 65]. Those generic methods are usually trained from a large annotated dataset which is, however, not available for our problem.

Based on the discussions so far, we may conclude that each of the image segmentation methods shows good properties to solve a generic problem but also has its own limitations. Therefore, it is reasonable to develop a method to take advantage of the good properties of each of the methods. Here we aim at an *efficient* and robust solution for our zebrafish segmentation problem from bright-field microscope images. Considering the intrinsic characteristics of bright-field images of zebrafish, we propose the HY method. This method applies an unsupervised learning method, i.e. mean shift algorithm, to obtain an overview shape description of the object. The edge based level set method takes the pre-segmentation as initialisation and detects the explicit boundary. Finally, the two segmentation candidates are incorporated to obtain a better shape representation of the zebrafish.

2.3 A hybrid method for zebrafish segmentation

In this section we develop the HY method by fusing the advantages of the MS algorithm and the edge based level set methods, i.e the ILS method, to obtain accurate segmentation for bright field microscope imaging of zebrafish. The term *hybrid* represents a dual semantics. We first refer to hybrid as the improved manner of initialisation for the level set method with the MS algorithm. Compared with the functional based models, the MS algorithm shows the advantage of fast convergence and robust discrimination of transparency and weak edges. In this manner a segmentation candidate representing an overview of the zebrafish shape can be obtained and used to initialise the LSF for the ILS method. The ILS method can obtain another segmentation candidate to retain the explicit contour of the zebrafish. Then we refer to hybrid as the hybrid operation of the two segmentation candidates.

This section describes three topics: mean shift algorithm and the first segmentation candidate (in Subsection 2.3.1), the hybrid of improved level set method and accelerated initialisation for a second segmentation candidate (in Subsection 2.3.3), and hybrid of the segmentation candidates (in Subsection 2.3.3).

2.3.1 Mean shift algorithm and the segmentation candidate

We present a short recap of the MS algorithm in the application of clustering. In principle, the MS algorithm can cluster the similar data points through the estimation of the maximal density distribution of each data point. It is a kernel based density estimator which is derived from a method known as Parzen window.

Given n data points $\mathbf{x}_i,\ i = 1, ..., n$, the density distribution of a data point of $\mathbf{x}$ can be approximated by a kernel density estimator

$$\hat{f}(\mathbf{x}) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{\mathbf{x} - \mathbf{x}_i}{h}\right), \tag{2.8}$$

where h is the size of the bandwidth; d is the feature dimension; and $K(\cdot)$ usually takes the form of multivariate Guassian kernel which can be written as $K(\mathbf{x}) = (2\pi)^{-d/2}\exp(-||\mathbf{x}||^2/2)$. From the definition of Eq. (2.8), one can find that a data point similar to $\mathbf{x}$ will contribute more to its density estimation. We take the profile notation $k(x) = \exp(-x/2)$ instead of the kernel representation of K and yields the a profile representation of Eq. (2.9).

$$\hat{f}_{h,K}(\mathbf{x}) = \frac{c_{k,d}}{nh^d} \sum_{i=1}^{n} k\left(\left|\left|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right|\right|^2\right). \tag{2.9}$$

If a function is defined as $g(x) = -k'(x)$, the negative gradient of the profile function k, the gradient of (2.9) can be computed and transformed into the form as follows:

$$\hat{\nabla} f_{h,K}(\mathbf{x}) = \frac{2c_{k,d}}{nh_{d+2}} \left[\sum_{i=1}^{n} g\left(\left|\left|\frac{\mathbf{x} - \mathbf{x}_i}{h}\right|\right|^2\right)\right] \left[\frac{\sum_{i=1:n} \mathbf{x}_i g\left(||\frac{\mathbf{x}-\mathbf{x}_i}{h}||^2\right)}{\sum_{i=1}^{n} g\left(||\frac{\mathbf{x}-\mathbf{x}_i}{h}||^2\right)} - \mathbf{x}\right]. \tag{2.10}$$

The second term in Eq. (2.10) inspired us to the definition of the *mean shift*

$$\mathbf{m}_h(\mathbf{x}) = \frac{\sum_{i=1}^{n} \mathbf{x}_i g\left(||\frac{\mathbf{x}-\mathbf{x}_i}{h}||^2\right)}{\sum_{i=1}^{n} g\left(||\frac{\mathbf{x}-\mathbf{x}_i}{h}||^2\right)} - \mathbf{x}, \tag{2.11}$$

which indicates that the density maximizer of the data point $\mathbf{x}$ directs from the current data point to the kernel weighted mean of all the training data within a bandwidth of h. The location of the maximal density distribution of data point $\mathbf{x}$ can be approximated by updating Eq. (2.11) until convergence.

We apply the MS algorithm in image texture augmentation which we refer to as the image filtering and smoothing. In our problem of segmentation in images of zebrafish, the texture augmentation serves to (1) improve the discrimination from the transparent object with respect to the background and (2) to enhance the weak boundary. Considering both the color and spatial features in images, two bandwidths should be defined separately for those two metrics. The kernel of K should combine those two feature spaces and is represented as follows:

$$K_{h_r,h_s}(\mathbf{x}) = \frac{C}{h_r^3, h_s^2} k\left(\left|\left|\frac{\mathbf{x}^r}{h_r}\right|\right|^2\right) k\left(\left|\left|\frac{\mathbf{x}^s}{h_s}\right|\right|^2\right), \tag{2.12}$$

where k keeps the form of profile as previous definition; $(\mathbf{x}^r, \mathbf{x}^s)$ denote color and spatial features, respectively; the pair (h_r, h_s) represents the bandwidth in the two feature spaces. We use the three-channel RGB image and represent the spatial feature as two-dimensional coordinates of the pixel location. According to Eq. (2.12), the pixels within a range domain contribute more, i.e. represented as higher weights, for the density estimation of the centre pixel when the neighbouring pixels and the centre pixel are similar in color and spatial space.

By determining a proper combination of the bandwidths for (h_r, h_s) and applying the MS algorithm on the images of zebrafish, the weak boundary of the specimen can be, to a certain extent, recovered by the neighbouring pixels. At the same time, the color inhomogeneous regions are smoothed. For our application, only one object is present in the image, so a segmentation candidate for the zebrafish is directly obtained by thresholding the texture augmented images and represented as $\mathcal{S}_M$.

2.3.2 Hybrid of level set method and accelerated initialisation

In this chapter, we apply the ILS method for two reasons: (1) the efficient implementation and (2) its tunable properties to a problem. The ILS method improves the GAC model by the employment of a "region based term". Its energy functional is defined in Eq. (2.13).

$$\widehat{E}(\phi) = \int_\Omega [\mu g(|\nabla I|)|\nabla H_\epsilon(\phi)| \;+\; \upsilon(I - m)H_\epsilon(\phi)]\, d\mathbf{x} \tag{2.13}$$

where m is a user provided value which is used to pre-process the images. We use a smooth approximation of the Heaviside function, here defined as

$$H_\epsilon(x) = \begin{cases} \frac{1}{2}(1 + \frac{x}{\epsilon} + \frac{1}{\pi}\sin(\frac{\pi x}{\epsilon})), & \text{if } |x| \leq \epsilon \\ 1, & \text{if } x > \epsilon \\ 0, & \text{if } x < -\epsilon, \end{cases} \tag{2.14}$$

and its derivative

$$\delta_\epsilon(x) = \begin{cases} \frac{1}{2\epsilon}\left[1 + \cos\left(\frac{\pi x}{\epsilon}\right)\right], & \text{if } |x| \leq \epsilon \\ 0, & \text{if } |x| > -\epsilon \end{cases} \tag{2.15}$$

According to Eq. (2.3), the gradient flow of the ILS method is derived as:

$$\frac{\partial \phi}{\partial t} = \delta_\epsilon(\phi)\left\{\mu\left[g(|\nabla I|)div\left(\frac{\nabla\phi}{|\nabla\phi|}\right) + \nabla g(|\nabla I|)\frac{\nabla\phi}{|\nabla\phi|}\right] + v(I-m)\right\}, \tag{2.16}$$

where *div* denotes the divergence operator.

In fact, the ILS method replaces the “area constraint” in the original GAC model by a region based term inferred from the image to make the solution more tunable. For the sake of fast implementation, the additive operator splitting (AOS) scheme [47, 66] is used.

In general, an LSF should be defined to initialise the level set methods. Multiple options are available to accomplish this; e.g. random initialisation. Application of a random initialisation for segmentation of zebrafish images has the risk of the enclosed contour of the zero level set converging at a local minimum which is presented as the noise. The segmentation candidate from the MS algorithm offers an overall shape representation of the zebrafish, which is a reasonable initialiser and can be fed to the ILS method. The LSF initialised by the MS algorithm is an approximation of the object, which imposes the curve evolution of the ILS method to be activated from a considerably good location. Based on this idea, we accomplish the first goal of the HY method and specify the curve evolution of Eq. (2.7) in two phases:

$$\begin{cases} \phi_1 = \phi_0^M + \Delta t_1 \frac{\partial\phi}{\partial t}, \ t = 0, \\ \\ \phi_{t+1} = \phi_t + \Delta t_1 \frac{\partial\phi}{\partial t}, \ t = 1 \text{ to } T_1 - 1, \end{cases} \tag{2.17}$$

where the notation ϕ_0^M denotes the shape constrained LSF by the MS algorithm. Compared to the random initialisation fashions, the proposed HY method leads the LSF to approach to the global minimum, such that the ILS method is accelerated and more robust with less iterations. We obtain the second segmentation candidate of the zebrafish, represented as $\mathcal{S}_L$ through searching for the non-negative level sets in the converging LSF of ϕ.

2.3.3 Hybrid of the segmentation candidates

In order to accomplish the second task of the HY method, we define a hybrid operator to obtain the hybrid for the two segmentation candidates. To that end, we first detect the orientation of the zebrafish. In general, the side close to the broadest part of a zebrafish is recognised as the head side. The hybrid operator includes multiple operations of splitting and fusing and is mathematically defined as

$$\begin{aligned} \mathcal{S} &= \mathcal{S}_L \oplus \mathcal{S}_M \\ &= (\mathcal{S}_L^H \cap \mathcal{S}_M^H) \cup (\mathcal{S}_L^T \cup \mathcal{S}_M^T), \end{aligned} \tag{2.18}$$

where $\mathcal{S}$, $\mathcal{S}_L$, and $\mathcal{S}_M$ represent the segmentations by the hybrid operation, the ILS method and the MS algorithm, respectively. $\mathcal{S}^H$ and $\mathcal{S}^T$ denote the segments from the *Head* and *Tail* sides of the zebrafish. For the segment close to the side of head in zebrafish, the ILS method offers more compact contour so we take the intersection of the corresponding segments from the two segmentation candidates. For the segment close to the side of tail, the MS algorithm offers an approximation for the natural shape of zebrafish, therefore we take the union of the corresponding segments. The splitter of the two segments is then defined as $F(\mathcal{S}) \to \{\mathcal{S}^H,\ \mathcal{S}^T\}$. The splitting factor is empirically chosen as 10% of the full length of the zebrafish with respect to the tip of the narrowest part.

From the observation of the initial result of the HY method, segmentation artifacts at the stitching point might occur. Therefore, we propose a refinement in the form of the second-phase curve evolution based on the LSF initialised by the initial hybrid segmentation result. We specify this idea in Eq. (2.19). Hereby we use u to define the LSF to distinguish from Eq. (2.7).

$$\begin{cases} u_1 = u_0^{HY} + \Delta t_2 \frac{\partial u}{\partial t},\ t = 0, \\ u_{t+1} = u_t + \Delta t_2 \frac{\partial u}{\partial t},\ t = 1 \text{ to } T_2 - 1. \end{cases} \tag{2.19}$$

Algorithm 1: The HY method for zebrafish segmentation in bright-field microscopy

```
Input: Bright-field zebrafish image $\mathbf{I}$
Setup: $\mu,\ \upsilon,\ \Delta t_1,\ \Delta t_2,\ T_1,\ T_2,\ h_r,\ h_s$
Begin:
    Pre-process the noise of capillary: $\mathbf{I}_d = detect\_capillary(\mathbf{I})$
    Apply the MS algorithm: $\mathbf{I}_M = meanshift(\mathbf{I}_d)$
    Extract the segmentation candidate: $\mathcal{S}_M = threshold(\mathbf{I}_M)$
    Initialise LSF: $\phi_0^M = distance\_transform(\mathcal{S}_M)$
    for $iterator = 1 : T_1$ do
        Compute Eq. (2.16) to obtain gradient flow $\frac{\partial \phi}{\partial t}$
        Compute Eq. (2.17) to update $\phi$
    Obtain segmentation candidate: $\mathcal{S}_L = \mathbf{1}(\phi \geq 0)$
    Factorise segmentations: $F(\mathcal{S}_L) \rightarrow \{\mathcal{S}_L^H,\ \mathcal{S}_L^T\},\ F(\mathcal{S}_M) \rightarrow \{\mathcal{S}_M^H,\ \mathcal{S}_M^T\}$
    Apply Eq. (2.18) to obtain hybrid result $\mathcal{S}$
    Initialise LSF: $u_0^{HY} = distance\_transform(\mathcal{S})$
    for $iterator = 1 : T_2$ do
        Compute Eq. (2.16) to obtain gradient flow $\frac{\partial u}{\partial t}$
        Compute Eq. (2.19) to update $u$
    Obtain the final hybrid segmentation: $\mathcal{S}_F = \mathbf{1}(u \geq 0)$
End
```

Through the aforementioned manner, we can obtain more accurate representation of the zebrafish contour which is embedded as the zero level set in the u. The step size Δt_2 of the gradient flow is set to be much smaller than the previous one of Δt_1, which prevents the occurrence of edge leakage. In order to clearly illustrate the proposed method, we summarise the whole procedure in Algorithm 1.

2.4 Experiments

In this section we first present two datasets of bright field axial-view images of zebrafish from the Vertebrate Automated Screening Technology (VAST BioImager) (`http://www.unionbio.com/vast/`) (Subsection 2.4.1). We apply our HY method as well as several popular segmentors on the datasets to compare performances (Subsection 2.4.2). We evaluate the methods in the forms of accuracy and efficiency. The visualisation of segmentation results show the limitations of the reference methods and the merit of the HY method for segmentation of bright-field microscope images of zebrafish (Subsection 2.4.3).

2.4.1 Data collection

The VAST BioImager is developed for high-throughput experiments with zebrafish; the device can be mounted on a microscopes; its main feature is the ability of manipulation of zebrafish in the field of view by loading them in capillary. The VAST camera detects the orientation and location of the object. Once the object is present in the field of view of the imager, a set of stepper motors holding the capillary rotate the specimen in a full revolution, so that images of the zebrafish can be acquired in any axial-view. In our experiments, 84 axial-views (images) are evenly sampled from a full revolution (around 4.3^o per view) for each specimen.

Dataset A - The VAST BioImager is equipped with a standard camera, the VAST camera, which is used to detect the object presence in the field of view. With this camera axial-view images for the specimen can also be acquired representing an overview of the object. These images are 1024×1024 in size with a pixel size of $5.5\ \mu m \times 5.5\ \mu m$. In Fig. 2.4 examples of the images acquired by the VAST camera are depicted. The partial transparency and weak edge are clearly visible in most of the images. The All images in *Dataset A* are collected with the VAST camera. *Dataset A* includes a range of developmental stages of the zebrafish, i.e. three, four and five days post fertilization (dpf). The dataset contains three groups with in 60 examples. With 84 views per sample, this results in over 5,000 images in total (84×60).

Dataset B - The images produced by the VAST BioImager are relatively low-resolution and are insufficient for detailed observations of the zebrafish. Our setup consists of a microscope on which the VAST BioImager is mounted to produce high-resolution images. The VAST BioImager manipulates the specimen and the camera mounted on the microscope acquires the high-resolution images. Therefore, as an extension to *Dataset A* a *Dataset B* is obtained. The same imaging protocol with respect to *Dataset A* is used, i.e. 84 evenly sampled axial-views are acquired in a full revolution. The image size of each is 1920×2560 with a pixel size of $3.4\ \mu m \times 3.4\ \mu m$. In Fig. 2.5 some the examples of these images are depicted. For better visualisation, both of the vertical sides of the images are cropped to centre of the object and the image size is cropped to 600×2560.

Table 2.1: Segmentation performance of different methods on *Dataset A*

Model	Runtime (seconds)	F-score
CV model	1.74 ± 0.31	0.758 ± 0.123
CV model+MS	1.32 ± 0.16	0.758 ± 0.123
LRLS	22.83 ± 3.70	0.956 ± 0.026
LRLS+MS	19.56 ± 0.15	0.968 ± 0.014
GAC model	3.34 ± 0.38	0.976 ± 0.006
GAC model+MS	1.72 ± 0.13	0.976 ± 0.007
ILS	2.65 ± 0.42	0.976 ± 0.007
ILS+MS	1.26 ± 0.32	0.978 ± 0.006
MS	0.63 ± 0.07	0.964 ± 0.006
HY	1.37 ± 0.22	$\mathbf{0.983 \pm 0.004}$

Table 2.2: Segmentation performance of different methods on *Dataset B*

Model	Runtime (seconds)	F-score
CV model	8.87 ± 1.78	0.838 ± 0.120
CV model+MS	6.96 ± 1.63	0.838 ± 0.120
LRLS	152.27 ± 1.06	0.968 ± 0.016
LRLS+MS	126.60 ± 1.76	0.977 ± 0.011
GAC model	21.92 ± 0.19	0.918 ± 0.068
GAC model+MS	8.95 ± 0.40	0.957 ± 0.034
ILS	14.53 ± 6.39	0.970 ± 0.015
ILS+MS	7.23 ± 1.73	0.973 ± 0.022
MS	2.32 ± 0.31	0.965 ± 0.023
HY	8.30 ± 0.98	$\mathbf{0.986 \pm 0.004}$

2.4.2 Evaluation of different methods

In the experiment, the efficiency and performance are evaluated for different segmentation methods. The abbreviations of *CV*, *GAC*, *LRLS* and *ILS* consistently represent the Chan-Vese model, geodesic active contours model, local region based level set model [42] and the improved level set method [43], respectively; *MS* denotes the mean shift algorithm. The representation of $*$+*MS* indicates the $*$ model with an initialiser from the MS algorithm and *HY* is the proposed HY method.

In order to have a groundtruth set, we manually segmented 336 images of 4 specimens (84 views per specimen) from *Dataset A*. In addition, a subset from *Dataset B* including 33 images selected from 3 objects, is also manually segmented to obtain groundtruth annotations.

We measure the accuracy represented as F-score and the efficiency as runtime for all the methods on the subsets. The F-score is defined as $F = (2 \cdot recall \cdot precision)/(recall + precision)$. The closer to one the F-score is, the better performance of a method is. The mean and standard deviation for the two measurements are computed.

In the experiment, we partially used the fast implementation from [43]. To justify different methods, we give the same setups. For the models initialised by the MS algorithm, we take the configuration of the kernel bandwidths (h_r, h_s) as $(20, 20)$. Besides, all the methods are configured with the same number of iterations.

(A) Performance evaluation on subsets of *Dataset A*

In Table 2.1, we show the performance of different methods, evaluated on the subset of *Dataset A* with groundtruth. One can see that the CV model obtains the lowest F-score. This can also be seen in the segmentation result visualisation depicted in Fig. 2.1(B). Due to intensity inhomogeneity of the zebrafish in the image, it is difficult for the CV model to estimate the general mean of the texture inside and outside the object. Consequently, the CV model almost completely fails to detect the zebrafish.

For the other methods, comparable performances are seen, but differences are still existing. It is obvious that the MS algorithm is the most efficient segmentor. This provides evidence for the fact that a segmentor equipped with an MS initialiser is always more efficient than the same model with the random initialisation. We may conclude that the hybrid of the MS initialisation with the functional based segmentation model is helpful to improve the efficiency of zebrafish segmentation. The reason is that, the MS initialiser can produce a good estimation of the overall shape of the zebrafish. This shape approaches the global minimum.

The LRLS model also achieves a good performance. However, we should make more effort for the configurations and post-processing to obtain a natural shape for the zebrafish in the LRLS model. We do not have the fast implementation for the LRLS model, so that we cannot reasonably given a justification of its efficiency. Nevertheless, we can appreciate the hybrid of the MS algorithm and the LRLS model for a fast curve evolution.

Both the ILS method and the GAC model can obtain better segmentation results than the aforementioned methods. We find that the ILS method works faster than the GAC model. So, we choose to use the ILS method in our HY method. Considering the accuracy, the proposed HY method has the best performance. This is reasonable as the HY method combines the advantages of the MS algorithm and the ILS method. The segmentation result preserves an overall shape and retains the original explicit contour of the zebrafish.

(B) Performance evaluation on subsets of *Dataset B*

In Table 2.2 we show the performances of the different methods as evaluated on the subset of *Dataset B*. We can directly see that the efficiency of all methods is lower as a result of the larger image size. In addition, similar to the experiment on *Dataset A* it can be seen that the methods equipped with the MS initialiser generally work faster than the methods with random initialisation. Although the LRLS model obtains slightly better results than the ILS, the latter usually works faster. We do not have equivalent implementation of the LRLS model, so for the runtime no justification can be given. Due to the employment of the hybrid operation and post-processing, the proposed HY method works a little bit slower than the ILS method with an MS initialiser, but the segmentation accuracy is clearly improved.

2.4.3 Inspection of results by visualisation

In this experiment, we have visualised some representative segmentation results of *Dataset A* and *Dataset B* in this experiment.

For *Dataset A*, we randomly selected one zebrafish specimen from each group in *Dataset A*. For each example, three typical axial-views (lateral, 45^o tilted and ventral) are selected and shown in Fig. 2.4. We can observe that for the images with the zebrafish positioned in the view of ventral (dorsal), all the methods result in an accurate segmentation; this is due to the fact that the image portrays an explicit boundary of zebrafish. In the images with a lateral view of the zebrafish, the GAC model, LRLS model and ILS method fail to detect the weak edges. This phenomenon of edge leakage commonly occurs. Although the MS algorithm can retain a natural shape for the zebrafish, it loses the edge sensitivity. The proposed HY method obtains more accurate segmentations.

In Fig. 2.5, a representative set of images from *Dataset B* is depicted. Compared to *Dataset A*, these images have a better contrast and the outline (contour) of the zebrafish specimen is more explicit. Consequently, the classical edge based segmentors such as the GAC model have less difficulty segmenting the zebrafish from these images; even for zebrafish from a lateral view. The risk of edge leakage, however, still exists. In Fig. 2.5(B) and (C), we can see the contours resulted from the GAC model, LRLS model and ILS method converging at the wrong regions. The MS algorithm results in a segmentation retaining the whole boundary of the object, but the shape as a whole is less compact. From our experiment, we may conclude that the proposed HY method is able to deal with the segmentation problem for zebrafish specimens in bright-field microscopy. It results in more accurate results and shows a good performance. Due to the illumination conditions in the microscope, the acquired images are sometimes less explicit; this is depicted in the third column of Fig. 2.1(A). A straightforward pre-processing solution such as color equalization can improve the image contrast of the object with respect to the background. More segmentation results in this experiment represented as animations can be found here: `http://bio-imaging.liacs.nl/galleries/VAST-Hybrid/`.

2.5 Chapter conclusions and future work

In this chapter, we have presented a hybrid method to accomplish the task of *efficient* and *robust* segmentation of zebrafish from the bright-field microscope images. This answers RQ 2: *To what extent is it possible to obtain an accurate 2D shape description for the zebrafish from the MM-HTAI architecture*? Below we specify this answer more precisely.

We propose to employ the mean shift algorithm to augment the color representation for the partial transparent regions and transform the ambiguous edges more separable, such that we can obtain a segmentation candidate which preserves an overview of the zebrafish shape. A distance regularized level set function is initialised from this segmentation candidate and fed to an improved level set method in order to obtain a more compact shape representation preserving the explicit object contours. This hybrid operation accelerates the curve convergence at the regions of interest. We intuitively fuse those two segmentation candidates and employ a refinement in order to obtain the accurate hybrid segmentation. The results of our segmentation method facilitate the visualisation and evaluation of gene expressions in zebrafish in both 2D and 3D. This is directly relevant for the success of experiments in which imaging is crucial. Such experiments are typical for applications in life-sciences, e.g. cancer and pharmacokinetics. Furthermore, the proposed method is very suitable for high-throughput applications with zebrafish.

Below we provide three future perspectives. (1) The proposed method can be generalised by taking images into consideration that contain multiple objects positioned in various orientations. For orientation detection and initialisation over multiple instances modules need be developed that constitute the generalisation. For the work presented in this paper, the single instance is the approach for high-throughput applications. (2) Moreover, bright-field microscopy is a standard component for this type of applications. Nevertheless, the proposed HY method can be evaluated for other imaging modalities, with other lenses and illumination architectures. In this manner the HY method is probed and challenged for other and different image qualities. As an example, we consider optical projection tomography (OPT) imaging [22]; bright-filed images are included in this imaging technique and the processing of these images might benefit from the application of the proposed HY method. (3) Application to other imaging techniques will contribute to a further development and evaluation of the HY method.

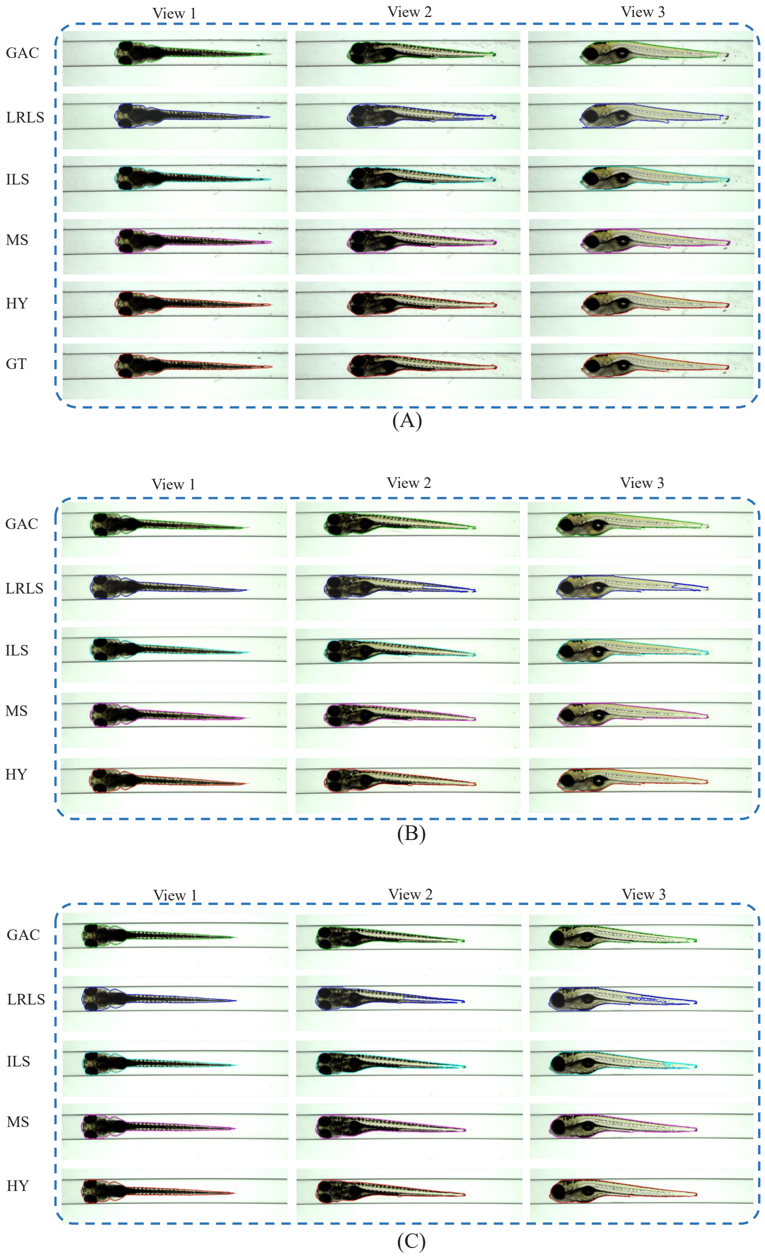
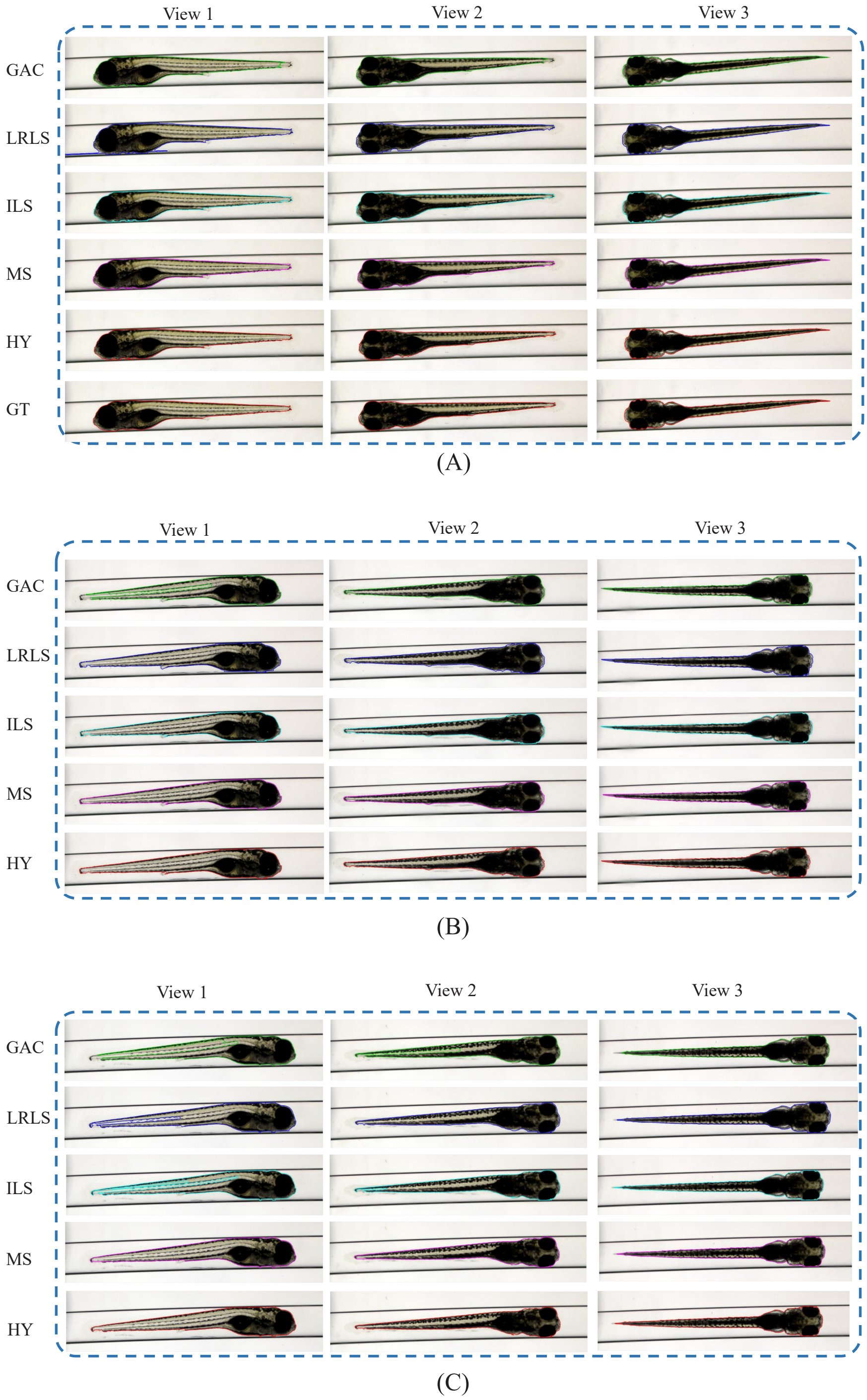


Figure 2.4: Segmentation results visualisation of different methods on *Dataset A*. GAC = geodesic active contours model [28]. LRLS = local region based level set model [42]. ILS = Improved Level Set method [43]. MS = mean shift algorithm [29]. HY = the proposed hybrid method. GT = groundtruth. A subset of the zebrafish larvae of 5 dpf are provided with manual annotations. For each example, three typical views (dorsal, tilted and lateral) are shown. (A), (B) and (C) correspond to three zebrafish larval stage of 5, 4 and 3 dpf respectively.



(A)

(B)

(C)

Figure 2.5: Segmentation results visualisation of different methods on *Dataset B*. (A), (B) and (C) correspond to three zebrafish examples, respectively.

Chapter 3

Shape-based 3D Reconstruction and 3D Measurements

Based on:

- Y. Guo, W.J. Veneman, H.P. Spaink & F.J. Verbeek, "Silhouette-based 3D model for zebrafish high-throughput imaging" in IEEE Conference on Image Processing Theory, Tools and Applications, Orleans, France, 2015, pp. 403-408.
- Y. Guo, W.J. Veneman, H.P. Spaink & F.J. Verbeek, "Three-dimensional reconstruction and measurements of zebrafish larvae from high-throughput axial-view *in vivo* imaging," Biomedical Optics Express, vol. 8, no. 5, pp. 2611-2634, 2017.

This chapter addresses RQ 1 and RQ 3.

> **RQ 1**: *To what extent is it possible to develop an MM-HTAI architecture for the zebrafish larvae?*
>
> **RQ 3**: *To what extent is it possible to obtain precise 3D shape description and derive accurate 3D measurements that are statistically relevant for the zebrafish from the MM-HTAI architecture?*

Abstract – High-throughput imaging is applied to provide observations for accurate statements on phenomena in biology and this has been successfully applied in the domain of cells, i.e., cytomics. In the domain of whole organisms, we need to take the challenge to ensure that the imaging can be accomplished with a sufficient throughput and reproducibility. For vertebrate biology, zebrafish is a popular model system for high-throughput applications (see Section 3.1). The development of the Vertebrate Automated Screening Technology (VAST BioImager), a microscope mounted system, enables the application of zebrafish high-throughput screening. For the VAST BioImager, fluorescence and/or confocal microscopes are used. Quantitation of a specific signal as derived from a label in one fluorescent channel requires insight into the zebrafish volume to be able to normalise quantitation to volume units. However, from the setup of the VAST BioImager, a specimen volume cannot be straightforwardly derived.

We first present the multi-modal high-throughput axial-view imaging architecture by the employment of different types of microscopes based on the VAST BioImager (see Subsection 3.3.1). We then propose shape-based 3D reconstruction to produce 3D volumetric representations for zebrafish larvae using the axial-views (see Subsection 3.3.3 and 3.3.4). Volume and surface area are then derived from the 3D reconstruction to obtain the shape characteristics in high-throughput measurements. In addition, we develop a calibration and a validation of our methodology (see Subsection 3.3.5). In Section 3.4, from our experiments, we show that with a limited amount of views, accurate measurements of volume and surface area for zebrafish larvae can be obtained. We have applied the proposed method on a range of developmental stages in zebrafish and produced real metric references for the volume and surface area of each stage (see Subsection 3.4.4).

3.1 3D shape description for zebrafish

The application of high-throughput imaging is wide-spread in modern molecular genetics based biology. Imaging takes an important position in a variety of high-throughput applications and the image based applications are found on the cellular, tissue-culture and organismal level. High-throughput is a demanding process and as throughput is required, the optics setup needs to cover for sufficient image quality in further processing of the samples.

Zebrafish is a popular model system to study a broad range of biological phenomena. A lot of these phenomena are rooted in molecular genetics, i.e., developmental biology, toxicology, cancer, infectious diseases, and drug targeting. In an experimental setting zebrafish are easy to breed and per crossing about 200 eggs can be obtained. The early embryonic stages are easily studied through microscopy as well as easy to manipulate. In the past decade, through genetic engineering, a large amount of transgenic zebrafish lines incorporating Green Fluorescent Protein (GFP, and the like) as a reporter gene have become available. These lines, and the genes they represent, support the research in that the gene-expression from the transgenic line helps indicating specific events in space and time. The GFP-like reporter genes are studied through fluorescent and confocal microscopy and the fluorescence facilitates measurability of the expression so that a numerical representation of experiments can be obtained.

Zebrafish, augmented with the large amount of reporter lines, are quite suitable for high-throughput experiments given the number of embryos that can be obtained for experimentation. Furthermore, the genomics of zebrafish is close to human. Therefore, zebrafish is now massively used as a model in disease studies which are often performed in a high-throughput setting. Initially, these experiments were designed to such a way that the read-out was realised using 2D images and subsequent processing of these images. In this manner a kind of mid-throughput could be accomplished. The space and time features of these images could, de facto, be used efficiently. In order to scale experiments to true high-throughput, other solutions need be probed [7]. Moreover, spatial acuity is required to enable the 3D image analysis. Apparently, confocal laser scanning microscopy (CLSM) seems the best option. So, in addition to the microscope and automated acquisition protocols, devices are required to make it possible to do accurate and fast imaging, both 2D and 3D, on life specimens and, if possible, in a time-lapse fashion. A solution has been presented through the adaptation

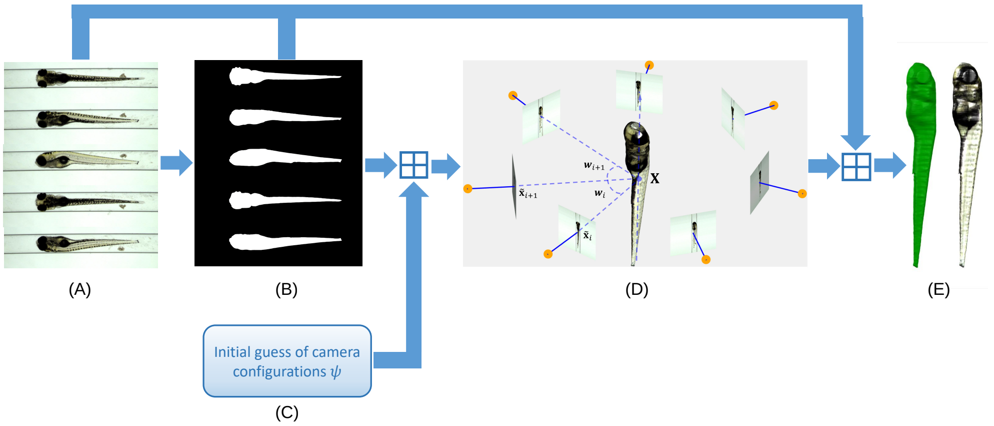


Figure 3.1: Flowchart of the 3D reconstruction and measurements of zebrafish larvae for the high-throughput axial-view imaging system. (A) Axial-view images of zebrafish as acquired from the VAST camera. (B) 2D shapes are obtained from the segmentation of axial views. (C) Initialisation of camera configurations estimated from the VAST BioImager. (D) Visualisation of parameterisation and calibration of the axial-view camera system. The bold dots represent centre of the camera lens; the dash lines represent the principal axis of the camera. (E) Reconstructed 3D zebrafish models with a volumetric representation shown on the left and a texture-mapping model on the right. The ⊞ symbol indicates the integration of different computational modules. For further details see Section 3.3.

of screening hardware from the invertebrate field and making it suitable for the vertebrate field, more specifically for zebrafish; i.e., the Vertebrate Automated Screening Technology (VAST BioImager) [26]. The VAST BioImager extends a microscope, equipped with (epi-)fluorescence and/or confocal imaging options. Through a capillary system, zebrafish embryos/larvae are delivered, transported and kept in a fixed position for imaging. The capillary is mounted such that it can rotate while the specimen is kept in its fixed position. In this manner, images from different angles, i.e., axial-view images, can be acquired and analysed. Some typical examples of the axial views obtained by the VAST camera are depicted in Fig. 3.1 (A).

In order to obtain successful image analysis of zebrafish, accurate and precise shape measurements are important, especially for providing a reference of the overall size and shape. In fluorescence, only the specific shape characteristics from specific labels can be analysed. In contrast, bright-field images will give all information on shape and size if they present an overview of the specimen. Therefore, we investigate how the axial-view images can be employed to obtain precise and accurate measurements on size and shape [40]. We first design a multi-modal high-throughput axial-view imaging (MM-HTAI) architecture based on the VAST system and light microscopy. This architecture can produce a full revolution of the specimen and acquire, with the VAST camera, a set of axial-view images. A potential of the MM-HTAI architecture is to produce detailed microscopic images both in bright-field and fluorescence. Subsequently, we will investigate how these images can be used to reconstruct a volume and obtain the size and shape measurements. These measurements then serve as a reference of size and shape of more specific features. A general overview of the proposed method is shown in Fig. 3.1. Given the microscope setup with the VAST BioImager, axial-view shapes of the zebrafish as shown in Fig.3.1 (B) are extracted from the original images as shown in Fig. 3.1 (A) and, henceforth, a 3D volume can be produced through combination of the shapes. The underlying idea of the 3D reconstruction from axial-views is similar to the problem of multi-view based 3D vision. To that end, a camera calibration model needs to be developed as framed in Fig. 3.1 (D). The 3D reconstruction can be accomplished once the segmentation and camera calibration are in place. An example of the 3D reconstructed zebrafish model is depicted in Fig. 3.1 (E). We need, however, to make sure that the result is a good representation of the shape; i.e., the measurements that we derive from the model can be considered correct. So, validation experiments and evaluation of the methodology on zebrafish larvae will be performed.

In order to clarify our method, the following issues are addressed in the next sections of this chapter. In Section 3.2, a number of 3D reconstruction methods are reviewed and the proposed method is motivated by the 3D reconstruction based on shapes, i.e., the binary masks obtained from segmentation of the object. In order to apply the method, an accurate estimation of the camera configuration is required. Therefore, we elaborate this issue for microscope optics and include some necessary optimisation steps for efficient computation in Section 3.3. In Section 3.4, we continue with a validation of the proposed method by introducing calibration beads that are used in an initial experiment testing whether our setup can produce accurate reconstructions. In addition, we test whether our measurement algorithms are sound. Subsequently, we have performed experiments on a large set of zebrafish larvae and shown successful applications of the proposed method for zebrafish larvae. The experiment deals with a range of developmental stages in zebrafish and the outcome of the measurements is interpreted. One of the results is a graph of the distribution for volume and surface area of the zebrafish larvae. We explain how this distribution is derived from the experimental results. We present the major conclusions from this research, answer the RQ 1 and RQ 3, and extrapolate to future work in Section 3.5.

3.2 Background and method motivation

In multi-view 3D reconstruction methods, two major categories can be discerned. The first category is based on epipolar geometry and aims estimating the depth information by matching the corresponding points from correlated images of one identical scene using geometrical clues [67, 68]. The methods use texture mapping to define point similarity or disparity, but require pixel-wise correspondence. The second category concerns the multi-view 3D reconstruction by producing a volumetric representation of the object from a series of binary masks of the shape, which sometimes is referred to as the silhouette based 3D reconstruction.

For the 3D measurements that we wish to extract from the zebrafish, the 3D shape plays an important role. Our method should, thus, have an innate shape preserving ability. Due to the partial transparency of the zebrafish embryos, it is difficult to determine point correspondences from the object surface within the multi-view images. Therefore, the silhouette based 3D reconstruction method is more suitable for this type of data. In this chapter, a binary mask is derived from an axial-view image which is preferably defined as a shape. As of now, we thereby use, in the remainder of this paper, the term shape-based 3D reconstruction to indicate our reconstruction method.

In the silhouette based method, the initial idea takes into account the parallel projections derived from each individual shape [69] to generate a 3D intersection. Various 3D volumetric representations are then proposed. Hierarchical resolution voxels are used for an efficient construction of the 3D object [70]. The voxels that are populating the object surface are represented with a high resolution grid, while the voxels inside the object are represented by much coarser grid. Such data structure is commonly referred as an Octree. Due to the characteristics of the camera imaging system, cone-shaped projections have been derived from the camera configurations [31, 71]. The silhouette based 3D reconstruction methods have become more feasible through the introduction of the concept of the visual hull [72, 73, 74]. The space carving algorithm [75, 76], as conceptualised in the space-carving theory, is an important implementation of this concept. In this method, based on the pinhole camera projection model, all the silhouettes are back-projected to the 3D space represented as a voxel-space. According to space carving theory [76], only the voxels visible to each of the silhouettes are preserved and used to reconstruct the final 3D model.

Other methods combine the information from texture mappings of the original images; one of which includes the surface reflectance to generate more natural 3D scene [77]. In recent approaches [78, 79, 80, 81], the 3D reconstruction is accomplished by searching for an optimal surface. The 3D surface is embedded in a higher dimensional space. Each voxel candidate included or excluded by the enclosed surface is assigned with a probability which indicates the possibility of the voxel candidate belonging to the object or the background. The probability is modelled by a multi-variate Gaussian kernel which is applied on the texture mappings interactively indicated by users. An energy function is subsequently defined by incorporating the total probability and the total surface area. Finally, the variational framework is taken to solve the optimisation problem. Recently, innovations [82, 83] were reported, which (1) are carefully formulated to preserve and refine the details across the projective rays on the object surface by introducing various constraints, and (2) achieve elaborate results on public datasets.

Our aim is to accomplish 3D measurements from our high-throughput imaging system. From the related work, we have taken our inspiration to further develop the multi-view based 3D reconstruction and introduce new ideas for the calibration of the optical setup in microscopy. Our point of departure is an existing opto-electronic configuration for high-throughput imaging to accomplish a reconstruction of zebrafish larvae. As a consequence, in order to be able to produce reconstructions from multi-view imaging, a calibration of the microscope imaging system needs be performed. To that end, a solution for the calibration problem is proposed using a revisited objective function in the optimisation. This function is defined as the voxel residual volume (VRV) maximization problem, which is a simplified formulation of the area coherence [84]. In terms of implementation, VRV is more efficient compared to the silhouette coherence [85]. In order to validate the proposed method, a large dataset of zebrafish larvae of different developmental stages is produced. For each of the instances in this dataset, a multi-view image set is acquired. From our method, 3D reconstructions are produced including the generation of dense surface points and triangulated meshes [86]. From the volumetric reconstructions and the meshes, volume and surface area can be computed. In addition, the method is applied to calibrated spheres in order to obtain metrical references. With a statistical analysis of the dataset, we have obtained a metrical reference for the most frequently used developmental stages of zebrafish in High-throughput imaging. This analysis will provide a baseline evaluation for future research in high-throughput imaging of zebrafish as well as a baseline for volume and surface estimations in zebrafish development providing reference values for a range of other features.

3.3 Materials and methods

Below we elaborate on the MM-HTAI system in Subsection 3.3.1. Our starting point is images that are acquired with the system; from these images, the shapes (silhouettes) need to be extracted using the method proposed in Chapter 2 which is also discerned in Subsection 3.3.2. We parameterise the camera model in Subsection 3.3.3, which is essential to the shape-based 3D reconstruction as elaborated in Subsection 3.3.4. For a good shape-based 3D reconstruction, the camera calibration should be operated and is addressed in Subsection 3.3.5.

3.3.1 MM-HTAI architecture

In this subsection, we describe the imaging system and procedures. Fig. 3.2 shows the so-called MM-HTAI architecture based on the VAST BioImager as designed for zebrafish high-throughput imaging. One at a time, the specimen is loaded into a dedicated capillary. A system of stepper-motors is used to fix and rotate the capillary; the positioning of the specimen in the field of view of the microscope is accomplished with an extra camera (Allied Vision Systems, Pro Silica GE 1050) which is part of the VAST BioImager system. We refer to this camera as the VAST camera which has a mounted lens that observes the capillary via a prism. Originally, the VAST camera is used for visual detection of the location and orientation of the specimen in the capillary so that the system will be able to manipulate the position and perspective of the specimen and initialise the imaging.

In practice, the VAST BioImager can operate with bright-field, fluorescence and confocal microscopy. In a synchronised acquisition protocol, images from different views, e.g, lateral, dorsal, and ventral, are acquired with the microscope mounted

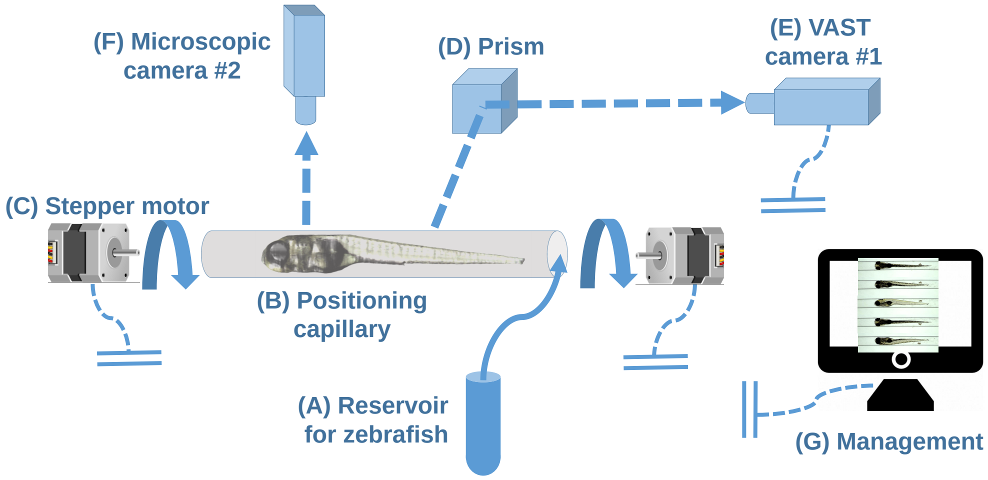


Figure 3.2: A schematic illustration of the MM-HTAI architecture based on the VAST BioImager. Zebrafish larvae are loaded from a reservoir (A) and delivered in to the capillary (B). The stepper motors (C) manipulate the view of the specimen. Through a prism (D), the VAST camera (E) detects the position and orientation of an object and then keeps the object in its field of view. The whole system is mounted on a microscope, of which an equipped high resolution camera (F) facilitates both of the organ- and cellular-level imaging. In addition, the VAST camera (E) manipulated by the management software (G) can acquire arbitrary axial-view images which always present an overview of the zebrafish.

camera or photomultiplier (in case of CLSM) in a full revolution of the capillary. The microscope view renders more detail, due to the characteristics of its lenses, but does not necessarily produce a complete view of the specimen. In addition, the VAST camera can also acquire images of the specimen, at a lower resolution, but always with a full view of the specimen. In fact, it can do so rather efficiently for a large amount of specimen. This is very suitable for our application of 3D reconstruction and measurements of the specimen. Therefore, we adopted the system to be the basis of our 3D reconstructions. It should be noted that the VAST BioImager is mounted on a microscope, of which the high resolution camera will be able to produce more detailed images of the specimen. In combination with the VAST camera, the further analysis of cellular-, organ- and overview-level of the specimen will become possible. This gives rise to the MM-HTAI architecture. In this chapter, we only use the modality of the VAST camera in our MM-HTAI to produce bright-field images.

3.3.2 Preprocessing and segmentation of the images

In our method, we require the binary representations of the object segmented from the colour images acquired from the VAST BioImager. In order to obtain a good and solid reconstruction, these binary representations should reflect the whole subject; in the case of transparent zebrafish, we have to ensure that the translucent parts are also included in the shape representation. We concluded that standard segmentation methods are not sufficient and therefore we have implemented a hybrid method that incorporates the mean shift algorithm (MS) [29] and an improved level-set method (ILS) [87]. In Chapter 2, one can observe that the MS preserves a whole shape representation for the object, but fails in edge sensitivity for the original clear boundaries. The ILS can obtain a more compact 2D shape but is hampered by the problem of edge leakage. So, we combine the two segmentation methods followed by a refinement to obtain accurate segmentations of the object, i.e., the zebrafish. For the rest of the chapter, the results from this segmentation method will be the basis for the 3D reconstructions. The segmentation of a set of axial views will be referred to as the shapes.

3.3.3 Camera model parameterisation

Key to the shape-based 3D reconstruction is a feasible and sufficiently accurate parameterisation of the camera imaging system. For our particular application, the camera position is static while the object rotates through the revolution of the capillary that holds the object, over a given profile axis. In order to make parameter parameterisation and visualisation feasible, the camera is represented as a range of fixed cameras around the object in a circular path as shown in Fig. 3.1(D). Now, for each of these cameras, the pinhole imaging principle can be used to interpret the imaging procedure. We start from a straightforward case of the camera model in which case the camera centre is defined as the world origin. Let $\mathbf{X} = (X, Y, Z, 1)^{\mathrm{T}}$ be a point in 3D space and $\mathbf{x} = (x, y, 1)^{\mathrm{T}}$ be the corresponding point in a 2D image plane, represented in homogeneous coordinates. Then, the mapping between the point and the corresponding image can be geometrically represented as a proportional projection of $x = (fX)/Z$ and $y = (fY)/Z$, where f denotes the focal length of the camera lens. The scaling factors k_x and k_y, image centre $(u_x, u_y)^{\mathrm{T}}$ and the skew factor s are, de facto, the intrinsic properties of the camera. Taking these intrinsic properties into consideration, the matrix representation of the projection can then be formulated as:

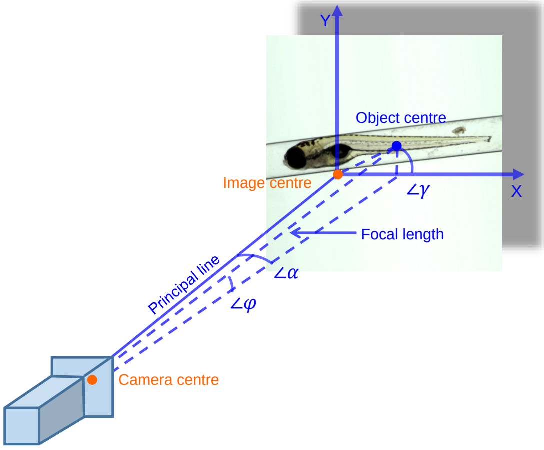


Figure 3.3: A visualisation of the parameterisation of 3D transformation from camera centre to object centre. We assume that the object is positioned in the focal plane and a good quality image is acquired, so the projection line from camera centre to the object centre is defined as the focal length. The $\angle\alpha$ and $\angle\gamma$ represent the 3D rotation angles of camera centre around Y and Z axis. The $\angle\varphi$ is defined as the "translation" angle from the object centre to the image centre. It models the 3D translations of camera centre over Y and Z axis.

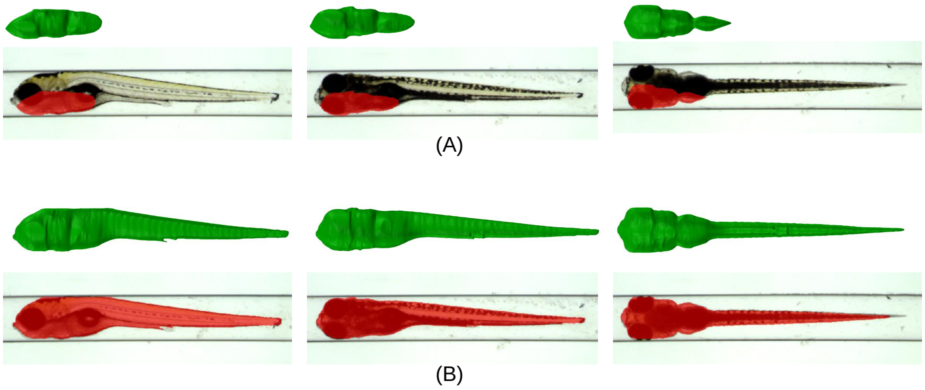


Figure 3.4: Illustration of the importance of camera system calibration. 3D reconstructions of one zebrafish, without (A) and with (B) camera calibration, are depicted. In the first row of (A) and (B) from three different viewpoints are shown. The overlap between the projected image from the 3D models and the original image are shown in red in the second row of (A) and (B). (A) Using an uncalibrated camera configurations results in poor 3D reconstruction and thus a relatively small overlap. (B) Using a calibrated camera system generates an accurate and natural 3D shape. It can be appreciated that projecting the 3D shape to the original axial view results in an almost perfect overlap with respect to the original object.

$$\tilde{\mathbf{x}} = \begin{pmatrix} fk_x & s & u_x \\ 0 & fk_y & u_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \mathbf{X}, \tag{3.1}$$

where $\tilde{\mathbf{x}} = \mathbf{x}Z$. To simplify the formulation, an implicit transform of Eq. (3.1) is denoted as $\tilde{\mathbf{x}} = \mathbf{P}\mathbf{X}$, where $\mathbf{P} = \mathbf{K}\ (\mathbf{I} \mid \mathbf{0})$, and $\mathbf{P}$ is defined as the camera projection matrix and $\mathbf{K}$ is defined as the intrinsic camera matrix. In practice, the more generic case for the camera model is to define the object centre as the world origin, which is more convenient for computation and parameterisation. The decomposed parameterisation for the camera model allows convenient incorporation of the extrinsic camera parameters which are the 3D transformation of the camera centre with respect to the world centre.

Suppose that for one object we have N views, the intrinsic camera matrix $\mathbf{K}$ is shared by all the cameras, however, a corresponding 3D transformation for each individual view should be formulated. This is illustrated in Fig. 3.1(C) and Fig. 3.3. For the i^{th} camera, the 3D translation of the camera centre from the object centre is parametrized as $\mathbf{C}_i = [0,\ f\cos(\varphi_i),\ f\sin(\varphi_i)]^{\mathrm{T}}$ following a general definition [85]. Now, let $\mathbf{R}_i = \mathbf{R}_{i,X}\mathbf{R}_{i,Y}\mathbf{R}_{i,Z}$ be the total 3D rotation, the revised formulation of the camera projection matrix is then represented by $\mathbf{P}_i = \mathbf{K}\ \mathbf{R}_i\ (\mathbf{I} \mid \mathbf{C}_i)$.

The "translation" angle of φ and the rotation angle of γ around Z axis are specified by each of the cameras. In this manner we intend solving the problem that the rotation axis is not exactly parallel to the object centre. In practice, with the VAST BioImager one can accurately align the object, so that these two angles can be shared by all N cameras. Finally, a camera parameter vector is generated through the concatenation of all parameters.

$$\psi = (\underbrace{f, k_x, k_y, u_x, u_y, s}_{intrinsic}, \underbrace{\alpha, \gamma, \varphi, \omega_{1:N-1}}_{extrinsic})^{\mathrm{T}}. \tag{3.2}$$

Given a specific camera parameter vector ψ, a series of correlated camera configurations can be generated, so that the multi-view 3D projections $\mathbf{P}$ can be modelled.

3.3.4 Shape-based zebrafish 3D reconstruction

We aim to recover the volumetric representation of the object from a range of 2D axial-view images. The shape-based 3D reconstruction method shows to be suitable for shape preservation in this application. Given a series of (object) shapes extracted from the axial-view images, combined with the corresponding camera projection matrices, the 3D representation can be generated by back-projecting the shapes to the 3D world frame so as to obtain the intersections.

Let the 3D point $\mathbf{X}^j \in \mathcal{X}$ ($j \in [1, M]$) be the index of the j^{th} voxel candidate centre included by a collection of voxels, where M is the total number of the voxels; if the collection of voxels is spatially constrained, then M will determine the 3D volumetric resolution. The larger the M is, the higher the resolution will be. Next, one can find the corresponding pixel of the j^{th} voxel in the i^{th} image plane by $\mathbf{x}_i^j = \mathbf{P}_i\mathbf{X}^j$. According to the space-carving theorem, if the voxel $\mathbf{X}^j$ is visible to the view of i^{th} camera, its corresponding image should be covered by the i^{th} shape $\mathbf{S}_i$. In mathematical terms this can be formulated as:

$$\mathcal{X}_i = \{\mathbf{X} \mid \mathbf{S}_i(\mathbf{x}_i^j) \neq 0, \mathbf{x}_i^j = \mathbf{P}_i\mathbf{X}^j, \mathbf{X}^j \in \mathcal{X}\}, \tag{3.3}$$

where $\mathcal{X}_i$ denotes the volumetric representation for the object generated from the i^{th} camera. When involving all the views, we can locate the pixel location in each image plane for each voxel. In general, only the voxels which are visible to all of the N shapes are preserved. Thus,

$$\mathcal{X}^* = \underset{i=1,\ldots,N}{\cap} \mathcal{X}_i \tag{3.4}$$

where $\mathcal{X}^*$ denotes the optimal 3D volumetric representation of the object. Eq. (3.4) can be efficiently implemented by iteratively discarding the invisible voxels from the original voxel collection. However, regarding the eventual imperfection in the shapes segmentation, some voxels which should be included in the 3D object are excluded during this space carving operation. Instead, we assign a score to each of the voxels indicating its visibility. This score is formalised as

$$\mathbf{V}^j = \sum_{i=1}^{N} \mathbf{1}\left[\mathbf{S}_i(\mathbf{x}_i^j) \neq 0\right] \tag{3.5}$$

where $\mathbf{1}[\cdot]$ denotes an indicator function which takes the value of 1 if the image of the j^{th} voxel lies in the i^{th} shape and 0 otherwise. The score function assigns a value between 0 and N for each individual voxel. If, in addition, a constrained error tolerance rate ϵ is defined, the problem of imperfection in the shape segmentation can be more or less solved by thresholding the membership function of a voxel. Accordingly, the 3D reconstruction model of Eq. (3.4) is reparameterised as:

$$\mathcal{X}^* = \{\mathbf{X} \mid \mathbf{V}^j \geqslant (1-\epsilon)N, \mathbf{X}^j \in \mathcal{X}\},\ \epsilon \in [0,1). \tag{3.6}$$

For our particular application in zebrafish imaging, we have empirically established that the parameter ϵ takes the value of 0.05, meaning that the voxels visible up to 95% of the shapes will contribute to the 3D reconstruction. With the reconstructed 3D volumetric representation, an initial 3D surface model can be obtained by the marching cubes algorithm; this surface model is formulated as dense surface points and a triangulated mesh, which can be further optimised [86].

3.3.5 Camera system optimisation

From the previous sections it becomes clear that the camera projection matrix, i.e., the camera configuration, plays an important role in the shape-based 3D reconstruction. A good estimation of the camera configuration will ensure accuracy in the 3D projective geometry. If a set of specific camera parameters as denoted in Eq. (3.2) are given, the camera projection matrix $\mathbf{P}$ can be correspondingly computed. The camera parameters, more specifically, the extrinsic camera properties, are usually unknown or not sufficiently accurate; e.g., through a drift of the image centre from the object centre. Therefore, an automated camera configuration estimation, i.e., a camera calibration, should be performed. The importance of a camera calibration is demonstrated in Fig. 3.4 by the results of applying the method to 3D reconstruction of zebrafish shapes. Standard camera calibration methods [88, 89] for our imaging environment are not available for the small scales of microscope imaging, i.e., the VAST camera. Therefore we have proposed [40] a method similar to ideas on texture registration and stitching [84, 85]. We define the area coherence as the area of overlap between the image projected from the

3D reconstructed model and the groundtruth shape. Accordingly, we can define an energy function as:

$$f(\psi) = \frac{1}{N}\sum_{i=1}^{N} C(\mathbf{S}_i, \mathbf{P}_i(\psi)\mathcal{X}), \tag{3.7}$$

where the camera matrix is reparameterised as $\mathbf{P}_i(\psi)$ meaning that $\mathbf{P}$ is a function of ψ. By evaluating the camera configuration space, the optimal suggestion of ψ can be obtained by maximising Eq. (3.7)

$$\psi^* = \underset{\psi \in \mathbf{\Psi}}{\arg\max}\ f(\psi). \tag{3.8}$$

Visualisations corresponding with this approach are depicted in the lower rows of Fig. 3.4 (A) and (B). We argue, however, that the expression of Eq. (3.7) is equivalent to the, so called, voxel residual volume (VRV) which is defined as

$$f(\psi) = |\mathcal{X}^*(\psi)|. \tag{3.9}$$

Eq. (3.9) denotes the total number of the voxel candidates of the 3D volumetric representation. Similarly, the reparameterisation of $\mathcal{X}^*(\psi)$ indicates the functional relation between $\mathcal{X}^*$ and ψ. In fact, keeping the maximal total overlap area between the projected images and the groundtruth shapes is equivalent with keeping the maximal VRV of the reconstructed object. The VRV is, however, computationally more efficient while the same performance can be obtained. This is because the computation of the area coherence is unnecessary. For such an unconstrained optimisation problem, multiple optimisers are available, for example, the Nelder-Mead simplex [90] and the Evolution Strategy (ES) [91, 92]. One should be aware that, although small, there always exists the risk of a local minimum of the optimisers which may lead to an inaccurate camera configuration estimation. However, feasible initialisations are possible so as to provide accelerations and improvements to the optimisation. The intrinsic camera parameters including the focal length and the scaling factors are provided by the camera and lens specifications. The extrinsic parameters including the 3D rotations and translations can be assigned through reasonable estimations; e.g., the parameter of ω_i is initialised as $\omega_i = (i-1) * 2\pi/N$. With these sensible initialisations, more efficient and accurate camera configurations become possible.

3.4 Experimental results

In this section we explore the method through validation experiments. In Subsection 3.4.1, we define the sampling for both the axial-view and spatial elements. In Subsection 3.4.2, we perform the experiments on calibration particles to determine the sampling size and validate the accuracy of proposed method. In Subsection 3.3.3, we evaluate the performance of our method on zebrafish larvae. This will result in numerical and statistical validity of the method so that it can be successfully applied in further research. In Subsection 3.3.4, we report the 3D measurements of volume and surface area using a statistical representation for the zebrafish in three larval stages.

3.4.1 Sampling of axial-views and volume for the experiments

With respect to the sampling, the physical limitations imposed by the hardware and the imaging scale are confined. There are two sampling modes that cooperate in the reconstruction process. The production of axial-view images is depending on the properties of the hardware, which is referred to as the Axial-view Sampling Density (ASD) and determined by the step size of the stepper motors that operate the capillary. For the experiments we will use a range of steps in order to find a good operational ASD from sparse to dense. For the shape-based 3D reconstruction, a volumetric representation is required. From the hardware and the spatial constraints of specimens, a confined volume of $5000 \times 1000 \times 1000$ μm^3 (width × height × depth) is determined, which can include the whole object. The sampling involves the variation of the size of the isotropic elements (voxels) in this volume, which can be inferred as the total number of the elements. The smallest sampling size will be equal to the sampling elements on the CCD sensor in which case the number of sampling elements is 30×10^6; this represents the largest number of sampling elements. The smallest number of elements is 1×10^6 which gives rise to a considerably different performance. The given sampling volume is used in all experiments described in this chapter.

3.4.2 Validation of the proposed method

In order to validate the accuracy of the proposed method, particles, i.e., beads, with a known size distribution are used (GP 500 μm Control Particles, High

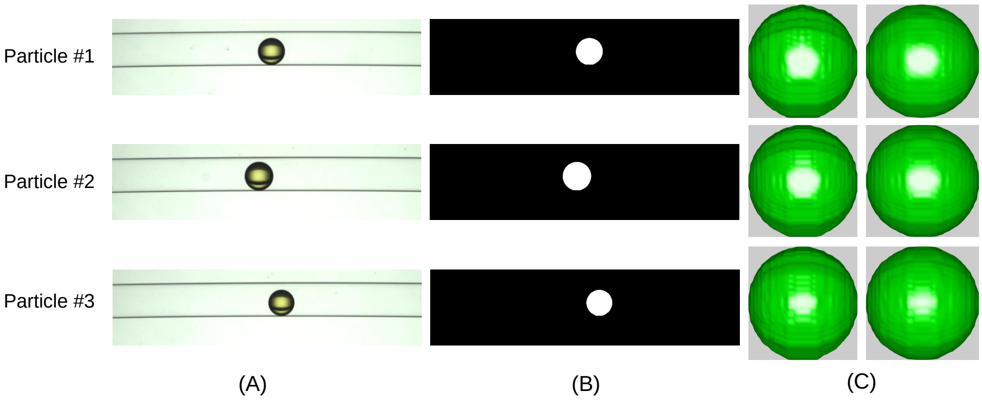


Figure 3.5: Illustration of example images of 3 calibration particles and their 3D reconstruction. Each row indicates one calibration particle example. (A) Original RGB images. (B) Extracted profiles. (C) Reconstructed 3D models by an ASD=7 (left column) and an ASD=21 (right column). Although the ASD=7 already produces acceptable 3D shapes for the calibration particles, while some carving artifacts are still visible. A better result is shown for the 3D models reconstructed with an ASD=21.

Fluorescence, lot 110701, Union Biometrica). The particles are used to calibrate the VAST BioImager system. Some typical examples are depicted in Fig. 3.4(A), from which we can observe some variation in the diameter of the particles. We assume that this variation in size is according to a Gaussian distribution as given by the manufacturer, i.e., $\mathcal{N}(500, 25)$ measured in μm. We have acquired image sets from the VAST BioImager for 25 calibration particles. Subsequently, the proposed 3D reconstruction method is applied to these image sets.

If the size measured from the reconstructions corresponds with the distribution of the size of the original calibration particles, it means that the proposed method can accurately recover the 3D object in terms of size and shape. There are two methods for evaluations: (A) To measure the diameter of the reconstructed calibration particles, a 3D sphere is estimated to fit the dense surface points by minimizing a least squares function. (B) To measure volume and surface area, the reconstructed 3D models can be used.

The result of the shape-based 3D reconstruction method can be affected by the Axial-view Sampling Density (ASD) and the voxel size (sampling density). As far as the ASD is concerned, more views may result in a more accurate and natural shape, similarly a higher sampling density of the voxels may result in a more accurate representation of the shape. Therefore, the experiments are designed to verify the accuracy according to these two parameters as well as establishing the correctness of the measurement itself by comparing to analytical representations or empirical observations.

The results are presented in Table 3.1 to 3.3; here the sampling density is denoted as SD given the size of the isotropic sampling element (in μm). For all tables the horizontal direction, from left to right, indicates an increasing sampling density from 17.10 μm per voxel to 5.85 μm per voxel. The vertical direction indicates an increase of the ASD from 4 to 84 views. We have configured our imaging system to take 84 images, corresponding with a step-size of 4.3° per axial-view, in a full revolution of the shape. Initial experiments have shown that a strategy of even axial-view sampling in a full revolution accomplishes the most stable performance. By resampling of the 84 views of 4.3°, a range of different ASDs is obtained for the experiments.

(A) Evaluations for diameter of the calibration particles

We interpret the measurements of the diameter in terms of the mean and standard deviation over the 25 calibration particles and present the results in Table 3.1. The results show that from the proposed method accurate estimations are obtained which are in range of the distribution (500 ± 25 μm) provided by the manufacturer. The results vary with different ASD and SD. More specifically, an increase in the sampling density results in an increase of the estimated diameter which asymptotically stabilises at the higher sampling densities. A high sampling density generates a smoother 3D surface, while for lower resolutions the volumetric representation tends to be a little bit expanded. An increase of the ASD results in an decrease of the estimated diameter while for higher ASD it stabilises. Using more views to reconstruct the 3D shapes causes less voxels to be preserved and the 3D shapes seem to be more compact compared to the ones reconstructed by smaller ASD. We tested whether our measurements were sufficiently accurate. In Table 3.1, in each second row, the T-test scores are given per ASD. We state in the null hypothesis that the measured diameter from the 3D reconstructed models equals to the values (mean, sigma) given by the manufacturer, i.e., 500 μm. For a significance level of 0.01 (two-sided T-test), a value of 2.787 is given from the table of selected values. From Table 3.1 one can appreciate that most T-test scores are smaller than the selected value; this means that our measurements for the diameter of the calibration particles from the 3D reconstructions are accurate. Especially, the values from the ASD of 21 combined an SD of 6.93 μmor smaller come out very well.

(B) Evaluations for volume and surface area of the calibration particles

The calibration particles have been provided with an indication of the size distribution, i.e., given diameter. However, the real distributions of the volume and surface area are unknown. In Table 3.2 and 3.3 the volume and surface area of the calibration particles are presented as computed from the reconstructed shapes in a range of parameters for ASD and SD. Volumes are computed directly from the voxels in the shape, i.e., integrating over the residual voxels after reconstruction. For each object a surface model is composed from a dense surface point cloud and a mesh triangulation from these points. The surface area can be obtained by integrating over the facets of the meshes using Heron's formula [86] $A = (s(s-a)(s-b)(s-c))^{(1/2)}$, where $s = (a+b+c)/2$ and $a,\ b,\ c$ represent the edge-length of a triangle. We assume that the shape of the calibration particles is a standard sphere, consequently the diameters from Table 3.1 are used to obtain the analytical computation of the volume and surface area by $V = (4/3) * \pi\,(d/2)^3$ and $S = 4\pi\,(d/2)^2$ for each set of parameters.

For each ASD, the values measured from the reconstructions are listed in each first row of Table 3.2 and 3.3, indicated by M. The second row for each ASD in Table 3.2 and 3.3 lists the analytical computations, indicated by A. The trends in Table 3.2 (Volume) and 3.3 (Surface area) are similar to those observed in Table 3.1. While the ASD increases, the reconstructed shapes are smoother and less voxels are preserved. Moreover, the variation in the results decreases at higher SD with voxels smaller than than 6.93 μm.

Taking the reconstruction effects shown in Fig. 3.5 (C) into account, it can be appreciated that a ASD of 21 is a reasonable choice to generate a smooth and natural shape with very few carving artifacts. Moreover, one can observe rather small differences between the measured and analytical values for an ASD of 21 as of the SD of 6.93 μm; i.e., the shape of the calibration particle is well preserved and reconstructed with such set of parameters. In comparison, a larger value of ASD requires more camera projections and the 3D reconstruction is therefore computationally more expensive; a sampling density smaller than 6.93 μm requires more memory and computation to ensure the precision. So, the results in this experiment indicate that the proposed method can obtain accurate 3D reconstructions measured in size, volume and surface area; the ASD of 21 and SD of 6.93 μm are empirically selected as the configurations for the further experiments.

3.4.3 Evaluations on zebrafish larvae

In order to obtain 3D reconstructions and evaluate the performance of the proposed method as applied to zebrafish specimens, we first assemble a large number of image sets of zebrafish larvae with the help of the axial-view imaging architecture presented in Subsection 3.3.1 and put the images in a dataset. We have chosen the most frequently used larval stages (for VAST BioImager) to be represented in the dataset; meaning 3 groups, i.e., 3, 4 and 5 days post fertilisation (dpf), containing 12, 24 and 24 instances, respectively. Although the VAST BioImager is suitable for live imaging, for this experiment, we have used fixed samples of one and the same strain, i.e., ABxTL wild type. All samples were fixed in 4% PFA/PBS (PFA: paraformaldehyde; PBS: Phosphate-Buffered saline). For the imaging, the samples were kept in a solution of PBS with 0.3% Tween-80.

For each specimen, the same imaging setup of the VAST BioImager is consistently used (cf. calibration experiment configuration). Specifically, in one full revolution, 84 views are captured with equal step size. Other ASDs, i.e., 42, 21, etc., are evenly sampled from the these 84 views. In the acquisition, the rotation steps between any two contiguous views are considered to be equal. However, due to the drift on the electro-mechanical parts of the VAST imager, we have to take into account that there can be a variation in the step size. If present, this variation can seriously affect the 3D reconstruction of the zebrafish specimens and therefore a calibration of the camera system is necessary.

In Table 3.4 and 3.5, we list the volume and surface area for the three groups of larval stages. For these tables the horizontal direction indicates an increase in ASD. In order to get a better idea of the data, we visualised Table 3.4 and 3.5 in Fig. 3.6 and 3.7, in which the voxel residual volume (VRV) and the surface area are plotted as a function of the ASD, respectively.

For all three groups the trend is similar, as the ASD increases the VRV decreases. This can be explained from the methodology as only voxels which are visible to most of the views are preserved. Consequently, less voxels will be preserved when more views are employed. The decrease of VRV rapidly changes in an asymptotic fashion of VRV for an ASD larger than 7. The asymptote indicates that we do not need an abundance of sampling views, an ASD of 42 or 84 does not result in a much better reconstruction and therefore these are unnecessary for the reconstruction of the zebrafish larvae in our application.

The VRV changes with the growth of the zebrafish for the different larval stages; i.e., as the zebrafish grow, the VRV increases. Similar trends can also be found in Fig. 3.6 which depicts the asymptotics of surface area for the 3D zebrafish models. External conditions influence growth as it also depends on the feeding policy. At some point in development the larvae have used all their yolk and are thereafter depending on nutrients in the medium. So, if the amount of nutrients is low, a delayed growth will occur.

From Fig. 3.6 and 3.7 it can be concluded that an ASD of 7 is already a good option for this application. However, integrating the results from all experiments, shows that an ASD of 21 achieves the best performance measured in terms of size, volume and area. In order to appreciate this, the reconstructed results are visualised for an identical zebrafish specimen in Fig. 3.8. For a ASD of 4 or 7, the 3D models exhibit a large number of artificialities. The space carving effects generate some flat regions on the surface. However, with a higher ASD value, such as 12 or 14, the 3D reconstruction results improve; the surface of the shape is still rough though. As of an ASD of 21 or larger, one can appreciate that results are more or less similar; the 3D shapes all appear smooth and natural. The visualisation quality does not improve much for an ASD larger than 21, while a large ASD is computationally much more expensive. The empirical ASD of 21 is therefore a motivated tradeoff over a range of considerations regarding the 3D reconstruction of zebrafish larvae and the measurement thereof.

A number of typical examples of models from 3D reconstructed zebrafish larvae are shown in Fig. 3.9. For each zebrafish developmental stage in our experiment two individuals have been selected; the 3D shapes are visualised with and without texture mapping. The texture mappings are just serving visualisation purposes, where texture is used to add some realism to the view. The specimens, in this stage, are still partially transparent, therefore, some artificial mapping was applied to obtain a realistic visualisation. For measurement the binary volume models are used.

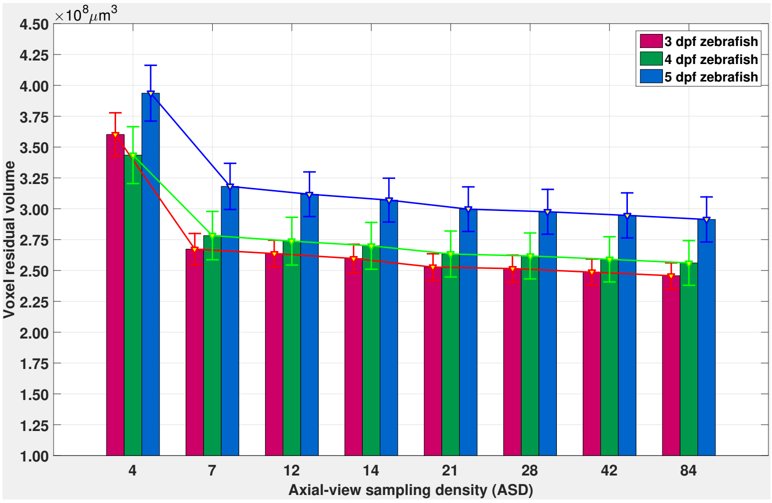


Figure 3.6: Voxel residual volume of zebrafish larval stages against ASD. The descending trends of the voxel residual volume are clearly demonstrated in the graph for all 3 groups of zebrafish larvae at increasing ASD. However, after an ASD of 7 the decrease tends to be asymptotic, especially after an ASD of 21. This means that dense axial-view sampling (42, 84, etc.) for our 3D zebrafish reconstruction and measurements is unnecessary. The consistent differences of the volume for the three larval stages can be observed.

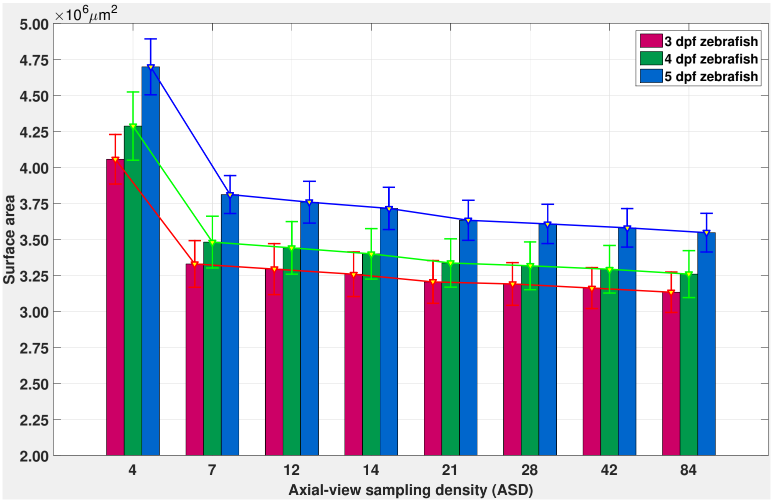


Figure 3.7: Surface area of zebrafish larval stages against ASD. The trend in surface area is similar to that of the volume as depicted in Fig. 3.6. The surface area is computed from a mesh, to suppress noise a mesh smoothing, i.e., implicit fairing, with 10 iterations was applied to all meshes [93].

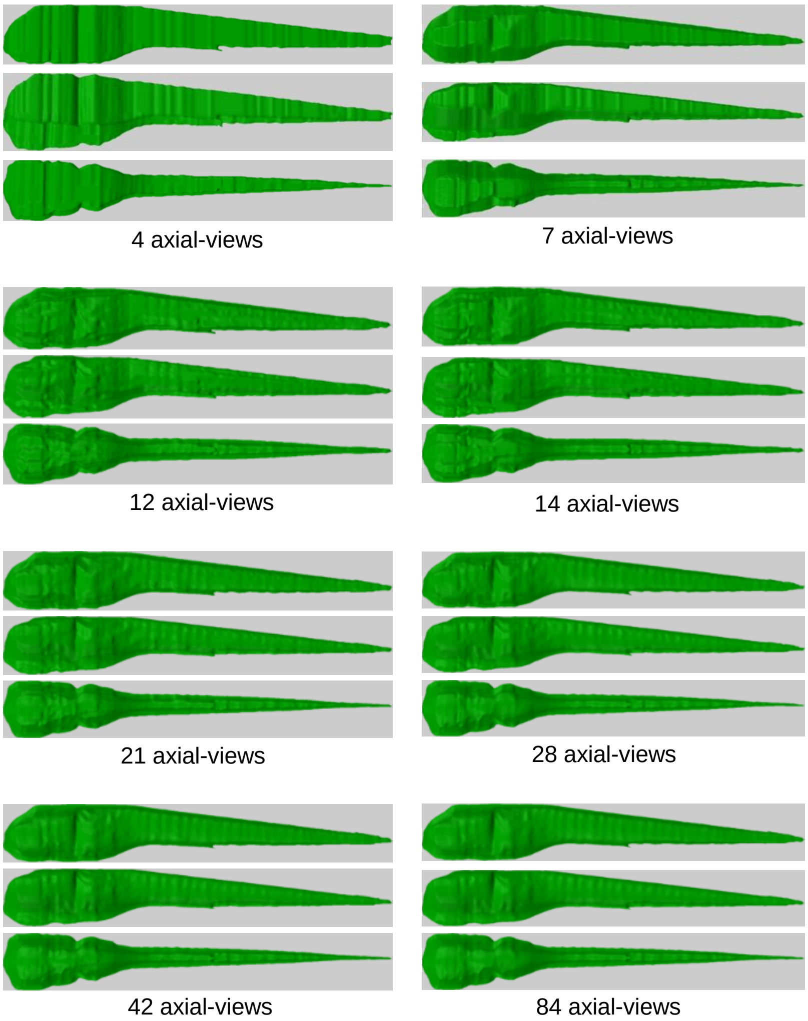
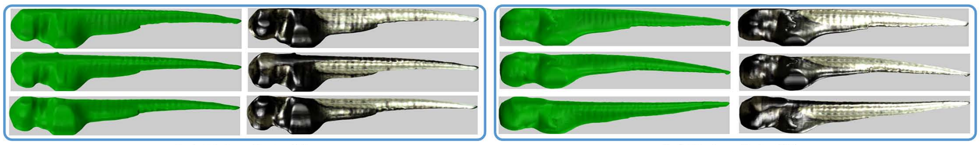
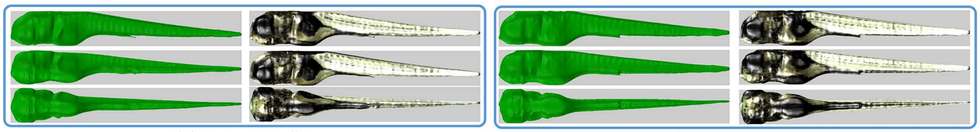


Figure 3.8: Visualisations of 3D reconstructions for a range of different ASD using the same zebrafish larva instance. The use of an ASD=4 (A) and an ASD=7 (B) results in the 3D reconstructions presenting a large number of sharp and flat surface elements which are generated by the carving effects of shape-based 3D reconstruction method. The use of an ASD=12 (C) and an ASD=14 (D) results in an improvement of the 3D models, but some carving artifacts remain. Using an ASD=21(E) to an ASD=84 (H) results in accurate and natural-shaped 3D models. For our particular problem domain, this suggests that both sparse and dense axial-view sampling are not the optimal configuration for zebrafish larvae 3D reconstruction. Sparse axial-view sampling produces poor 3D models while dense axial-views sampling requires more computation time whilst the effects do not improve accordingly.



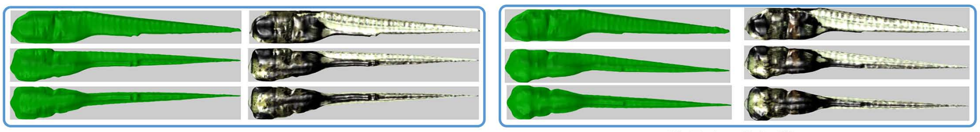
(a) Zebrafish #1 (b) Zebrafish #2

(A) 3D model examples of 3 dpf zebrafish



(c) Zebrafish #3 (d) Zebrafish #4

(B) 3D model examples of 4 dpf zebrafish



(e) Zebrafish #5 (f) Zebrafish #6

(C) 3D model examples of 5 dpf zebrafish

Figure 3.9: Visualisation of 3D models of 3 zebrafish larval stages (3 dpf, 4 dpf and 5 dpf). Each box represents a reconstructed 3D model for one specific zebrafish larvae visualised from three different viewpoints. The 3D volumetric representations are shown on the left side in each box. The models with texture-mapping are in the right side of each box. (A) 3D models of two selected 3 dpf zebrafish larvae. (B) 3D models of two selected 4 dpf zebrafish larvae. (C) 3D models of two selected 5 dpf zebrafish larvae. Variation in size and shape between stages and within stages (interclass and intraclass) can be appreciated from the visualisations. A remarkable intraclass discrimination originates from the size and colour of the yolk. In addition, animations of the 3D zebrafish models are available at: `http://bio-imaging.liacs.nl/galleries/VAST-3Dimg/`

.

3.4.4 3D measurements on zebrafish larvae

From the aforementioned experiments we may conclude that (1) an accurate reconstruction can be obtained from our method and that (2) from the 3D reconstruction accurate measurements can be made. In order to assess the shape and size in a high-throughput setup we use the database of our samples to find the distributions per larval stage.

From our empirically established parameters we reconstruct our images to 3D models using an ASD of 21 and a voxel sampling density of 6.93 μm. On these models, the computations are performed, the volume and surface area expressed in μm^3 and μm^2 respectively. Over the measurements per group, the mean and

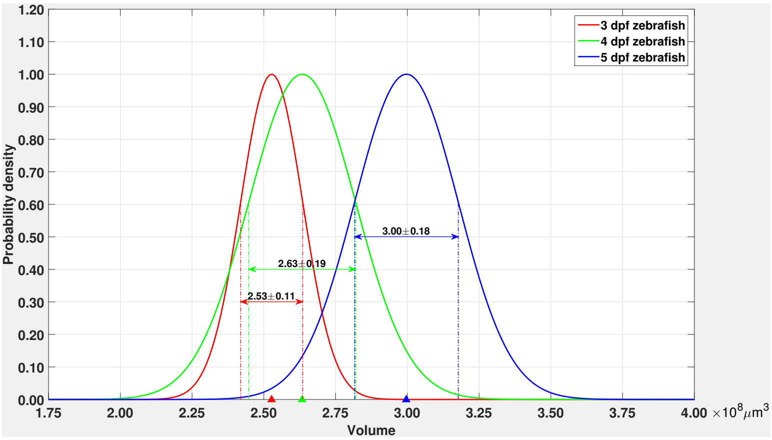


Figure 3.10: Distribution of volume of zebrafish larval stages (3 - 5 dpf). X axis: volume; Y axis: normalised probability density. The colour-filled triangles on the Volume-axis indicate the locations of the mean for each of the 3 distributions. The numerical values of the mean and standard deviation are indicated with corresponding double-sided coloured arrows. One can see that the volume of older zebrafish larvae is always larger than the younger ones. The growth rate increases more from 4 dpf to 5 dpf compared to that from 3 dpf to 4 dpf.

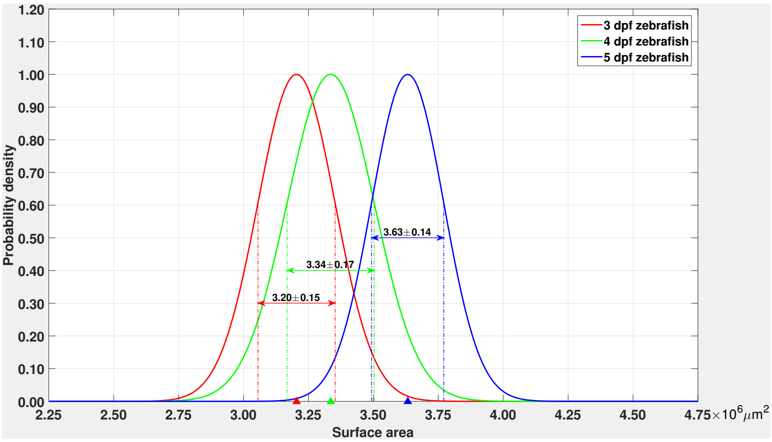


Figure 3.11: Distribution of surface area of zebrafish larval stages (3 - 5 dpf). X axis: surface area; Y axis: normalised probability density.

standard derivation are calculated and a Gaussian distribution is applied to model the data. In order to compare the 3 groups of larval stages the density distributions are normalised.

The distribution of the volume is shown in Fig. 3.10. The trend as seen in Fig 3.6 is reflected in the results in that the volume of 5 dpf zebrafish is the largest, and that of 3 dpf zebrafish is the smallest. The volume is more discriminative for the 3 dpf and 5 dpf zebrafish. The 3 dpf and 4 dpf zebrafish appear to be very similar in volume, which can be seen from the large overlap of the two distributions. Indeed, these larval stages are quite similar in shape and appearance and this is also illustrated in Fig. 3.9. The distribution of surface area is shown in Fig. 3.11 and can be interpreted in similar fashion to the volume measurement. A plot for the joint distribution of volume and surface area is shown in Fig. 3.12. Although the distributions of 4 and 5 dpf zebrafish are close to each other we can clearly separate the 3 centres.

This experiment provides significant quantitative information on zebrafish larvae because from the distributions, given the larval stage, the regularity of the shape can be directly assessed. The distributions provide an absolute reference baseline of volume and surface area for zebrafish in 3 larval stages. This is very important for applications in toxicology and drug targeting in which quantities of specific substances need to be related to volume and shape and phenotypical information on dose and effect. These need to be derived from measurements on the specimen. Moreover, as the readout of effects of substances is often measured with fluorescent probes, a relation to volume is even more important. In addition to these basic measurements other 3D features, i.e., moments, wavelets, shape descriptors, etc., can be considered. This will further strengthen the phenotypical description and assessment of zebrafish larvae in high-throughput imaging applications.

In order to make sure that we could measure the specific characteristics of the given larval stages, we used fixed specimens. In a control experiment we have used living larvae staged by specialists and imaged with the VAST BioImager under exactly the same conditions. We used 7 examples from 4 dpf and 8 examples from 5 dpf. The results of the measurements are shown in Table 3.6. From these results one can appreciate that they are all within the range of the fixed specimens.

We validated our methodology for the 3D reconstruction and measurements and at the same time showed the importance of the baseline measurements for the 3 larval stages; the major stages in compound screens with zebrafish. The reference values for volume and surface area as well as their distributions for the different larval stages can be directly used in experiments. Consequently, for phenotype analysis an assessment for relative size and shape can be made. Moreover,

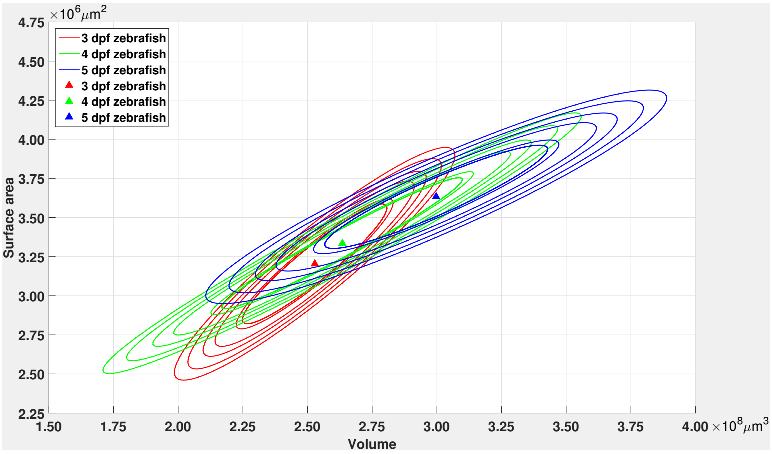


Figure 3.12: Joint distribution of volume and surface area of zebrafish larval stages (3 - 5 dpf). X axis: volume; Y axis: surface area. The 3 joint distributions have overlap with respect to each other, especially for the 4 dpf zebrafish larvae. Nevertheless, individual distribution centres can still be separated. The colour schema is similar to that of Fig. 3.10 and Fig. 3.11

features that are frequently used for phenotyping zebrafish and thereby evaluating anomalies can be related to size and shape to better assess the effect under study.

3.5 Chapter conclusions and future work

In this chapter, we presented the multi-modal high-throughput axial-view imaging (MM-HTAI) architecture based on the VAST BioImager in which images of zebrafish larvae are acquired. It answers the RQ 1: *To what extent is it possible to develop an MM-HTAI architecture for the zebrafish larvae*? We presented the shape-based 3D reconstruction using the acquired axial-view images to reconstruct 3D models so that 3D shape description of the zebrafish are obtained. Subsequently, 3D measurements can be applied on the whole specimen. It answers the RQ 3: *To what extent is it possible to obtain precise 3D shape description and derive accurate 3D measurements that are statistically relevant for the zebrafish from the MM-HTAI architecture*?

More concrete, the research presented in this chapter integrates the MM-HTAI architecture, image segmentation, camera system optimisation and the shape-based 3D reconstruction method. The 3D measurements of volume and surface area, as well as the 3D visualisation become available through the proposed method on zebrafish larvae using high-throughput imaging. The important conclusions are that (1) accurate 3D measurements for the zebrafish larvae can be made in high-throughput; (2) as of now a metrical reference of shape descriptors for 3 important larval stages is produced and (3) the methodology for these measurements is validated.

Below, we propose four directions for future research. (1) The results are directly applicable for a range of different *in vivo* applications, such as toxicology, drug targeting and infection studies. Further research will be directed to 3D shape alignment so that an even better match between different zebrafish instances can be given. Subsequently, other 3D features will be explored to accomplish elaborate classification of subtle differences between different treatment groups of zebrafish in experiments. (2) The VAST camera provides good overview images which have been employed for this study. The same methodology can be applied with the microscope camera. The higher numerical aperture and the magnification of the lenses in the microscope will then provide 3D details of the specimen in similar fashion. The tradeoff of using a confocal or wide field approach for such applications needs to be assessed from future results. (3) Alternative to the axial-reconstruction with the VAST BioImager, Optical Projection Tomography (OPT) imaging can be probed. This imaging approach has been successfully applied in zebrafish research [21] and, recently, successful high-throughput applications in zebrafish have been reported [22]. (4) As indicated, the measurements and their distributions provide reference values for volume and surface area (size and shape). This supports the assessment of other observations as we now can relate the effect of a change in size to a norm. In a similar manner, the VAST system can be used to provide distributions of other morphometrical landmarks in zebrafish development; to this end the microscope can be used to acquire axial-views at higher resolution and, in that manner, provide 3D features. These features can be combined with volume/surface area features to enable relative comparisons between experiments. Such analysis will further enhance our understanding of phenotypes in zebrafish and their exposure to experimental conditions which is a necessity for high-throughput analysis.

Table 3.1: Diameter statistics of calibration particles (M) measured from 3D reconstructed models (μm) and the corresponding (T) T-test scores.

SD / **ASD**		**17.10**	**10.00**	**7.94**	**6.93**	**6.30**	**5.85**	**5.50**
4	M	507.20±26.21	511.07±25.94	511.74±25.83	512.08±25.88	512.33±25.85	512.42±25.89	512.42±25.86
	T	1.34	2.09	2.23	2.29	2.34	2.35	2.35
7		494.60±25.69	499.00±25.39	499.85±25.33	500.26±25.34	500.50±25.31	500.64±25.32	500.68±25.30
		1.03	0.19	0.03	0.05	0.10	0.12	0.13
12		502.83±25.85	505.71±25.55	506.14±25.50	506.35±25.54	506.49±25.49	506.56±25.50	506.57±25.57
		0.31	0.94	1.02	1.06	1.09	1.10	1.10
14		499.70±25.80	502.98±25.49	503.48±25.42	503.75±25.44	503.90±25.42	504.00±25.42	504.02±25.46
		0.06	0.57	0.67	0.72	0.75	0.77	0.77
21		493.57±25.64	497.96±25.36	498.70±25.28	**499.03±25.29**	499.22±25.27	499.32±25.27	499.32±25.34
		1.23	0.39	0.25	**0.19**	0.15	0.13	0.13
28		491.39±25.63	496.52±25.35	497.52±25.27	498.00±25.29	498.29±25.25	498.46±25.28	499.00±25.37
		1.65	0.67	0.48	0.39	0.33	0.30	0.29
42		491.93±25.61	496.88±25.34	497.84±25.26	498.27±25.27	498.52±25.26	498.68±25.25	498.73±25.34
		1.54	0.60	0.42	0.34	0.29	0.26	0.25
84		490.82±25.61	496.05±25.35	497.09±25.27	497.56±25.27	497.84±25.25	498.00±25.26	498.05±25.36
		1.76	0.76	0.56	0.47	0.42	0.39	0.38

SD = sampling density; ASD = axial-view sampling density.

Table 3.2: Volume statistics of (M) measurements from the 3D reconstructed models and (A) analytic calculation from the fitted diameter for calibration particles ($\times 10^7$ μm^3).

SD / ASD		**17.10**	**10.00**	**7.94**	**6.93**	**6.30**	**5.85**	**5.50**
4	M	7.073±1.085	7.076±1.080	7.070±1.075	7.072±1.079	7.074±1.078	7.074±1.080	7.070±1.079
	A	6.884±1.077	7.042±1.083	7.069±1.081	7.083±1.083	7.093±1.084	7.097±1.086	7.097±1.084
7		6.651±1.021	6.648±1.015	6.644±1.014	6.646±1.015	6.647±1.014	6.647±1.015	6.644±1.018
		6.385±1.005	6.554±1.010	6.588±1.011	6.604±1.013	6.613±1.013	6.619±1.014	6.621±1.017
12		6.530±1.002	6.534±0.999	6.530±0.997	6.530±0.999	6.532±0.997	6.531±1.000	6.528±1.004
		6.708±1.029	6.822±1.030	6.839±1.029	6.848±1.031	6.853±1.031	6.856±1.031	6.857±1.033
14		6.531±1.002	6.527±0.999	6.522±0.996	6.523±0.997	6.524±0.997	6.524±0.998	6.519±1.004
		6.584±1.029	6.712±1.030	6.732±1.029	6.743±1.031	6.748±1.031	6.752±1.031	6.754±1.033
21		6.518±1.000	6.517±0.998	6.514±0.995	**6.514±0.997**	6.515±0.997	6.515±0.998	6.512±1.002
		6.345±0.998	6.514±1.005	6.542±1.004	**6.555±1.006**	6.563±1.006	6.567±1.007	6.567±1.009
28		6.449±0.993	6.449±0.992	6.443±0.988	6.443±0.989	6.446±0.990	6.444±0.991	6.439±0.998
		6.282±0.991	6.472±1.000	6.509±1.000	6.525±1.002	6.535±1.003	6.542±1.003	6.544±1.007
42		6.441±0.991	6.442±0.990	6.436±0.985	6.436±0.986	6.438±0.986	6.437±0.989	6.432±0.994
		6.262±0.989	6.458±0.999	6.496±1.000	6.515±1.002	6.526±1.002	6.533±1.003	6.535±1.007
84		6.416±0.988	6.419±0.988	6.413±0.983	6.411±0.984	6.413±0.984	6.413±0.988	6.407±0.995
		6.240±0.986	6.439±0.997	6.480±0.998	6.498±0.999	6.509±1.000	6.515±1.001	6.517±1.005

SD = sampling density; ASD = axial-view sampling density.

Table 3.3: Surface area statistics of (M) measurements from the 3D reconstructed models and (A) analytic calculation from the fitted diameter for calibration particles ($\times 10^5$ μm^2).

SD ASD		**17.10**	**10.00**	**7.94**	**6.93**	**6.30**	**5.85**	**5.50**
4	M	8.171±0.854	8.353±0.855	8.396±0.854	8.423±0.857	8.442±0.859	8.446±0.859	8.443±0.855
	A	8.102±0.842	8.226±0.839	8.247±0.837	8.258±0.839	8.266±0.838	8.269±0.840	8.269±0.839
7		7.708±0.807	7.873±0.806	7.911±0.807	7.935±0.808	7.951±0.810	7.955±0.810	7.954±0.809
		7.705±0.805	7.842±0.802	7.869±0.801	7.882±0.802	7.889±0.802	7.893±0.802	7.895±0.805
12		7.960±0.824	8.075±0.822	8.103±0.822	8.123±0.825	8.139±0.826	8.143±0.826	8.144±0.826
		7.925±0.820	8.028±0.816	8.041±0.815	8.049±0.816	8.053±0.816	8.055±0.816	8.056±0.817
14		7.861±0.817	7.984±0.815	8.010±0.814	8.031±0.817	8.045±0.819	8.049±0.818	8.049±0.818
		7.865±0.816	7.967±0.812	7.983±0.810	7.992±0.811	7.996±0.811	7.999±0.811	8.000±0.812
21		7.669±0.802	7.822±0.802	7.852±0.801	**7.871±0.803**	7.883±0.803	7.885±0.803	7.883±0.803
		7.673±0.801	7.810±0.799	7.832±0.798	**7.843±0.799**	7.849±0.798	7.852±0.799	7.852±0.801
28		7.600±0.798	7.773±0.799	7.808±0.798	7.830±0.800	7.844±0.801	7.847±0.801	7.845±0.803
		7.606±0.797	7.764±0.797	7.796±0.796	7.810±0.797	7.820±0.797	7.825±0.798	7.826±0.801
42		7.617±0.798	7.784±0.799	7.818±0.798	7.838±0.800	7.850±0.801	7.853±0.801	7.851±0.802
		7.622±0.798	7.776±0.797	7.806±0.796	7.819±0.797	7.827±0.797	7.832±0.797	7.833±0.800
84		7.582±0.796	7.757±0.798	7.792±0.797	7.812±0.799	7.825±0.799	7.827±0.800	7.824±0.801
		7.588±0.796	7.750±0.796	7.782±0.795	7.797±0.796	7.806±0.796	7.811±0.796	7.812±0.799

SD = sampling density; ASD = axial-view sampling density.

Table 3.4: Volume statics of zebrafish of the 3 larval stages under study as derived from the 3D reconstructed models ($\times 10^8$ μm^3).

ASD / **dpf**	**4**	**7**	**12**	**14**	**21**	**28**	**42**	**84**
3	3.60±0.18	2.67±0.13	2.64±0.11	2.60±0.11	**2.53±0.11**	2.52±0.11	2.49±0.11	2.46±0.10
4	3.43±0.23	2.78±0.20	2.74±0.19	2.70±0.19	**2.63±0.19**	2.62±0.19	2.59±0.18	2.56±0.18
5	3.94±0.23	3.18±0.19	3.12±0.18	3.07±0.18	**3.00±0.18**	2.98±0.18	2.95±0.18	2.91±0.18

Table 3.5: Surface area statics of zebrafish of the 3 larval stages under study as derived from the 3D reconstructed models ($\times 10^6$ μm^2).

ASD / **dpf**	**4**	**7**	**12**	**14**	**21**	**28**	**42**	**84**
3	4.06±0.17	3.33±0.16	3.29±0.18	3.26±0.16	**3.20±0.15**	3.19±0.15	3.16±0.14	3.13±0.14
4	4.29±0.24	3.48±0.18	3.44±0.18	3.40±0.18	**3.34±0.17**	3.32±0.17	3.29±0.17	3.26±0.16
5	4.70±0.19	3.81±0.13	3.76±0.15	3.71±0.15	**3.63±0.14**	3.61±0.14	3.58±0.13	3.55±0.13

ASD = axial-view sampling density; dpf = days post fertilisation.

Table 3.6: Statics of volume ($\times 10^8$ μm^3) and surface area ($\times 10^6$ μm^2) for living zebrafish larvae

	4 dpf	5 dpf
V	2.69±0.17	3.01± 0.16
S	3.48±0.15	3.74±0.13

V = volume; S = surface area.

Chapter 4

A Novel 3D Reconstruction Approach

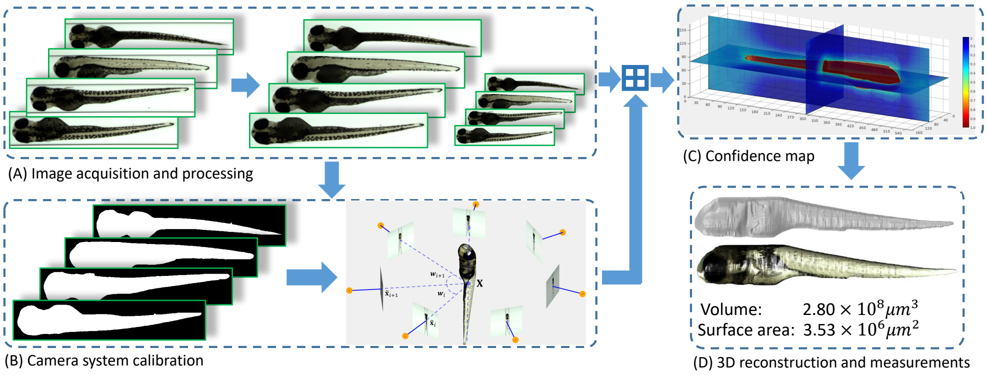
Based on:

- Y. Guo, Y. Zhang & F.J. Verbeek, “A two-phase 3D reconstruction approach for light microscopy axial-view imaging,” in IEEE Journal of Selected Topics in Signal Processing, 2017.

This chapter addresses RQ 4.

> **RQ 4**: *How can we efficiently deal with the translucency and transparency of specimen in light microscopy and still obtain a good 3D shape description from the MM-HTAI architecture*?

Abstract – Three-dimensional representations in light microscopy are important for accurate shape assessment of model systems in biosciences (see Section 4.1). The computational multi-view 3D reconstruction seems feasible in obtaining the 3D representations, in particular for high-throughput. The specimen for imaging can have properties, i.e., transparency and translucency, that impede the detection of well-defined boundaries (see Section 4.2). Consequently, 3D reconstruction and measurements, i.e., volume and surface area, will be inaccurate. The motivation in this chapter is to develop a two-phase 3D reconstruction approach for light microscopy axis-view imaging that can deal with these properties (see Section 4.4). In phase I of this approach, we develop an improved 3D volumetric representation defined as the confidence map. It is derived from texture-augmented axial-view images of the specimen. In phase II, the 3D reconstruction is accomplished by searching the optimal surface for the specimen over the confidence map. Subsequently, from the obtained 3D reconstruction, 3D measurements can be extracted. We apply our MM-HTAI architecture presented in Chapter 3 and propose three typical datasets with different imaging modalities, including (1) standard RGB images, (2) the bright-field images of zebrafish larvae, and (3) zebrafish liver in fluorescence (see Section 4.3). In the experiments, we have applied our approach on these datasets. We find that our approach yields a precise 3D shape representation and a natural visualisation (see Section 4.5). In comparison with a groundtruth setup, we have obtained accurate 3D measurements both for the organism and the organ, which holds a promising shape assessment for model systems in biosciences (see Section 4.6).



(A) Image acquisition and processing

(B) Camera system calibration

(C) Confidence map

(D) 3D reconstruction and measurements

Figure 4.1: A schematic of the 2-3DLA approach. Phase I: (A) A series of axial-view images of a specimen are acquired. The textures are augmented by the mean shift filtering, and the multi-scale images are used as input of the system. (B) 2D shape approximations of the specimen are obtained from the augmented textures, by which the camera calibration can be accomplished. (C) An improved 3D volumetric representation in the form of a confidence map is derived from the textures through the volume intersection model. Phase II: (D) Region based level set method is adapted and applied to the confidence map to estimate the optimal enclosed surface retaining a natural shape and smooth appearance of the specimen, from which the volume and surface area are derived.

4.1 Improved 3D reconstruction

In modern life-science research, large volumes of microscopy data, i.e., representative images of a specimen from cellular to whole-mount scale are used. Robust and reproducible methods for data acquisition, image processing, and analysis are essential for further handling of the data. As data volumes get larger, comprehensive visualisations summarising the data content also become important.

Images from samples, i.e., specimens, acquired from light microscopy carry a great deal of information that can be expressed by features such as shape and texture. Variations in these features provide information for a classification according to a specific condition. Such an approach is typically employed in imaging applications in the fields of cytology, toxicology, oncology and others. At a glance, experts are able to observe variations in colour, though subtle differences of size and shape are more difficult to capture. Moreover, as the volume of the data increases, it is not possible to classify the differences objectively. Therefore, these features need to be extracted in an automated manner. So, more advanced systems need to be developed [38]. For high-throughput systems this is especially true. Manual inspection is not feasible and in order to make inferences from the data, robust methods are required that produce accurate and reproducible measurements for shape analysis.

For the study of whole specimens, we should acknowledge that the nature of the shape is, de facto, in 3D. Thus, from a 3D shape representation of the specimen one will be able to produce an informative shape description, e.g., volume, surface area and 3D shape factors. With such features, we can accurately assess and compare shape variations in applications. These descriptors are also important for phenotypical systems evaluation that requires volume normalisation into metric measurements [39].

3D images can be generated by different techniques. With a confocal laser scanning microscope (CLSM) 3D images can be acquired from which 3D representations can be derived. For high-throughput imaging, CLSM is less efficient and also larger objects are not feasible for CLSM imaging. The post-processing of the images is not straightforward [94]. Alternative to CLSM is the optical projection tomography (OPT) [95]. However, a serious limitation of this technique is that it does not work for live specimens. Moreover, the specimen preparation is rather time consuming. In light field microscopy [96], microlenses are configured between the main lens and image sensor. Through 3D deconvolution a larger focal range is addressed and in this manner focal images are combined into a 3D image [97]. This technique is suitable for semi-transparent objects. In our research, we deal with specimens that consist of opaque, specular and transparent regions. A second limitation of light-field microscopy is the trade-off between spatial and angular resolution.

In this chapter, we present an axial-view imaging architecture based on light microscopy using the Vertebrate Automated Screening Technology (VAST BioImager) [26]. From this imaging architecture, we can generate a series of 2D axial-view images for the specimen. In fact, this imaging belongs to the class of multi-view imaging techniques, which is also referred to as turn-table sequence [31]. In the field of computer vision, multi-view stereo (MVS) approaches have been developed to recover a 3D scene, or an object, from a range of 2D multi-view images. In these approaches, surface points, or a depth map, of an object can be estimated through the matching of correlated images [30]. However, for light microscopy imaging, objects of interest can be (partially) translucent and/or transparent; this holds for the microscope modalities that we are using, i.e., bright-field and fluorescence. These qualities are in favour of the observation of internal structures of the specimen but prevent the feasibility of surface points matching which is required for a MVS approach.

In Chapter 3, we have obtained a binary 3D volumetric representation from the shape-based 3D reconstruction which is intuitive and suitable for the goal of shape analysis through 3D measurements [98]. This method requires precise 2D shape segmentation. However, the accurate shape segmentation is sometimes difficult to obtain from images depicting specimens with poor-defined boundaries as a consequence of translucent and transparent properties of the specimen. Instead, a probabilistic framework for the 3D volumetric representation generates more flexible but not always accurate 3D shape description. To address the problems, we propose a **two**-phase approach for the **3D** reconstruction and measurements from **light** microscopy **axial**-view imaging; this is abbreviated as 2-3DLA. A schematic representation of this approach is shown in Fig. 4.1. The system takes a series of axial-view images as input and reconstructs a precise 3D model of the object from which accurate 3D measurements, i.e., volume and surface area, can be derived.

Specifically, in **Phase I** of 2-3DLA, an improved 3D volumetric representation in the form of a confidence map is constructed. First, the mean shift algorithm [29] is applied to improve the texture representation of the original images. Thus, the translucent and transparent regions of the specimen are enhanced so that they become more separable. Subsequently, approximations of the 2D shapes of the specimen are obtained and used for camera system calibration. Together with all the data, based on the camera projection intersection model, we obtain the confidence map by imposing a score, instead of a binary value or a probability, to each voxel element in the 3D space. The score is obtained by integrating two probabilistic models. These probabilistic models are jointly estimated from all the axial-view images in a multi-scale fashion and aims to generalise the texture distributions of the object and background. The confidence map indicates the likelihood of each voxel in 3D space to be part of the object. This is a more flexible 3D representation for the optimisation of the surface. In **phase II**, we present the assumption that the optimal surface which includes the specimen will be able to (1) maximally separate the voxels from the object to the background as well as (2) retain a smooth appearance of the object. The 3D reconstruction and measurements are accomplished by searching for such an enclosed surface over the confidence map. Consequently, we formulate the 3D reconstruction as a

3D segmentation problem which can be solved by employing a region based level set method.

In modern life-sciences, e.g., developmental biology and pharmacokinetics, zebrafish are widely used as model systems in various experimental settings [18]. Zebrafish are small in embryonic stages and can be easily studied with different types of microscopes; zebrafish are transparent in the early stages. In recent years, zebrafish are augmented with a large amount of reporter lines and these lines are extensively used in disease studies. We use our light microscopy axial-view imaging architecture to acquire three representative datasets including bright-field images of zebrafish and fluorescence images of zebrafish liver. We use the zebrafish and the corresponding axial-view images as a case study for the evaluation of 2-3DLA. We have found that the 2-3DLA can be successfully applied in this research field; we have achieved promising results for 3D reconstruction and measurements of zebrafish and its liver.

We summarise our major contributions as follows.

I We present a multi-modal axial-view imaging architecture using light microscopy.

II We present three representative datasets of light microscopy axial-view imaging including the zebrafish in bright-field and the zebrafish liver in fluorescence.

III We propose a computational and automated system named 2-3DLA to solve the problem of accurate 3D reconstruction and measurements.

IV The proposed 2-3DLA is applicable in the datasets to obtain detailed 3D shape description both for the zebrafish and its liver. This makes the 2-3DLA generic for shape analysis on the level of the organism as well as on the level of the organ.

The remainder of the chapter is structured as follows. In Section 4.2, the background and related topics of our approach is introduced and the proposed system is motivated. In Section 4.3, we introduce three different complementary datasets using the MM-HTAI architecture. In Section 4.4, we describe the prior knowledge of the light microscope camera calibration and the binary 3D volumetric representation. This is followed by a detailed description of the proposed system. In Section 4.5 we present experiments and discuss the results. Finally, in Section

4.6 we summarise the answer to RQ 4, present conclusions and describe future work.

4.2 Background and related work

Given a sequence of calibrated multi-view images of an unknown scene, one category of the multi-view stereo (MVS) approaches aims to estimate a depth map for each view by matching each pixel of a binocular image pair [99, 100, 101]. The matching criterion is usually defined within a support window, i.e., a local neighbourhood, centred around a target pixel [102, 103]. Another approach formulates the depth map estimation as a continuous optimisation problem of an energy functional integrating colour, spatial consistency of neighbouring views and a global smoothness constraint [104]. A merging strategy is developed to integrate the multi-view depth maps. A patchmatch method [105] takes the slanted planar surface into account for a better depth map estimation in binocular stereo. An extension of the patchmatch method is developed for massive parallelisation and integration of multi-view depth maps [106].

Alternatively, a conventional MVS approach directly estimates the surface points of an object using a sparse-to-dense strategy [107]. A novel approach [108] extends the patch-based method into multi-view 4D reconstruction, solving the problem of temporally consistent 3D modelling in videos.

Regarding MVS, an accurate estimation of epipolar geometry is essential, which could significantly reduce the search space on the epipolar line and thereby increase the matching quality [67]. Structure from motion (SFM) enables the estimation of camera poses from the cooperation of salient point detection [109, 110] and bundle adjustment [89, 111]. However, both SFM and MVS rely on the quantity and quality of the salient point detection and matching from correlated images. In our light-microscopy imaging, the challenge is that the object surface and boundaries are not always well-defined. In most cases, volume instead of surface for an object is observable.

For image based 3D reconstruction, the volumetric representation approach is an important category. The goal is to estimate a convex hull in the 3D space represented as discrete voxels according to their visibility to each view [31, 73, 74, 112]. Accordingly, the shape-based method defines the 3D object through the intersection of a set of projections exposed from the 2D shapes of the objects in the images. A space carving algorithm aims to recover the 3D object by wiping out the voxels which are consistently invisible to the views [75, 76]. These methods, however, require accurate image segmentations which are not always available in light microscopy imaging.

In addition, the textures from the multi-view images can be used to optimise the 3D volumetric representation based 3D reconstruction [79, 80, 81]. Instead of using a binary representation, the probabilities indicating the membership of each voxel are estimated by applying multivariate Gaussian kernels on the textures of foreground and background, respectively. However, this requires user specification. Consequently, the quality of user input will, to a certain extent, determine the quality of the 3D reconstruction. Other methods attempt to fuse the shape and texture consistency in an integral deformable framework [113].

There are also other well-designed 3D reconstruction methods. Some in particular solve the problem of specular and transparent object reconstruction [114]. These methods work well on a macroscopic scale as they position special patterns behind the object to enable the shape-from-distortion or set up various lighting to collect surface reflective-highlights [115, 116, 117]. A recent study [118] presents a semantic reconstruction as a convex-relaxation formulation which combines a data term and a regularisation constraint, achieving elaborate results on public datasets [100]. Nowadays, deep learning is used in the MVS to improve the matching quality [119, 120]. This type of method requires a large volume of training data, which, in our case, is not available.

From the study of related work, we may conclude that the volumetric representation based approaches, such as shape-based 3D reconstruction with binary volumetric representation are most promising in addressing the challenge of 3D reconstruction and measurements for light microscopy imaging. However, this conclusion is not sufficient in itself, as in some cases accurate 2D shapes cannot be obtained. Therefore, we have developed the 2-3DLA method to solve the problem. In our approach, we first estimate a confidence map using the augmented textures from the axial-view images. The confidence map is actually an improved 3D representation. Subsequently, we accomplish the 3D reconstruction using the region based level set method on the confidence map. A validation has shown that our method yields accurate 3D measurements.

4.3 Dataset collection

In Chapter 3, we have developed our MM-HTAI architecture. In this section, we provide a comprehensive depiction of the dataset collection and show how the multi-modal images are produced.

The VAST BioImager is specifically set up to work for high-throughput imaging of zebrafish. For the remainder of the chapter, the concept of specimen and zebrafish are both used. Specimen is used in a generic context while zebrafish is used in the particular application of the zebrafish high-throughput imaging. In the VAST BioImager, the specimens, i.e., zebrafish larvae, are positioned along their longitudinal axis as in this manner the most important features can be readily observed; it is also related to the manipulation of the specimen in the capillary. In this orientation we obtain images from all axial-views along the profile axis from one full revolution. A schema of the imaging architecture is depicted in Subsection 3.3.1 (Fig. 3.2). A positioning module consists of a capillary (the holder of specimen) and a set of stepper motors that accomplish the rotation manipulation of the specimen.

A VAST camera (#1) is mounted with the device and used to detect the location and orientation of the specimen so as to keep it in the field of view for image acquisition. An object is revolved over 360 degrees by the stepper motors so that bright-field images for the specimen in the axial-views can be acquired. These image always depict the whole specimen. The VAST camera (#1) is an Allied Vision Systems, Pro Silica GE 1050 CCD (pixel size 5.5 μm × 5.5 μm.). This camera acquires images of 1024×1024 pixels. The resolving power for this camera system is about 13.4 μm.

The VAST unit is mounted on a microscope, of which the microscope-camera (#2) is able to acquire detailed microscopic images both in bright-field and fluorescence from arbitrary axial positions; this setup uses the Leica DFC450C CCD (pixel size 3.4 $\mu m \times 3.4$ μm) and it acquires images of 1920×2560 pixels (5 Megapixel). We use two objectives in this case. The resolving power for a 2.5× objective is about 4.8 μm and for a 4× lens (red fluorescence) it is about 3.1 μm.

The image acquisition protocol is as follows. In a full revolution of the capillary, 84 axial-view images are acquired for the specimen. The step size between adjacent axial-views is about 4.3° ($360° \div 84 \simeq 4.3°$). We have shown that 21 evenly sampled axial-views ($N = 21$) are sufficient to obtain accurate results [98]. We apply this acquisition protocol to obtain 3 different datasets that we will use as case-studies for the evaluation of our 2-3DLA approach under different imaging conditions.

Dataset A is obtained by the VAST camera. A number of zebrafish larvae during different development stages, i.e., 3, 4 and 5 days post fertilisation (dpf) are acquired. The dataset finally contains 12 examples for 3 dpf, 24 examples for 4 dpf and 24 examples for 5 dpf, respectively, and 60 examples in total. We will use the images from this dataset as example to interpret our 2-3DLA system. Examples of the images in this dataset are shown in Fig. 4.3. This dataset is also used in Chapter 3.

Dataset B is obtained using the microscope-camera; the VAST unit manipulates the position of the specimen while, in this case, the microscope-camera acquires bright-field images. For the acquisition the objective $2.5 \times /0.07NA$ is used. *Dataset B* consists of a set of representative examples, i.e., 3 specimens of 5 dpf. Examples of the images are depicted in Fig. 4.6.

Dataset C represents a collection of images of an internal structure of the zebrafish, i.e., an organ. These images are obtained using the fluorescence imaging modality as the organ is specifically visible with fluorescence. For fluorescence we use the objective $4 \times /0.12NA$ to collect the dataset for the zebrafish liver. We used 7 zebrafish samples of 3 dpf. The microscope is equipped for fluorescence and the images are acquired with the microscope-camera. Examples of the images are depicted in Fig. 4.7; in Appendix A the preparation of these samples is given.

4.4 Two-phase 3D reconstruction from axial-views

In this section we elaborate on the 2-3DLA approach. For the 3D reconstruction using axial-views, a microscope camera calibration is necessary. This can be solved by the algorithm of voxel residual volume maximization. As a baseline method, the shape-based 3D reconstruction is used for comparison with our method. For the details of the microscope camera calibration and the shape-based 3D reconstruction method, we refer to Section 3.3. Below we discuss: the improved 3D volumetric representation as confidence map in Subsection 4.4.1 and 3D reconstruction as objective function optimisation in Subsection 4.4.2.

4.4.1 Improved 3D volumetric representation as confidence map

In shape-based 3D reconstruction we need accurate segmentations of the original images. These are, however, not always satisfactory. One of the important complications is the translucency and transparency of the specimen in the images. We, therefore, propose the improved 3D volumetric representation in the form of a confidence map which is derived from the probabilistic models as estimated from the object and background presented in the axial-view images.

We observe that the partial transparency of the specimen is difficult to recover. Severe texture variation within the object challenges the generality of the probabilistic models. We therefore need to collect prior knowledge on the approximate locations of the objects of interest and background. It can be implemented in an interactive manner. In this case, the results of the 3D reconstruction will rely on the quality of the user input; insufficient user input tends to produce an underestimated model.

In following, we propose to apply the MS algorithm to obtain the approximations of the 2D shapes for the specimen. The MS algorithm improves the texture representation of the object so that the translucent and transparent regions of the specimen become more separable from the background. We threshold the texture-augmented images to obtain the 2D shape approximations. This also results in an augmentation of the texture representation in the transparent parts of the specimen. The approximations of the 2D shapes can roughly separate the object and background. Although such shape approximations are inaccurate, morphological operations to the envelope of the shape (dilation and erosion) enforce the discrimination of the object and background. In this manner, we have solved the problem of texture augmentation for the partial transparent specimens. In addition, the obtained 2D shape approximations can serve as the shape constraints, such that almost all texture information included by the object can be sampled. This improves the generalisation of the probabilistic models and contributes to the successful automation of the whole system.

We combine the multi-scale textures interpreted as image pyramids to further enhance the probabilistic models. Now, let $\mathcal{I}_{sc} = \{\mathbf{I}_{(1,sc)}, \mathbf{I}_{(2,sc)}, ..., \mathbf{I}_{(N,sc)}\}$ and $\mathcal{S} = \{\mathbf{S}_1, \mathbf{S}_2, ..., \mathbf{S}_N\}$ denote the images (after mean shift filtering) in scale of sc and the approximations of 2D shapes in original scale, where N represents the number of views. In any of the views, the foreground and background can be approximately discriminated and indexed by $\mathbf{x}_f$ and $\mathbf{x}_b$, where $\mathbf{S}(\Omega) \approx \{\mathbf{x}_f, \mathbf{x}_b\}$

and Ω represents the whole image domain. Then, the multi-scale textures are represented as $\mathbf{I}_{sc}(\mathbf{x}_f)$ and $\mathbf{I}_{sc}(\mathbf{x}_b)$, where $\mathbf{I}_{sc}(\mathbf{x}) = (r,\ g,\ b)^T$ denote the RGB values in colour space for any pixel in scale sc. Next, a multivariate Gaussian is applied to estimate the probabilistic models of the range images for each specific scale. Thus:

$$\begin{aligned} p_{(f,sc)}(\mathbf{x}_f) &= \frac{1}{\sqrt{(2\pi)^3|\mathbf{\Sigma}_f|}} e^{-\frac{1}{2}(\mathbf{I}_{sc}(\mathbf{x}_f)-\mu_f)^T\mathbf{\Sigma_f^{-1}}(\mathbf{I}_{sc}(\mathbf{x}_f)-\mu_f)} \\ p_{(b,sc)}(\mathbf{x}_b) &= \frac{1}{\sqrt{(2\pi)^3|\mathbf{\Sigma}_b|}} e^{-\frac{1}{2}(\mathbf{I}_{sc}(\mathbf{x}_b)-\mu_b)^T\mathbf{\Sigma_b^{-1}}(\mathbf{I}_{sc}(\mathbf{x}_b)-\mu_b)} \end{aligned} \tag{4.1}$$

where $p_{(\bullet,sc)}$ represents the probabilistic model estimated for the textures in scale sc; and $(\mu_f,\ \mathbf{\Sigma}_f)$ and $(\mu_b,\ \mathbf{\Sigma}_b)$ denote the mean vectors and the covariance matrices estimated for the foreground and background, respectively. To avoid overestimation of the probabilistic models, we randomly select a subset of the axial-views for computation.

A voxel $\mathbf{X} \in \mathcal{X}$ in 3D space corresponds to a pixel location in $\mathbf{I}_{(i,sc)}$ (the i^{th} image in scale sc) through the pinhole camera model $\mathbf{x} = \mathbf{P}_i\mathbf{X}$, where $\mathbf{P}_i \in \mathcal{P}$ is the camera projection matrix estimated for each view. We can obtain the probabilities of this voxel being the foreground and background by $p_f(\mathbf{x})$ and $p_b(\mathbf{x})$. Suppose that all the N axial-view images $\mathcal{I}_{sc}$ in scale sc are independent, the joint probabilities of a voxel $\mathbf{X}$ indicating its visibility to the foreground and background can then be modelled as follows:

$$\begin{aligned} P_{(f,sc)}(\mathbf{X}) &= \left(\prod_{i=1:N} p_{(f,sc)}(\mathbf{x})\right)^{\frac{1}{N}} \\ P_{(b,sc)}(\mathbf{X}) &= 1 - \left(\prod_{i=1:N} (1 - p_{(b,sc)}(\mathbf{x})\right)^{\frac{1}{N}} \end{aligned} \tag{4.2}$$

We can interpret Eq. (4.2) as the intersection of the N camera projections for each voxel. We use the logarithm to re-scale the two probabilities. The multi-scale probabilistic models are then fused to obtain the confidence map defined as:

$$\mathbf{I}^* = \sum_{sc=1}^{M} \beta_{sc} \left[log(P_{(f,sc)}) - log(P_{(b,sc)}) \right] \tag{4.3}$$

where $\beta = \{\beta_{sc}\}_{sc=1:M}$ specifies the weight for each scale and $\mathbf{1}^T \cdot \beta = 1$. In our specific implementation, a uniform distribution of β is used, which simply averages the multi-scale probabilities. We use three-scale textures ($M = 3$) in a hierarchical fashion, such that the higher scales can be generated by halving the images in the respective lower scales. The confidence map takes a confidence score as the entry for each voxel to indicate its likelihood to be the object (or background). The larger the confidence score of a voxel is, the more possible that voxel is classified as the object.

4.4.2 3D reconstruction as objective function optimisation

With the confidence map, we can accomplish the 3D reconstruction by searching for an enclosed surface which is able to correctly classify all the voxels into the object and background. The marching cubes algorithm [121] straightforwardly estimates a specific surface. However, it is difficult to determine which is the optimal one; and serious fractal behaviour of the estimated surface will occur. This will result in inaccurate 3D measurements of volume and surface area. An empirical solution can be applied to cascade a surface refinement module which needs to be carefully validated. In this work, we apply the region based level set method (CV model) [41] to accomplish this task. The employment of the CV model transfers the 3D reconstruction problem into a 3D segmentation. The optimal enclosed surface retaining a smooth appearance is achieved.

According to the CV model, an enclosed 3D surface $\mathcal{C}$ is embedded in a distance regularised level set function (LSF) $\mathcal{C} := \{\mathbf{X} \in \Omega; \phi(\mathbf{X}) = c\}$, where Ω and $\mathbf{X}$ are consistently defined as the image domain (3D in this case) and the spatial locations of the voxels in world frame. To separate the object and background according to the similarity in the confidence map we define the objective function Eq. (4.4). This includes the external force as the confidence map and the internal force as the smooth appearance of the surface. Thus:

$$
\begin{aligned}
E(\phi) \;=\;& \underbrace{\mu \int_{\Omega} |\nabla H(\phi)| d\mathbf{X}}_{\text{internal force}} \\
&+ \underbrace{\upsilon \int_{\Omega} \left[|\mathbf{I}^{*} - u_{in}|^{2} H(\phi) \;+\; |\mathbf{I}^{*} - u_{out}|^{2} (1 - H(\phi)) \right] d\mathbf{X}}_{\text{external forece}}
\end{aligned}
\tag{4.4}
$$

where, u_{in} and u_{out} represent the mean intensity of the confidence map in and outside of the surface; μ and υ are constants used to balance different terms. H is the Heaviside function defined as follows:

$$
H_{\epsilon}(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases} \tag{4.5}
$$

Using the Euler-Lagrange equation, as:

$$
\frac{\partial \phi}{\partial t} = -\frac{\partial E}{\partial \phi}, \tag{4.6}
$$

one can implement the surface evolution by computing the gradient flow for the objective function defined as:

$$
\frac{\partial \phi}{\partial t} = \delta(\phi) \Big\{ \mu \; div \left(\frac{\Delta \phi}{|\Delta \phi|} \right) + \upsilon \left[\left(|\mathbf{I}^{*} - u_{out}|^{2} - |\mathbf{I}^{*} - u_{in}|^{2} \right) \right] \Big\}, \tag{4.7}
$$

where $\delta(\phi)$ is the derivative of $H(\phi)$. By setting a proper step size Δt, we can complete the 3D reconstruction via gradient descent, defined as:.

$$
\phi_{t+1} = \phi_{t} + \Delta t \frac{\partial \phi}{\partial t} \tag{4.8}
$$

3D measurements inference Given the 3D reconstruction represented as the optimal surface embedded in ϕ, we can derive the 3D measurements. The volume is obtained by the integration over all the voxels which are included in the object ϕ: $V = \sum_{\phi} \mathbf{1}[\phi > c^{*}] \times v$, where $\mathbf{1}[\bullet]$ keeps its representation as an indicator function; c^{*} is the optimal c-level set which we will investigate in following sections; v is the unit volume for the voxels which is pre-defined. A set of surface points are generated from the c^{*}-level set of ϕ, which will be used to produce a triangulated mesh. The surface area is obtained by the integration over all the facets in the triangulated meshes using Heron's formula $A = (s(s-a)(s-b)(s-c))^{1/2}$, where $s = (a+b+c)/2$ and $a,\ b,\ c$ represent the edges of a triangle.

4.5 Experiments

In this section, we describe our experiments to evaluate the performance of the proposed approach for our datasets. In Subsection 4.5.1, we perform a visual inspection of the results on *Dataset A*. In Subsection 4.5.2, we discuss how to obtain accurate 3D measurements for *Dataset A*. In Subsection 4.5.3, we perform visualisation and 3D measurements on *Dataset B* & *C*. In Subsection 3.5.4, we evaluate the proposed method by runtime.

4.5.1 Visual inspection of the results on *Dataset A*

We first apply (A) different methods on the *Dataset A* to compare the obtained 3D models through visualisation. We next (B) apply our 2-3DLA approach on the whole *Dataset A* to demonstrate its performance on the zebrafish in various developmental stages.

(A) Comparison with various methods using visualisation

We implemented an interactive method [81] which requires user input to indicate the approximate locations of the object and background. So, the quality of the 3D models will rely on the user input. The shape-based 3D reconstruction is also used in this comparison.

In Fig. 4.2 (A), the results obtained by the interactive method are shown. Two types of user input are employed to estimate the probabilistic models for each specimen. One input collects less textures by drawing shorter scribbles on the object. The other input covers a larger region of the object, thus collecting more texture information. In Fig 4.2 (A), the corresponding 3D reconstructions are shown next to the input drawings. Our assumption is confirmed that variations in user input impact the quality of the 3D reconstructions. Sufficient texture sampling usually results in better 3D reconstructions preserving most of the shape of the specimen. However, even with "good" user input, parts of the zebrafish, mostly at the tail area, still fail to be faithfully reconstructed. The main reason is the considerable texture inhomogeneity of the specimen which hampers the

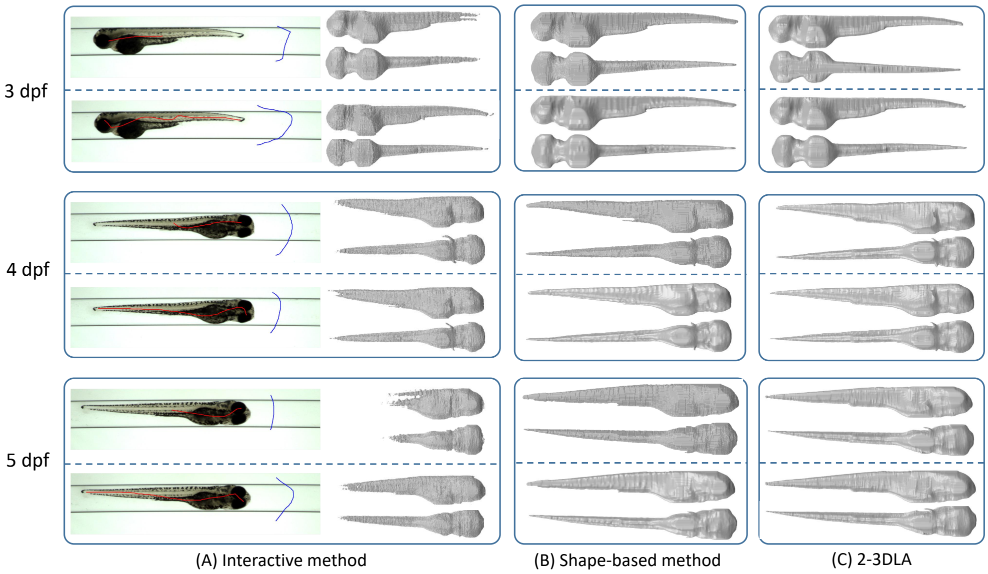


Figure 4.2: Comparison of visualised results of different methods. We select one example from each larval stage, i.e., 3, 4 and 5 dpf. For each 3D model, we visualise the same perspectives, i.e., lateral and ventral. (A) The 3D reconstruction of the interactive method. The users randomly draw on the object of interest (red scribble) and background (blue scribble), respectively. Two types of user input are considered, as shown on the left. Corresponding 3D reconstructed models are shown on the right. (B) The 3D reconstruction of the shape-based 3D reconstruction. The raw 3D models are shown in the top two rows. The 3D models with extra surface refinement are shown below. (C) The 3D reconstructions of the 2-3DLA using only the original-scale textures and multi-scale textures are separately shown in each box above and below.

generality of the probabilistic models and thereby results in a misclassification of the translucent part of the specimen as background. From the results of the interactive method, we can also observe the fractal behaviour of the surface. It is apparent that inaccurate 3D reconstruction will result in inaccurate 3D measurements.

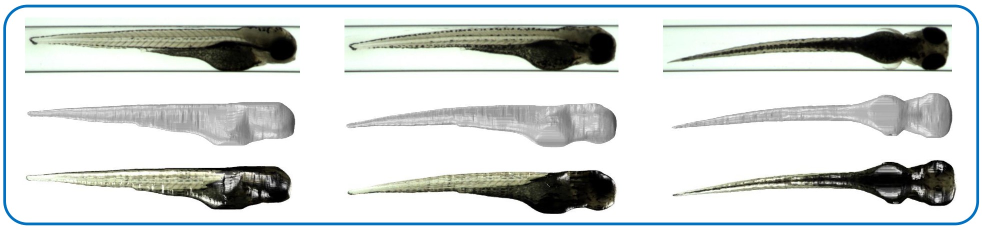
In Fig. 4.2 (B) we show results of the shape-based method with a set of accurate 2D shapes, which generates very accurate 3D shape representation of the specimen. Each shape shows a natural and detailed surface, but fractal behaviour also occurs due to the carving effects of the method. This is depicted in the top two rows of Fig. 4.2 (B). After extra surface refinement [93], accurate 3D measurements can be available. The refined 3D models are shown in the bottom two rows of Fig. 4.2 (B). We need to accept that the shape-based method requires accurate 2D segmentations from the original images which are sometimes difficult to obtain in light microscopy imaging.

In Fig. 4.2 (C) we show the 3D models obtained by the 2-3DLA system. In the top two rows, the 3D reconstructed models are obtained by the confidence map derived with one-scale texture. In the bottom two rows, the results obtained with three-scales texture are shown. One can observe a better 3D visualisation for both of the results. In the 2-3DLA, the employment of the MS algorithm ensures improved discrimination between the object and background. Furthermore, the shape constrained texture collection augments the generality of the confidence map. These improvements cooperate in obtaining a better volumetric representation. The CV model aims to search for the optimal surface which can largely separate the object and background in the confidence map and preserves a smooth appearance of the specimen, which subsequently results in accurate 3D measurements. However, a little fractal behaviour is still showing in the results with one-scale texture, resulting in a slightly deficient tail region of the zebrafish.

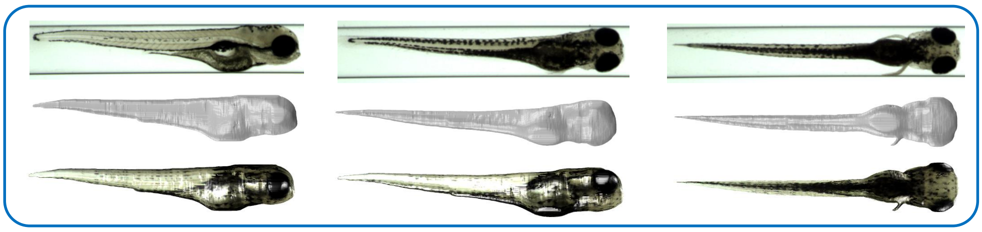
(B) 3D visualisation for *Dataset A* using 2-3DLA

Here we apply the 2-3DLA on the whole *Dataset A*. In Fig. 4.3 some results are depicted. For each developmental stage of the zebrafish, we randomly select one specimen in the dataset. To create additional value for 3D models, we map the textures from the original images onto the surface of the 3D models to produce texturised 3D shapes. For practical reasons, the texturing of the partial transparent regions is artificial. We separately show the binary and the texturised 3D models in the middle and bottom rows for each specimen. We select three typical axial-views, i.e., lateral, tilted, and ventral.

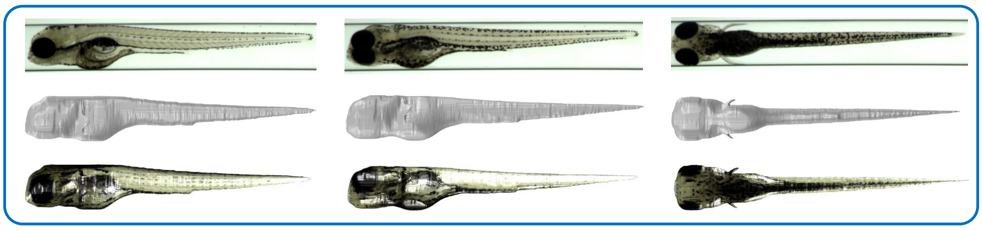
From the visualisation, we can assess that the 2-3DLA obtains quite natural and vivid 3D shapes for the specimens. The problem arising from the translucency and transparency seems to be solved. From the 3D shapes it is obvious that older zebrafish are larger compared to the younger ones. The size of yolk in the various stages is different. During the development, the yolk of the zebrafish provides nutrient for its growth. Consequently, older zebrafish have smaller yolk, simply as a result of their growth. Furthermore, the 3D shape obtained by our approach offers a good framework for the visualisation of the biological system by introducing organ-level 3D systems, the circulatory system, the liver, the brain, etc.. Using specific staining, the 3D systems can be incorporated and visualised. The whole



(A) 3 dpf



(B) 4 dpf



(C) 5 dpf

Figure 4.3: 3D models visualisation of the zebrafish in various larval stages from *Dataset A*. Each bounding box denotes a staged zebrafish. For each specimen, three typical perspectives are shown in different columns. The original axial-view images are shown in the top row; the pure 3D models are shown in the middle row; the texturised models are shown in the bottom row.

process is reviewed in a 2-3DLA spotlight including animated visualisations of the 3D models at `http://bio-imaging.liacs.nl/galleries/VAST-2-3DLA/`.

4.5.2 3D measurements for *Dataset A*

We investigate how to obtain accurate 3D measurements from the 3D models obtained by the 2-3DLA. To this end, the hyperparameter of c-level set needs to be tuned, because an unvalidated c-level set may lead to inaccurate 3D measurements. We present two interpretations, (A) the 2D shape coherence and (B) the 3D shape coherence, for the inference of accurate 3D measurements. We compute the (C) 3D measurements based on the inference for *Dataset A*.

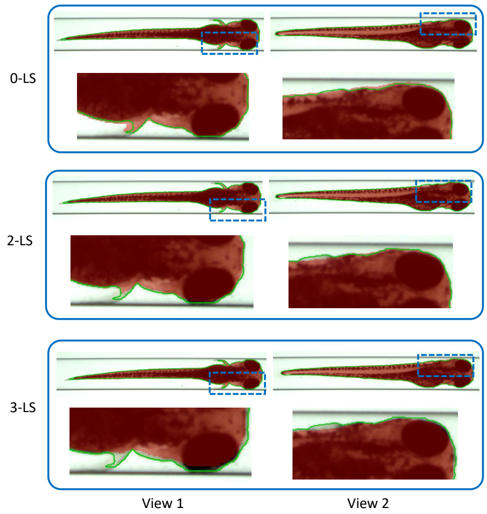


Figure 4.4: 2D shape coherence (red) between the projections from 3D models to axial-view images and the groundtruth 2D shapes (green contours) with various c-level sets. A subsection of the original 2D shape coherence (dashed bounding box) is zoomed in and shown at the bottom of each row.

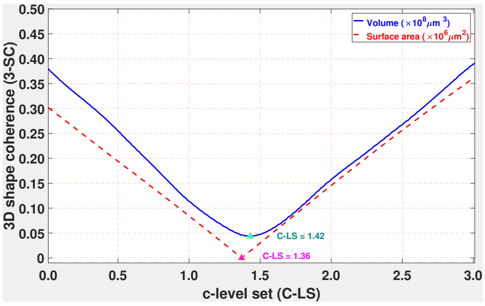


Figure 4.5: 3D shape coherence with various c-level sets. The triangles indicate the minimum of the measurements, which are found around 1.4-level set.

(A) 2D shape coherence

A good 3D reconstruction must approximate the real shape of the object. We evaluate and interpret this as shape coherence. We define the 2D shape coherence as a disparity map from the projected shape of a 3D model to a "groundtruth 2D shape" in each original axial-view. We have developed a robust method for the segmentation of zebrafish larvae. Using manual segmentations, the method is validated and turned out to be accurate. We thereby use the segmentation results as the approximation of the groundtruth 2D shape for the *Dataset A*. Accordingly, we interpret the 2D shape coherence as the overlap from the projected shape to the groundtruth 2D shape. We chose three different c-level sets (c-LS), i.e., 0-LS, 2-LS, and 3-LS, to extract the corresponding 3D models. The 2D shape coherence that are obtained in this manner are shown in Fig. 4.4. We depict two typical axial-views of the object, i.e., ventral and lateral. We represent the groundtruth 2D shapes as green contours. In order to appreciate the results, a subsection of the 2D shape coherence is zoomed in and its visualisation is shown in the bottom row for each view.

We observe that a small c-LS, i.e., 0-LS, produces an overestimation of the 3D model, which presents a large 2D shape coherence. This is illustrated as the red area exceeds the green contours in the top row of Fig. 4.4. A large c-LS, i.e., 3-LS, results in a more compact 3D shape of the object, which obtains a smaller 2D shape coherence. It is illustrated as the red area separating from the green contours in the bottom row of Fig. 4.4. In contrast, a reasonable c-LS, i.e., 2-LS, obtains highly correlated 2D shape coherence producing a better 3D reconstruction.

(B) 3D shape coherence

The shape-based 3D reconstruction method results in a binary volumetric representation. If accurate 2D shapes are used in this method, the obtained 3D model can serve as the "groundtruth". The 3D shape coherence is then defined as the disparity from the 3D models obtained by the 2-3DLA to the "groundtruth 3D model". We propose two measurements to evaluate the 3D shape coherence: volume disparity ($\overline{V}$) and surface area disparity ($\overline{A}$). They are separately defined as $\overline{V} = [(1/M)\sum_i (V_i^* - V_i^m)^2]^{(1/2)}$ and $\overline{A} = [(1/M)\sum_i (A_i^* - A_i^m)^2]^{(1/2)}$, where V^* and A^* are the volume and surface area measured from the 3D model groundtruth; V^m and A^m are the volume and surface measured from the 3D models obtained using a specific c-LS; and M is the number of specimens. We use half of the *Dataset A* for the validation of this experiment. In Fig. 4.5, we show a graph of the 3D shape coherence for various c-LS. We illustrate that the choice of a small c-LS, e.g., 0-LS, tends to produce an overestimation of the measurements. The choice of a large c-LS, e.g., 2-LS, will lead to an underestimation. This corresponds nicely with the findings in the 2D shape coherence as shown in Fig. 4.4. Finally, we find that the 1.4-LS is a good compromise and gives the best estimation of the volume and surface area. For the next paragraph (4.5.2 (C)), we use the 1.4-LS to obtain the deterministic 3D models from which we can derive the 3D measurements.

(C) 3D measurements for *Dataset A*

We use the remaining half of the *Dataset A* to compute the 3D measurements of the zebrafish in various developmental stages, i.e., 3, 4, and 5 dpf. In this experiment, we compute the statistics for the volume and surface area of the zebrafish and show the results in Table 4.1 and 4.2. The 3D measurements obtained from

Table 4.1: Volume statistics ($\times 10^8 \mu m^3$) of the 3D models for the zebrafish in various developmental stages

	Shape	2-3DLA
3 dpf	2.53 ± 0.11	2.56 ± 0.12
4 dpf	2.63 ± 0.19	2.66 ± 0.19
5 dpf	3.00 ± 0.18	2.96 ± 0.18

Table 4.2: Surface area statistics ($\times 10^6 \mu m^2$) of the 3D models for the zebrafish in various developmental stages

	Shape	2-3DLA
3 dpf	3.20 ± 0.15	3.21 ± 0.14
4 dpf	3.34 ± 0.17	3.36 ± 0.17
5 dpf	3.63 ± 0.14	3.61 ± 0.14

the shape-based 3D reconstruction method are used to assess the performance of the 2-3DLA.

From the tables we may conclude that the 2-3DLA can obtain accurate 3D measurements of volume and surface area comparable to the groundtruth method. This holds for the zebrafish in the three developmental stages. The results of the 2-3DLA are slightly larger, but the difference is small and acceptable. One should note that the 2-3DLA will be advantageous for the cases in which accurate segmentations are not available. Furthermore, the increasing trends of the volume and surface area for the zebrafish staged from 3 dpf to 5 dpf is consistent with the expected physical growth of the zebrafish. Importantly, Table 4.1 and 4.2 offer baseline metrics for volume and surface area of the zebrafish in various stages; this facilitates quantitative and qualitative analysis using the zebrafish as model system.

4.5.3 3D reconstruction and measurements of *Dataset B & C*

We have shown that the 2-3DLA can obtain an accurate 3D shape description of the whole specimen offering the baseline reference for the 3D measurements. This is, however, not sufficient to evaluate a specimen on the level of organs. From the optics used, the resolving power of *Dataset B & C* (see Section 4.3) is much higher. Moreover, different microscope modalities are used for these datasets, i.e., bright-field and fluorescence. As a case study, in addition to the whole specimen, we use the zebrafish liver as a model organ system. The liver size is important to pharmacokinetics, as it is an crucial organ for drug metabolism. For *Dataset B & C*, 3D measurements of volume and surface area are separately reported in Table 4.3 and 4.4. As a comparison, we still use the results obtained from the shape-based 3D reconstruction approach.

Table 4.3: Volume ($\times 10^8 \mu m^3$) and surface area ($\times 10^6 \mu m^2$) of the 3D models for the zebrafish in *Dataset B*

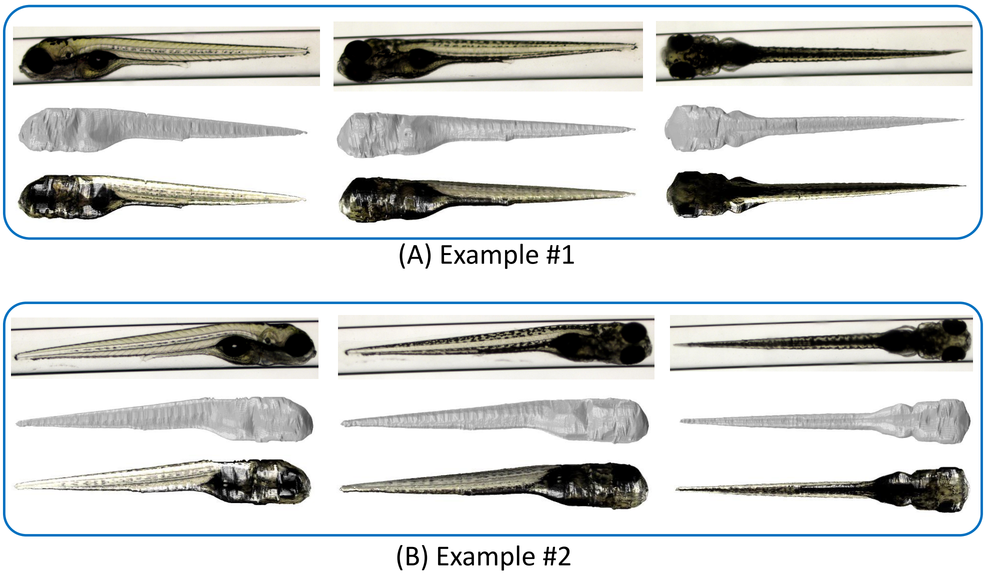
	Volume		Surface area	
	Shape	2-3DLA	Shape	2-3DLA
#1	3.00	3.05	3.61	3.65
#2	2.95	3.04	3.52	3.60
#3	3.06	3.13	3.69	3.74

Table 4.4: Volume ($\times 10^5 \mu m^3$) and surface area ($\times 10^4 \mu m^2$) of the 3D models for zebrafish liver in *Dataset C*

	Volume		Surface area	
	Shape	2-3DLA	Shape	2-3DLA
#1	7.70	7.56	4.67	4.57
#2	5.38	4.15	3.61	3.15
#3	8.10	8.47	4.91	5.04
#4	9.06	9.28	5.20	5.18
#5	11.6	9.77	6.07	5.49
#6	15.1	16.6	6.89	7.69
#7	6.55	6.79	4.28	4.31

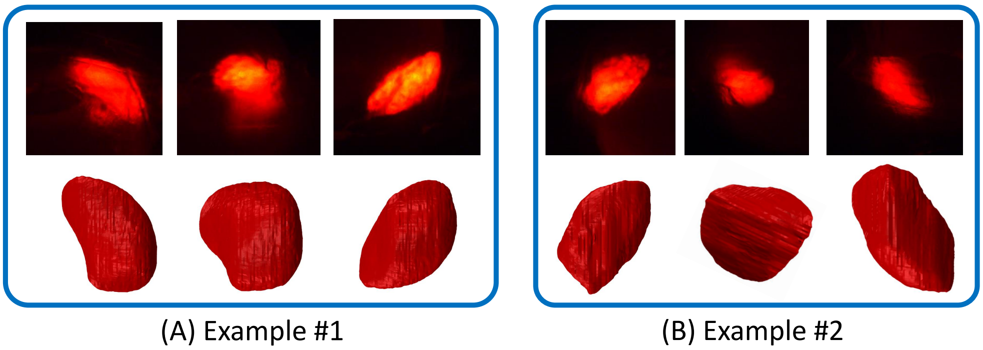
In Table 4.3 we observe only a small difference of the volume and surface area produced by the 2-3DLA compared to baseline method. This difference is acceptable. It is apparent that both of the measurements are within the distributions given in Table 4.1 and 4.2. This suggests that the 2-3DLA system shows stable performance for bright-field microscopy axial-view imaging. Subsequently, it can obtain accurate 3D measurements for the specimen as a whole. In Fig. 4.6, we show the 3D models for two specimens. The results are presented in the similar fashion as used in Fig. 4.3. An interesting phenomenon shown in the Example #1 is that the tail of the 3D model is thinner than the original specimen. The reason is that, in this experiment, the specimen were alive and anesthetised. During the imaging, the motion of the positioning capillary introduces slight shape deformations of the specimen. This impedes accurate 3D reconstruction which requires static objects.

In Table 4.4 we can find that most of the zebrafish livers are reconstructed well and comparable 3D measurements are obtained. Significant difference for the 3D measurements are found in example #2. Both the volume and surface obtained by the 2-3DLA are smaller than those obtained by the baseline method. This


(A) Example #1

(B) Example #2

Figure 4.6: Visualisation of 3D models of the zebrafish in bright-field microscopy from *Dataset B*. For each specimen, three typical axial-view images are shown on the top; the corresponding axial-views of the pure and texturised 3D models are shown in the middle and bottom, respectively.


(A) Example #1

(B) Example #2

Figure 4.7: Visualisation of 3D models of the zebrafish liver in fluorescent microscopy from *Dataset C*. We show three axial-view images and the corresponding 3D perspectives in the top and bottom row for each specimen.

is caused by the image quality for the example #2. We visualise the 3D model for this example in Fig. 4.7 (B). We observe that the original axial-view images do not clearly represent the specimen. A reason for the obfuscation may be the relatively small external force in Eq. (4.7) for all the specimens, which produces compact 3D models. In fact, using a different configuration for the parameter in example #2 will result in accurate 3D measurements. However, in order to ensure

Table 4.5: Performance evaluation on runtime (s=second,min=minute)

	Step A	Step B	Step C	Step D	Total
Dataset A	$36 \pm 1.4(s)$	$12.6 \pm 0.8(s)$	$35.9 \pm 0.9(s)$	$2.2 \pm 0.5(min)$	$\sim 3.6(min)$
Dataset B	$4.3 \pm 0.2(min)$	$56.4 \pm 0.9(s)$	$37.2 \pm 0.4(min)$	$7.9 \pm 0.5(min)$	$\sim 50.3(min)$
Dataset C	$5.5 \pm 0.6(min)$	$39.0 \pm 7.4(s)$	$8.8 \pm 1.2(min)$	$12.1 \pm 1.1(min)$	$\sim 27.1(min)$

- Step A: Image capture
- Step B: 2D shape approximation
- Step C: Camera system calibration
- Step D: 3D reconstruction

a justified evaluation of this experiment, we reported the results obtained with the same parameters. In future work, we can validate the parameter specifications in relation to image quality.

4.5.4 Evaluation on efficiency

For the evaluation of the computational efficiency of our approach, we first specify the setup. We represent the process of (1) image capture, (2) 2D shape approximation, (3) camera system calibration, and (4) the 3D reconstruction as Step A, B, C, and D, respectively. In Step A, we use the VAST BioImager to automatically obtain *Dataset A*. A well-trained biologist used the microscope setup (Leica DMRB) to acquire *Dataset B & C*. In this data acquisition, the object is manipulated by the VAST BioImager to present in the view of the microscope. The remaining steps are implemented with the Matlab platform (and partial C/C++ implementation) on a desktop equipped with a CPU i7 and 16G RAM. In Step B, we heuristically configure the filter width in spatial, colour feature space as $(h_s, h_r) = (10, 20)$ for *Dataset A*; and $(h_s, h_r) = (20, 30)$ for *Dataset B & C*. In Step C, we find that the Nelder-Mead simplex method works much faster for *Dataset A & C*. The evolution strategy achieves robust performance for *Dataset B*, though it is not efficient. We take a compromise on the optimisers and we apply the Nelder-Mead simplex method for *Dataset A & C* and the evolution strategy for *Dataset B*. In phase II of our 2-3DLA, we set the parameters $\mu = 10$, $\upsilon = 1$, and $\Delta t = 0.5$ for all cases.

In Table 4.5, for each step from A to D, we report the performance evaluation in terms of runtime in minutes or seconds for the three datasets. We compute the average and standard deviation of the runtime for each specimen in the datasets. First, from the results, we may conclude that the performance of our approach on *Dataset A* is most efficient. In contrast, the larger size of the images in the *Dataset B* and *C* results in an expensive computation for our approach. Second, Step B performs very efficiently in the three datasets due to our fast implementation of the MS algorithm. Third, Step C in *Dataset B* and *C* requires a large amount of computations. The reason is that during the camera system optimisation, massive evaluations for the cost function are executed. However, a better optimiser and a high-performance scheme such as parallelisation or GPU programming can be employed to improve the efficiency of this step. Fourth, Step D in the three datasets seems be computationally expensive. One reason is the iterative 3D surface evolution. We set up a restricted termination criterion that requires the change of the cost function to be close to zero. So, this results in more iterations. In addition, we set a small grid size for the voxel space. This leads to massive projections and demands a large amount of memory to fulfil a precision guarantee. Again, this can, to a certain extent, be solved by the introduction of high-performance computing strategies.

4.6 Chapter conclusions and future work

Imaging of partially transparent objects impedes the visualisation of a continuous surface which renders it difficult to estimate a depth map of the object. To handle this problem, we first present three typical datasets using our MM-HTAI architecture. We then have developed the 2-3DLA approach for 3D reconstruction and measurement in light microscopy axial-view imaging. It answers RQ 4: *How can we efficiently deal with the translucency and transparency of specimen in light microscopy and still obtain a good 3D shape description from the MM-HTAI architecture*?

More concrete, in phase I of our 2-3DLA approach, we estimate an improved 3D volumetric representation as a confidence map from a range of texture-augmented images by applying the MS algorithm on the original images. In phase II, we adapt the region based level set method to estimate the optimal enclosed surface for the object which balances the 3D shape integrity and the smoothness of appearance for the 3D model. Compared with the shape-based method, the proposed 2-3DLA approach obtains comparable 3D reconstructions and measurements, but it does not require accurate 2D segmentations of the original images. As segmentation of partial transparent objects can be difficult, this system provides a good solution to this shortcoming. Unlike the shape-based 3D reconstruction method, the 2-3DLA does not need extra refinement to obtain a natural and smooth 3D surface. We have successfully applied our approach for zebrafish analysis in multi-modal light microscopy axial-view imaging including bright-field and fluorescence. The obtained results can be directly used for the model system evaluation in the fields of toxicology, infectious diseases and oncology.

Still, we admit that the efficiency of our approach needs improvement with respect to larger images. A possible solution can be a dynamic programming scheme for high-performance parallelisation. As is well known, the zebrafish model is intensively used in the field of life-sciences. Therefore, in our 2-3DLA we can learn a good probabilistic model for the zebrafish either in supervised or semi-supervised way [122]. Yet, our approach needs the support from 2D shape approximation in the process of camera system calibration. In future research, we thus aim for a method which is independent on image segmentation. In addition, other multi-view imaging modes are investigated by our method. However, the axial-view imaging is the most commonly used modality, so the 2-3DLA can be considered generic for other specimen observations.

Appendix A

The zebrafish used in *Dataset C* are from a transgenic line Tg(lfabp:dsRed; elaA:EGFP) with 2 fluorescent colours for liver and pancreas (2CLIP). The eggs were kept in 60 $\mu g/mL$ Instant Ocean Sea Salts (Sera Marin, Heinsberg Germany) in demineralised water and treated with 0.003% 1-phenyl-2-thiourea (PTU, Sigma-Aldrich, Zwijndrecht, The Netherlands) to prevent pigmentation.

Chapter 5

Multi-modal 3D Reconstruction

Based on:

- Y. Guo, R.C. van Wijk, E.H.J. Krekels, H.P. Spaink, P.H. van der Graaf & F.J. Verbeek, “Multi-modal 3D reconstruction and measurements of zebrafish larvae and its organs using axial-view microscopy,” in IEEE Conference on Image Processing, Beijing, China, 2017.

This chapter addresses RQ 5.

> **RQ 5**: *How can we obtain a multi-modal 3D description and the corresponding measurements for the zebrafish from the MM-HTAI architecture*?

Abstract – In life sciences, light microscopy is used to study specimens. On the organism-level a bright-field representation presents an overview for the whole shape of a specimen; the organ-level fluorescent staining representation insightfully supports in the interpretation of the detailed intrinsic structures (see Section 5.1). We apply our MM-HTAI architecture presented in Chapter 3 (see Subsection 3.3.1) to acquire axial-view images for the organism and organs of zebrafish larvae (See Subsection 5.2.1). We obtain multi-modal 3D reconstruction using the shape-based method, from which we can derive the 3D measurements of volume and surface area (see Subsection 5.2.2). In this method, we employ a microscope camera calibration using voxel residual volume maximization algorithm. We intuitively align and fuse the obtained multi-models (see Subsection 5.2.3). Experimental results show natural visualisation both for the whole organism and organ of zebrafish larvae; subsequently accurate 3D measurements are obtained (see Section 5.3). The method is very suitable for high-throughput research in which knowledge on size and shape is relevant to the understanding for development, effects of compounds or drugs (see Section 5.4).

5.1 Multi-modal 3D reconstruction

In modern life-sciences research, e.g., developmental biology, (patho)physiology, toxicology, and pharmacology, light microscopy is commonly used to produce 2D colour representations of biological phenomena. Zebrafish is a popular vertebrate model organism in biomedical research because of its many advantages, among which is optical transparency at early stages [11, 123]. The organs of the transparent larvae can be studied *in vivo* through microscopy. Transgenic lines are available that express a fluorescent reporter gene in specific organs, tissues, or cell types [16]. In this way, organ development of the genetically engineered zebrafish can be visualised and monitored over time by fluorescence microscopy. It makes the zebrafish particularly suitable for large scale screening experiments using light microscopy. For the screening of large libraries of compounds for organ toxicity, such as hepatotoxicity [23], quantitative endpoints like organ size or growth retardation are required. In order to accurately evaluate the shape and size of an organ like the liver, 3D modelling of both the liver and the whole organism are required. Compared with 2D imaging, 3D measurements, e.g., size, volume and surface area, are more reliable. The organism-level imaging is an overview of the shape of the object, serving as shape reference to normalise the 3D measurements of liver into unit metrics. Our aim is to develop a method for 3D reconstruction and measurements of zebrafish larvae and its organs using axial-view microscopy.

In stereo vision, a 3D scene can be recovered by matching correlated multi-view images [30]. This is implemented by pixel-level searching [102] or salient point detection and matching [124]. However, these methods are challenged for our typical application at hand, the zebrafish. Here, partial transparency of the zebrafish complicates straightforward application of these methods as it is difficult to match the implicit surface points on the object. Based on the concept of visual hull [125], the space carving algorithm reconstructs the 3D shape from a range of 2D binary shapes [126], which is also referred to as the silhouette-based 3D reconstruction [80, 81]. In a sample population of zebrafish, a rather large colour variation exists; this holds both for bright-field and fluorescence. Therefore, we propose to use 2D binary representations. Accordingly, we have developed the profile-based 3D zebrafish reconstruction method based on a series of 2D axial-view shapes of the object, obtaining a precise 3D representation and accurate 3D measurements in various larval stages [98].

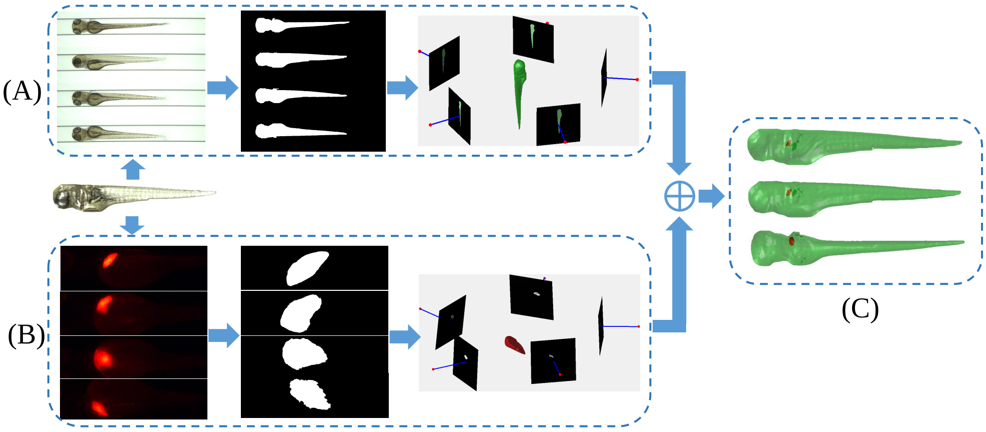
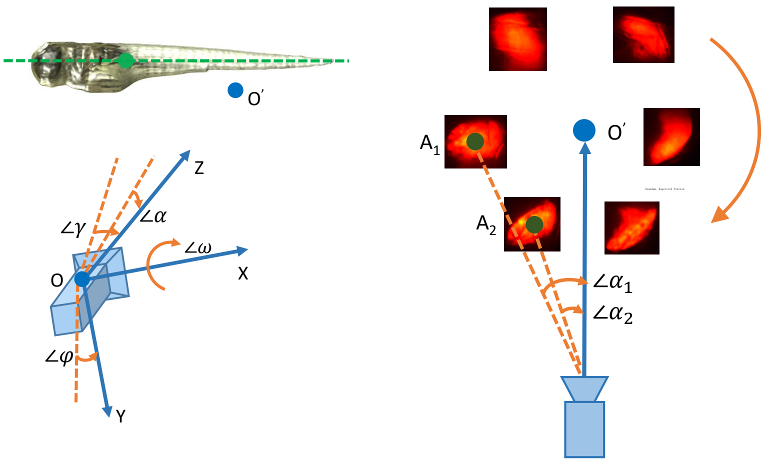


Figure 5.1: A schema of the proposed method. One zebrafish larva presents in the two imaging pipelines. (A) The organism-level 3D reconstruction. (B) The organ-level 3D reconstruction. (C) The multi-modal 3D reconstruction fusion and visualisation. In (A) and (B), the first columns show the original axial-view images; the second columns illustrate the 2D binary shapes; the third columns denote the camera system calibration and profile-based 3D reconstruction.

In Chapter 3 (see Subsection 3.3.1), we have implemented our MM-HTAI architecture. In this chapter, using the axial-view images acquired from this multi-modal imaging modality, we propose a multi-modal 3D reconstruction method as presented in Fig. 5.1. A zebrafish larva with fluorescent marker expressed in the liver is captured in two imaging modes. The VAST camera enables the organism-level imaging and a microscope camera facilitates the organ-level imaging. From the axial-view images, 2D binary shapes are obtained through modern segmentation algorithms [28, 29]. We estimate the camera projection geometry for the two camera systems using the voxel residual volume maximization algorithm. We produce the multi-modal 3D reconstruction for the organism (zebrafish) and organ (liver) using the profile-based 3D reconstruction method; we use heuristics to fuse the two models acquired from different imaging modalities. From the obtained 3D models, 3D measurements, e.g., volume and surface area, are derived.

5.2 Our approach

In this section, we present a new dataset containing the zebrafish on both organism- and organ-level imaging (see Subsection 5.2.1). We introduce the microscope camera calibration and elaborate shape-based 3D reconstruction (see Subsection



(a) (b)

Figure 5.2: (a) The VAST camera pose is modelled as the 3D transformation from the camera centre O to the object centre (green circle). The green line denotes the profile-axis along which the object revolves. $\angle\omega$, $\angle\varphi$ and $\angle\gamma$ represents the 3D rotation angles of the camera along the X, Y and Z directions, respectively. $\angle\alpha$ is modelled as the "translation angle". (b) The centre of the zebrafish liver is not aligned with the zebrafish centre which results in its rotation and revolution with respect to the zebrafish centre, such that the "translate angle" $\angle\alpha$ of the microscope camera is specified for each view.

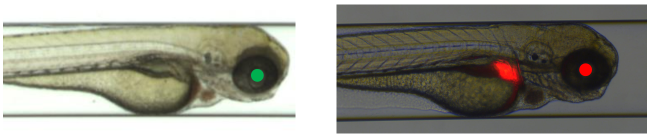


Figure 5.3: The alignment of the multi-modal 3D reconstruction is implemented according to the iris centre of the zebrafish. On the left-side is a bright-field image of organism-level imaging. We crop this image to fit the space. On the right-side is a bright-field image overlaid with a fluorescent image of organ-level imaging.

5.2.2). We present an interactive method to align and fuse the multi-modal 3D reconstruction (see Subsection 5.2.3).

5.2.1 Dataset collection

In Chapter 4, we have collected three datasets using our MM-HTAI architecture. We extend *Dataset C* in this chapter. We use the VAST camera in our MM-

HTAI architecture to obtain bright-field images presenting the whole shape of the zebrafish larvae. For the same specimens, we use the imaging modality of fluorescent microscopy in our MM-HTAI architecture to obtain images presenting the zebrafish liver. We used the imaging protocol in accordance with Chapter 4. In this manner, we obtain multi-modal images for the same group of specimens. This further facilitates our multi-modal 3D reconstruction.

5.2.2 Shape-based 3D reconstruction

In this subsection, we apply our shape-based 3D reconstruction approach to obtain the multi-modal 3D reconstruction for our specimen, i.e., the zebrafish larvae and its livers. More details for the methodology can be found in Subsection 3.3.3 and 3.3.4. We have learned that a good shape-based 3D reconstruction is conditioned to a feasible microscope camera calibration. In Subsection 3.3.5, this has been done by the voxel residual volume maximisation (VRV) algorithm. However, in this chapter, a proper adaption for the method should be made. As shown in Fig. 5.2 (a), the revolution centre of the specimen is aligned with its centre of the profile-axis. This benefits the camera system parameterisation, because "translation angle" can be shared by all the virtual cameras. In Fig. 5.2 (b), one can see that the motion of the zebrafish liver is a mixture of the rotation along the centre of the zebrafish profile-axis and the revolution along its own centre. Accordingly, we just need to specify the "translation angle" for each view in the microscope camera calibration. The optimisation for the VRV algorithm again can be solved by the Nelder-Mead simplex method [90]. In this manner, we separately perform the camera calibration in the two imaging modalities, i.e., the VAST-camera and the microscope camera.

5.2.3 3D multi-models alignment and fusion

The 3D models are obtained from different imaging modes. So, we need to align the resulting multi-modal 3D reconstruction as part of the fusion operation. During the organ-level imaging, we acquired bright-field images in register with the fluorescent images. These images only partially depict the object, but provide a good reference for alignment. We have obtained the camera poses for both models from camera calibration. So, we choose the same axial-view for the zebrafish and localise the iris centre as shown in Fig. 5.3. We use the organ-level 3D model as a template. The organism-level 3D model is scaled, rotated, and shifted to align with the former according to the camera pose and the position of the iris centre.

For a view of the zebrafish liver, we use its corresponding bright-field microscopic image to localise the iris centre for the specimen represented as $\mathbf{l}_1$; we choose the same axial-view for VAST image for the same zebrafish and localize the iris centre represented as $\mathbf{l}_2$. In this subsection, we use index 1 and 2 to separately represent the organ-level and organism-level coordinate frame. These are shown in Fig. 5.3.

According to the camera projection model $\mathbf{P}$, we can find the pixel location $\mathbf{c}$ for the world centre $\mathbf{C}$ defined in the two cases as follows.

$$\begin{cases} \mathbf{c}_1 = \mathbf{P}_1 \times \mathbf{C}_1 \\ \mathbf{c}_2 = \mathbf{P}_2 \times \mathbf{C}_2 \end{cases} \tag{5.1}$$

We can compute the relative displacement for our reference point, i.e., the zebrafish iris centre, with respect to the world centre in the image plane as $\mathbf{d}_1 = \mathbf{l}_1 - \mathbf{c}_1$ and $\mathbf{d}_2 = \mathbf{l}_2 - \mathbf{c}_2$.

Through camera calibration, we have known the accurate focal length f of the camera and the translation F along the principal line from world centre to the camera centre. Using 3D geometry, we can find the coordinates of the two reference points in 3D world frame as:

$$\begin{cases} \mathbf{L}_1 = \mathbf{d}_1 \times \frac{F_1}{f_1} \\ \mathbf{L}_2 = \mathbf{d}_2 \times \frac{F_2}{f_2} \end{cases} \tag{5.2}$$

We use $\mathbf{V}_1$ and $\mathbf{V}_2$ represent the vertices in the triangulated mesh for the 3D modellings of the organ and the organism, respectively. From the calibrated images, we can extract accurate 3D rotation $\mathbf{R}$ for the object with respect to the world centre in our chosen reference view. If we choose the world frame as the template, we need to rotate and translate the world frame for the organism to align the former frame. This is formulated as follows.

$$\mathbf{V}_{2\rightarrow 1} = \mathbf{V}_2 \times \mathbf{R}^{-1} + (\mathbf{L}_1 - \mathbf{L}_2) \tag{5.3}$$

Where $\mathbf{R}^{-1}$ denotes the inverse rotation for the organism.

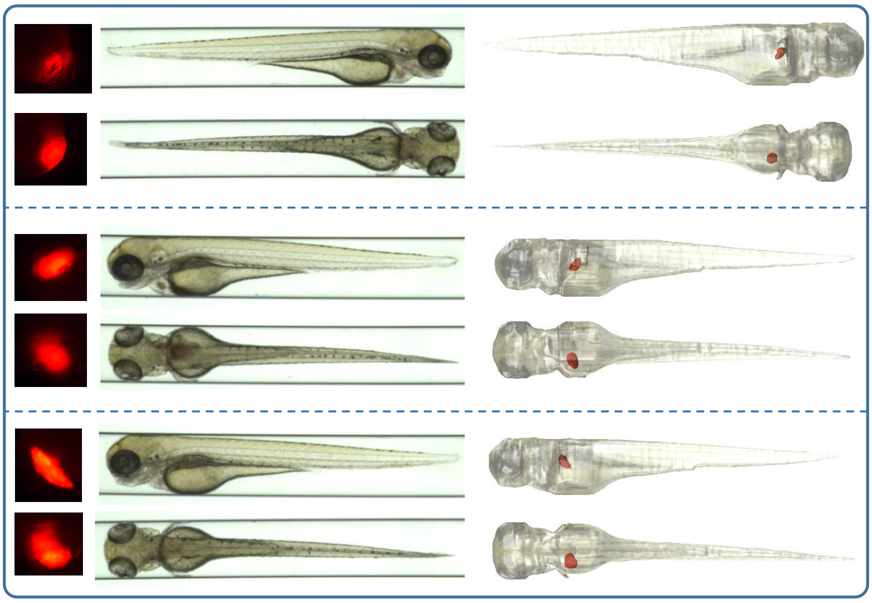


Figure 5.4: Multi-modal 3D reconstruction visualisation. We selected three examples from our dataset and for each example we visualise two typical views (lateral and dorsal). The first column represents fluorescent liver images. The middle column is the zebrafish image in bright-field. The last column visualises the fusion of the multi-modal 3D reconstruction. One can zoom in for a better observation.

Through the above process, we have obtained the aligned 3D model for the organism with respect to the organ. We then visualise the 3D multi-models $\mathbf{V}_1$ and $\mathbf{V}_{2\to1}$ in the same world frame to accomplish the multi-modal 3D reconstruction fusion.

5.3 Experiments

In this section, we apply our method on the zebrafish dataset for performance evaluation. We first visualise some examples of the multi-modal 3D reconstruction in Subsection 5.3.1, and subsequently report on the 3D measurements of volume and surface area for the zebrafish larvae and its liver in Subsection 5.3.2.

Table 5.1: 3D measurements of the 3D reconstructed models for the zebrafish (ZF) and its liver (Liver)

	Volume		Surface area	
	ZF ($\times 10^8 \mu m^3$)	Liver ($\times 10^5 \mu m^3$)	ZF ($\times 10^6 \mu m^2$)	Liver ($\times 10^4 \mu m^2$)
#1	2.74	7.70	3.33	4.67
#2	2.59	5.38	3.24	3.61
#3	2.50	8.01	3.13	4.91
#4	2.91	9.06	3.44	5.20
#5	2.67	11.60	3.31	6.07
#6	2.80	15.15	3.47	6.89
#7	2.80	6.55	3.41	4.28

5.3.1 Results visualisation

From our dataset, we select three examples for visualisation as shown in Fig. 5.4. Two typical axial-views are shown, i.e., lateral and dorsal, and each example is separated by blue lines. The first column shows the original organ-level fluorescent images. Those images depict the natural shape of the zebrafish liver. One can observe a variation in image quality from the different examples. This is caused by strength of the fluorescent marker. The middle column shows the organism-level bright-field images. The zebrafish are partially transparent but retains explicit contours for the shape. The last column visualises the fusion of the 3D models. For a natural appearance, we map the texture from the zebrafish to the 3D model. We clearly observe the shape of both the zebrafish and its liver. The visual and spatial discernibility of the models are emphasised from the multi-modal fusion. It is interesting that, although the liver is not completely visible in a all views (the first view of the first example), our method still recovers a good 3D shape by imposing a threshold to the confidence score to estimate a 3D model allowing a range of errors. An animated visualisation of the results can be found at `http://bio-imaging.liacs.nl/galleries/VAST-3Dorgan/`.

5.3.2 3D measurements for 3D multi-models

From the multi-modal 3D reconstruction, we derive 3D measurements, i.e., volume and surface area. The volume is obtained by the integration over all the voxels included in the object. A set of surface points is generated from the voxels by the marching cubes algorithm [121], from which a triangulated mesh can be produced. The obtained 3D surface is further refined [93]. Subsequently, the surface area is obtained by the integration of all the facets in the triangulated mesh using Heron's formula [92]. In Table 5.1 we report on the computed 3D measurements of volume and surface area for both zebrafish and its liver.

In previous work [98], we have reported accurate 3D measurements for the 3 dpf zebrafish, from which we obtained the volume statistics as 2.53 ± 0.11 $(\times 10^8 \mu m^3)$ and the surface area as 3.20 ± 0.15 $(\times 10^6 \mu m^2)$. In this experiment, the statistics of the 3D measurements for the zebrafish are 2.72 ± 0.14 $(\times 10^8 \mu m^3)$ for the volume and 3.33 ± 0.12 $(\times 10^6 \mu m^2)$ for the surface area. The phenomenon that the specimens in this experiment are larger compared to our reference set is due to the fact that we did not accurately time the development for this experiment. We also computed the statistics of the 3D measurements for the liver as 9.06 ± 3.33 $(\times 10^5 \mu m^3)$ for the volume and 5.09 ± 1.10 $(\times 10^4 \mu m^2)$ for the surface area. The shape variation of the zebrafish liver is large for different individuals, but we can observe that a larger zebrafish tends to have a larger liver [127].

We implemented our method using Matlab programming on a desktop with an Intel i7 CPU and a 16G RAM. Subsequently, we evaluated the efficiency as runtime for the 3D reconstruction of the zebrafish and its liver separately as $22.0 \pm 0.4(s)$ $26.3 \pm 1.3(s)$. The results of this experiment can be directly used for establishing physiological values of a healthy liver of a 3 dpf zebrafish. The method can be more generically used to assess all observable effects of any compound on the shape and size of an organ.

5.4 Chapter conclusions and future work

In this chapter, we have presented a method for multi-modal 3D reconstruction and fusion on both organism- and organ-level through light microscopy axial-view imaging. It answers RQ 5: *How can we obtain a multi-modal 3D description and the corresponding measurements for the zebrafish from the MM-HTAI architecture*? We applied our MM-HTAI architecture to extend *Dataset C* presented in Chapter 4 (see Subsection 4.3). We then applied our multi-modal 3D reconstruction method on the dataset. Within the reconstructed 3D models, we observe an overview shape for the object on the organism-level and the detailed structure on the organ-level. The former provides a good shape reference to normalise and evaluate the organ development in phenotypical research. The experimental results show a natural visualisation of the multi-modal fusion images. Additionally, accurate 3D measurements are obtained, which can be directly used for the evaluation of the biological system with compound screening.

We believe our method is adequate to address RQ 5. However, in near future, we think it can be further developed for the determination of size and shape of other fluorescently labelled organs and objects, such as pathogens or tumour cells. In order to improve this work, a larger sample size of our subjects should be considered to get better statistics for the 3D measurements. High-throughput imaging would also be a good approach for this task.

Chapter 6

Case Study: Image Features and Classification Models

Based on:

- Y. Guo, H. Dibeklioglu & L. van der Maaten, “Graph-based kinship recognition,” in IEEE Conference on Pattern Recognition, Stockholm, Sweden, 2014, pp. 4287-4292.
- Y. Guo, C. Liang, F. Lens, R. Vos & F.J. Verbeek, “Image based taxonomy using convolutional neural networks,” publication in preparation.

This chapter addresses RQ 6.

RQ 6: *To what extent is it possible that the classification models (or regression models) are able to validate the performance of the image features to characterise the phenotypes in support of shape analysis*?

Abstract – It is difficult to characterise the phenotypes from high-magnification and high-resolution only through the shape analysis. For example, the variation of the local structures of cells and tissues is difficult to represent by the shape description as a whole. Therefore, we propose to, additionally, apply image features to extract the phenotypes encoded in the textures and local structures for the objects in images. Consequently, we use classification models to validate the performance of the applied features on phenotype characterisation. Rather than departing from zebrafish, in this chapter, we use a set of annotated datasets of images, i.e., human faces, a family of butterflies, a family of orchids and an public source for wood species. We aim to develop methods to estimate a structured taxonomy for each of these datasets. For the dataset of human faces, kinship is carefully labelled for pairwise faces and using this dataset, we propose a graphical model to recognise the kinship among a group of people in a family photo (see Section 6.1). In fact, the kinship can be considered as a particular example of taxonomy in which the parents and the children respectively correspond to a parent- and child-node in the hierarchy. For the other datasets, a two-level taxonomy, i.e., the genius and species, are used in the annotations. With the development of feature engineering such the feature learning using a supervised manner, the performance of image classification has been impressively improved. Therefore, we want to investigate representative features for the task of image based taxonomy using the convolutional neural networks (CNN) (see Section 6.2). Experimental results show that our proposed methods have improved the recognition accuracy in both cases. This results in a good understanding of the behaviour of our methods which can be applied in the applications with zebrafish as model system.

6.1 Graphical model for kinship recognition

Genetic correlation among family members is formally represented as kinship, which can be straightforwardly modelled using facial appearance similarity, a particular phenotype. However, due to the diversity of human faces, this phenotype similarity is weak and subsequently presents a challenge to image-based kinship recognition which plays an important role in the application of phenotype characterisation. It is difficult to estimate the kinship from paired faces only through shape analysis. Some prior studies solve the problem of pairwise kinship verification, i.e., on the question of whether two people are kin, through the assessment of the similarity of visual features on images of faces. Such approaches fail to exploit the fact that a global assessment on a group of family members may provide more clues for an accurate kinship recognition; for instance, the probability of two people being brothers increases when both people are recognized to have the same father. In this work, we propose a graphical model that integrates a local kinship confidence, i.e., facial similarity for all pairwise family members in an image, and a global kinship estimation which is represented as a series of reasonable semantic kinship graphs. For a complete and feasible kinship graph, we present an annotated dataset for the kinship of siblings to extend the existing kinship datasets; we also present a dataset of the images with group family members (more than 1) for the performance evaluation of our approach. In our experiments, we have found that the visual features such as Local Binary Patterns can well represent the facial appearance similarity for kinship recognition. The proposed graphical model has improved the accuracy of kinship recognition in group faces.

6.1.1 Kinship recognition using faces

Kinship can be expressed as physiological similarity among family members. For example, parents and children tend to show similar facial appearance and behaviours. In life-sciences, kinship research will support to track genetic evolution of a species. With respect to human beings, facial appearance as an important phenotype can be used as evidence to recognise kinship among different individuals. The image-based kinship recognition has become popular due to its efficiency and reproductivity, which tries to recognise kinship between people based solely on photographs of their faces. Such application benefits the phenotype characterisation from a large volume of facial images. This may be further helpful in uncovering and analysing social networks, and has applications in surveillance and in criminal investigation. Image-based kinship recognition is a challenging problem: it is a hard task even for humans to recognise kinship among people based on facial similarities. It is encouraging that some recent studies have demonstrated the possibility of kinship verification by means of image-based approaches [128, 129] identifying facial patterns that people may have inherited from their parents. In particular, siblings have the same gene sources which results in the presence of similar facial features. Facial cues that are informative for kinship recognition include the colour and shape of the eyes, eyebrows, nose, and mouth [130].

Prior work on image-based kinship recognition has three main limitations. First, prior studies only consider kinship verification: they try to determine whether kinship exists between a pair of faces, but they do not aim at recognising the exact type of kinship [129, 131, 132, 133]. Second, current kinship datasets are insufficient for the evaluation of existing kinship recognition algorithms, in particular, because existing datasets do not contain examples of siblings. Third, prior studies only consider settings in which kinship needs to be verified between pairs of people. This does not correspond to the typical setting encountered on social network websites, on which people often upload photographs that contain more than two family members. One may deal with this problem by separately classifying all pairs of faces in the family picture, but such an approach fails to share information between the pairs of people and may produce classifications that are inconsistent (*e.g.*, two people may be classified as sisters whilst they are also classified as having different parents).

Motivated by the aforementioned problems of prior work in kinship recognition, we study image-based kinship recognition in photographs that contain several family members. Specifically, this section makes three main contributions. First, we focus on kinship *recognition* instead of kinship *verification*: we aim to recognise the type of kinship relations between people. Second, we introduce two new datasets: (a) an annotated dataset containing photographs of siblings and (b) an annotated dataset of family photographs. The latter dataset and part of the former dataset is made publicly available. Third, we propose a novel graph-based algorithm that performs joint kinship recognition of all faces in a family picture.

The general framework of this algorithm is illustrated in Figure 6.1. The key advantage of our graph-based algorithm is that it exploits the fact that in a normal family, the recognised kinship of a particular pair of faces provides evidence for (non)kinship between other pairs of people. For example, in a family, two siblings should have the same father and mother[1]: if A and B are brothers and C is the father of A, then C must also be the father of B. Our graph-based algorithm constructs a fully connected graph in which faces are represented by vertices and kinship relations between pairs of faces are represented as edges. Using a few simple kinship rules (that are shown in Table 6.1), we can generate all valid kinship graphs. For each new test image, the predicted kinship graph is the one that obtains the highest score when we sum all scores of the pairwise classifiers that correspond to the edges. Because our graph-based algorithm shares information between the pairwise classifiers, ambiguities in the pairwise kinship classifications may be resolved, which may lead to improved performance. The results of our experiments demonstrate that the proposed algorithm can substantially improve kinship recognition accuracy.

6.1.2 Previous work

Most prior studies on image-based kinship recognition aim to solve the kinship verification problem using computer vision and machine learning techniques [129, 131, 132, 133]. All these approaches extract facial features and train a kinship verification classifier on a collection of annotated examples. In the seminal paper on automatic kinship detection [129], facial resemblance is represented by the difference between facial features. The extracted features include face colour, the position and shape of face parts, as well as gradient histograms. Face parts are localised using a pictorial structures model [134]. Classification is performed using a k-nearest neighbour classifier. [129] presents experiments in which the performance of an automatic kinship verification system is compared with human performance; the results show that the proposed algorithm performs 4.9% better than human accuracy on this task. [132] improves over this method by dropping the assumption that kinship examples have higher feature similarities than non-kinship examples. They learn a distance metric that aims to repel non-kinship samples as far as possible, whilst kinship samples are pulled close. The method of

---

[1] In this study, step relationships are not considered.

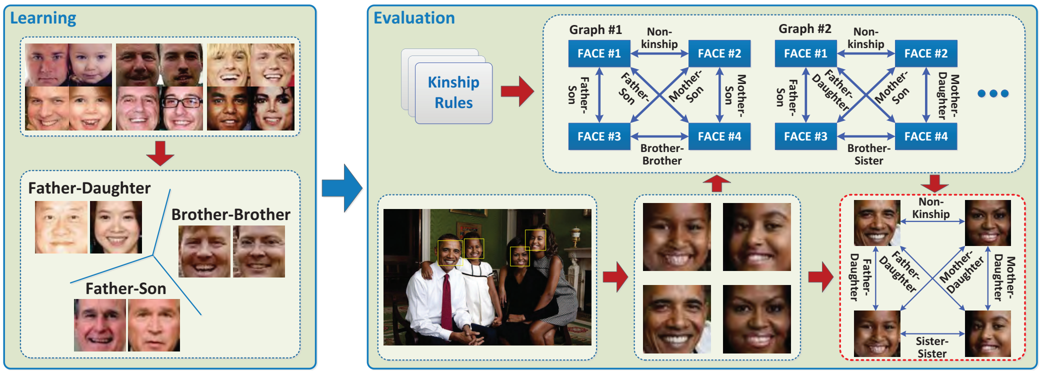


Figure 6.1: Overview of the proposed kinship recognition system. In the learning phase, a multi-class kinship classifier is jointly trained on different kinship relations. In the evaluation phase, the faces in family photographs are detected, cropped, and normalised. The set of all valid kinship graphs is generated according to the constraints on kinship relations. For each resulting candidate graph, the classifier scores are summed to obtain an overall score. The kinship graph with the highest overall score is selected as the prediction.

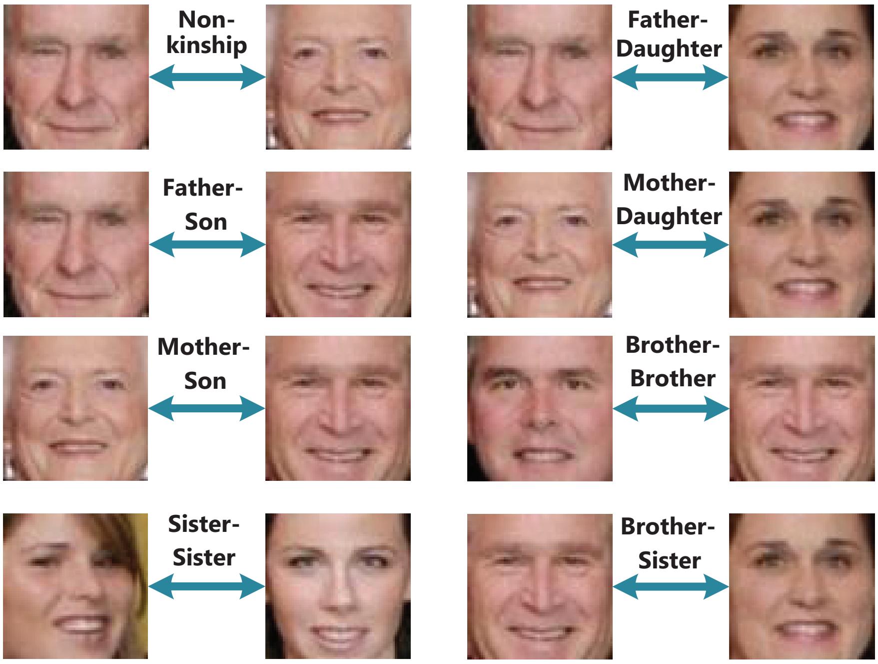


Figure 6.2: Normalised face pairs (from the Group-Face dataset) showing different kinship relations.

[132] also combines different types of feature descriptors by learning a multiview distance metric.

In [131] and [135], Xia *et al.* propose to use transfer subspace learning methods for kinship verification. They exploit the idea that the kinship verification between children and their parents is easier when the parents are young. The method learns a subspace in which old parents and their children are projected close together; the subspace model can then be used to make images of parents look younger. Recently, Dibeklioğlu *et al.* have proposed a method that uses facial expression dynamics combined with spatio-temporal appearance features to verify kinship in videos [133]. This method is based on the observation that the dynamics of facial expressions are informative for kinship recognition based on videos of people.

In contrast to the aforementioned methods, [136] does not focus on kinship verification but aims at recognising whether a group picture is a family picture. The method estimates the gender and age of every face in the group picture. An image graph is constructed by fitting a minimum spanning tree based on the face locations. Subsequently, the image is represented as a bag of image subgraphs. The resulting bag-of-image-subgraph features are then used to determine whether the group picture is a family picture. The method, however, does not recognise the types of kinship that are present within the family picture.

Our work has several differences in comparison to prior studies. First of all, instead of verifying kin relationships, our study focusses on recognising the exact type of kinship relations. Additionally, our study is the first attempt to generate complete kinship graphs for family photographs.

6.1.3 Graphical model for kinship recognition

Here, we propose an automatic kinship recognition system that relies on graph-based optimization of multi-class kinship classification. This work does not consider kinship verification between face pairs but focusses on classifying the type of kin relations. Assuming that kin pairs are known in a given group photograph (or predicted by an existing kinship verification system), our system predicts a kinship graph that describes the kinship relations between the family members.

(A) Feature extraction

Table 6.1: Kinship graph generation rules

Definition	Instance
• One child can at most have one father and one mother.	(A-B:Father-Daughter/Son) ⇒ ¬ (C-B:Father-Daughter/Son) (A-B:Mother-Daughter/Son) ⇒ ¬ (C-B:Mother-Daughter/Son)
• Siblings have the same parents.	[(A-B:Father/Mother-Daughter/Son) ∧ (A-C:Father/Mother-Daughter/Son)] ⇒ (B-C:Sister/Brother-Sister/Brother)
• Siblings have the same siblings.	[(A-B:Sister/Brother-Sister/Brother) ∧ (A-C:Sister/Brother-Sister/Brother)] ⇒ (B-C:Sister/Brother-Sister/Brother)
• There should not be kinship between father and mother.	[(A-B:Father-Daughter/Son) ∧ (C-B:Mother-Daughter/Son)] ⇒ (A-C:Non-kinship)

For the reliability of similarity analysis, face images need to be aligned before the feature extraction step. To this end, eye corners are located using the facial landmarking method proposed in [137]. Based on the eye locations, faces are aligned (in terms of roll rotation, translation, and scale) and cropped. The size of the resulting images are 64×64 pixels. Figure 6.2 shows samples of the normalised faces.

To describe the facial appearance, we use Local Binary Pattern (LBP) features [138]. Following [133], LBP features are extracted from each cell in a 7×5 grid that is imposed over the normalised face. In addition to LBP appearance features, we also extract gender and age features from the face images.

In order to estimate a gender feature $f_{\text{gender}}(I_i) \in \{-1, +1\}$ for a given face image I_i, we classify LBP and bio-inspired features (BIF) [139] using a binary support vector machine (SVM) classifier (with radial basis function kernel). Additionally, we extract an age feature $f_{\text{age}}(I_i, I_j) \in \{-1, 0, +1\}$ that describes the relative age of the given face images I_i and I_j:

$$f_{\text{age}}(I_i, I_j) = \begin{cases} -1: & a(I_i) < a(I_j) \\ 0: & a(I_i) \cong a(I_j) \\ +1: & a(I_i) > a(I_j) \end{cases}, \tag{6.1}$$

where a denotes the true age of the given subject. For the estimation of f_{age}, we employ a three-class SVM classifier using BIF features. To obtain the final feature vector for a pair of face images (I_i, I_j), all features are concatenated:

$$\mathbf{x}_{ij} = [f_{\text{LBP}}(I_i), f_{\text{LBP}}(I_j), f_{\text{gender}}(I_i), f_{\text{gender}}(I_j), f_{\text{age}}(I_i, I_j)] \,.$$

(B) Pairwise kinship classification

We model the resulting feature vectors to be able to distinguish between different kinship types. Moreover, we aim to predict the direction of these relations. For instance, the estimation for the given images will be that I_i is the father of son I_j (father→son), instead of just indicating that I_i and I_j have father-son relation. To this end, we define 12 types of directional kinship relations such as father→daughter, father←daughter, father→son, father←son, mother→daughter, mother←daughter, mother→son, mother←son, brother→sister, brother←sister, brother-brother, and sister-sister. By using these kinship types, more distant kinship relationships such as grandparents↔grandchildren, cousins, and uncle/aunt-nephew/niece may also be inferred if the family picture also contains the "intermediate" people.

We use a multi-class linear logistic regressor (LR) as the classifier in our system. For a pair of face images, the predicted label $\mathbf{y}^*$ is thus given by:

$$\mathbf{y}^* = \underset{\mathbf{y}}{\operatorname{argmax}}\ \mathbf{y}^\top \left(\mathbf{W}^\top \mathbf{x} + \mathbf{b}\right), \tag{6.2}$$

where $\mathbf{y}$ is a 1-of-K label vector. $\mathbf{W}$ and $\mathbf{b}$ denote the classifier weights and bias, respectively. To train the multi-class logistic regressor, we define the class-conditional probability:

$$p(\mathbf{y}|\mathbf{x}) = \frac{\exp(\mathbf{y}^\top(\mathbf{W}^\top\mathbf{x} + \mathbf{b}))}{\sum_{\mathbf{y}'} \exp(\mathbf{y}'^\top(\mathbf{W}^\top\mathbf{x} + \mathbf{b}))}. \tag{6.3}$$

In our application, this probability represents the likelihood of the kinship type given a pair of faces. We aim to minimize the penalized conditional log-likelihood $\mathcal{L}$:

$$\mathcal{L}(\mathbf{W}, \mathbf{b}) = \underset{\mathbf{W}}{\operatorname{argmax}} \left(\sum_{\mathbf{x}} \log p(\mathbf{y}|\mathbf{x}) - \lambda \|\mathbf{W}\|_2^2 \right). \tag{6.4}$$

Herein, the second term is an L2-norm regulariser that is employed to prevent overfitting. The value of the regularisation λ is set based on the error measured on a small, held-out validation set.

(C) Kinship graphs

A straightforward way to recognise kinship relations in a family photograph is to classify each pair of faces individually. However, this approach does not share information between the pairwise classifications: if the classifier doubts between two kinship types, individual classification cannot exploit the other kinship relations in the photo to resolve this ambiguity. Individual classification may even produce infeasible kinship graphs. For example, it may predict that two people are brothers whilst predicting that they have different parents. The graph-based algorithm we propose aims to resolve these two problems by: (1) generating all feasible kinship graphs and (2) selecting the kinship graph that obtains the highest score.

A kinship graph can be defined as $G = (V, E)$ in which faces correspond to vertices and edges to kinship relations. In other words, each edge $(i, j) \in E$ has an associated label $\mathbf{y}_{ij}$. Two examples of kinship graphs using three faces are shown in Figure 6.3. Note that the graph shown in Figure 6.3(b) is actually infeasible since it violates the constraints on kinship relations that are given in Table 6.1. In the first step, all possible kinship graphs that satisfy these constraints are generated. It is important to note that the candidate graphs can actually be generated offline. The resulting set of candidate kinship graphs are denoted by $\mathscr{G}$. Afterwards, we assign a score to each of the candidate kinship graphs that measures the (log)likelihood of that kinship graph for the observed family picture. Specifically, we define the kinship graph score as the sum of the kinship classifier scores that correspond to each of the edges in the graph:

$$s(G|\mathcal{I}) = \sum_{(i,j)\in E} \mathbf{y}_{ij}^{\top} \left(\mathbf{W}^{\top}\mathbf{x}_{ij} + \mathbf{b}\right), \tag{6.5}$$

where $\mathcal{I}$ is the family photo, $G = (V, E)$ is the kinship graph that we are scoring, $\mathbf{x}_{ij}$ is the feature vector extracted from the pair of faces associated to edge $(i, j) \in E$, and $\mathbf{y}_{ij}$ is the corresponding kinship label. We perform kinship graph prediction for family photo $\mathcal{I}$ by maximising the graph score over the set of all candidate kinship graphs:

$$G^{*} = \underset{G\in\mathscr{G}}{\operatorname{argmax}}\ s(G|\mathcal{I}), \tag{6.6}$$

where graph G^{*} is the predicted kinship graph.

6.1.4 New datasets

To evaluate our approach, we gathered two new kinship recognition datasets: (A) a dataset with image pairs of siblings and (B) a dataset with family photographs.

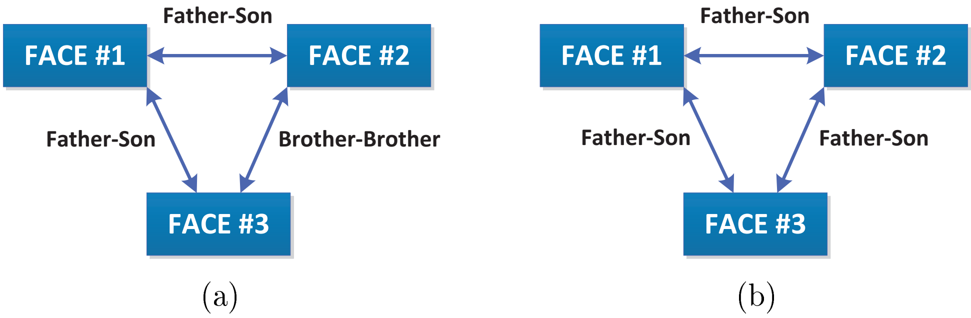


Figure 6.3: Samples of (a) feasible and (b) infeasible kinship graphs.

Figure 6.4: Sample images from the Group-Face dataset.

Both datasets (except some copyrighted images in the first dataset) are made available to the research community. One can find the dataset at `https://pan.baidu.com/s/1nvPxQ8D` (pincode: e8if). Both datasets are described separately below.

(A) Sibling-Face dataset

Existing large-scale kinship datasets (such as the KFW-II dataset [132]) do not

Table 6.2: Distribution of kin pairs (image pairs) in the KFW-II, Sibling-Face and Group-Face datasets.

	KFW-II	Sibling-Face	Group-Face
Father-Daughter	250	-	69
Father-Son	250	-	69
Mother-Daughter	250	-	70
Mother-Son	250	-	62
Brother-Brother	-	232	40
Sister-Sister	-	211	32
Brother-Sister	-	277	53

include sibling pairs. The UvA-NEMO dataset [133, 140] contains sibling pairs, but it has a small number of subjects. We have gathered a new dataset that contains more than 200 image pairs for each of three possible sibling relations (brother-brother, sister-sister, and brother-sister). All sibling images have been collected from websites such as Flickr; the sibling relations have been determined based on the tags or descriptions of the images. The sibling faces have been processed in the same way as done for the images in the KFW-II dataset: they are aligned according to the position of eyes, and resized to a fixed size of 64×64 pixels. In our experiments, the Sibling-Face dataset is combined with the KFW-II dataset to train kinship classifiers. The distribution of kin pairs in the KFW-II and Sibling-Face datasets is given in Table 6.2.

(B) Group-Face dataset

We have also gathered a collection of group photographs from publicly available sources such as Flickr. Specifically, we have selected group pictures in which the people are all frontally facing the camera. Some samples from the collected dataset are shown in Figure 6.4. The dataset consists of 106 group photographs, of which 82 contain group(s) of family members. To facilitate labelling of the kinship relations, we have selected photographs of famous families (royalty, presidents, Hollywood stars, etc.) and photographs of regular families with reliable kin labels. The Group-Face dataset contains father-daughter (FD), father-son (FS), mother-daughter (MD), mother-son (MS), brother-brother (BB), sister-sister (SS) and brother-sister (BS) pairs. Table 6.2 shows the number of image pairs in each kinship class. All the faces in the dataset have been cropped and aligned in the same way as the faces in the Sibling-Face dataset.

6.1.5 Experimental Results

In our experiments, the KFW-II and Sibling-Face datasets are combined and used for training. We employ the family photos in our Group-Face dataset as the test set. It is assumed that we know which pairs of faces in the family pictures have kinship and which pairs of faces do not, i.e., we assume that we have access to a perfect kinship verification algorithm and focus solely on recognising what type of kinship exists between two people. In our experiments, the maximum number of family members is limited to four because, in our current (naive) implementation, the total number of candidate kinship graphs and the required amount of memory drastically increases when more than four faces are used. Specifically, when a family photo contains two parents and four children, we manually split the family into two groups which both have parents and two children. In this way, we obtained 98 kinship groups (16 groups with two faces, 40 groups with three faces, and 42 groups with four faces) that we use in our kinship recognition experiments. The test set we used in our experiments is made publicly available (as part of the Group-Face dataset).

As a baseline approach, we individually perform pairwise classification on each edge of the kinship graph to determine the type of kinship. We set the regularisation parameter λ (see Equation 6.4) of the kinship classifier by cross-validating over a small held-out validation set.

To test the reliability and efficacy of the proposed graph-based kinship recognition, we perform two different experiments. In the first experiment, kinship recognition performances of the graph-based and pairwise approaches are compared. In the second experiment, we investigate the effect of age/gender estimation accuracy on the robustness of the graph-based and pairwise methods. To this end, we systematically perturb the gender and age features which are extracted from the test data. The details and results of these experiments are given below.

(A) Graph-based versus pairwise classification

In this experiment, the correct classification rates of the graph-based and pairwise approaches are compared. As shown in Table 6.3, the graph-based method proposed in our study outperforms the pairwise kinship classification by 16.77% (absolute) on average. This result demonstrates the efficacy of the graph-based kinship recognition. The highest performance of the graph-based method is achieved for the sister-sister relationship with an accuracy of 76.92%.

Table 6.3: Kinship recognition accuracy of the pairwise (baseline) and the graph-based approaches.

Relationship	Pairwise (%)	Graph-based (%)	# Test Pairs
Father-Daughter	66.15	67.69	65
Father-Son	51.72	65.52	58
Mother-Daughter	57.81	71.88	64
Mother-Son	48.15	72.22	54
Brother-Brother	43.33	63.33	30
Sister-Sister	34.62	76.92	26
Brother-Sister	44.00	68.00	25
All	52.48	69.25	322

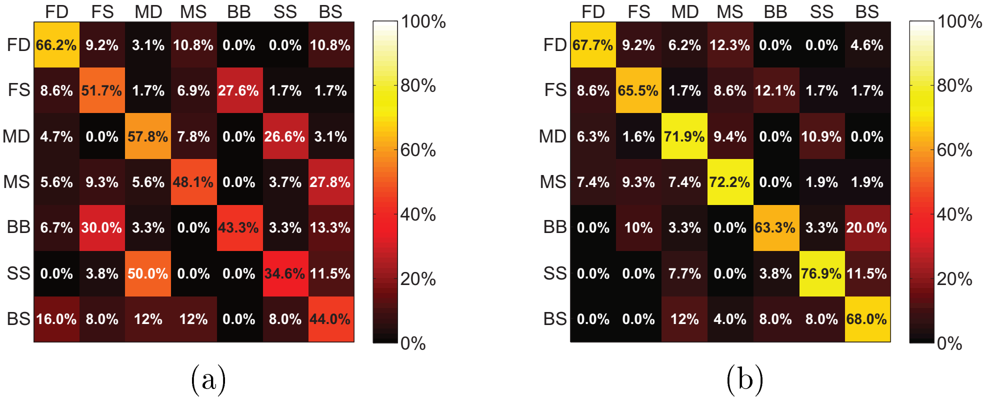
For further exploration of the results, the confusion matrices for both methods are given in Figure 6.5. The results suggest that, unlike the pairwise classification, the graph-based approach is able to recover from errors in the age/gender estimations. For instance, the baseline approach often confuses the father-son relation with the brother-brother relation, presumably due to errors in the relative age estimation[1]. By contrast, the graph-based approach corrects most of such misclassifications by incorporating other relations in the graph, and by ensuring that the predicted kinship graph is feasible. This is confirmed by the number of kinship graphs which are correctly predicted (*completely*) on the Group-Face dataset. Whilst the graph-based approach correctly predicts 56 of 98 kinship graphs, only 29 kinship graphs are correctly recognised by the baseline method.

(B) Effect of age and gender estimation accuracy

The results presented in the previous subsection illustrate the potential merits of our graph-based algorithm, which mainly stem from its ability to correct errors in the age and gender estimations. We further investigate the effect of age and gender estimation accuracy in our method. To this end, we randomly generate labels for the relative age classes and gender by systematically changing the error rate. Both the graph-based and pairwise methods are tested using these labels.

---

[1] The correct classification rate of the gender classifier, used in our experiments, is approximately 90% based on 10-fold cross-validation. Combination of the KFW-II, Sibling-Face, and UvA-NEMO datasets is used for the evaluation. 10-fold cross-validation accuracy of the relative age estimator is approximately 65% on the combination of KFW-II and Sibling-Face datasets. Higher error rate in age estimation is mostly due to small size (low resolution) of the face images, which makes facial wrinkles nearly invisible.



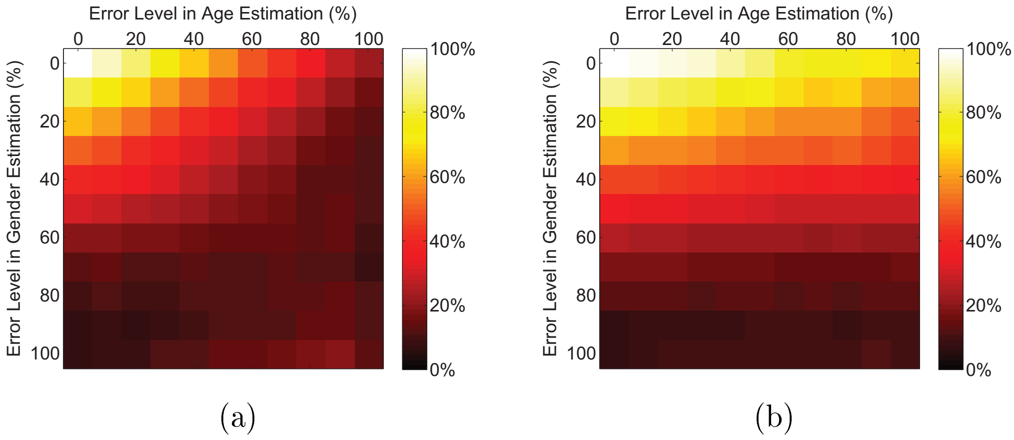
(a)

	FD	FS	MD	MS	BB	SS	BS
FD	66.2%	9.2%	3.1%	10.8%	0.0%	0.0%	10.8%
FS	8.6%	51.7%	1.7%	6.9%	27.6%	1.7%	1.7%
MD	4.7%	0.0%	57.8%	7.8%	0.0%	26.6%	3.1%
MS	5.6%	9.3%	5.6%	48.1%	0.0%	3.7%	27.8%
BB	6.7%	30.0%	3.3%	0.0%	43.3%	3.3%	13.3%
SS	0.0%	3.8%	50.0%	0.0%	0.0%	34.6%	11.5%
BS	16.0%	8.0%	12%	12%	0.0%	8.0%	44.0%

(b)

	FD	FS	MD	MS	BB	SS	BS
FD	67.7%	9.2%	6.2%	12.3%	0.0%	0.0%	4.6%
FS	8.6%	65.5%	1.7%	8.6%	12.1%	1.7%	1.7%
MD	6.3%	1.6%	71.9%	9.4%	0.0%	10.9%	0.0%
MS	7.4%	9.3%	7.4%	72.2%	0.0%	1.9%	1.9%
BB	0.0%	10%	3.3%	0.0%	63.3%	3.3%	20.0%
SS	0.0%	0.0%	7.7%	0.0%	3.8%	76.9%	11.5%
BS	0.0%	0.0%	12%	4.0%	8.0%	8.0%	68.0%

Figure 6.5: Confusion matrices for (a) the pairwise and (b) graph-based approaches.



(a) (b)

Figure 6.6: Kinship recognition accuracy (%) as a function of the error level in age and gender estimation for (a) pairwise and (b) graphical model of kinship recognition.

Figure 6.6 shows the kinship recognition accuracy as a function of the error level in age and gender estimation. As shown in Figure 6.6, both methods achieve 100% classification accuracy when the age and gender ground truths are used: age and gender completely determine the type of kinship relation between two people, if we assume that the given pair has kinship.

The results show that both pairwise and graph-based approaches perform worse when the perturbation rate is increased for gender and age. However, our graph-based method is more robust to gender and age estimation errors than to the pairwise approach. In particular, the graph-based algorithm is less sensitive to incorrect age prediction. This is beneficial because age estimation is a difficult task in real-life conditions, in particular, because age estimates are strongly influenced by changes in resolution, illumination, gender [141], and facial expression [142]. Our graph-based algorithm is more robust to the resulting errors in the age estimates. As shown in Figure 6.6 (see top right side of the accuracy maps), graph-based approach performs much better than the pairwise classifier in such conditions.

6.1.6 Section conclusions and future work

In this section, we have proposed a novel graph-based method to recognise kinship relations in family photos. It partially answers RQ 6: *To what extent is it possible that the classification models (or regression models) are able to validate the performance of the image features to characterise the phenotypes in support of shape analysis*? Our approach models the kin relationships using a fully connected graph in which faces are represented by vertices and edges represent kinship relations. The overall score of each feasible kinship graph is computed by summing classifier scores over the edges of the graph. The graph with the highest overall score is selected as the prediction. The results of our experiments demonstrate that our graph-based outperforms the pairwise kinship classification approach. Moreover, the proposed method guarantees consistency of the predicted kinship graphs.

We consider that RQ 6 is partially answered that the graphical model and a classification model, i.e., the logistic regression, have cooperated to validate the performance of the LBP features in the application of image based kinship recognition. It turns out that the well-designed image features will be able to characterise the subtle variation of the phenotypes such as shape and texture.

As a future direction, we aim to develop a graph-based method to train our kinship classifier as well by framing the task as a structured prediction problem. Also, we aim to improve the speed of our current (naive) implementation by exploiting redundancies in the score computations (like in dynamic programming). Moreover, we plan to include a kinship verification step prior to the classification of relations. Finally, we will apply the method in the applications which use the zebrafish as model system.

6.2 Image based taxonomy using CNN

Phenotypes including shape and texture represented in appearance are essential in image based taxonomic classification of biological specimens. This presents a challenge to the choice of features to generalise these phenotypes. We are motivated to investigate representative features for the task of image based taxonomy using the convolutional neural networks (CNN). We first present three dataset with a taxonomic structure, which include orchids, butterflies as well as introduce an open source for wood species (in fact, the kinship addressed in Section 6.1 is a special category of the taxonomic structure). We adapt a popular CNN architecture, the VGGNet-16, to learn representative features for these tasks in a supervised manner. We implement a multi-output layer of which each output corresponds to a flat classifier for each level in the taxonomy. In this manner, we can introduce multi-supervision to the training time of the networks. This avoids to learn individual classifiers on each level or each node which is commonly used in conventional hierarchical classification. We use a fine-tuning strategy to accelerate and stabilise the training process. Experimental results show that the proposed approach achieves better performance compared to the methods using hand-crafted features and pre-trained networks. From our observation, representative features are of great importance to a well-performing recognition system for taxonomy. Importantly, in our method the prediction for each level in the taxonomy can be performed in one forward pass.

6.2.1 Image based taxonomy

A feasible and convenient manner for categorisation gives rise to digitization, reuse and efficient management for the large amount of the collection of cultural heritage. Under these circumstances, taxonomic categories are commonly used, which formally use a hierarchical ranking i.e., Kingdom, Phylum, Class, Order, Family, Genus and Species to categorise and annotate the specimens [1, 143]. This manner also facilitates an efficient top-to-bottom data retrieval. In practice, a taxonomic recognition system will also facilitate many applications such as recognition of endangered species [144].

Using imaging of specimens makes image based taxonomy possible. It aims to learn a model to recognise each rank in the taxonomy for a specimen using images which represent that specimen as a whole or microscopic structure. In practice, researchers in life-sciences make use of their expertise to identify the species of a specimen [145]. However, some species, for example, the ones in the same genus, present rather similar shapes; subsequently, their textures such as special patterns on the specimen surface should be emphasised. Therefore, the image based taxonomy requires comprehensive investigation of phenotypes including shape and texture in the whole appearance of the specimen.

Image based taxonomy is, in fact, a typical hierarchical classification problem [146]. Each level in the hierarchy represents a rank in the taxonomy. In a task of image based taxonomy, it is usually easy to recognise a higher rank due to the remarkable dissimilarity of appearance for the specimens from different classes; and it is usually difficult to recognise a lower rank due to the dramatic similarity of appearance for the specimens of which the classes share the same parent rank. Therefore, a proper choice of feature representation for phenotypes can result in a well-performed taxonomic recognition system. For example, experts can accurately recognise a wood species through a careful investigation on microscopic features such as shape and size of vessels and fibrous structure of tissues [147].

In practice, we have multiple options of image features. For the last decades, many local features have been increasingly used for image recognition, such as Histograms of Oriented Gradients (HOG) [24], Local Binary Patterns (LBP) [138] and Scale Invariant Feature Transform (SIFT) [109]. There are also many available shape features, such as shape context [148], the angular radial transform [149] and projective invariant contexts [150]. Some of these features are generic and suitable to the problem of image based taxonomy; some are well-designed for a particular domain. With the fast development of deep convolutional neural networks (CNN) [25, 151, 152, 153], successful applications have been made in many fields like computer vision [64, 154, 155] and gaming [156]. The readout of a deep CNN architecture is, in fact, a feature engineering which learns discriminative features from images in a supervised manner. This makes the deep CNN architecture very flexible for learning representative features for corresponding applications. The current development has inspired us to apply the CNN architecture in the image based taxonomy due to the diversity of taxonomic categorisation as for each specific taxonomic category, different features should be emphasised.

Here, we first present three taxonomic structured datasets with expert tags; Javanese butterflies, slipper orchids and wood species. The first one is obtained from a collection of Dutch National Natural History Museum (Naturalis Biodiversity Center `http://www.naturalis.nl/`) for the family of Papilionidae. The second one is obtained from some public sources such as ImageNet [157] for the family of Cypripedioideae. Both datasets are labelled by a two level taxonomy: genus and species. The third dataset is a public source for microscope images of wood species [158]. This dataset contains a more specific taxonomic structure from class to species. In this work, we only employ two level annotations including class and species.

We are motivated to present a CNN architecture based on the VGGNet [151] which is extended with a multi-output layer for the image based taxonomy. Each output corresponds to a local classifier for each level in the taxonomy. We train the whole networks considering all the taxonomic annotations for each example to be trained. This means that the multi-supervision jointly contributes to the training phase. We use a fine-tuning strategy for the training of the networks. We first introduce a pre-trained model using a large dataset such as ImageNet and then use our datasets to enhance the representability of the networks for our application. This operation largely stabilises and accelerates the training of the network.

In an hierarchical classification, the proposed method can be categorised as a local classification per level approach, which is also referred to as top-down strategy [159]. This method may introduce the problem of label inconsistency. For instance, a testing example may be assigned labels that not refer to a reasonable parent-child routine. This procedure can be improved by post processing. Another possible solution is to use the flat classification approach. Such an approach only trains a classifier for the bottom level and a bottom-top strategy can be used to back-propagate the labels on higher level according the deterministic property of the parent-child mode [160]. Other attempts concern global classification models [161, 162, 163, 164]. We should note that all these methods mainly focus on a classification model to generalise the hierarchical classification problem. The method can be considered to improve the output layer in our application of image based taxonomy. Here, we would like to focus on the contribution of features in our particular problem. Therefore, we first use a simple classification strategy e.g., softmax [33], as a local classifier for each level to validate the performance of our representative features.

Actually, the hierarchical classification problem can be considered as a special case of multi-label classification [165]. Many deep CNN architectures have been reported to solve this problem [166, 167, 168]. Recently, a hierarchical deep CNN (HD-CNN) architecture has been reported, which presented a coarse-to-fine strategy for a large scale of visual recognition [169]. This enables the so-called local classifier per node approach with a CNN architecture. Here, we stress the importance of features in our image based taxonomy of biological specimens. So, we propose to extend the CNN architecture with a multi-output layer, of which each output corresponds to a level in the taxonomy. This will provide a good understanding of the performance of the features for each level. In future work, we can consider to introduce a dedicated architecture such as the HD-CNN in our problem.

6.2.2 Image based taxonomy using CNN architecture

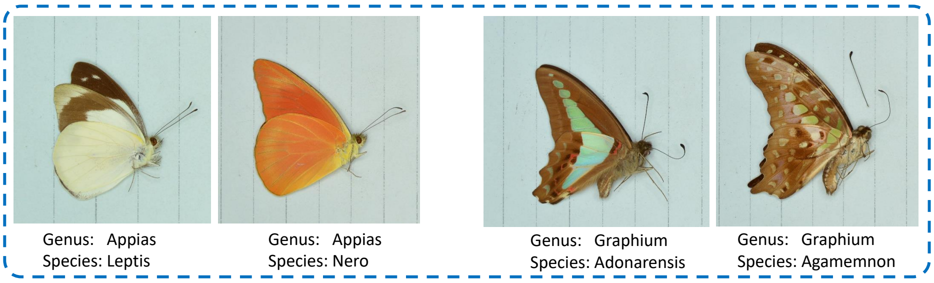
We first present (A) the datasets used in this work, and (B) elaborate in details the CNN architecture we have adapted.

(A) Datasets

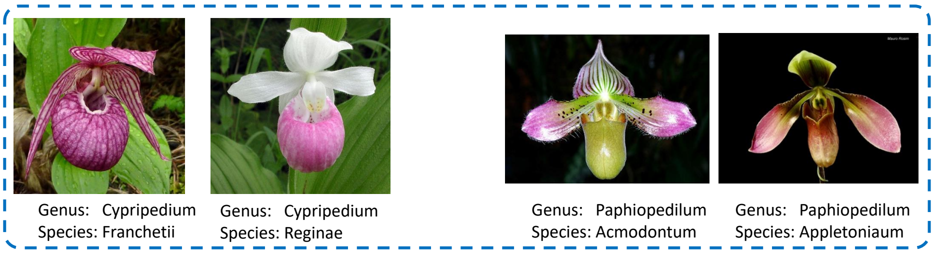
Below we briefly discuss the datasets, i.e., Butterflies, Orchids and Woods.

The dataset of *Butterflies* are obtained from a large collection for the family of Papilionidae, a category of Javanese butterflies caught in the 1930s. With the development of digitalisation of cultural heritage, images have been made for these specimens. Some examples can be seen in Fig. 6.7 (A). In the images, the specimens are well-positioned on their profile-view and most of the features such as the texture and patterns on their wings are clearly presented. In this manner, we avoid the effects of shape misalignment and scaling. This dataset is structured in a two taxonomic categories, i.e., genus and species. Until now, the dataset consists of 1829 images which are from 18 genera and 45 species.

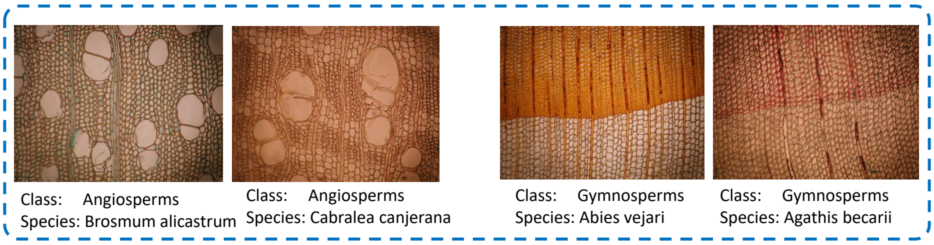
The dataset of *Orchids* are obtained for the family of Cypripedioidea. The orchid experts annotated 1117 images with 5 genera and 116 species [170]. Examples can be seen in Fig. 6.7 (B). One should note that the datasets of *Butterflies* and *Orchids* have the problem of data imbalance. Some classes only contains a small number of examples and some others contain much more. This will present a challenge to a classification model which may result in overfitting for the classes with a large number of examples.



(A) Examples of Butterflies



(B) Examples of Orchids



(C) Examples of Woods

Figure 6.7: Examples of the images from dataset (A) Butterflies (B) Orchids and (C) Woods. For each dataset, we select four examples from four species, two of which are from the same genus (class). One can observe that, the phenotypes, such as the colour and patterns on butterfly's wings, the shape and the texture of orchid's pedals, the shape and structure of wood's vessel and tissue, show significant similarity from the species which share the same genus (class).

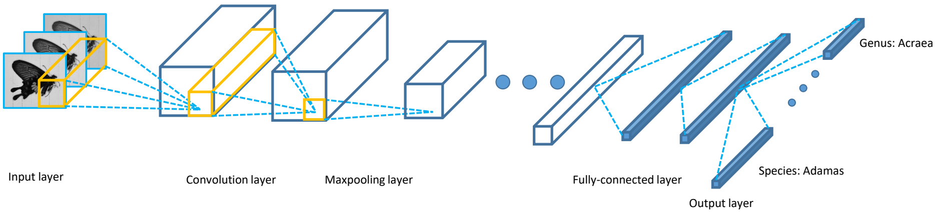


Figure 6.8: The CNN architecture with a multi-output layer. Only one filter is shown for the convolution and pooling layers.

The dataset of *Woods* was originally presented in [158]. It contains 2240 wood images from 2 classes and 112 species. This dataset has an even distribution as each species contains 20 examples. The images from stained wood slices are acquired using a microscope (Olympus Cx40) with a 100x objective. The size of the acquired RGB images is 1024x768. Examples of the images can be found in Fig. 6.7 (C).

From the datasets, we can see a remarkable diversity of the task to design the image based taxonomy. It is difficult to design a type of generalised feature for this application with diverse objectives. For example, the HOG features works fine in the butterflies and orchids but totally fails in the woods. This requires a generalised framework for feature engineering to obtain more discriminative and representative features for each task. We can also observe that the datasets of *Butterflies* and *Woods* are produced in specific imaging conditions. The acquired images are all standard, for example, the orientation and position of the specimen represented in the images is stable. The dataset of *Orchids* is more challenging because the examples are collected in a natural setting.

(B) VGGNet with a multi-output layer

A standard CNN architecture consists of one input layer, a set of convolution layers, several pooling layers, one or two fully-connected layers and one or multiple outputs layer. In Fig. 6.8, we show a schematic representation of the CNN architecture with a multi-output layer.

The input layer is also referred to as data layer which converts the input image into the format a CNN architecture requires. The convolution layer generates a set of feature maps through convolving the previous feature maps using different filters. The weights of a filter are shared by the whole convolution which produces one feature map. This means each element in a feature map corresponds to a receptive field from the original image. One should note that a non-linear operation of rectification such as ReLU [25] is performed after each convolution layer. The pooling layer aims to subsample a feature map, to an extent holding good spatial property in the feature representation. Similar to the conventional multi-layer perceptron [171], the fully-connected layer connects all the elements in previous layer to each of the neurons in the fully-connected layer. This operation converts the feature map into a one-dimensional feature vector. The output layer can be a fully-connected layer which can be followed by a loss in training time. The supervision of the network training is implemented in this process. In this supervised manner, the parameters in a CNN architecture can be obtained using a standard algorithm such as gradient back-propagation. As a result, a well-trained CNN architecture can largely fit the training data and the extracted feature maps can be discriminative and representative to our task, i.e., image based taxonomic recognition.

We have adapted the VGGNet-16 in our CNN architecture. In order to produce taxonomic categories, we adapted the last layer of VGGNet-16, i.e., the output layer, with a multi-output layer. Each output corresponds to a level in the taxonomy. The networks consists of 14 convolutional layers, 5 pooling layers, 2 fully-connected layers and 1 multi-output layer. More concrete, the configuration is depicted as follows: Input image $(224 \times 224 \times 3) \rightarrow 2$ (3×3) convolution layers (64 feature maps) → maxpooling layer → 2 (3×3) convolution layers (128 feature maps) → maxpooling layer → 3 (3×3) convolution layers (256 feature maps) → maxpooling layer → 3 (3×3) convolution layers (512 feature maps) → maxpooling layer → 3 (3×3) convolution layers (512 feature maps) → maxpooling layer → fully-connected layer (4096) → fully-connected layer (4096) → multi-output layer (softmax).

We use the library of Caffe [172] in our implementation. Both for the training and testing, we re-scale all the images into a size of 256×256 pixels. At training time, we use a pre-trained model to initialise the weight layers including all the convolution layers and the 2 fully-connected layers. For the last fully-connected layer, i.e., the multi-output layer, we initialise the weights using a Gaussian distribution with the mean as 0 and the standard deviation as 0.01. We use the statistical gradient descent strategy to train the networks and we set the batch size as 64. We set the total iterations as 2000 and the learning rate as 5×10^{-3}. We decay the learning rate as half of the original value after 1000 iterations. We train and test our model using two NVIDIA TITAN X GPUs.

6.2.3 Experiments

In this subsection, we apply our CNN architecture on the datasets to evaluate the its performance in the task of image based taxonomy. (A) We compare the performance by different methods. (B) We discuss the classification results using confusion matrix and visualise the representative features from our CNN architecture using the t-SNE map [173].

(A) Performance evaluation with different features

In this experiment, we perform cross validation on the datasets using different methods. Due to the data imbalance in the datasets of *Butterflies* and *Orchids*, we leave out the classes with less than 3 examples and finally we use 3-fold cross validation. For the dataset of *Woods*, we use 5-fold cross validation. We randomize the partition of the folds and repeat the whole process for 5 times to obtain a statistical representation for the accuracy. In Table 6.4, we separately report the accuracy for the two levels in the taxonomy of the datasets. In each row of Table 6.4, the upper value corresponds to the accuracy for the prediction of genus (class), and the lower value represents the accuracy for the prediction of species.

A1. Configuration We use two popular features, the rotation-invariant uniform LBP and HOG, as comparisons in this experiment. In order to obtain the identical feature dimensions in each dataset, we rescale the images from *Butterflies* and *Orchids* to 256×256 pixels; and we keep the original image size for the *Woods*, i.e., 1024×768.

For the LBP, we configure the sampling radius and the number of sampling points as (2,8) for the *Butterflies* and *Orchids*; (3,24) for the dataset of *Woods* due to its large image size. The former results in a 59-dimensional feature vector; the latter produces a 555-dimensional feature vector.

For the HOG, we configure the cell size and block size as (32,4) for the *Butterflies* and *Orchids*; (8,2) for the *Woods* to capture its microscopic structure. Due to the high-dimensional of the obtained HOG features, we apply principal component analysis (PCA) for feature dimensionality reduction. We keep 99% components of the decomposed principal components, which dramatically reduce the obtained feature size.

For the classification model, we use the polynomial kernel SVM. We set the regularisation term as 10 to prevent overfitting of the model.

We also use a shallow CNN architecture i.e., the AlexNet [25] for comparison. In Table 6.4, a CNN architecture without an indication of $*$ denotes that we use a pre-trained network based on a large image datasets which does not include our datasets to extract features. We consequently use the polynomial SVM for classification. The notation $*$ means that we use the strategy presented in this

Table 6.4: Accuracy (%) of different methods on taxonomic datasets

	LBP	HOG	AlexNet	AlexNet*	VGGNet	VGGNet*
Butterflies	91.6±0.3	96.9±0.3	97.8±0.1	99.5±0.1	98.2±0.2	**99.6±0.1**
	82.2±0.2	93.5±0.2	95.1±0.2	98.7±0.1	95.5±0.2	**98.9±0.2**
Orchids	86.0±0.4	88.4±0.5	91.6±0.2	98.4±0.2	92.4±0.2	**98.8±0.2**
	9.4±0.5	41.8±0.7	51.1±1.0	82.7±0.7	50.8±0.4	**86.1±0.5**
Woods	97.2±0.2	75.7 ±0.4	99.2±0.2	100 ±0.0	99.8±0.02	100±0.0
	88.4±0.4	30.1±0.5	85.9±0.4	95.6±0.4	90.7±0.2	**95.6±0.3**

chapter. Namely, we use our datasets to fine-tune the pre-trained network and extend the network with a multi-output layer for prediction.

A2.Results First, from the results, we can see that the taxonomic recognition on a higher rank is relatively more easy than that of a lower rank. This is reflected by a much higher classification accuracy on the level of genus(class) than that on the level of species for all the datasets, using different methods.

Second, if we focus on the well-designed features in the first two columns, we can observe that the LBP can obtain higher recognition accuracy for the *Woods* and the HOG can obtain higher recognition accuracy for the *Butterflies* on both levels. The LBP is advantageous in capturing textural structures and the HOG is capable of holding the whole appearance in an image. Accordingly, the characteristics of the woods are represented as important patterns on the shape and structure of the vessels and tissues; the characteristics of the butterflies are represented in larger scale patterns on butterfly's wings. Those can be separately stressed by the LBP and HOG. For the *Orchids*, the LBP and HOG features obtain similar results on level 1, both of which, however, failed in the species recognition. This is caused by the diverse patterns for the orchids. One should integrate colour, texture, shape as well imaging conditions to characterise orchid's patterns. It is difficult for the LBP and HOG to generalise all these characteristics.

Third, we can find that a simple and pre-trained CNN architecture like AlexNet can obtain better performance on the three datasets than the well-designed features, but it fails to compete with the LBP features on the species recognition of the *Woods*. This is because the pre-trained CNN architectures do not have sufficient training images similar to the *Woods*. This leads the CNN cannot sufficiently generalise the microscopic tissular patterns.

Fourth, after a fine-tuning, both of the shallow and deeper CNN architectures can obtain very accurate recognition on the three datasets. This again illustrates the power of the CNN architecture on representative feature learning for image based taxonomy. In addition, although a small different performance can be found for the two CNN architectures on *Butterflies* and *Woods*, a large improvement is made by the VGGNet for the species classification on the *Orchids*. From the observation, we may conclude that for a relatively simple image base taxonomy which introduces less variant conditions can be solved by a simple CNN architecture such as AlexNet, while the complication of a task, e.g., the taxonomic recognition for the *Orchids*, requires a deeper CNN architecture.

(B) Results visualisation

In this experiment, we further explore the results obtained by the proposed method with the manners of confusion matrix and feature visualisation.

B1. Confusion matrix In Fig. 6.9 (a1) to (a3), we present the confusion matrix for the genus recognition of each dataset obtained from the proposed method, i.e., the VGGNet with a multi-output layer.

In each confusion matrix of Fig 6.9, we use orange lines to indicate the grouping of the species. The species separated by the lines are from the same genus (class). We have left out the species with less than 3 examples and the corresponding result is shown as zero on the diagonal in the confusion matrix. Due to the limited space, we show the names of some selected species.

First, one can observe in the confusion matrices that the recognition accuracy for the three datasets is high which corresponds to the result shown in Table 6.4. Although data imbalance is occurring in the datasets of *Butterflies* and *Orchids*, it is hardly to see serious overfitting for the species with more examples. This can be reflected by the high recall and precision for all the species.

Second, we can find an important phenomenon that the classification errors of the species are mainly distributed within the same genus. One can see the squares associated with the diagonals in Fig. 6.9 (a2) and (a3) for this message. According to this observation, we can conclude that, in the image based taxonomy for biological specimens, it is more difficult to recognise the species which share the same genus. The prediction on a higher level, e.g. genus and class, probably does not help to improve the recognition accuracy on the level of species.

B2. Feature visualisation In Fig. 6.9 (b1) to (b3), we produce the so-called t-SNE map [173] for the visualisation of the representative features obtained in our method. From this visualisation, we can clearly see the separation among different species in each dataset according to the representative features. This is shown as the separated clusters. A relatively sparse t-SNE map is obtained for the *Orchids* and *Woods* due to their large number of species. Another reason resulting in the sparse t-SNE map is that the learned representative features in the same species are very similar for different specimens. This produces rather dense overlap among specimens from the same genus (class). In fact, in each clustering center a dense overlapped with the features extracted from different specimens. Yet, obvious clustering centers can be found for each species in these two datasets.

6.2.4 Section conclusions and future work

For the task of image based taxonomy, we have presented a CNN architecture which extends the conventional VGGNet with a multi-output layer. This makes the prediction on each level in the taxonomy possible. We have proposed to apply the fine-tuning strategy to accelerate and stabilise the training of the networks. We also present two taxonomic structured datasets of biological specimens. Compared to the well-designed image features, i.e., LBP and HOG, the proposed method can obtain discriminative and representative features for each task, yielding much better taxonomic recognition accuracy.

This section answers RQ 6: *To what extent is it possible that the classification models (or regression models) are able to validate the performance of the image features to characterise the phenotypes in support of shape analysis*? It conveys us the message that the CNN architecture is very helpful to characterise the phenotypes including shape and texture from macroscopic to microscopic imaging scale. Importantly, we find that a good estimation on a higher level in a taxonomy probably is not helpful to improve the recognition accuracy on the level of species. In order to further explore this, we can apply a structured prediction model such as the popular CNN+RNN architecture [174]. Regarding the application of phenotype characterisation using microscopy, we need to solve the problem of limited availability of annotated training data. In this context, semi-supervised or weakly supervised learning algorithms should be taken into account. Moreover, an increasing size of the dataset will also help.

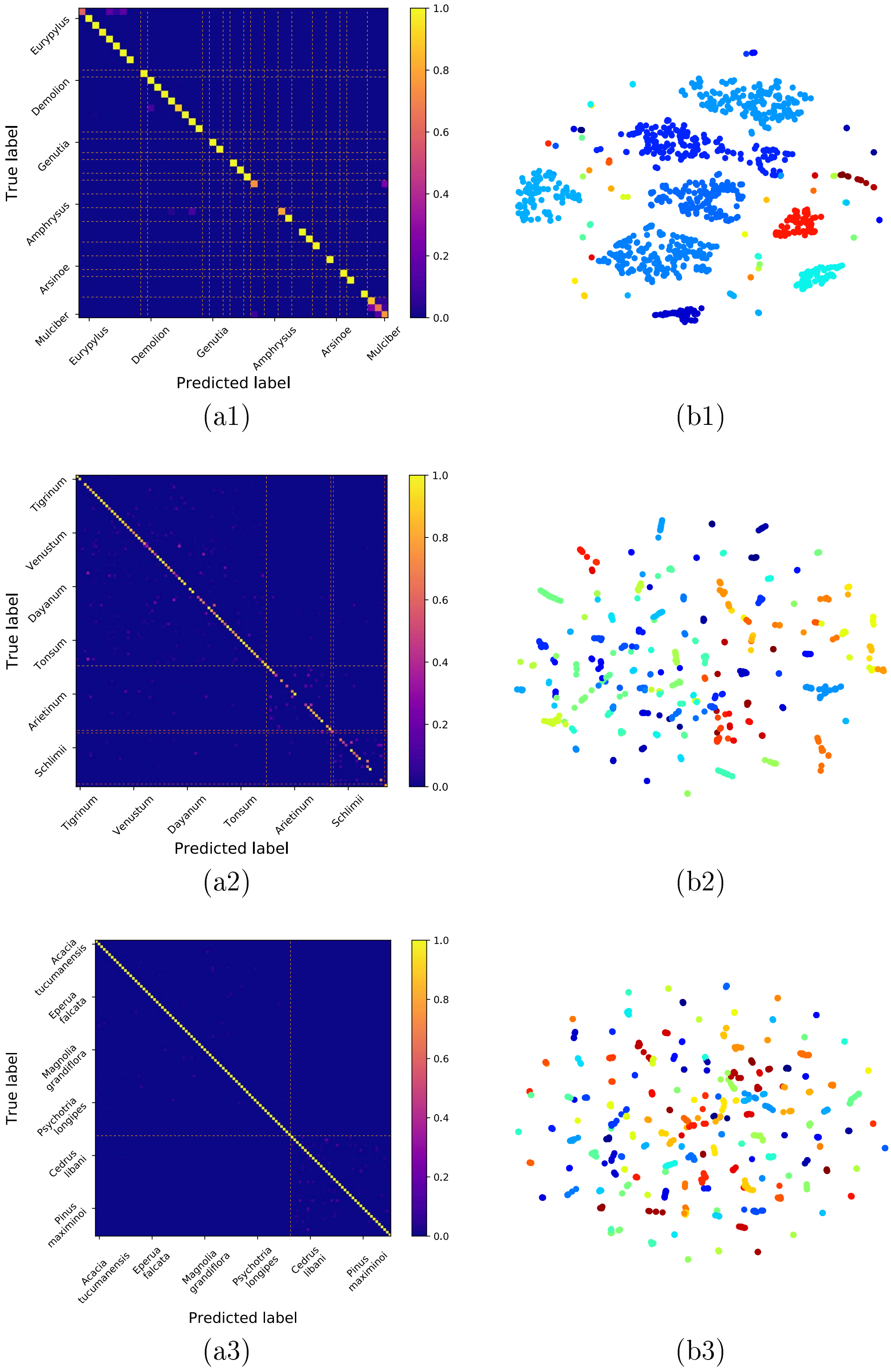


Figure 6.9: (a1)-(a3) Confusion matrix and (b1)-(b3) t-SNE map of *Butterflies*, *Orchids* and *Woods* obtained from the proposed method.

Chapter 7

Conclusions and Discussion

In this thesis we have developed an architecture for multi-modal high-throughput axial-view imaging (MM-HTAI) including six new computational approaches for shape analysis in support of phenotype characterisation in life-sciences research. The shape analysis is conditioned by good 2D and 3D shape descriptions. The 3D shape description are further interpreted as 3D measurements represented as volume and surface area. Specifically, the 3D measurements are derived from our 3D modelling approaches and can serve as an assessment for size and shape in a biological model system, e.g., zebrafish. The proposed approaches are developed for high-throughput (HT) applications such as HT compound screening which requires massive and reproducible evaluations. In addition, we have developed a pipeline which incorporates well-designed image features and a graphical model to predict the kinship. This represents a particular example of taxonomy applied to a group of faces. We have extended our analysis with a CNN architecture for accurate taxonomy prediction in different datasets. These provide an insight into the behaviour of our system that can be transferred to shape analysis using popular model systems such as zebrafish. In this chapter, we summarise our answers to the six research questions (RQs) (Section 7.1). In the after we address the problem statement (PS) (Section 7.2). Next we discuss limitations of our current methods and subsequently propose possible solutions to address these new challenges (Section 7.3). Finally, we formulate recommendations for future research (Section 7.4).

7.1 Answers to the six research questions

In Chapter 2, we answered RQ 2: *To what extent is it possible to obtain an accurate 2D shape description for the zebrafish from the MM-HTAI architecture*? We stressed that an accurate 2D shape description for a zebrafish larva is of importance for both shape and phenotype analysis as well as for the subsequent 3D reconstruction method. A good 2D shape representation should clearly present the object as a whole. In this manner, we can accurately evaluate the shape variations of the object and identify anomalies. However, in the case of the zebrafish, the transparent part of an object especially challenges almost all the current segmentation methods. So, in this chapter, we focussed on the development of an efficient and robust hybrid method for zebrafish segmentation. With the developed method we are able to obtain a whole shape representation for the zebrafish from bright-field microscopy. Instead of investigating very complex systems, we combined the merits of the unsupervised learning method i.e., mean shift and the edge based level set method. The mean shift algorithm is able to obtain an approximation for the whole shape of the zebrafish whereas the edge based level set method are able to retain the clear contour. The 2D shape approximation obtained by the mean shift algorithm also provides the level set method with a good initialization thereby accelerating the convergence for curve evolution. The implementation of this idea made our method suitable in bright field microscopy in HTI. Furthermore, we developed a process to split, align and stitch the two segmentation candidates. In addition, we also developed an efficient refinement on the hybrid result and obtained better 2D shape representations suitable for axial-view zebrafish imaging.

In Chapter 3, we answered RQ 1: *To what extent is it possible to develop an MM-HTAI architecture for the zebrafish larvae*? and RQ 3: *To what extent is it possible to obtain precise 3D shape description and derive accurate 3D measurements that are statistically relevant for the zebrafish from the MM-HTAI architecture*? Taking the 3D nature of the shape for an organism, the 3D shape analysis using 3D measurements of volume and surface area can obtain a more robust and stable assessment. This is only available with the help of a good 3D shape representation. The conventional 3D imaging modalities can obtain 3D images, while the low imaging efficiency of and complicated post processing should be addressed. So, we have implemented the MM-HTAI architecture based on the VAST BioImager and light microscopy to acquire axial-view images for the zebrafish. Next we have developed the shape-based 3D reconstruction method using a few amounts of axial-view images. This method is inspired by the multi-view stereo, and as such the 3D reconstruction was efficiently implemented. To guarantee a good shape-based 3D reconstruction, we have presented the voxel residual volume maximisation algorithm for camera calibration. From the 3D modelling, we have obtained the 3D measurements for the zebrafish larvae in three larval stages and reported the 3D measurements as statistical representations. The first merit of this work is providing an accurate shape reference to normalise the assessment in phenotype analysis. In addition, the statistical representations for the 3D measurements enables rapid shape screening for applications using zebrafish larvae. The other merit in this work is that we have obtained natural 3D visualisations for the zebrafish larvae which can be used as a shape basis for an integrated zebrafish atlas [175].

In Chapter 4, we answered RQ 4: *How can we efficiently deal with the translucency and transparency of specimen in light microscopy and still obtain a good 3D shape description from the MM-HTAI architecture*? We addressed the challenge that the shape-based 3D reconstruction method requires accurate 2D shapes for the zebrafish. However, in some cases, the boundaries for the specimen are not well-defined, which prevents the production of an accurate 2D shape representation. In Chapter 2, we have developed a new approach for accurate segmentation of zebrafish, though, this is not always feasible in all applications. In addition to our earlier work, we therefore have developed a two-phase method to address the problem. We first developed an improved volumetric representation as a confidence map which takes a confidence score for each voxel in 3D space. The confidence map is estimated from the textures of the original axial-view images. Next we have applied the region based level set method to explore the optimal 3D shape description over the confidence map. In comparison with the 3D measurements obtained from the shape-based 3D reconstruction which can be regarded as approximations of groundtruth, we have found that the proposed two-phase method can produce sufficiently accurate 3D measurements. We also have shown the feasibility of the method in high-resolution imaging settings.

In Chapter 5, we answered RQ 5: *How can we obtain a multi-modal 3D description and the corresponding measurements for the zebrafish from the MM-HTAI architecture*? We indicated that an accurate 3D shape description for organ development provides important measurements for toxicology. For example, quantitative endpoints like organ size or growth retardation are very much desired for a good assessment. As a result, we have developed a multi-modal 3D reconstruction for the zebrafish larvae combining whole-mount bright-field with organ scale modelling. With the help of our MM-HTAI architecture described in Chapter 3, we acquired the bright-field images representing the whole zebrafish and the fluorescent images representing the organs under study (e.g. liver). We have adapted our previous shape-based 3D reconstruction method to obtain the multi-modal 3D reconstruction and developed an alignment to fuse the 3D multi-models. We have reported the 3D measurements for the zebrafish and its liver and found a trend that a larger organism tend to have a larger liver.

In Chapter 6, we answered RQ 6: *To what extent is it possible that the classification models (or regression models) are able to validate the performance of the image features to characterise the phenotypes in support of shape analysis*? We have demonstrated that if we use higher magnifications and resolution in our imaging e.g., tissue or cellular scale, the textures in images can convey informative features for shape analysis. Therefore, we made a pipeline which integrates various types of image features and classification models to validate feasibility of the image features in our application. In Subsection 6.1, we used local binary patterns (LBP) on human faces as present in images. From the extracted features, we trained a multi-class logistic regressor for kinship recognition according to the facial appearance similarity. A set of semantic kinship graphs were learned offline and applied at testing time to estimate the kinship in a group of people. In Subsection 6.2, we have proposed a CNN architecture which is suitable for taxonomy prediction for different datasets including butterflies, orchids and wood species. In order to accelerate the training process, we have presented a fine-tune strategy using the CNN models trained on a large scale of images. The results show the accuracy of the proposed methods compared to the baseline methods. This provides a good understanding for the performance of our methods in a large scale of texture based classification problems.

7.2 Answers to the general problem statement

From the answers to the six RQs, we will address the PS.

> **PS**: *To what extent can we develop a stable HTI architecture and produce a robust and accurate shape analysis for phenotype characterization from the HTI architecture?*

A feasible HTI architecture is necessary to ensure an efficient and sufficient sampling size for model system. However, in life-sciences, efficient acquisition of images is limited by the complicated manipulation of a small specimen like zebrafish larvae. The zebrafish are always positioned along their longitudinal axis; multiple axial-views are commonly used to depict the zebrafish larvae. The VAST BioImager has been developed for the purpose that one can easily manipulate a zebrafish larva in any arbitrary axial-view. This leads the HTI architecture to the HTAI architecture. In practice, the VAST BioImager can be mounted on various types of microscopes, such as bright-field, fluorescence and confocal, so we can use the VAST BioImager to manipulate the zebrafish and have the microscope produce the images. If we employ different types of microscopes, the HTAI architecture is translated into the MM-HTAI architecture. This architecture is able to obtain multi-modal images for a specimen presented to the observer in an efficient manner. Of course, proper adaptions for the imaging software are required. *Therefore, we have handled the problem of the development of a stable HTI architecture by constraining the imaging in axial-view. In addition, the employment of multi-modal microscopy extends the architecture to the MM-HTAI architecture.*

Actually, a complete shape analysis requires delicate shape features, such as shape context [148], statistical shape models [60] and other features including convexity, compactness, curvature, moments etc. [176]. All the shape features should be investigated from the geometrical primitives of a shape. As a result, the prerequisite for shape analysis lies in available and accurate 2D/3D shape description. With the help of the MM-HTAI, we have sufficient data for unbiased shape analysis. However, existing methods fail to generalise the shape description for the zebrafish. So, we have developed the six new approaches towards a reliable shape analysis both in 2D and 3D: (1) the hybrid segmentation method for zebrafish segmentation, (2) the shape-based 3D reconstruction method, (3) the two-phase 3D reconstruction method in light microscopy, (4) the multi-modal 3D reconstruction, (5) the graphical model for kinship recognition, and (6) the adapted CNN architecture for image based taxonomy.

Now the question is: How can we validate the robustness and accuracy of the obtained shape descriptions? We have designed three strategies to answer this question. In Chapter 2, the first strategy is the employment of the groundtruth shape description as manually annotated contours in 2D. We have found that the shape description obtained by our method matches well with its groundtruth counterpart. In Chapter 3, the second strategy is to compute the scalar primitives for a shape including volume and surface area in 3D. We have introduced a known-size calibration particle and our method have yielded very accurate diameter, volume and surface area for these particles. This knowledge can be transferred to validate the 3D shape description of the zebrafish. In Chapter 6, the last strategy is developed to validate the performance of visual features through classification models. This resembles the behaviours of the visual features for the phenotype characterisation. *Therefore, we have handled the problem of robust and accurate shape analysis for phenotype characterisation by the production of robust and accurate shape descriptions and 3D measurements. The research community could use our results for further shape analysis as required in their fields.*

7.3 Limitations and possible solutions

We believe that the methods proposed in this thesis will be able to handle the six RQs and address PS. However, we have to concern the limitations of our approaches.

(1) From the hybrid segmentation method, we can obtain very accurate zebrafish segmentation results in bright-field imaging and we have measured such with segmentation accuracy and F1 score. We have to realise that in our application, there is always only one subject which is oriented in its longitudinal direction. It is also important for our method that the imaging condition should be well controlled such that the zebrafish in the images is depicted as a whole. This is an example which is difficult to be generalised by conventional methods. In this context, we can consider our method as a dedicated exploration for the conventional methods. However, we have not yet validated the performance of our approach under more challenged circumstance, such as the images (A) with serious lighting variation and (B) with multiple objects which are positioned in different orientations. To address these new challenges, we may incorporate orientation detection and multi-initialisation to our current method.

(2) We have indicated that the shape-based 3D reconstruction approach depends on good 2D shape representations. The visibility of a point on the zebrafish surface to an image plane e.g. profile-view is ambiguous. This will complicate the segmentation and subsequently result in an inaccurate 3D shape. We should note that in the case of zebrafish, for some axial-views, e.g., ventral and dorsal, the 2D shapes are more observable. In our shape-based 3D reconstruction method, we have to investigate a proper threshold for the visibility of each 3D point to estimate the optimal 3D surface. (A) The first limitation of the method lies in a trivial investigation for a proper threshold. Furthermore, the camera calibration model is somewhat sensitive to initialisation. From our imaging architecture, we can obtain good estimations for the intrinsic configurations including focal length, CCD sensor size and pixel size from the camera of the camera specification. We can also provide an approximate estimation for the extrinsic configurations i.e., the camera poses according to the pinhole camera projection model. Indeed, the operations have sufficiently improved our camera calibration which has been successfully applied to our setup. (B) We have admit that if a good initial estimation for the camera parameters is unavailable, especially for the camera pose, we cannot any more guarantee a good performance of the method. In our applications,

the shape of our organism system is longitudinal and holds a convex surface. Our method can obtain an accurate convex-hull for an organism. (C) However, we have to realise that the method cannot deal with estimations for concave parts on the object surface. Although most of the biological models own a convex surface, we should consider to address this challenge in our method.

(3) In addition to our previous shape-based 3D reconstruction method, the improved two-phase method is less dependent on 2D shape representation. We should realise that this approach needs a 2D shape approximation for a sufficient texture sampling for the object and background in the original axial-view zebrafish images. This can, however, easily be addressed by the mean shift algorithm in our application. (A) We still need to test our method in diverse model systems which probably are difficult to handle with the mean shift algorithm. Another important issue is to investigate a proper c-level set for the optimal 3D surface estimation. We know that the sampling size of our dataset is sufficiently large i.e., 60 subjects using our imaging condition. We can use this dataset for a good estimation of the optimal c-level set. When a various lighting condition is employed, the colour distribution will be quite different. (B) It will be difficult to generalise the texture distribution using our current dataset. To solve this problem, we have to sample more subjects in more complex imaging condition to enlarge our dataset, so that we may obtain a more generalise estimation. In Chapter 4, we have evaluated the method on our dataset and a supplement dataset produced from the microscope. Both evaluations show accurate 3D measurements in comparison with the baseline method. (C) However, we have not validated the method for other model organisms, i.e., daphnia, etc..

(4) In principle, the confocal laser scanning microscope (CLSM) is widely used to acquire 3D images in plan-parallel slices. However, with standard equipment, this imaging method is very time-consuming and the image quality is subjected to the strength of the fluorescent markers. Our multi-modal 3D reconstruction method takes the shape-based 3D reconstruction as the basis and produces natural 3D modelling on the scale of organs using a regular fluorescence microscope. The developed method is efficient, which is, however, to a certain extent hampered by the quality of the fluorescent images. Some axial-view images for the organ, zebrafish liver in this work, fail to depict a whole shape. This is caused by by the self-occlusion (the thick yolk occludes the liver from the view of dorsal). As stated in previous section, we have to investigate a proper threshold for the estimation of the optimal 3D surface in our method.

(5) We have proved the feasibility of the features and graphical models in our current taxonomical datasets on detailed texture scale. (A) We, however, realise that for each dataset the amount of examples was still limited. On the one hand, we need more data to ensure a more general and robust fitting model; on the other hand, we probably need to consider a weakly supervised strategy to obtain a good model from a small number annotated instances. This is significant in life-science research in which collecting carefully labelled dataset is very expensive and sometimes cumbersome. (B) In addition, further application of the approaches will require an evaluation of the experimental settings of our method with respect to zebrafish imaging which is, at the moment, not available yet.

7.4 Future research

Based on the discussion of our current work, we provide six recommendations for future research.

We have separately developed several new approaches for corresponding tasks i.e., 2D shape representation acquisition, 3D shape reconstruction, multi-modal 3D reconstruction on multiple scales. Our first recommendation is to integrate all the individual modules into one framework (software) in support of shape and phenotype analysis. The framework should be able to communicate with the imaging architecture. In fact, the user only needs to prepare specimens and load the specimens in the imaging device. The whole system will capture axial-view images, pre-process the acquired images, obtain 2D shape for the subject, optimise the camera configuration and generate 3D shape representation for the whole-mount and organ scale of the zebrafish larvae. In the end, accurate 3D measurements of volume and surface area are done for each 3D modelling task.

In our current platform, the imaging process for the zebrafish larvae is accomplished in a sequential fashion. Once the imaging is done, the 3D reconstruction can be performed offline. Our second recommendation is to accelerate the whole pipeline by parallelisation. A straightforward manner is to accelerate the computation by distributing the computations for each subject to different CPU cores. Importantly, the most computationally expensive process is to keep track of the projection for each voxel in 3D space to each of the axial-view images. This is densely operated especially for the camera system calibration due to the massive evaluations of the objective function. Our third recommendation is to employ a parallelization scheme by accelerating the computation under the condition that all the voxels in 3D space are independent. This can also be accomplished by the employment of GPU.

Our fourth recommendation is to apply our unified framework on diverse model organisms and subsequently evaluate its performance of a generalization. We have shown the successful application of our methods on the zebrafish larvae. We believe that similar results may be achieved on different model systems.

Currently, we have acquired the 3D models for whole-mount zebrafish larvae and some of its organs, i.e., the liver and cartilage. Our fifth recommendation is to apply our method for more organ systems 3D modelling like the zebrafish blood vessels and heart. We hope to finally integrate all the 3D models to comply with the zebrafish [19] and with other modalities like OPT [22]. Then, the zebrafish atlas can be used for accurate modelling and visualisation of gene expression and the development of various diseases.

In addition, our last recommendation refers to the imaging for the zebrafish under experimental conditions. With the control group we have collected, we will evaluate features and classification models on the zebrafish images enabling real texture based phenotype analysis.

References

[1] Thain, M., Hickman, M.: Penguin dictionary of biology. Penguin books (2004)

[2] Nezhinsky, A.: Pattern recognition in high-throughput zebrafish imaging. PhD thesis, Leiden University (2013)

[3] Cao, L.: Biological model representation and analysis. PhD thesis, Leiden University (2014)

[4] Tleis, M.: Image analysis for gene expression based phenotype characterization in yeast cells. PhD thesis, Leiden University (2016)

[5] Thompson, D.: On growth and form. On growth and form. (1942)

[6] Koch, B., Stougaard, J., Spaink, H.: Keeping track of the growing number of biological functions of chitin and its interaction partners in biomedical research. Glycobiology **25** (2015) 469–482

[7] Veneman, W., Marín-Juez, R., de Sonneville, J., Ordas, A., Jong-Raadsen, S., Meijer, A., Spaink, H.: Establishment and optimization of a high throughput setup to study staphylococcus epidermidis and mycobacterium marinum infection as a model for drug discovery. Journal of Visualized Experiments (2014)

[8] Sang, J.: Drosophila melanogaster: the fruit fly. Encyclopedia of Genetics (2001) 157–162

[9] White, J., Southgate, E., Thomson, J., Brenner, S.: The structure of the nervous system of the nematode Caenorhabditis elegans. Philos Trans R Soc Lond B Biol Sci **314** (1986) 1–340

[10] Rosenthal, N., Brown, S.: The mouse ascending: perspectives for human-disease models. Nature Cell Biology **9** (2007) 993–999

[11] Rennekamp, A., Peterson, R.: 15 years of zebrafish chemical screening. Current Opinion in Chemical Biology **24** (2015) 58–70

[12] Howe, K., et al.: The zebrafish reference genome sequence and its relationship to the human genome. Nature **496** (2013) 498–503

[13] Kimmel, C., Ballard, W., Kimmel, S., Ullmann, B., Schilling, T.: Stages of embryonic development of the zebrafish. Developmental Dynamics **203** (1995) 253–310

[14] Ali, S., Champagne, D., Spaink, H., Richardson, M.: Zebrafish embryos and larvae: A new generation of disease models and drug screens. Birth Defect Research (Part C) **93** (2011) 115–133

[15] Deo, R., MacRae, C.: The zebrafish: scalable in vivo modeling for systems biology. Wiley Interdiscip Rev. Syst. Biol. Med. **3** (2011) 335–346

[16] Moro, E., et al.: Generation and application of signaling pathway reporter lines in zebrafish. Mol. Genet. Genomics **288** (2013) 231–242

[17] He, S., Lamers, G., Beenakker, J.M., Cui, C., Ghotra, V., Danen, E., Meijer, A.H., Spaink, H.P., Snaar-Jagalska, B.: Neutrophil-mediated experimental metastasis is enhanced by VEGFR inhibition in a zebrafish xenograft model. Journal of Pathology **227** (2012) 431–445

[18] Sacco, A., Roccaro, A., Ma, D., Shi, J., Mishima, Y., Moschetta, M., Chiarini, M., Munshi, N., Handin, R., Ghobrial, I.: Cancer cell dissemination and homing to the bone marrow in a zebrafish model. Cancer Research **76** (2016) 463–471

[19] Verbeek, F., Boon, P.: High-resolution 3d reconstruction from serial sections: microscope instrumentation, software design, and its implementations. In: International Symposium on Biomedical Optics, International Society for Optics and Photonics (2002) 65–76

[20] Welten, M.: Spatio-temporal gene expression analysis from 3D in situ hybridization images. PhD thesis, Leiden University (2007)

[21] Pardo-Martin, C., Allalou, A., Medina, J., Eimon, P., Wählby, C., Yanik, M.: High-throughput hyperdimensional vertebrate phenotyping. Nature Communications **4** (2013) 1467

[22] Tang, X., van't Hoff, M., Hoogenboom, J., Guo, Y., Cai, F., Lamers, G., Verbeek, F.: Fluorescence and bright-field 3D image fusion based on sinogram unification for optical projection tomography. In: IEEE Conference on Bioinformatics and Biomedicine, IEEE (2016) 403–410

[23] McGrath, P., Li, C.: Zebrafish: a predictive model for assessing drug-induced toxicity. Drug Discovery Today **13** (2008) 394–401

[24] Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: IEEE Conference on Computer Vision and Pattern Recognition. Volume 1., IEEE (2005) 886–893

[25] Krizhevsky, A., Sutskever, I., Hinton, G.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems. (2012) 1097–1105

[26] Pardo-Martin, C., Chang, T., Koo, B., Gilleland, C., Wasserman, S., Yanik, M.: High-throughput in vivo vertebrate screening. Nature Methods **7** (2010) 634–636

[27] Szeliski, R.: Computer vision: algorithms and applications. Springer Science & Business Media (2010)

[28] Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. International Journal of Computer Vision **22** (1997) 61–79

[29] Comaniciu, D., Meer, P.: Mean shift: A robust approach toward feature space analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (2002) 603–619

[30] Hartley, R., Zisserman, A.: Multiple view geometry in computer vision. Cambridge University Press (2003)

[31] Fitzgibbon, A., Cross, G., Zisserman, A.: Automatic 3D model construction for turn-table sequences. In: European Workshop on 3D Structure from Multiple Images of Large-Scale Environments, Springer (1998) 155–170

[32] Friedman, J., Hastie, T., Tibshirani, R.: The elements of statistical learning. Volume 1. Springer Series in Statistics Springer, Berlin (2001)

[33] Bishop, C.: Pattern recognition. Machine Learning **128** (2006) 1–58

[34] Van Wijk, R., et al.: The zebrafish as model for translational systems pharmacology: expanding the allometric scale in vertebrates with five orders of magnitude, (ISSN 1871-6032)

[35] Van Wijk, R., et al.: A parent-metabolite pharmacokinetic model of paracetamol in zebrafish, (ISSN 1871-6032)

[36] Pham, L., Kanther, M., Semova, I., Rawls, J.: Methods for generating and colonizing gnotobiotic zebrafish. Nature Protocols **3** (2008) 1862–1875

[37] Ordas, A., et al.: Testing tuberculosis drug efficacy in a zebrafish high-throughput translational medicine screen. Antimicrobial Agents and Chemotherapy **59** (2015) 753–762

[38] Nezhinsky, A., Verbeek, F.: Pattern recognition for high throughput zebrafish imaging using genetic algorithm optimization. In: IEEE Conference on Pattern Recognition in BioInformatics. Springer (2010) 301–312

[39] Kantae, V., Krekels, E., et al.: Pharmacokinetic modeling of paracetamol uptake and clearance in zebrafish larvae: Expanding the allometric scale in vertebrates with five orders of magnitude. Zebrafish **13** (2016) 504–510

[40] Guo, Y., Veneman, W., Spaink, H., Verbeek, F.: Silhouette-based 3D model for zebrafish high-throughput imaging. In: IEEE Conference on Image Processing Theory, Tools and Applications, IEEE (2015) 403–408

[41] Chan, T., Vese, L.: Active contours without edges. IEEE Transactions on Image Processing **10** (2001) 266–277

[42] Li, C., Huang, R., Ding, Z., Gatenby, J., Metaxas, D., Gore, J.: A level set method for image segmentation in the presence of intensity inhomogeneities with application to MRI. IEEE Transactions on Image Processing **20** (2011) 2007–2016

[43] Zhang, Y., Matuszewski, B., Shark, L., Moore, C.: Medical image segmentation using new hybrid level-set method. In: Conference on BioMedical Visualization, IEEE (2008) 71–76

[44] Bai, P., Liu, Q., Li, L., Teng, S., Li, J., Cao, M.: A novel region-based level set method initialized with mean shift clustering for automated medical

image segmentation. Computers in Biology and Medicine **43** (2013) 1827–1832

[45] Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. International Journal of Computer Vision **1** (1988) 321–331

[46] Osher, S., Sethian, J.: Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. Journal of Computational Physics **79** (1988) 12–49

[47] Goldenberg, R., Kimmel, R., Rivlin, E., Rudzsky, M.: Fast geodesic active contours. IEEE Transactions on Image Processing **10** (2001) 1467–1475

[48] Chen, Y., Tagare, H., Thiruvenkadam, S., Huang, F., Wilson, D., Gopinath, K., Briggs, R., Geiser, E.: Using prior shapes in geometric active contours in a variational framework. International Journal of Computer Vision **50** (2002) 315–328

[49] Chan, T., Zhu, W.: Level set based shape prior segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition. Volume 2., IEEE (2005) 1164–1170

[50] Cremers, D., Rousson, M., Deriche, R.: A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. International Journal of Computer Vision **72** (2007) 195–215

[51] Schoenemann, T., Cremers, D.: Introducing curvature into globally optimal image segmentation: Minimum ratio cycles on product graphs. In: IEEE Conference on Computer Vision, IEEE (2007) 1–6

[52] Schoenemann, T., Kahl, F., Masnou, S., Cremers, D.: A linear framework for region-based image segmentation and inpainting involving curvature penalization. International Journal of Computer Vision **99** (2012) 53–68

[53] Cohen, L., Kimmel, R.: Global minimum for active contour models: A minimal path approach. International Journal of Computer Vision **24** (1997) 57–78

[54] Chen, D., Mirebeau, J., Cohen, L.: Global minimum for a Finsler elastica minimal path approach. International Journal of Computer Vision (2016) 1–26

[55] Ray, S., Turi, R.: Determination of number of clusters in k-means clustering and application in colour image segmentation. In: Conference on Advances in Pattern Recognition and Digital Techniques, Calcutta, India (1999) 137–143

[56] Chuang, K., Tzeng, H., Chen, S., Wu, J., Chen, T.: Fuzzy c-means clustering with spatial information for image segmentation. Computerized Medical Imaging and Graphics **30** (2006) 9–15

[57] Ren, X., Malik, J.: Learning a classification model for segmentation. In: IEEE Conference on Computer Vision, IEEE (2003) 10–17

[58] Ning, J., Zhang, L., Zhang, D.and Wu, C.: Interactive image segmentation by maximal similarity based region merging. Pattern Recognition **43** (2010) 445–456

[59] Cheng, Y.: Mean shift, mode seeking, and clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence **17** (1995) 790–799

[60] Cootes, T., Taylor, C.J., Cooper, D., Graham, J.: Active shape models-their training and application. Computer Vision and Image Understanding **61** (1995) 38–59

[61] Lafferty, J., McCallum, A., Pereira, F.: Conditional random fields: Probabilistic models for segmenting and labeling sequence data. **1** (2001) 282–289

[62] Zhang, Y., Brady, M., Smith, S.: Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Transactions on Medical Imaging **20** (2001) 45–57

[63] Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition. (2014) 580–587

[64] Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition. (2015) 3431–3440

[65] Zheng, S., Jayasumana, S., Romera-Paredes, B., Vineet, V., Su, Z., Du, D., Huang, C., Torr, P.: Conditional random fields as recurrent neural networks. In: IEEE Conference on Computer Vision. (2015) 1529–1537

[66] Weickert, J., Romeny, B., Viergever, M.: Efficient and reliable schemes for nonlinear diffusion filtering. IEEE Transactions on Image Processing **7** (1998) 398–410

[67] Xu, G., Zhang, Z.: Epipolar geometry in stereo, motion and object recognition: a unified approach. Volume 6. Springer Science & Business Media (2013)

[68] Zhang, Z., Deriche, R., Faugeras, O., Luong, Q.: A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. Artificial Intelligence **78** (1995) 87–119

[69] Martin, W., Aggarwal, J.: Volumetric descriptions of objects from multiple views. IEEE Transactions on Pattern Analysis and Machine Intelligence (1983) 150–158

[70] Szeliski, R.: Rapid octree construction from image sequences. CVGIP: Image Understanding **58** (1993) 23–32

[71] Laurentini, A.: The visual hull concept for silhouette-based image understanding. IEEE Transactions on Pattern Analysis and Machine Intelligence **16** (1994) 150–162

[72] Matusik, W., Buehler, C., Raskar, R., Gortler, S., McMillan, L.: Image-based visual hulls. In: Annual Conference on Computer graphics and Interactive Techniques, ACM (2000) 369–374

[73] Franco, J., Boyer, E.: Exact polyhedral visual hulls. In: British Machine Vision Conference. (Volume 1.) 329–338

[74] Lazebnik, S., Furukawa, Y., Ponce, J.: Projective visual hulls. International Journal of Computer Vision **74** (2007) 137–165

[75] Furukawa, Y., Ponce, J.: Carved visual hulls for image-based modeling. In: European Conference on Computer Vision, Springer (2006) 564–577

[76] Kutulakos, K., Seitz, S.: A theory of shape by space carving. In: IEEE Conference on Computer Vision. Volume 1. (1999) 307–314

[77] Vogiatzis, G., Hernandez, C., Cipolla, R.: Reconstruction in the round using photometric normals and silhouettes. In: IEEE Conference on Computer Vision and Pattern Recognition. Volume 2. (2006) 1847–1854

[78] Kolev, K., Brox, T., Cremers, D.: Robust variational segmentation of 3D objects from multiple views. In: Joint Pattern Recognition Symposium, Springer (2006) 688–697

[79] Kolev, K., Klodt, M., Brox, T., Cremers, D.: Continuous global optimization in multiview 3D reconstruction. International Journal of Computer Vision **84** (2009) 80–96

[80] Cremers, D., Kolev, K.: Multiview stereo and silhouette consistency via convex functionals over convex domains. IEEE Transactions on Pattern Analysis and Machine Intelligence **33** (2011) 1161–1174

[81] Kolev, K., Brox, T., Cremers, D.: Fast joint estimation of silhouettes and dense 3D geometry from multiple images. IEEE Transactions on Pattern Analysis and Machine Intelligence **34** (2012) 493–505

[82] Galliani, S., Lasinger, K., Schindler, K.: Massively parallel multiview stereopsis by surface normal diffusion. In: IEEE Conference on Computer Vision. (2015) 873–881

[83] Savinov, N., Ladicky, L., Hane, C., Pollefeys, M.: Discrete optimization of ray potentials for semantic 3D reconstruction. In: IEEE Conference on Computer Vision and Pattern Recognition. (2015) 5511–5518

[84] Lensch, H., Heidrich, W., Seidel, H.P.: A silhouette-based algorithm for texture registration and stitching. Graphical Models **63** (2001) 245–262

[85] Hernández, C., Schmitt, F., Cipolla, R.: Silhouette coherence for camera calibration under circular motion. IEEE Transactions on Pattern Analysis and Machine Intelligence **29** (2007) 343–349

[86] Cao, L., Verbeek, F.J.: Analytical evaluation of algorithms for point cloud surface reconstruction using shape features. Journal of Electronic Imaging **22** (2013) 043008–043008

[87] Zhang, Y., Matuszewski, B., Shark, L., Moore, C.: Medical image segmentation using new hybrid level-set method. (In: IEEE Conference on BioMedical Visualization) 71–76

[88] Zhang, Z.: A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence **22** (2000) 1330–1334

[89] Brown, M., Lowe, D.G.: Automatic panoramic image stitching using invariant features. International Journal of Computer Vision **74** (2007) 59–73

[90] Lagarias, J., Reeds, J., Wright, M., Wright, P.: Convergence properties of the Nelder–Mead simplex method in low dimensions. SIAM Journal on Optimization **9** (1998) 112–147

[91] Hansen, N., Ostermeier, A.: Completely derandomized self-adaptation in evolution strategies. Evolutionary Computation **9** (2001) 159–195

[92] Cao, L., Verbeek, F.: Evaluation of algorithms for point cloud surface reconstruction through the analysis of shape parameters. In: IS&T/SPIE Electronic Imaging, International Society for Optics and Photonics (2012) 82900G–82900G

[93] Desbrun, M., Meyer, M., Schröder, P., Barr, A.: Implicit fairing of irregular meshes using diffusion and curvature flow. In: Conference on Computer Graphics and Interactive Techniques, ACM Press/Addison-Wesley Publishing Co. (1999) 317–324

[94] Xiong, Z., Verbeek, F.: Segmentation of zebrafish larva inhomogeneous 3D images using the level-set method. Electronic Imaging **2016** (2016) 1–9

[95] Arridge, S.: Optical tomography in medical imaging. Inverse Problems **15** (1999) R41

[96] Levoy, M., Ng, R., Adams, A., Footer, M., Horowitz, M.: Light field microscopy. ACM Transactions on Graphics **25** (2006) 924–934

[97] Broxton, M., Grosenick, L., Yang, S., Cohen, N., Andalman, A., Deisseroth, K., Levoy, M.: Wave optics theory and 3D deconvolution for the light field microscope. Optics Express **21** (2013) 25418–25439

[98] Guo, Y., Veneman, W., Spaink, H., Verbeek, F.: Three-dimensional reconstruction and measurements of zebrafish larvae from high-throughput axial-view in vivo imaging. Biomedical Optics Express **8** (2017) 2611–2634

[99] Zhang, Z., Deriche, R., Faugeras, O., Luong, Q.: A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. Artificial Intelligence **78** (1995) 87–119

[100] Seitz, S., Curless, B., Diebel, J., Scharstein, D., Szeliski, R.: A comparison and evaluation of multi-view stereo reconstruction algorithms. In: IEEE

Conference on Computer Vision and Pattern Recognition. Volume 1., IEEE (2006) 519–528

[101] Wei, J., Resch, B., Lensch, H.: Multi-view depth map estimation with cross-view consistency. In: British Machine Vision Conference. (2014)

[102] Yoon, K., Kweon, I.: Adaptive support-weight approach for correspondence search. IEEE Transactions on Pattern Analysis and Machine Intelligence (2006) 650–656

[103] Rhemann, C., Hosni, A., Bleyer, M., Rother, C., Gelautz, M.: Fast cost-volume filtering for visual correspondence and beyond. In: IEEE Conference on Computer Vision and Pattern Recognition, IEEE (2011) 3017–3024

[104] Liu, Y., Cao, X., Dai, Q., Xu, W.: Continuous depth estimation for multi-view stereo. In: IEEE Conference on Computer Vision and Pattern Recognition, IEEE (2009) 2121–2128

[105] Bleyer, M., Rhemann, C., Rother, C.: Patchmatch stereo-stereo matching with slanted support windows. In: British Machine Vision Conference. Volume 11. (2011) 1–11

[106] Galliani, S., Lasinger, K., Schindler, K.: Massively parallel multiview stereopsis by surface normal diffusion. In: IEEE Conference on Computer Vision. (2015) 873–881

[107] Furukawa, Y., Ponce, J.: Accurate, dense, and robust multiview stereopsis. IEEE Transactions on Pattern Analysis and Machine Intelligence **32** (2010) 1362–1376

[108] Huang, C., Cagniart, C., Boyer, E., Ilic, S.: A Bayesian approach to multi-view 4D modeling. International Journal of Computer Vision **116** (2016) 115–135

[109] Lowe, D.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision **60** (2004) 91–110

[110] Rublee, E., Rabaud, V., Konolige, K., Bradski, G.: ORB: an efficient alternative to SIFT or SURF. In: IEEE Conference on Computer Vision, IEEE (2011) 2564–2571

[111] Wu, C., Agarwal, S., Curless, B., Seitz, S.: Multicore bundle adjustment. In: IEEE Conference on Computer Vision and Pattern Recognition, IEEE (2011) 3057–3064

[112] Djelouah, A., Franco, J., Boyer, E., Le Clerc, F., Pérez, P.: N-tuple color segmentation for multi-view silhouette extraction. In: European Conference on Computer Vision, Springer (2012) 818–831

[113] Esteban, C., Schmitt, F.: Silhouette and stereo fusion for 3D object modeling. Computer Vision and Image Understanding **96** (2004) 367–392

[114] Ihrke, I., Kutulakos, K., Lensch, H., Magnor, M., Heidrich, W.: Transparent and specular object reconstruction. In: Computer Graphics Forum. Volume 29., Wiley Online Library (2010) 2400–2426

[115] Ye, J., Ji, Y., Li, F., Yu, J.: Angular domain reconstruction of dynamic 3D fluid surfaces. In: IEEE Conference on Computer Vision and Pattern Recognition, IEEE (2012) 310–317

[116] Liu, D., Chen, X., Yang, Y.: Frequency-based 3D reconstruction of transparent and specular objects. In: IEEE Conference on Computer Vision and Pattern Recognition. (2014) 660–667

[117] Qian, Y., Gong, M., Yang, Y.: 3D reconstruction of transparent objects with position-normal consistency. In: IEEE Conference on Computer Vision and Pattern Recognition. (2016) 4369–4377

[118] Savinov, N., Häne, C., Ladicky, L., Pollefeys, M.: Semantic 3D Reconstruction with Continuous Regularization and Ray Potentials Using a Visibility Consistency Constraint, IEEE (2016)

[119] Han, X., Leung, T., Jia, Y., Sukthankar, R., Berg, A.: MatchNet: unifying feature and metric learning for patch-based matching. In: IEEE Conference on Computer Vision and Pattern Recognition. (2015) 3279–3286

[120] Zbontar, J., LeCun, Y.: Stereo matching by training a convolutional neural network to compare image patches. Journal of Machine Learning Research **17** (2016) 1–32

[121] Lorensen, W., Cline, H.: Marching cubes: A high resolution 3D surface construction algorithm. In: ACM Siggraph Computer Graphics. Volume 21., ACM (1987) 163–169

[122] Rother, C., Kolmogorov, V., Blake, A.: Grabcut: Interactive foreground extraction using iterated graph cuts. In: ACM Transactions on Graphics. Volume 23., ACM (2004) 309–314

[123] van Wijk, R., Krekels, E., Hankemeier, T., Spaink, H., van der Graaf, P.: Systems pharmacology of hepatic metabolism in zebrafish larvae. submitted (2017)

[124] Brown, M., Lowe, D.: Automatic panoramic image stitching using invariant features. International Journal of Computer Vision **74** (2007) 59–73

[125] Laurentini, A.: The visual hull concept for silhouette-based image understanding. IEEE Transactions on Pattern Analysis and Machine Intelligence **16** (1994) 150–162

[126] Furukawa, Y., Ponce, J.: Carved visual hulls for image-based modeling. International Journal of Computer Vision **81** (2009) 53–67

[127] Field, H., Ober, E., Roeser, T., Stainier, D.: Formation of the digestive system in zebrafish. i. liver morphogenesis. Developmental Biology **253** (2003) 279–290

[128] Maloney, L., Dal Martello, M.: Kin recognition and the perceived facial similarity of children. Journal of Vision **6** (2006)

[129] Fang, R., Tang, K., Snavely, N., Chen, T.: Towards computational models of kinship verification. In: IEEE Conference on Image Processing. (2010) 1577–1580

[130] Dal Martello, M., Maloney, L.: Where are kin recognition signals in the human face? Journal of Vision **6** (2006)

[131] Xia, S., Shao, M., Fu, Y.: Kinship verification through transfer learning. In: International Joint Conference on Artificial Intelligence. Volume 3. (2011) 2539–2544

[132] Lu, J., Zhou, X., Tan, Y., Shang, Y., Zhou, J.: Neighborhood repulsed metric learning for kinship verification. IEEE Transactions on Pattern Analysis and Machine Intelligence **36** (2014) 331–345

[133] Dibeklioğlu, H., Salah, A., Gevers, T.: Like father, like son: Facial expression dynamics for kinship verification. In: IEEE Conference on Computer Vision. (2013) 1497–1504

[134] Felzenszwalb, P., Huttenlocher, D.: Pictorial structures for object recognition. International Journal of Computer Vision **61** (2005) 55–79

[135] Xia, S., Shao, M., Luo, J., Fu, Y.: Understanding kin relationships in a photo. IEEE Transactions on Multimedia **14** (2012) 1046–1056

[136] Chen, Y., Hsu, W., Liao, H.: Discovering informative social subgraphs and predicting pairwise relationships from group photos. In: ACM Conference on Multimedia. (2012) 669–678

[137] Xiong, X., De la Torre, F.: Supervised descent method and its applications to face alignment. In: IEEE Conference on Computer Vision and Pattern Recognition. (2013) 532–539

[138] Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence **24** (2002) 971–987

[139] Guo, G., Mu, G., Fu, Y., Huang, T.: Human age estimation using bio-inspired features. In: IEEE Conference on Computer Vision and Pattern Recognition. (2009) 112–119

[140] Dibeklioğlu, H., Salah, A., Gevers, T.: Are you really smiling at me? Spontaneous versus posed enjoyment smiles. In: European Conference on Computer Vision. (2012) 525–538

[141] Guo, G., Dyer, C., Fu, Y., Huang, T.: Is gender recognition affected by age? In: IEEE Conference on Computer Vision Workshops. (2009) 2032–2039

[142] Guo, G., Wang, X.: A study on human age estimation under facial expression changes. In: IEEE Conference on Computer Vision and Pattern Recognition. (2012) 2547–2553

[143] Wheeler, Q., Raven, P., Wilson, E.: Taxonomy: impediment or expedient? Science **303** (2004) 285–285

[144] Lammers, Y., Peelen, T., Vos, R.A., Gravendeel, B.: The HTS barcode checker pipeline, a tool for automated detection of illegally traded species from high-throughput sequencing data. BMC bioinformatics **15** (2014) 44

[145] Francoy, T., Wittmann, D., Drauschke, M., Müller, S., Steinhage, V., Bezerra-Laure, M., De Jong, D., Gonçalves, L.: Identification of Africanized

honey bees through wing morphometrics: two fast and efficient procedures. Apidologie **39** (2008) 488–494

[146] Silla Jr, C., Freitas, A.: A survey of hierarchical classification across different application domains. Data Mining and Knowledge Discovery **22** (2011) 31–72

[147] Wheeler, E., Baas, P., Gasson, P.: IAWA list of microscopic features for hardwood identification. (1989)

[148] Belongie, S., Malik, J., Puzicha, J.: Shape context: A new descriptor for shape matching and object recognition. In: NIPS. Volume 2. (2000)

[149] Ricard, J., Coeurjolly, D., Baskurt, A.: Generalizations of angular radial transform for 2D and 3D shape retrieval. Pattern Recognition Letters **26** (2005) 2174–2186

[150] Jia, Q., Fan, X., Liu, Y., Li, H., Luo, Z., Guo, H.: Hierarchical projective invariant contexts for shape recognition. Pattern Recognition **52** (2016) 358–374

[151] Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)

[152] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: IEEE Conference on Computer Vision and Pattern Recognition. (2015) 1–9

[153] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition. (2016) 770–778

[154] Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhudinov, R., Zemel, R., Bengio, Y.: Show, attend and tell: Neural image caption generation with visual attention. In: International Conference on Machine Learning. (2015) 2048–2057

[155] Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: Towards real-time object detection with region proposal networks. In: Advances in Neural Information Processing Systems. (2015) 91–99

[156] Silver, D., Huang, A., Maddison, C., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M.,

et al.: Mastering the game of go with deep neural networks and tree search. Nature **529** (2016) 484

[157] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A., Li, F.: Imagenet large scale visual recognition challenge. International Journal of Computer Vision **115** (2015) 211–252

[158] Martins, J., Oliveira, L., Nisgoski, S., Sabourin, R.: A database for automatic classification of forest species. Machine Vision and Applications **24** (2013) 567–578

[159] Costa, E., Lorena, A., Carvalho, A., Freitas, A., Holden, N.: Comparing several approaches for hierarchical classification of proteins with decision trees. In: Brazilian Symposium on Bioinformatics, Springer (2007) 126–137

[160] Barbedo, J., Lopes, A.: Automatic genre classification of musical signals. EURASIP Journal on Applied Signal Processing **2007** (2007) 157–157

[161] Dekel, O., Keshet, J., Singer, Y.: Large margin hierarchical classification. In: International Conference on Machine Learning, ACM (2004) 27

[162] Cesa-Bianchi, N., Gentile, C., Zaniboni, L.: Hierarchical classification: combining Bayes with SVM. In: International Conference on Machine Learning, ACM (2006) 177–184

[163] Cesa-Bianchi, N., Gentile, C., Zaniboni, L.: Incremental algorithms for hierarchical classification. Journal of Machine Learning Research **7** (2006) 31–54

[164] Wang, H., Shen, X., Pan, W.: Large margin hierarchical classification with mutually exclusive class membership. Journal of Machine Learning Research **12** (2011) 2721–2748

[165] Tsoumakas, G., Katakis, I.: Multi-label classification: An overview. International Journal of Data Warehousing and Mining **3** (2006)

[166] Wei, Y., Xia, W., Huang, J., Ni, B., Dong, J., Zhao, Y., Yan, S.: CNN: Single-label to multi-label. arXiv preprint arXiv:1406.5726 (2014)

[167] Zhao, F., Huang, Y., Wang, L., Tan, T.: Deep semantic ranking based hashing for multi-label image retrieval. In: IEEE Conference on Computer Vision and Pattern Recognition. (2015) 1556–1564

[168] Wang, J., Yang, Y., Mao, J., Huang, Z., Huang, C., Xu, W.: CNN-RNN: A unified framework for multi-label image classification. In: IEEE Conference on Computer Vision and Pattern Recognition. (2016) 2285–2294

[169] Yan, Z., Zhang, H., Piramuthu, R., Jagadeesh, V., DeCoste, D., Di, W., Yu, Y.: HD-CNN: hierarchical deep convolutional neural networks for large scale visual recognition. In: IEEE Conference on Computer Vision. (2015) 2740–2748

[170] Pereira, S., Gravendeel, B., Wijntjes, P., Vos, R.: OrchID: a generalized framework for taxonomic classification of images using evolved artificial neural networks. bioRxiv (2016)

[171] Ruck, D., Rogers, S., Kabrisky, M., Oxley, M., Suter, B.: The multilayer perceptron as an approximation to a bayes optimal discriminant function. IEEE Transactions on Neural Networks **1** (1990) 296–298

[172] Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional architecture for fast feature embedding. arXiv preprint arXiv:1408.5093 (2014)

[173] van der Maaten, L., Hinton, G.: Visualizing data using t-SNE. Journal of Machine Learning Research **9** (2008) 2579–2605

[174] Guo, Y., Liu, Y., Bakker, E., Guo, Y., Michael, L.: CNN-RNN: A large-scale hierarchical image classification framework. Multimedia Tools and Applications (2017)

[175] Verbeek, F., et al.: Visualization of complex data sets over Internet: 2D and 3D visualization of the 3D digital atlas of zebrafish development. In: Electronic Imaging, International Society for Optics and Photonics (2001) 20–29

[176] Verbeek, F.: Lecture notes of Image Analysis in Microscopy. Leiden University (2017)

List of Abbreviations

2-3DLA	**Two**-phase approach for the **3D** reconstruction and measurements from **light** microscopy **axial**-view imaging
AOS	Additive operator splitting
ASD	Axial-view Sampling Density
c-LS	c-level set
CLSM	Confocal laser scanning microscope
CNN	Convolutional neural networks
CV	Chan-Vese
dpf	Days post fertilisation
ES	Evolution Strategy
GAC	Geodesic active contours
GFP	Green fluorescent protein
HD-CNN	Hierarchical deep CNN
hdf	Hours post fertilisation
HOG	Histograms of Oriented Gradients
HT	Hight-throughput
HTAI	High-throughput axial-view imaging
HTI	Hight-throughput imaging
HY	Hybrid

ILS	Improved level set
LBP	Local Binary Patterns
LR	Logistic regressor
LRLS	Local region based level set
LSF	Level set function
MM-HTAI	Multi-modal high-throughput axial-view imaging
MS	Mean shift
MVS	Multi-view stereo
OPT	Optical projection tomography
PBS	Phosphate-Buffered saline
PCA	Principal component analysis
PFA	Paraformaldehyde
PS	Problem statement
ReLU	Rectified linear units
RNN	Recurrent neural networks
RQ	Research question
SD	Sampling density
SFM	Structure from motion
SIFT	Scale Invariant Feature Transform
SVM	Support vector machine
t-SNE	t-Distributed Stochastic Neighbor Embedding
VAST	Vertebrate Automated Screening Technology
VRV	Voxel residual volume

List of Figures

List of Tables

Summary

In this thesis we have studied shape with a particular focus on the zebrafish model system. The *shape* is an essential appearance of the phenotype of a biological specimen and it can be used to read out a current state or response or to study gene expression. Therefore, accurate shape analysis requires a precise shape description of a model system such as the zebrafish. Moreover, a sufficiently large sampling size of the specimens is necessary to ensure a justified and unbiased shape analysis. The latter is, for instance, very important for high-throughput in compound screening. All in all, top performance in zebrafish analysis requires high-throughput imaging (HTI).

In order to deal with high-throughput imaging, we aim to design an elaborate and well-performing HTI architecture. For the essential operations we need computational approaches to obtain the 2D/3D shape representations that are precise and yet can be acquired fast. The quality of the obtained shape descriptions will be validated in a straightforward manner with scalar primitives, i.e., the volume and surface area of a 3D shape. These primitives serve as 3D measurements for a robust primary shape assessment in the phenotype characterisation. Using only shape description is not sufficient, e.g., for high-resolution imaging on tissue and cellular level, so texture should be considered to complement and enhance the shape analysis.

The work in this thesis is divided in 5 research chapters that each have their own research question. The overall problem we are addressing is:

> *To what extent can we develop a stable HTI architecture and produce a robust and accurate shape analysis for the phenotype characterisation from the HTI architecture?*

In Chapter 2, we focus on methods to obtain accurate 2D shape information from microscope images. For our particular case these images result from high-

throughput imaging. So we need to extract the object, i.e. the zebrafish, from these images and this must be done as precise as possible, as the shape from the 2D images is required to construct the 3D image. In this manner we can provide the basis for a fast and accurate 3D measurement. The method will be embedded in the high-throughput axial-view imaging (HTAI) architecture that we propose. So, a hybrid segmentation method is developed which integrates (a) the mean shift algorithm and (b) the improved level set method. This method enables us to achieve an accurate 2D shape description of the zebrafish larvae.

In Chapter 3 we elaborate the 3D reconstruction of shape from axial views of the object. We investigate the architecture for axial-view imaging and question if accurate 3D measurements can be obtained from the imaging architecture. The imaging architecture, the VAST-BioImager, is the basis for the development of a new shape-based 3D reconstruction method. We demonstrated that with this method we can obtain accurate 3D shape descriptions in an efficient manner. From the 3D models, we obtain the volume and surface area. This is applied in an experiment with a large collection of zebrafish larvae of different developmental stages. For three larval stages we have produced a statistical representation of shape from the 3D measurements of the zebrafish.

In Chapter 4, the same input as in Chapter 3 is considered. However, we specifically focus on some characteristics with the objects which complicate 3D reconstruction in a direct manner. In our specific case the objects, i.e. zebrafish larvae, are partially transparent and translucent. To that end we elaborate a probabilistic approach with probabilistic models from the image textures. In this manner we are less limited by the accuracy of the segmentation; this is especially true for some sub-optimal illumination conditions. This new approach for 3D reconstruction from axial views is referred to as the two-phase 3D reconstruction approach (2-3DLA). The evaluations demonstrate a good performance at the cost of a higher computation time.

Initially, the 3D reconstructions were built from bright-field microscopy images. In high-throughput imaging other modalities are equally important. In particular fluorescence microscopy, as it allows to specifically visualise parts of the object. Therefore, in Chapter 5, we further develop the imaging architecture to be able to obtain 3D reconstructions from different imaging modalities and fuse these modalities in one model. We demonstrate this with an application of modelling of the zebrafish liver using fluorescence while the shape is reconstructed from bright-field microscopy. Our results demonstrate a multi-modal 3D reconstruction from the fusion of 3D models on the organism- and organ-level.

Besides pure shape analysis, in Chapter 6, we investigate the application of classification models (or regression models) with the help of image features in annotated datasets. We question if in this manner we can be able to validate the performance of the features in shape analysis. We use completely different material compared to the previous chapters. We use four annotated datasets including human faces, butterflies, orchids and woods. From the human faces we develop graphical model for the kinship recognition of a group of faces in images. For the butterfly, orchid and wood datasets we have adapted a convolutional neural networks (CNN) architecture, a.k.a. deep learning, for learning representative features and developing a classification for prediction of the taxonomy of the species in the datasets. For all datasets we have demonstrated very good results.

Finally, in Chapter 7 we enumerate the conclusions of the research presented in this thesis by summarising the answers of the research questions that we have introduced in Chapter 1. We then provide a balanced discussion on the proposed approaches. Finally, we offer recommendations for further research.

Samenvatting

In dit proefschrift wordt beschreven hoe we vorm hebben bestudeerd, in het bijzonder van het zebravis model systeem. Vorm is een essentiële uiting van het fenotype van een organisme en het kan worden gebruikt als meting met betrekking tot de status van het organisme na een ingreep of voor de studie van de expressie van genen. Daarom is het van belang voor de meting van vorm een exacte vormbeschrijving van het model systeem, zoals de zebravis, te hebben. Bovendien is een voldoende monstergrootte van belang om er zeker van te zijn een evenwichtige en juiste vormanalyse te kunnen doen. Dit laatste is vooral van groot belang voor "high-throughput" analyse van bijvoorbeeld grote hoeveelheden chemische componenten. Alles welbeschouwd, voor superieure prestaties in zebravis (beeld-) analyse is een "high-throughput" imaging systeem nodig.

Om goed met high-throughput imaging te kunnen werken, stellen we ons ten doel een system te ontwerpen en te implementeren dat goede prestaties heeft voor "high-througput" imaging. Voor de kern-operaties is een computationele aanpak nodig waarmee 2D en 3D representaties van de vorm op een precize en snelle wijze kunnen worden verkregen. De kwaliteit van vorm representaties zal worden gevalideerd op een eenvoudige wijze door gebruik te maken van scalaire primitieven, te weten het volume en de oppervlakte van de 3D vorm. Deze primitieven dienen als 3D metingen voor een robuuste eerste inschatting van de vorm voor de karakterisering van het fenotype. Beperking tot alleen het gebruik van vormbeschrijving is niet voldoende; voor beeldvorming op een hogere resolutie, bijvoorbeeld op weefsel- en celniveau, zou ook textuur in aanmerking moeten worden genomen teneinde de vormanalyse te complementeren en verfijnen.

Het onderoek in dit proefschrift is verdeeld in 5 hoofdstukken met een verschillende onderzoekfocus, elk hoofdstuk heeft daarmee een eigen onderzoeksvraag. Het omvattende thema dat wordt onderzocht in dit proefschrift is:

> *Kunnen we een stabiel systeem voor high-throughput imaging ontwikkelen waarmee op robuuste en accurate wijze vormanalyse voor fenotype karakterisering kan worden uitgevoerd?*

In hoofdstuk 2 ligt de nadruk op methoden waarmee accurate 2D vorm informatie uit microscoopbeelden kan worden verkregen. Voor onze specifieke toepassing komen deze beelden van een high-throughput imaging systeem. Het onderzoeksobject, de zebravis, moet uit deze beelden worden gehaald en dit moet zo precies mogelijk omdat de vorm die uit deze beelden wordt gehaald nodig is voor het construeren van een 3D beeld. Op deze wijze voorzien we in de basis voor snelle en nauwkeurige 3D meting. De methode zal worden ingebed in een, door ons voorgesteld, systeem voor high-throughput axiale beeldvorming. Een hybride segmentatie methode is ontwikkeld dewelke twee benaderingen integreert, (a) het “mean-shift” algoritme en (b) een verbeterde versie van de “level-set” methode. Dit hybride algoritme stelt ons in staat een nauwkeurige 2D beschrijving van het zebravis object te verkijgen.

In hoofdtuk 3 wordt de 3D-reconstructie van een object vanuit axiale aanzichten uitgewerkt. We onderzoeken een specifieke architectuur voor beeldvorming uit axiale aanzichten en stellen de vraag of nauwkeurige 3D metingen met dit systeem kunnen worden verkregen. Het beeldvormende systeem, de VAST-BioImager, vormt de basis voor de ontwikkeling van een nieuwe vorm-gebaseerde 3D reconstructie methode. We laten zien dat met deze methode nauwkeurige 3D vormbeschrijvingen kunnen worden verkregen op een efficiente wijze. Van de 3D modellen verkrijgen we metingen voor volume en oppervlakte. Dit is toegepast in een experiment met een grote verzameling zebravis larven van verschillende ontwikkelingsstadia. Voor 3 larvale ontwikkelingsstadia is een statische representatie van de vorm gemaakt uit de 3D metingen van de zebravissen.

In hoofdstuk 4 wordt dezelfde invoer als hoofdstuk 3 gebruikt. Echter, hier richten we ons specifiek op eigenschappen van de objecten die de 3D reconstructie op een directe wijze compliceren. In ons specifieke geval zijn de objecten, de zebravis larven, gedeeltelijk transparant en doorzichtig. Om die reden ontwikkelen we een probabilistische benadering waarbij gebruik gemaakt wordt van probabilistische modellen van de textuur uit het beeld. Op deze wijze zijn we minder beperkt door de nauwkeurigheid van de segmentatie; dit is specifiek het geval bij suboptimale belichtingscondities. Deze nieuwe benadering voor 3D reconstructie uit axiale aanzichten wordt de twee-fase 3D reconstructie genoemd (2-3DLA). De evaluaties laten zien dat de methode goed presteert waarbij wel meer rekentijd nodig is.

De 3D reconstructies zijn initieel gemaakt uit helderveld beelden. In high throughput imaging zijn andere beeldmodaliteiten evenzo belangrijk. Fluorescentie microscopie in het bijzonder, daar dit toestaat specifieke delen van het object te visualiseren. In hoofdstuk 5 besteden we daarom aandacht aan het verder ontwikkelen van het beeldvormende systeem teneinde in staat te zijn 3D reconstructies van verschillende beeldmodaliteiten te verkrijgen en deze te fuseren in een model. We laten dit zien aan de hand van een toepassing waarin de lever van de zebravis wordt gemodelleerd met gebruik van fluorescentie terwijl de gehele vorm wordt gereconstrueerd uit helderveld beelden. Onze resultaten demonstreren een multi-modale 3D reconstructie door fusie van 3D modellen op orgamismeen orgaanniveau.

Naast strikte vormanalyse onderzoeken we in hoofdstuk 6 de toepassing van classificatie modellen (regressie modellen) met behulp van kernmerken uit beelden van geannoteerde datasets. We stellen ons de vraag of langs deze wijze de kracht van de kenmerken uit de vormanalyse kan worden gevalideerd. In vergelijking met de voorgaande hoofdstukken gebruiken we geheel verschillend materiaal; te weten vier geannoteerde datasets waaronder menselijke gezichten, vlinders, orchideeën en hout. Vanuit de dataset met menselijke gezichten hebben we een grafisch model ontwikkeld voor verwantschapsherkenning van groepen gezichten binnen een beeld. Voor de vlinder-, orchidee- en hout-datasets hebben we een convolutioneel neuraal netwerk (CNN) architectuur, ook wel bekend als “deep learning”, aangepast. Hiermee kunnen we inzicht krijgen in de representatieve kenmerken van de respectievelijke datasets en een classificatie ontwikkelen voor de voorspelling van de taxonomie van de soorten (afgebeeld in het beeld) in de dataset. Voor alle datasets laten we zien dat we met deze aanpak zeer goede resultaten hebben bereikt.

In hoofdstuk 7 sluiten we af met het opsommen van de conclusies uit het onderzoek gepresenteerd in dit proefschrift door het geven van antwoorden op de onderzoeksvragen die we geïntroduceerd hebben in hoofdstuk 1. Vervolgens voorzien we in een evenwichtige discussie van de voorgestelde benaderingen. Tenslotte geven we onze aanbevelingen voor verder onderzoek.

Curriculum Vitae

Yuanhao Guo was born in Jinan City in China on December 25, 1986. In 2009, He received his Bachelors degree from Qingdao University of Science and Technology in Electronics Information Science and Technology. In that year, he was awarded the title “Shandong Province Outstanding Graduate Student” and “Shandong Province Government Scholarship”. In 2012, he received his Masters Degree from Shandong University in Pattern Recognition and Image Processing under the supervision of Professor Wei Jiang. In that year, he was awarded the “Shandong University Guanghua Scholarship”.

In September 2012, he started his PhD supported by the China Scholarship Council. In April 2014, he continued his research in Imaging and Bioinformatics Group, Leiden Institute of Advanced Computer Science, Leiden University, under the supervision of prof.dr.ir. Fons J. Verbeek. His research focused on computational methods in bio-imaging for phenotype analysis from high-throughput imaging using light microscopy. His research interests also cover the applications of machine learning and computer vision in life-sciences. In particular, he collaborated with biologists and pharmacists to investigate 3D modelling for accurate shape assessment of model systems, i.e., the zebrafish larvae. He is also interested in deep learning, for example, the deep convolutional neural networks and recurrent neural networks.

SIKS Dissertation Series (2011-2017)

2011 01 Botond Cseke (RUN), Variational Algorithms for Bayesian Inference in Latent Gaussian Models

02 Nick Tinnemeier (UU), Organizing Agent Organizations. Syntax and Operational Semantics of an Organization-Oriented Programming Language

03 Jan Martijn van der Werf (TUE), Compositional Design and Verification of Component-Based Information Systems

04 Hado van Hasselt (UU), Insights in Reinforcement Learning; Formal analysis and empirical evaluation of temporal-difference

05 Bas van der Raadt (VU), Enterprise Architecture Coming of Age - Increasing the Performance of an Emerging Discipline.

06 Yiwen Wang (TUE), Semantically-Enhanced Recommendations in Cultural Heritage

07 Yujia Cao (UT), Multimodal Information Presentation for High Load Human Computer Interaction

08 Nieske Vergunst (UU), BDI-based Generation of Robust Task-Oriented Dialogues

09 Tim de Jong (OU), Contextualised Mobile Media for Learning

10 Bart Bogaert (UvT), Cloud Content Contention

11 Dhaval Vyas (UT), Designing for Awareness: An Experience-focused HCI Perspective

12 Carmen Bratosin (TUE), Grid Architecture for Distributed Process Mining

13 Xiaoyu Mao (UvT), Airport under Control. Multiagent Scheduling for Airport Ground Handling

14 Milan Lovric (EUR), Behavioral Finance and Agent-Based Artificial Markets

15 Marijn Koolen (UvA), The Meaning of Structure: the Value of Link Evidence for Information Retrieval

16 Maarten Schadd (UM), Selective Search in Games of Different Complexity

17 Jiyin He (UVA), Exploring Topic Structure: Coherence, Diversity and Relatedness

18 Mark Ponsen (UM), Strategic Decision-Making in complex games

19 Ellen Rusman (OU), The Mind's Eye on Personal Profiles

20 Qing Gu (VU), Guiding service-oriented software engineering - A view-based approach

21 Linda Terlouw (TUD), Modularization and Specification of Service-Oriented Systems
22 Junte Zhang (UVA), System Evaluation of Archival Description and Access
23 Wouter Weerkamp (UVA), Finding People and their Utterances in Social Media
24 Herwin van Welbergen (UT), Behavior Generation for Interpersonal Coordination with Virtual Humans On Specifying, Scheduling and Realizing Multimodal Virtual Human Behavior
25 Syed Waqar ul Qounain Jaffry (VU), Analysis and Validation of Models for Trust Dynamics
26 Matthijs Aart Pontier (VU), Virtual Agents for Human Communication - Emotion Regulation and Involvement-Distance Trade-Offs in Embodied Conversational Agents and Robots
27 Aniel Bhulai (VU), Dynamic website optimization through autonomous management of design patterns
28 Rianne Kaptein (UVA), Effective Focused Retrieval by Exploiting Query Context and Document Structure
29 Faisal Kamiran (TUE), Discrimination-aware Classification
30 Egon van den Broek (UT), Affective Signal Processing (ASP): Unraveling the mystery of emotions
31 Ludo Waltman (EUR), Computational and Game-Theoretic Approaches for Modeling Bounded Rationality
32 Nees-Jan van Eck (EUR), Methodological Advances in Bibliometric Mapping of Science
33 Tom van der Weide (UU), Arguing to Motivate Decisions
34 Paolo Turrini (UU), Strategic Reasoning in Interdependence: Logical and Game-theoretical Investigations
35 Maaike Harbers (UU), Explaining Agent Behavior in Virtual Training
36 Erik van der Spek (UU), Experiments in serious game design: a cognitive approach
37 Adriana Burlutiu (RUN), Machine Learning for Pairwise Data, Applications for Preference Learning and Supervised Network Inference
38 Nyree Lemmens (UM), Bee-inspired Distributed Optimization
39 Joost Westra (UU), Organizing Adaptation using Agents in Serious Games
40 Viktor Clerc (VU), Architectural Knowledge Management in Global Software Development
41 Luan Ibraimi (UT), Cryptographically Enforced Distributed Data Access Control
42 Michal Sindlar (UU), Explaining Behavior through Mental State Attribution
43 Henk van der Schuur (UU), Process Improvement through Software Operation Knowledge
44 Boris Reuderink (UT), Robust Brain-Computer Interfaces
45 Herman Stehouwer (UvT), Statistical Language Models for Alternative Sequence Selection
46 Beibei Hu (TUD), Towards Contextualized Information Delivery: A Rule-based Architecture for the Domain of Mobile Police Work
47 Azizi Bin Ab Aziz (VU), Exploring Computational Models for Intelligent Support of Persons with Depression

48 Mark Ter Maat (UT), Response Selection and Turn-taking for a Sensitive Artificial Listening Agent

49 Andreea Niculescu (UT), Conversational interfaces for task-oriented spoken dialogues: design aspects influencing interaction quality

2012 01 Terry Kakeeto (UvT), Relationship Marketing for SMEs in Uganda

02 Muhammad Umair (VU), Adaptivity, emotion, and Rationality in Human and Ambient Agent Models

03 Adam Vanya (VU), Supporting Architecture Evolution by Mining Software Repositories

04 Jurriaan Souer (UU), Development of Content Management System-based Web Applications

05 Marijn Plomp (UU), Maturing Interorganisational Information Systems

06 Wolfgang Reinhardt (OU), Awareness Support for Knowledge Workers in Research Networks

07 Rianne van Lambalgen (VU), When the Going Gets Tough: Exploring Agent-based Models of Human Performance under Demanding Conditions

08 Gerben de Vries (UVA), Kernel Methods for Vessel Trajectories

09 Ricardo Neisse (UT), Trust and Privacy Management Support for Context-Aware Service Platforms

10 David Smits (TUE), Towards a Generic Distributed Adaptive Hypermedia Environment

11 J.C.B. Rantham Prabhakara (TUE), Process Mining in the Large: Preprocessing, Discovery, and Diagnostics

12 Kees van der Sluijs (TUE), Model Driven Design and Data Integration in Semantic Web Information Systems

13 Suleman Shahid (UvT), Fun and Face: Exploring non-verbal expressions of emotion during playful interactions

14 Evgeny Knutov (TUE), Generic Adaptation Framework for Unifying Adaptive Web-based Systems

15 Natalie van der Wal (VU), Social Agents. Agent-Based Modelling of Integrated Internal and Social Dynamics of Cognitive and Affective Processes.

16 Fiemke Both (VU), Helping people by understanding them - Ambient Agents supporting task execution and depression treatment

17 Amal Elgammal (UvT), Towards a Comprehensive Framework for Business Process Compliance

18 Eltjo Poort (VU), Improving Solution Architecting Practices

19 Helen Schonenberg (TUE), What's Next? Operational Support for Business Process Execution

20 Ali Bahramisharif (RUN), Covert Visual Spatial Attention, a Robust Paradigm for Brain-Computer Interfacing

21 Roberto Cornacchia (TUD), Querying Sparse Matrices for Information Retrieval

22 Thijs Vis (UvT), Intelligence, politie en veiligheidsdienst: verenigbare grootheden?

23 Christian Muehl (UT), Toward Affective Brain-Computer Interfaces: Exploring the Neurophysiology of Affect during Human Media Interaction

24 Laurens van der Werff (UT), Evaluation of Noisy Transcripts for Spoken Document Retrieval

25 Silja Eckartz (UT), Managing the Business Case Development in Inter-Organizational IT Projects: A Methodology and its Application

26 Emile de Maat (UVA), Making Sense of Legal Text

27 Hayrettin Gurkok (UT), Mind the Sheep! User Experience Evaluation & Brain-Computer Interface Games

28 Nancy Pascall (UvT), Engendering Technology Empowering Women

29 Almer Tigelaar (UT), Peer-to-Peer Information Retrieval

30 Alina Pommeranz (TUD), Designing Human-Centered Systems for Reflective Decision Making

31 Emily Bagarukayo (RUN), A Learning by Construction Approach for Higher Order Cognitive Skills Improvement, Building Capacity and Infrastructure

32 Wietske Visser (TUD), Qualitative multi-criteria preference representation and reasoning

33 Rory Sie (OUN), Coalitions in Cooperation Networks (COCOON)

34 Pavol Jancura (RUN), Evolutionary analysis in PPI networks and applications

35 Evert Haasdijk (VU), Never Too Old To Learn – On-line Evolution of Controllers in Swarm- and Modular Robotics

36 Denis Ssebugwawo (RUN), Analysis and Evaluation of Collaborative Modeling Processes

37 Agnes Nakakawa (RUN), A Collaboration Process for Enterprise Architecture Creation

38 Selmar Smit (VU), Parameter Tuning and Scientific Testing in Evolutionary Algorithms

39 Hassan Fatemi (UT), Risk-aware design of value and coordination networks

40 Agus Gunawan (UvT), Information Access for SMEs in Indonesia

41 Sebastian Kelle (OU), Game Design Patterns for Learning

42 Dominique Verpoorten (OU), Reflection Amplifiers in self-regulated Learning

43 Withdrawn

44 Anna Tordai (VU), On Combining Alignment Techniques

45 Benedikt Kratz (UvT), A Model and Language for Business-aware Transactions

46 Simon Carter (UVA), Exploration and Exploitation of Multilingual Data for Statistical Machine Translation

47 Manos Tsagkias (UVA), Mining Social Media: Tracking Content and Predicting Behavior

48 Jorn Bakker (TUE), Handling Abrupt Changes in Evolving Time-series Data

49 Michael Kaisers (UM), Learning against Learning - Evolutionary dynamics of reinforcement learning algorithms in strategic interactions

50 Steven van Kervel (TUD), Ontologogy driven Enterprise Information Systems Engineering

51 Jeroen de Jong (TUD), Heuristics in Dynamic Sceduling; a practical framework with a case study in elevator dispatching

2013 01 Viorel Milea (EUR), News Analytics for Financial Decision Support

02 Erietta Liarou (CWI), MonetDB/DataCell: Leveraging the Column-store Database Technology for Efficient and Scalable Stream Processing

03 Szymon Klarman (VU), Reasoning with Contexts in Description Logics

04 Chetan Yadati (TUD), Coordinating autonomous planning and scheduling

05 Dulce Pumareja (UT), Groupware Requirements Evolutions Patterns

06 Romulo Goncalves (CWI), The Data Cyclotron: Juggling Data and Queries for a Data Warehouse Audience

07 Giel van Lankveld (UvT), Quantifying Individual Player Differences

08 Robbert-Jan Merk (VU), Making enemies: cognitive modeling for opponent agents in fighter pilot simulators

09 Fabio Gori (RUN), Metagenomic Data Analysis: Computational Methods and Applications

10 Jeewanie Jayasinghe Arachchige (UvT), A Unified Modeling Framework for Service Design.

11 Evangelos Pournaras (TUD), Multi-level Reconfigurable Self-organization in Overlay Services

12 Marian Razavian (VU), Knowledge-driven Migration to Services

13 Mohammad Safiri (UT), Service Tailoring: User-centric creation of integrated IT-based homecare services to support independent living of elderly

14 Jafar Tanha (UVA), Ensemble Approaches to Semi-Supervised Learning Learning

15 Daniel Hennes (UM), Multiagent Learning - Dynamic Games and Applications

16 Eric Kok (UU), Exploring the practical benefits of argumentation in multi-agent deliberation

17 Koen Kok (VU), The PowerMatcher: Smart Coordination for the Smart Electricity Grid

18 Jeroen Janssens (UvT), Outlier Selection and One-Class Classification

19 Renze Steenhuizen (TUD), Coordinated Multi-Agent Planning and Scheduling

20 Katja Hofmann (UvA), Fast and Reliable Online Learning to Rank for Information Retrieval

21 Sander Wubben (UvT), Text-to-text generation by monolingual machine translation

22 Tom Claassen (RUN), Causal Discovery and Logic

23 Patricio de Alencar Silva (UvT), Value Activity Monitoring

24 Haitham Bou Ammar (UM), Automated Transfer in Reinforcement Learning

25 Agnieszka Anna Latoszek-Berendsen (UM), Intention-based Decision Support. A new way of representing and implementing clinical guidelines in a Decision Support System

26 Alireza Zarghami (UT), Architectural Support for Dynamic Homecare Service Provisioning

27 Mohammad Huq (UT), Inference-based Framework Managing Data Provenance

28 Frans van der Sluis (UT), When Complexity becomes Interesting: An Inquiry into the Information eXperience

29 Iwan de Kok (UT), Listening Heads

30 Joyce Nakatumba (TUE), Resource-Aware Business Process Management: Analysis and Support

31 Dinh Khoa Nguyen (UvT), Blueprint Model and Language for Engineering Cloud Applications

32 Kamakshi Rajagopal (OUN), Networking For Learning; The role of Networking in a Lifelong Learner's Professional Development
33 Qi Gao (TUD), User Modeling and Personalization in the Microblogging Sphere
34 Kien Tjin-Kam-Jet (UT), Distributed Deep Web Search
35 Abdallah El Ali (UvA), Minimal Mobile Human Computer Interaction
36 Than Lam Hoang (TUe), Pattern Mining in Data Streams
37 Dirk Börner (OUN), Ambient Learning Displays
38 Eelco den Heijer (VU), Autonomous Evolutionary Art
39 Joop de Jong (TUD), A Method for Enterprise Ontology based Design of Enterprise Information Systems
40 Pim Nijssen (UM), Monte-Carlo Tree Search for Multi-Player Games
41 Jochem Liem (UVA), Supporting the Conceptual Modelling of Dynamic Systems: A Knowledge Engineering Perspective on Qualitative Reasoning
42 Léon Planken (TUD), Algorithms for Simple Temporal Reasoning
43 Marc Bron (UVA), Exploration and Contextualization through Interaction and Concepts

2014 01 Nicola Barile (UU), Studies in Learning Monotone Models from Data
02 Fiona Tuliyano (RUN), Combining System Dynamics with a Domain Modeling Method
03 Sergio Raul Duarte Torres (UT), Information Retrieval for Children: Search Behavior and Solutions
04 Hanna Jochmann-Mannak (UT), Websites for children: search strategies and interface design - Three studies on children's search performance and evaluation
05 Jurriaan van Reijsen (UU), Knowledge Perspectives on Advancing Dynamic Capability
06 Damian Tamburri (VU), Supporting Networked Software Development
07 Arya Adriansyah (TUE), Aligning Observed and Modeled Behavior
08 Samur Araujo (TUD), Data Integration over Distributed and Heterogeneous Data Endpoints
09 Philip Jackson (UvT), Toward Human-Level Artificial Intelligence: Representation and Computation of Meaning in Natural Language
10 Ivan Salvador Razo Zapata (VU), Service Value Networks
11 Janneke van der Zwaan (TUD), An Empathic Virtual Buddy for Social Support
12 Willem van Willigen (VU), Look Ma, No Hands: Aspects of Autonomous Vehicle Control
13 Arlette van Wissen (VU), Agent-Based Support for Behavior Change: Models and Applications in Health and Safety Domains
14 Yangyang Shi (TUD), Language Models With Meta-information
15 Natalya Mogles (VU), Agent-Based Analysis and Support of Human Functioning in Complex Socio-Technical Systems: Applications in Safety and Healthcare
16 Krystyna Milian (VU), Supporting trial recruitment and design by automatically interpreting eligibility criteria
17 Kathrin Dentler (VU), Computing healthcare quality indicators automatically: Secondary Use of Patient Data and Semantic Interoperability

18 Mattijs Ghijsen (UVA), Methods and Models for the Design and Study of Dynamic Agent Organizations
19 Vinicius Ramos (TUE), Adaptive Hypermedia Courses: Qualitative and Quantitative Evaluation and Tool Support
20 Mena Habib (UT), Named Entity Extraction and Disambiguation for Informal Text: The Missing Link
21 Kassidy Clark (TUD), Negotiation and Monitoring in Open Environments
22 Marieke Peeters (UU), Personalized Educational Games - Developing agent-supported scenario-based training
23 Eleftherios Sidirourgos (UvA/CWI), Space Efficient Indexes for the Big Data Era
24 Davide Ceolin (VU), Trusting Semi-structured Web Data
25 Martijn Lappenschaar (RUN), New network models for the analysis of disease interaction
26 Tim Baarslag (TUD), What to Bid and When to Stop
27 Rui Jorge Almeida (EUR), Conditional Density Models Integrating Fuzzy and Probabilistic Representations of Uncertainty
28 Anna Chmielowiec (VU), Decentralized k-Clique Matching
29 Jaap Kabbedijk (UU), Variability in Multi-Tenant Enterprise Software
30 Peter de Cock (UvT), Anticipating Criminal Behaviour
31 Leo van Moergestel (UU), Agent Technology in Agile Multiparallel Manufacturing and Product Support
32 Naser Ayat (UvA), On Entity Resolution in Probabilistic Data
33 Tesfa Tegegne (RUN), Service Discovery in eHealth
34 Christina Manteli (VU), The Effect of Governance in Global Software Development: Analyzing Transactive Memory Systems.
35 Joost van Ooijen (UU), Cognitive Agents in Virtual Worlds: A Middleware Design Approach
36 Joos Buijs (TUE), Flexible Evolutionary Algorithms for Mining Structured Process Models
37 Maral Dadvar (UT), Experts and Machines United Against Cyberbullying
38 Danny Plass-Oude Bos (UT), Making brain-computer interfaces better: improving usability through post-processing.
39 Jasmina Maric (UvT), Web Communities, Immigration, and Social Capital
40 Walter Omona (RUN), A Framework for Knowledge Management Using ICT in Higher Education
41 Frederic Hogenboom (EUR), Automated Detection of Financial Events in News Text
42 Carsten Eijckhof (CWI/TUD), Contextual Multidimensional Relevance Models
43 Kevin Vlaanderen (UU), Supporting Process Improvement using Method Increments
44 Paulien Meesters (UvT), Intelligent Blauw. Met als ondertitel: Intelligence-gestuurde politiezorg in gebiedsgebonden eenheden.
45 Birgit Schmitz (OUN), Mobile Games for Learning: A Pattern-Based Approach
46 Ke Tao (TUD), Social Web Data Analytics: Relevance, Redundancy, Diversity
47 Shangsong Liang (UVA), Fusion and Diversification in Information Retrieval

2015 01 Niels Netten (UvA), Machine Learning for Relevance of Information in Crisis Response

02 Faiza Bukhsh (UvT), Smart auditing: Innovative Compliance Checking in Customs Controls

03 Twan van Laarhoven (RUN), Machine learning for network data

04 Howard Spoelstra (OUN), Collaborations in Open Learning Environments

05 Christoph Bösch (UT), Cryptographically Enforced Search Pattern Hiding

06 Farideh Heidari (TUD), Business Process Quality Computation - Computing Non-Functional Requirements to Improve Business Processes

07 Maria-Hendrike Peetz (UvA), Time-Aware Online Reputation Analysis

08 Jie Jiang (TUD), Organizational Compliance: An agent-based model for designing and evaluating organizational interactions

09 Randy Klaassen (UT), HCI Perspectives on Behavior Change Support Systems

10 Henry Hermans (OUN), OpenU: design of an integrated system to support lifelong learning

11 Yongming Luo (TUE), Designing algorithms for big graph datasets: A study of computing bisimulation and joins

12 Julie M. Birkholz (VU), Modi Operandi of Social Network Dynamics: The Effect of Context on Scientific Collaboration Networks

13 Giuseppe Procaccianti (VU), Energy-Efficient Software

14 Bart van Straalen (UT), A cognitive approach to modeling bad news conversations

15 Klaas Andries de Graaf (VU), Ontology-based Software Architecture Documentation

16 Changyun Wei (UT), Cognitive Coordination for Cooperative Multi-Robot Teamwork

17 André van Cleeff (UT), Physical and Digital Security Mechanisms: Properties, Combinations and Trade-offs

18 Holger Pirk (CWI), Waste Not, Want Not! - Managing Relational Data in Asymmetric Memories

19 Bernardo Tabuenca (OUN), Ubiquitous Technology for Lifelong Learners

20 Lois Vanhée (UU), Using Culture and Values to Support Flexible Coordination

21 Sibren Fetter (OUN), Using Peer-Support to Expand and Stabilize Online Learning

22 Zhemin Zhu (UT), Co-occurrence Rate Networks

23 Luit Gazendam (VU), Cataloguer Support in Cultural Heritage

24 Richard Berendsen (UVA), Finding People, Papers, and Posts: Vertical Search Algorithms and Evaluation

25 Steven Woudenberg (UU), Bayesian Tools for Early Disease Detection

26 Alexander Hogenboom (EUR), Sentiment Analysis of Text Guided by Semantics and Structure

27 Sándor Héman (CWI), Updating compressed colomn stores

28 Janet Bagorogoza (TiU), Knowledge Management and High Performance; The Uganda Financial Institutions Model for HPO

29 Hendrik Baier (UM), Monte-Carlo Tree Search Enhancements for One-Player and Two-Player Domains

30 Kiavash Bahreini (OU), Real-time Multimodal Emotion Recognition in E-Learning

31 Yakup Koç (TUD), On the robustness of Power Grids
32 Jerome Gard (UL), Corporate Venture Management in SMEs
33 Frederik Schadd (TUD), Ontology Mapping with Auxiliary Resources
34 Victor de Graaf (UT), Gesocial Recommender Systems
35 Jungxao Xu (TUD), Affective Body Language of Humanoid Robots: Perception and Effects in Human Robot Interaction

2016 01 Syed Saiden Abbas (RUN), Recognition of Shapes by Humans and Machines
02 Michiel Christiaan Meulendijk (UU), Optimizing medication reviews through decision support: prescribing a better pill to swallow
03 Maya Sappelli (RUN), Knowledge Work in Context: User Centered Knowledge Worker Support
04 Laurens Rietveld (VU), Publishing and Consuming Linked Data
05 Evgeny Sherkhonov (UVA), Expanded Acyclic Queries: Containment and an Application in Explaining Missing Answers
06 Michel Wilson (TUD), Robust scheduling in an uncertain environment
07 Jeroen de Man (VU), Measuring and modeling negative emotions for virtual training
08 Matje van de Camp (TiU), A Link to the Past: Constructing Historical Social Networks from Unstructured Data
09 Archana Nottamkandath (VU), Trusting Crowdsourced Information on Cultural Artefacts
10 George Karafotias (VUA), Parameter Control for Evolutionary Algorithms
11 Anne Schuth (UVA), Search Engines that Learn from Their Users
12 Max Knobbout (UU), Logics for Modelling and Verifying Normative Multi-Agent Systems
13 Nana Baah Gyan (VU), The Web, Speech Technologies and Rural Development in West Africa - An ICT4D Approach
14 Ravi Khadka (UU), Revisiting Legacy Software System Modernization
15 Steffen Michels (RUN), Hybrid Probabilistic Logics - Theoretical Aspects, Algorithms and Experiments
16 Guangliang Li (UVA), Socially Intelligent Autonomous Agents that Learn from Human Reward
17 Berend Weel (VU), Towards Embodied Evolution of Robot Organisms
18 Albert Meroño Peñuela (VU), Refining Statistical Data on the Web
19 Julia Efremova (Tu/e), Mining Social Structures from Genealogical Data
20 Daan Odijk (UVA), Context & Semantics in News & Web Search
21 Alejandro Moreno Célleri (UT), From Traditional to Interactive Playspaces: Automatic Analysis of Player Behavior in the Interactive Tag Playground
22 Grace Lewis (VU), Software Architecture Strategies for Cyber-Foraging Systems
23 Fei Cai (UVA), Query Auto Completion in Information Retrieval
24 Brend Wanders (UT), Repurposing and Probabilistic Integration of Data; An Iterative and data model independent approach
25 Julia Kiseleva (TU/e), Using Contextual Information to Understand Searching and Browsing Behavior

26 Dilhan Thilakarathne (VU), In or Out of Control: Exploring Computational Models to Study the Role of Human Awareness and Control in Behavioural Choices, with Applications in Aviation and Energy Management Domains
27 Wen Li (TUD), Understanding Geo-spatial Information on Social Media
28 Mingxin Zhang (TUD), Large-scale Agent-based Social Simulation - A study on epidemic prediction and control
29 Nicolas Höning (TUD), Peak reduction in decentralised electricity systems - Markets and prices for flexible planning
30 Ruud Mattheij (UvT), The Eyes Have It
31 Mohammad Khelghati (UT), Deep web content monitoring
32 Eelco Vriezekolk (UT), Assessing Telecommunication Service Availability Risks for Crisis Organisations
33 Peter Bloem (UVA), Single Sample Statistics, exercises in learning from just one example
34 Dennis Schunselaar (TUE), Configurable Process Trees: Elicitation, Analysis, and Enactment
35 Zhaochun Ren (UVA), Monitoring Social Media: Summarization, Classification and Recommendation
36 Daphne Karreman (UT), Beyond R2D2: The design of nonverbal interaction behavior optimized for robot-specific morphologies
37 Giovanni Sileno (UvA), Aligning Law and Action - a conceptual and computational inquiry
38 Andrea Minuto (UT), Materials that Matter - Smart Materials meet Art & Interaction Design
39 Merijn Bruijnes (UT), Believable Suspect Agents; Response and Interpersonal Style Selection for an Artificial Suspect
40 Christian Detweiler (TUD), Accounting for Values in Design
41 Thomas King (TUD), Governing Governance: A Formal Framework for Analysing Institutional Design and Enactment Governance
42 Spyros Martzoukos (UVA), Combinatorial and Compositional Aspects of Bilingual Aligned Corpora
43 Saskia Koldijk (RUN), Context-Aware Support for Stress Self-Management: From Theory to Practice
44 Thibault Sellam (UVA), Automatic Assistants for Database Exploration
45 Bram van de Laar (UT), Experiencing Brain-Computer Interface Control
46 Jorge Gallego Perez (UT), Robots to Make you Happy
47 Christina Weber (UL), Real-time foresight - Preparedness for dynamic innovation networks
48 Tanja Buttler (TUD), Collecting Lessons Learned
49 Gleb Polevoy (TUD), Participation and Interaction in Projects. A Game-Theoretic Analysis
50 Yan Wang (UVT), The Bridge of Dreams: Towards a Method for Operational Performance Alignment in IT-enabled Service Supply Chains

---

2017 01 Jan-Jaap Oerlemans (UL), Investigating Cybercrime

02 Sjoerd Timmer (UU), Designing and Understanding Forensic Bayesian Networks using Argumentation

03 Daniël Harold Telgen (UU), Grid Manufacturing; A Cyber-Physical Approach with Autonomous Products and Reconfigurable Manufacturing Machines

04 Mrunal Gawade (CWI), Multi-core Parallelism in a Column-store

05 Mahdieh Shadi (UVA), Collaboration Behavior

06 Damir Vandic (EUR), Intelligent Information Systems for Web Product Search

07 Roel Bertens (UU), Insight in Information: from Abstract to Anomaly

08 Rob Konijn (VU) , Detecting Interesting Differences:Data Mining in Health Insurance Data using Outlier Detection and Subgroup Discovery

09 Dong Nguyen (UT), Text as Social and Cultural Data: A Computational Perspective on Variation in Text

10 Robby van Delden (UT), (Steering) Interactive Play Behavior

11 Florian Kunneman (RUN), Modelling patterns of time and emotion in Twitter #anticipointment

12 Sander Leemans (TUE), Robust Process Mining with Guarantees

13 Gijs Huisman (UT), Social Touch Technology - Extending the reach of social touch through haptic technology

14 Shoshannah Tekofsky (UvT), You Are Who You Play You Are: Modelling Player Traits from Video Game Behavior

15 Peter Berck (RUN), Memory-Based Text Correction

16 Aleksandr Chuklin (UVA), Understanding and Modeling Users of Modern Search Engines

17 Daniel Dimov (UL), Crowdsourced Online Dispute Resolution

18 Ridho Reinanda (UVA), Entity Associations for Search

19 Jeroen Vuurens (UT), Proximity of Terms, Texts and Semantic Vectors in Information Retrieval

20 Mohammadbashir Sedighi (TUD), Fostering Engagement in Knowledge Sharing: The Role of Perceived Benefits, Costs and Visibility

21 Jeroen Linssen (UT), Meta Matters in Interactive Storytelling and Serious Gaming (A Play on Worlds)

22 Sara Magliacane (VU), Logics for causal inference under uncertainty

23 David Graus (UVA), Entities of Interest — Discovery in Digital Traces

24 Chang Wang (TUD), Use of Affordances for Efficient Robot Learning

25 Veruska Zamborlini (VU), Knowledge Representation for Clinical Guidelines, with applications to Multimorbidity Analysis and Literature Search

26 Merel Jung (UT), Socially intelligent robots that understand and respond to human touch

27 Michiel Joosse (UT), Investigating Positioning and Gaze Behaviors of Social Robots: People's Preferences, Perceptions and Behaviors

28 John Klein (VU), Architecture Practices for Complex Contexts

29 Adel Alhuraibi (UvT), From IT-BusinessStrategic Alignment to Performance: A Moderated Mediation Model of Social Innovation, and Enterprise Governance of IT"

31 Ben Ruijl (Uvt), Advances in computational methods for QFT calculations

32 Thaer Samar (RUN), Access to and Retrievability of Content in Web Archives

33 Brigit van Loggem (OU), Towards a Design Rationale for Software Documentation: A Model of Computer-Mediated Activity

34 Maren Scheffel (OU), The Evaluation Framework for Learning Analytics

35 Martine de Vos (VU), Interpreting natural science spreadsheets