Computational optimisation of optical projection tomography for 3D image analysis
Tang, X.

Citation
Tang, X. (2020, June 10). *Computational optimisation of optical projection tomography for 3D image analysis*. Retrieved from https://hdl.handle.net/1887/106088

Cover Page

Universiteit Leiden

The handle http://hdl.handle.net/1887/106088 holds various files of this Leiden University dissertation.

Author: Tang, X.
Title: Computational optimisation of optical projection tomography for 3D image analysis
Issue Date: 2020-06-10

Computational optimisation of Optical Projection Tomography for 3D image analysis

PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit Leiden,
op gezag van de Rector Magnificus Prof. Mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties,
op woensdag 10 juni 2020
klokke 16:15 uur

door

Xiaoqin Tang

geboren te Sichuan, China, in 1988

Promotiecommissie

Promotor: Prof. Dr. Ir. F.J. Verbeek

Overige Leden: Prof. Dr. J.B.L. Bard   University of Edinburgh
                                                          University of Oxford
                  Prof. Dr. R.E. Poelmann   Institute of Biology Leiden
                  Prof. Dr. F.S. de Boer   CWI, the Netherlands
                  Dr. K.F.D. Rietveld
                  Prof. Dr. A. Plaat

Contents

Chapter 1

Introduction

Chapter summary

In this chapter, we focus on presenting the general idea of our research interest and topics. It starts with the importance of three-dimensional imaging techniques in biomedical research, providing a broad view of the selection of imaging instruments according to the research requirements. Subsequently, we introduce the three-dimensional imaging system used in our research, i.e. optical projection tomography, and thoroughly elaborate it from the imaging schema, experimental setup, imaging software and sample preparation. Beside the imaging aspects, the computational approaches of optical projection tomography are explained, including three-dimensional reconstruction, segmentation and quantification. Following, seven research questions are formulated from our research perspective.

1.1 Importance of three-dimensional imaging in biomedical research

Microscopes are our eyes for things beyond our sight. Therefore, in research this instrument is indispensable. The standard microscope was designed to produce a two-dimensional (2D) image of a sample that would otherwise remain unobservable by the eyes. In this manner, an intuitive representation of details and structures can be easily transformed to an image. In the life sciences the requirement for imaging is far beyond just a qualitative description. Quantitative approaches for measurements are required.

Advances in molecular genetics have enabled molecular imaging to visualize processes in cells, cell cultures, tissues, organs and organisms with a resolution from less than a micrometer to centimeters. These possibilities are making a tremendous impact on biology and medical research. Genetic engineering technologies such as in situ hybridization as well as fluorescence staining permit the qualitative, quantitative and localization analysis of protein and gene expression patterns in animals and plants.

With respect to cells and cell cultures, the application of genetic engineering to cells allows studying signaling processes in cells and mono-layer cell cultures. However, these mono-layer cell cultures may exhibit non-physiological behavior within their artificial planar environment. Hence, there is a trend from *in vitro* to *in vivo* experimentation and thus a trend of understanding biology at the level of the scale of tissue or whole organism. Imaging modalities need to support this trend [1], [2].

In order to understand spatial organization and gene expression, three-dimensional (3D) imaging is required. On the level of tissues and organisms this has been accomplished by making physical thin sections and producing a 3D image through reconstructing from these physical sections. This technique, referred to as invasive imaging, is laborious and sometimes complicates an understanding of the sample through artefacts that are introduced in the process of sample preparation and imaging.

In the past decades, studies on disease mechanisms and drug discovery have also benefited from the high-resolution fluorescence microscopy techniques such as confocal laser scanning microscopy (CLSM) or multiphoton laser scanning microscopy (MLSM), enabling the visualization of parts of the cell signaling network [3], [4]. The sample in 3D is scanned in a plan parallel fashion using the optics in a smart manner. This approach, of non-invasive imaging, works well with cellular mono layers and relatively thin samples, i.e. in the range of tens of micrometers to a millimeter. Samples that are larger and thicker are less suitable for this kind of approach, i.e. samples larger than 2 millimeters to one centimeter. As indicated, one approach for imaging would be an invasive technique like serial sectioning. However, there are other options.

Classical non-invasive *in vivo* imaging tools such as computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET) or single-photon computed tomography (SPECT) provide a spatial resolution in the millimeter scale at organ-level in living specimens and patients. These imaging techniques make significant contribution especially for disease models of the central nervous system

(CNS), such as brain tumours, Alzheimer's disease, or multiple sclerosis.

In terms of biology and optical imaging an efficient non-invasive instrument for imaging biological tissue, organ and organism is optical projection tomography (OPT). In the recent years, cancer progression [5]–[7], drug discovery [8] and development studies such as skeleton, teeth and blood vessels have successfully used OPT imaging [9]–[12].

Figure 1.1 briefly summarizes the range of resolution at which imaging techniques in bio-medical research operates. It ranges from nanometers to centimeters with imaging scale from protein to whole organism. The white color shown between two different types of imaging indicates some overlapping in scale where both types are being applied, depending on the experimental setup and research requirement.

With respect to OPT imaging, signal acquired covers 2D information. For this device, 3D imaging is archived by rotating the samples over a full revolution (360°) and at each step capture an image. The collection of these images is known as the tomogram. From the tomogram a 3D image is reconstructed. This reconstruction process is a computational process and requires design of smart algorithms and efficient computation strategies. An example of such reconstruction algorithm is the filtered back projection (FBP) algorithm [13].

The research in this thesis focusses on the application of OPT in biomedical research. Therefore, it deals with design and implementation of algorithms and computational strategies to deal with data, i.e. images that are acquired with an OPT microscope.
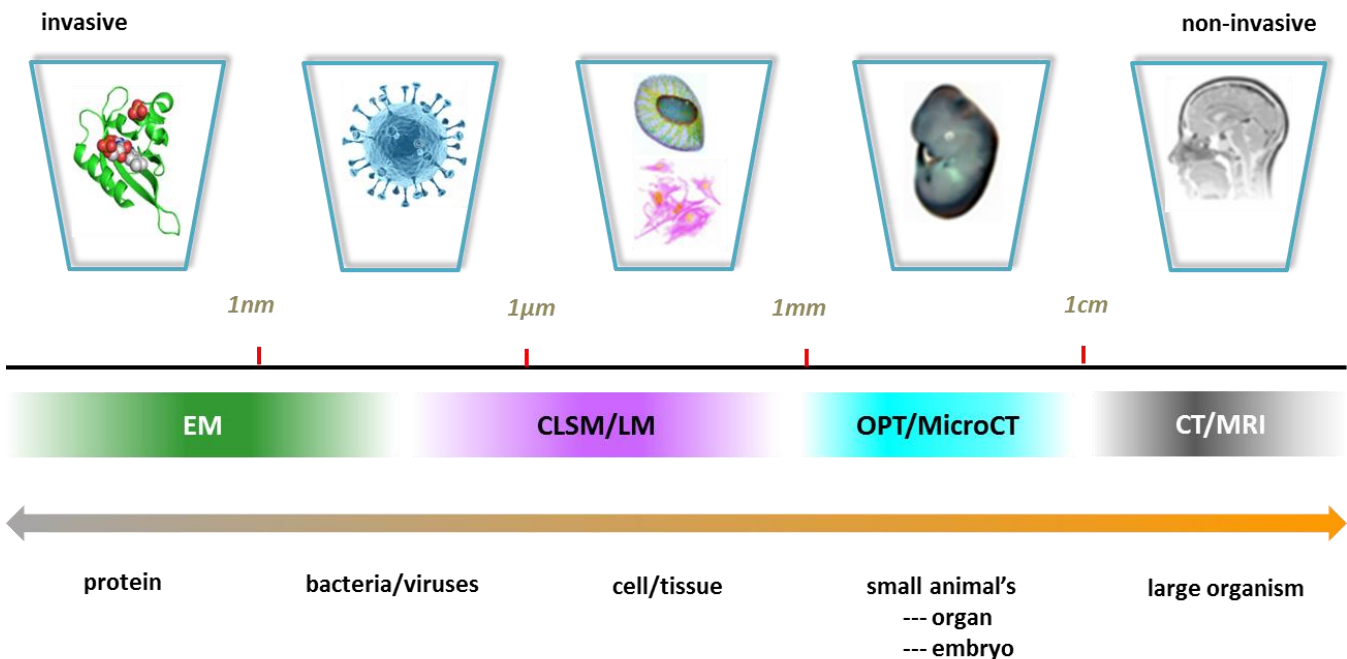
Figure 1.1. Overview of sample scale and the corresponding microscope applicable. The representations of different imaging levels are shown on the top from protein to organism, inspired by Alanis *et al.* [14].

1.2 Introduction of OPT imaging system

With respect to CLSM [15] we are confronted with a limitation of size of the specimen for whole-mount imaging. With MRI [16] the strength of the magnetic field determines the resolution that can be obtained for a whole-mount imaging. The OPT technique [17] for that matter, can overcome these shortcomings. It can visualize gene expression or specific staining in bright-field or fluorescence channel, while the specimen as a whole can be imaged. In that manner, OPT adds an important range of scale that can be imaged. It allows for the acquisition of high resolution full body images of animal/plant tissues as well as organs/organisms [18], [19]. It has been studied for the capability of imaging with good spatial resolution and contrast and minimal shadowing artefacts produced after reconstruction of a tomogram.

1.2.1 Introduction of OPT imaging schema

In Figure 1.2, we describe our imaging system conform the original set up as presented in [17]. We will focus on the OPT of zebrafish as this is the main specimen that we work with. The specimen is fixed for imaging in an agarose cylinder that can be rotated, mounted to a step motor unit using a magnet. Light transmitted (purple lines) is focused by the lenses onto the digital camera, i.e. Charge-Coupled Device (CCD). The apparatus is adjusted so that light emitted from a slice that is perpendicular to the rotation axis (red ellipse), is focused onto a single row of pixels on the CCD (red line). The slice highlighted as a red ellipse in (A) is seen as a red circle in (B). Different from CT imaging, in optical imaging system a pixel on the CCD contains the information of the specific slice of the specimen in the cone-shaped region (purple region). Points far from the focal plane will not be focused and will, as a consequence, not produce a sharp image. Contrarily, only the point closer to focal plane yields a high-quality image. The depth of the cone-shaped region is defined as depth of focus (DoF) which is determined by the property of lenses. For good OPT imaging the cone is ideally as narrow as possible and the DoF is expected to be large enough to include the whole specimen. But in practice, there is a trade-off between them; i.e, large DoF corresponds to wide cone. One typical way to solve this problem is to adjust the position of rotation centre so that only the front half of the specimen is in focus. This ensures that every part of the specimen is imaged in focus during a full revolution (360◦ rotation of the specimen). The sampled regions (C) from adjacent pixels are distributed across the section as an approximation of parallel line integrals.
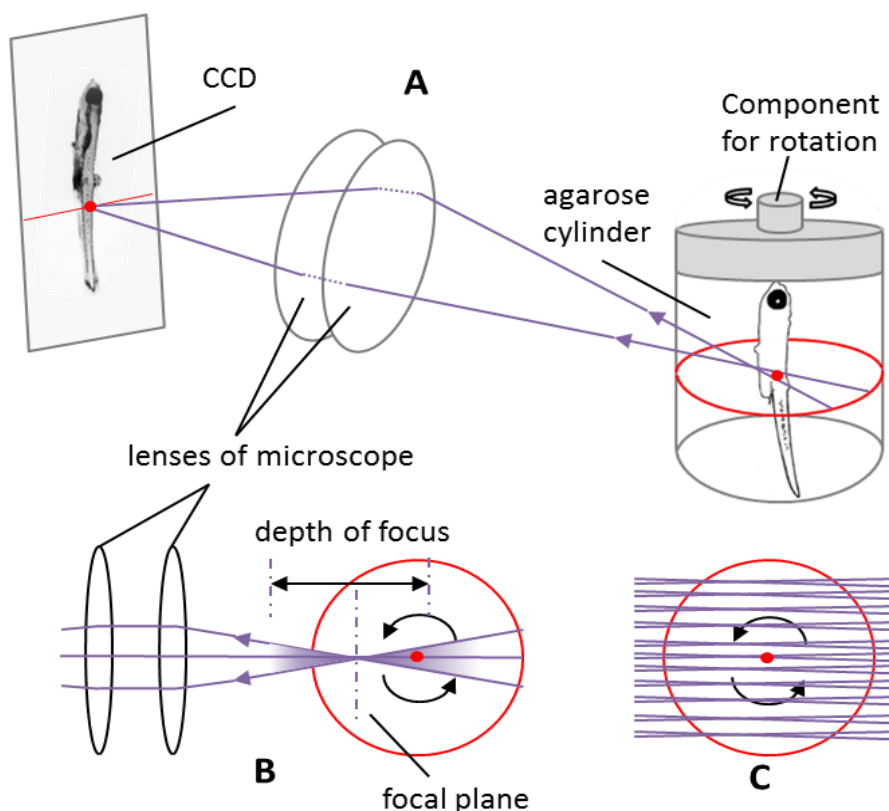
Figure 1.2. The schema of OPT microscopy imaging system. (A) The general imaging schema of a single pixel on a specific slice (red circle). (B) The optical explanation of a pixel of the image on CCD. (C) The formation of adjacent pixels of the image on a specific slice, i.e. the red line in (A).

1.2.2 Experimental OPT imaging setup

The flow diagram of our OPT imaging system is illustrated in Figure 1.3. It consists of a Leica MZ16 FA stereomicroscope with a Plan 0.5 and $135mm$ working distance objective lens (Leica 10446157). Images are acquired by a thermos electronically cooled Retiga Exi CCD camera with a chip size of 1036×1360. The images are saved as 16 bit tiff files but are effectively 12-bit from the Analog to Digital (AD) conversion. A full revolution of 400 images over 360° results in a 1.13Gb tiff-file. The acquisition is realized by a rotation of the specimen driven by a stepper unit, meaning that the stepper accomplishes a step of 0.9°. The OPT imaging system has imaging modules: for the bright-field imaging, specimen is illuminated with a LED and for fluorescence imaging a 100W mercury lamp attached to the microscope housing is used in combination with a filter block. The optical path of the bright-field channel is shown in yellow while that of the fluorescence channel is illustrated in blue and green depending on configuration of the filter block for green fluorescent protein (GFP), e.g. as presented in Figure 1.3. In our imaging system there are three fluorescence channels available including GFP, YFP and TRX. The switch between the bright-field and fluorescence module is controlled by software through an Arduino controller that is integrated in the system.
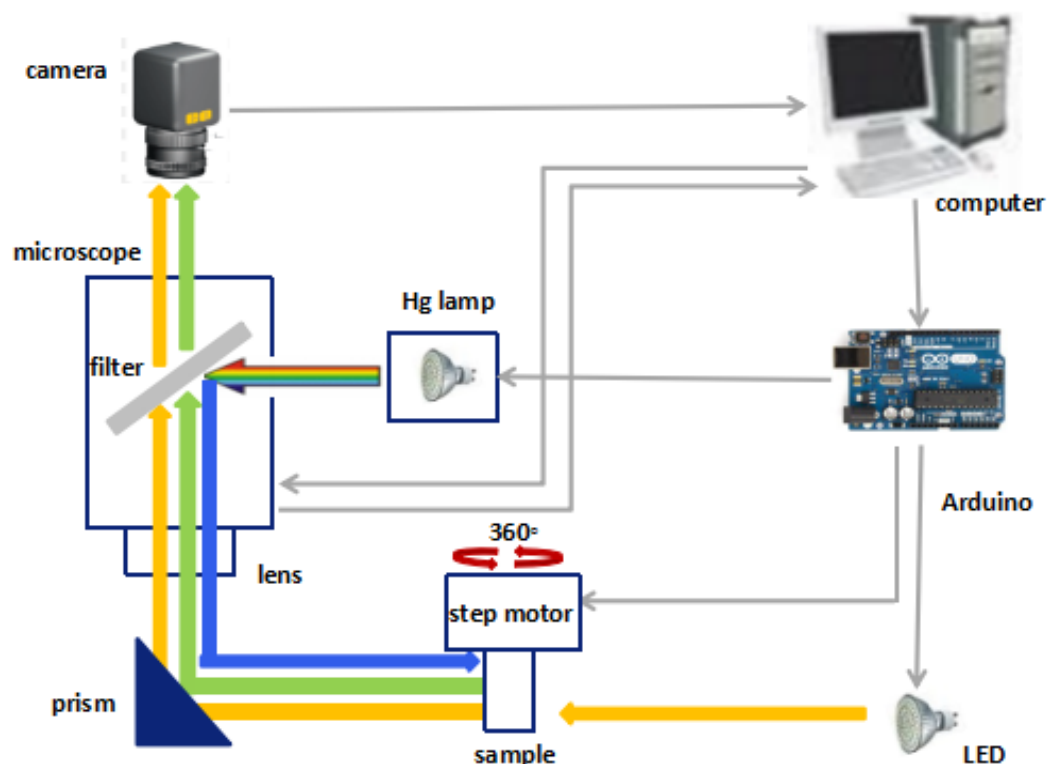
Figure 1.3. The diagram of the homemade OPT imaging system. The optical path of the bright-field channel is illustrated in yellow. The fluorescence optical path from Hg lamp goes through the filter block producing the required excitation wavelength (in blue). The emission wavelength (in green) is produced when the excitation light is excited by fluorescence protein such as GFP.

The details of the OPT imaging system are explained in Figure 1.4. As follows: (A) presents the overall view of the imaging units which consist of control unit (blue line), light path unit (purple line) and sample unit (red line). The control unit includes microcontroller Arduino and manual controller that are described in (B). The manual controller can be used for manual magnification adjustment by rolling zoom wheel. The focus wheel is used for focusing at a specific magnification. The filter change and shutter are combined to determine and change the filter for working in fluorescence mode. Set button stores the current positions of magnification/focus/filter changer/double iris diaphragm for five combinations. (C) and (D) elaborate the sample unit in which the sample is mounted and imaged. The sample, e.g. zebrafish, is embedded in agarose prior to a clearing process and after clearing glued to a cylindrical plastic stub with a metal ending. As shown in (C) the metal ending can be mounted to a magnet with a stepper motor, through a full revolution of the sample over 360∘. Taking into account the refraction of light, the specimen is located in a clearing solution inside a cuvet which has the same refractive index as the sample after clearing. When the specimen is mounted to the stub its position can be adjusted by using the position bar. The sample unit is contained in a black casing to block off ambient light. To further avoid ambient light in fluorescence mode, a plastic shield is used.
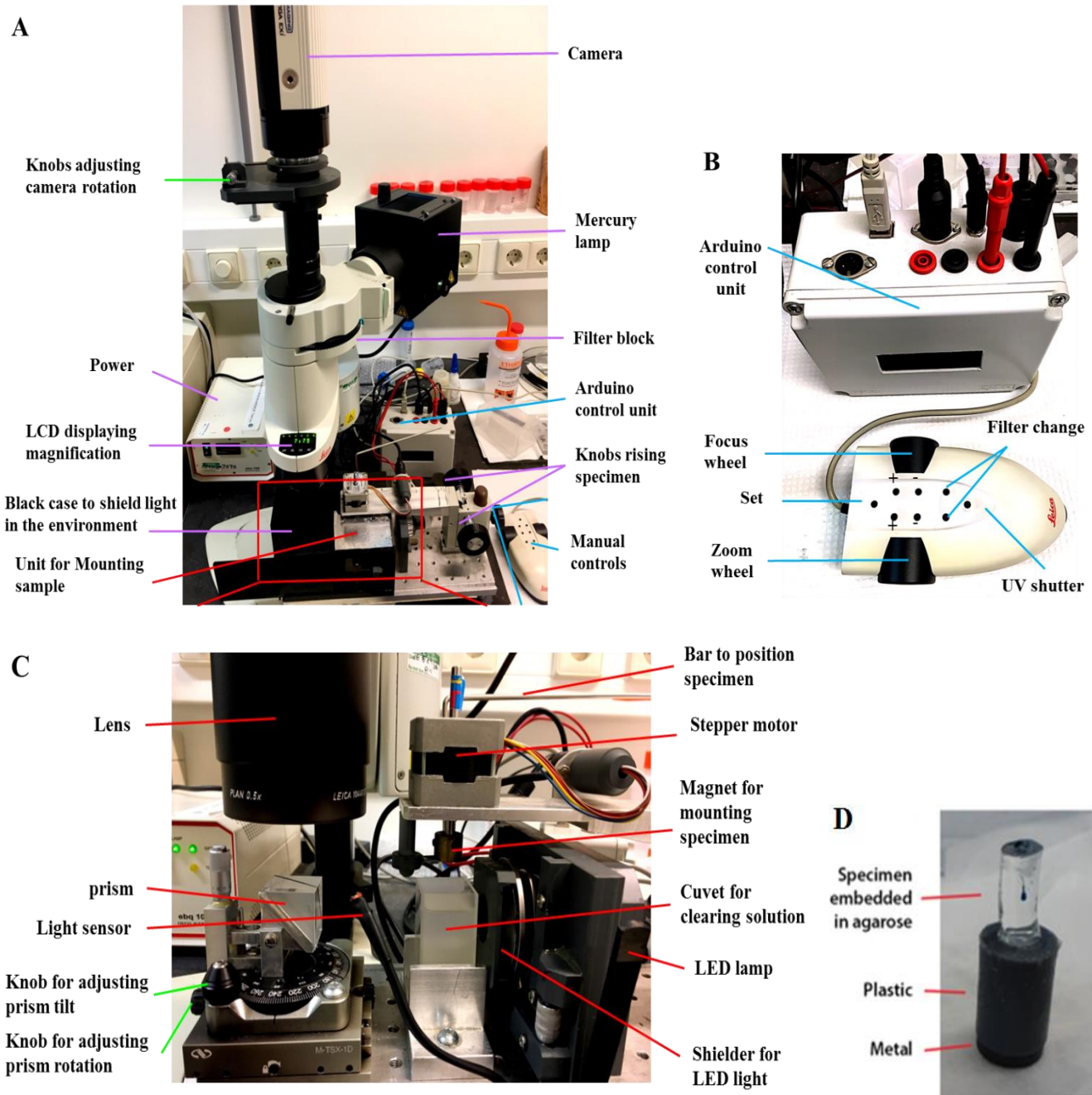
Figure 1.4. Explanation of experimental setup for the OPT imaging system. (A) The whole view of the imaging system excluding computer. (B) Unit of microcontroller chip (Arduino) and manual control. (C) Unit of imaging environment in which the specimen is located. (D) Unit for mounting specimen.

1.2.3 OPT imaging software

The graphical user interface (GUI) of the OPT imaging system contains three parts, i.e. calibration, experiment settings and imaging as depicted in Figure 1.5. The calibration parameters include camera rotation, prism rotation and prism tilt. They are combined to determine the position of the specimen in the Field of View (FoV). The camera rotation

(A)

(B)

(C)
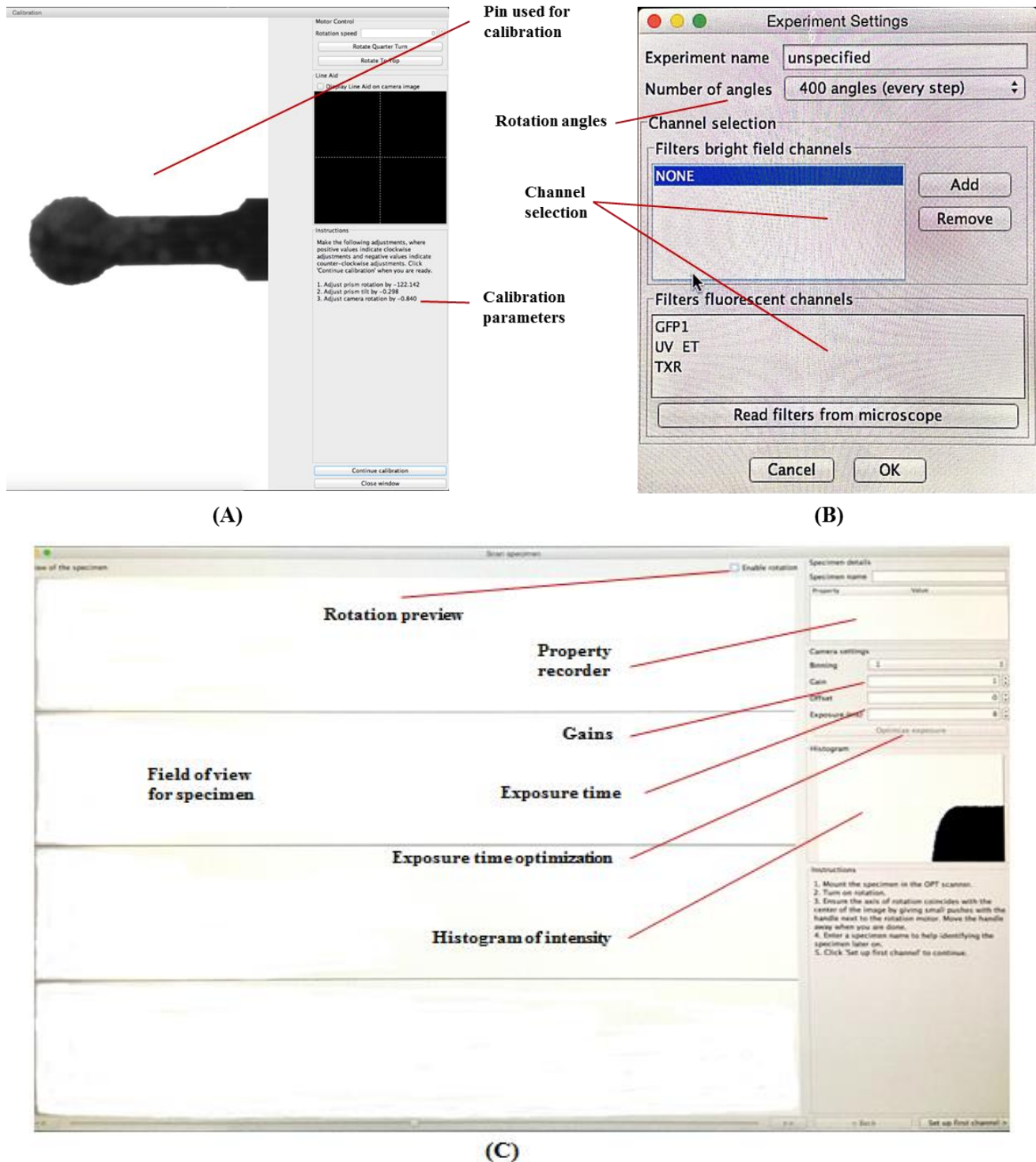
Figure 1.5. OPT imaging software. (A) The calibration user interface. (B) GUI for the experimental settings. (C) GUI for the bright-field and fluorescence imaging.

can be adjusted by the knob near the camera as annotated in Figure 1.4 (A), and the prism rotation and tilt adjustment can be completed by the knob as shown in Figure 1.4 (C). They can be adjusted separately, but the parameters calculated after each adjustment, are dependent on each other. The GUI for experimental setting specifies the number of rotation steps for a full revolution which is determined by the stepper motor. It starts with the selection for imaging mode or channel. In fluorescence mode GFP1, TXR and UV are

supported. Regarding to the imaging GUI, it first works for fluorescence imaging if both modes are selected. Each tomogram includes background correction, which means the specimen should be removed from the FoV by using two knobs as shown in Figure 1.4 (A). Figure 1.5 (C) elaborates the parameters that are required in the specimen imaging process and some useful information with respect to the image quality such as a life histogram informing on the intensity range that is employed with the current illumination settings.

1.2.4 Experimental sample preparation

Sample preparation refers to the protocols used on the specimen to assure the acquisition of useful information with OPT microscopy. It is the most time-consuming process in the OPT imaging workflow, allowing preparation of a few samples per day. Researchers, in general, need to image a lot of specimens to obtain good and statistically valid observations. We, therefore, have to considerably speed up this process.

Therefore, a protocol for efficient sample preparation including counterstaining, embedding of specimen in agarose, and optical clearing is essential to make the OPT suitable. The optimisation of the sample preparation step has been thoroughly studied [20]. This optimisation is out of the scope of our research but it significantly contributes to the quality of our image acquisition and data. Counterstaining (toluidine blue), cylindrical agarose and clearing agents (benzyl alcohol: benzyl benzoate, BABB) are mostly used for sample preparation of samples presented in this thesis.

1.3 Computational approaches of OPT imaging

From proper OPT sample preparation and imaging, we acquire the tomogram data that is, in fact, a collection of axial 2D images. For further processing and visualization we need to have the data on a regular 3D grid. Therefore computational approaches are required. These computational approaches cover 3D reconstruction, 3D segmentation, optimisation of iterative reconstruction based on segmentation performance and fluorescence quantification. For this thesis our main sample will be zebrafish.

1.3.1 3D reconstruction

The process of going from a tomogram to a 3D image defined on a regular rectangular grid is referred to as 3D reconstruction. There is no standardized manner of specimen mounting and nor do we have full control over the experimental environment; therefore a framework is needed including both fast reconstruction and accurate reconstruction to comply for different requirements. The fast reconstruction provides an efficient solution for specimens with dispersed signals, while the accurate reconstruction produces a more precise reconstruction with lesser artefacts at the expense of computation time. Fast reconstruction solves the reconstruction problem but fails to reduce artefacts. In tomography, a known category is the streak artefacts that due to local intensity loss of the

beam as a result of dense materials such are metal components. If the fast reconstruction is satisfactory, the reconstruction framework is completed. If accuracy is required and streak artefact need be removed, iterative reconstruction must be considered. Our workflow for 3D reconstruction is illustrated in Figure 1.6, with the top and bottom of the diagram showing the two different reconstruction methods. Optimisation in iterative reconstruction including initialization and iterative step, and GPU-based implementation will be explored in the research presented in this thesis, and upon evaluation, integrated in our reconstruction framework.
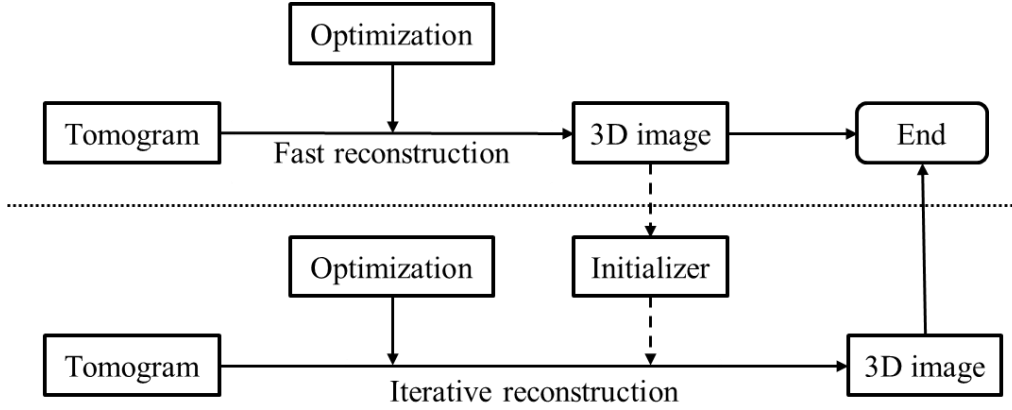
Figure 1.6. The scope of reconstruction and optimisation for OPT 3D imaging.

1) *Fast reconstruction and optimisation*

For fast reconstruction a framework is set up that takes into account, reconstruction, artefact reduction, 3D deconvolution and parallelization implementation.

The Radon transform [21] is widely applied to tomography which is produced from the projection associated with cross-sectional scans of an object, we first explore the applicability of Radon transform to our OPT imaging. The Radon transform represents, de facto, the projection data obtained as the scan from the OPT and the output is the tomogram. The inverse Radon Transform can be used to reconstruct to the initial object. The process of reconstruction with inverse Radon Transform is called back projection. The most frequently applied reconstruction algorithm for back projection is known as Filtered Back Projection (FBP) by Nygren and Anders [13]. In our fast reconstruction framework FBP will employed because of its wide applicability and efficiency.

In a previous study [22], it has been acknowledged that artefacts occurring in 3D images in OPT imaging systems are mainly introduced through two reasons. One is from the imaging setup such as non-uniform illumination, signal decay and CCD imperfection. Another sort of artefacts comes from the computational process including the centre of rotation (CoR) misalignment and limitations in the number of tomographic images that are available for the reconstruction algorithm. The artefacts from the second group, however, are more noticeable and salient; these will be studied in the following sections.

Limited by the diffraction of light in the optical imaging system, the images are unavoidably blurred in imaging process. Unlike a confocal imaging system, in tomographic images this imperfection cannot be deblurred directly from a systematic point spread function (PSF). An OPT imaging system has variant PSFs for points that are located within different distances to focal plane. How we can model the variant PSFs and take them into account in 3D image for deblur is one of our research topics of interest for optimisation in fast reconstruction.

The core for fast reconstruction is to apply parallel and distributed computing. The parallelization can efficiently decrease the computing time of the reconstruction process. We accomplish this by distributing all the slices to different processors of a computer cluster, ensuring that the 3D data is processed in a parallel manner.

2) Iterative reconstruction and optimisation

Iterative reconstruction refers to iterative algorithms used to reconstruct 2D or 3D images from tomographic imaging techniques such as CT and OPT. Iterative reconstruction was developed for CT imaging in order to improve the noise profiles and suppress streak artefacts that commonly show up with FBP. These algorithms are also considered superior when there is a lack of uniform angular projections or when projections are sparse. There is a large variety of iterative reconstruction algorithms, but they all have in common that it starts with an assumed initial image, computes projections from the image via a project function and updates the image according to the difference between calculated and actual projections. According to the updating strategy for the image, iterative algorithms can be categorized into four different approaches, i.e. algebraic reconstruction technique (ART) [23], iterative sparse asymptotic minimum variance (SAMV) [24], statistical reconstruction [25] and learned iterative reconstruction [26], [27]. Among all categories, statistical reconstruction and learned iterative reconstruction show relatively better performance with respect to a combination of effectiveness and robustness.

In general, iterative reconstruction can lead to a more accurate reconstruction compared to FBP. However, a large number of iterations may be required to generate an acceptable reconstruction and each of the iteration may take about the same amount of time as one FBP reconstruction does. Thus, to some extent the effectiveness of iterative reconstruction is achieved at the expense of huge computation time. One approach to reduce the number of iterations is to organize the projection data into a series of ordered subsets of evenly spaced projections and update the current estimate of the object after each subset rather than after the complete set of projections. The most commonly used algorithm that employs the subset strategy is referred to as ordered subset expectation maximization (OSEM) [28]. It improves the efficiency of iterative reconstruction with respect to computation time.

Iterative reconstruction methods have superior performance which shows by resistance to noise and streak artefacts in CT reconstruction. These algorithms, therefore are

supposed to have promising results on our OPT reconstruction where the streak artefacts are limiting the results. There are two sources for the streak artefacts with FBP reconstruction in our OPT imaging system. One is the relative lack of projections compared to perfect reconstruction without streak artefacts. This means that 400 samples over a full revolution are still not enough for FBP reconstruction in OPT because it has stronger light attenuation, comparing to CT. Another source for streak artefacts with FBP normally exists in emission OPT when the fluorescence signal is relatively small and highly concentrated, which is similar to the artefacts produced by metal components in the specimen with FBP in CT. With the aim to eliminate the streak artefacts in OPT, we will implement one of the iterative reconstruction algorithms, i.e. statistical reconstruction, for our data and optimize the results based on parameters required. Notably, there are two customary parameters to optimize. One is the initial image, in this thesis defined as initialization. It can be either none (meaning zero) or the result of fast reconstruction in our workflow (cf. dashes in Figure 1.6). The other one is the number of iteration steps for deciding the endpoint of iterative reconstruction. The impact of different parameters on reconstruction will be studied in Chapter 4.

In terms of evaluation, reconstruction, as a typical inverse problem, is characterized by the lack of benchmarks for a real imaging data. Researchers often measure the reconstruction performance by qualitatively comparing the results from different reconstruction approaches. Specifically, the qualitative measurement could either be less noises and artefacts or better image quality in terms of sharpness. With a lack of quantitative measurements for reconstruction performance, we take a first step to explore the possibility of transferring the problem of reconstruction evaluation to the segmentation evaluation. This evaluation is particularly applied to optimize the parameters for iterative reconstruction. We assume that in iterative reconstruction, reconstructions of the type of same data (e.g. zebrafish) from different parameters provide different inputs for segmentation, thus resulting in different segmentation performance for the resulting model, e.g. zebrafish. The reason why segmentation evaluation is employed as a reference for the optimisation of iterative reconstruction is that we need a good reconstruction of the specimen for a segmentation model and the segmentation performance can give us an intuitive and quantitative feedback about how good the reconstruction is. This framework works under the assumption that the same segmentation model is used. In terms of parameters in iterative reconstruction, there are several ones required but number of iterations and initialization are seen as the most customary ones. Theoretically, by applying this alternative reconstruction measurement, i.e. the segmentation performance, to the parameter optimisation of the iterative reconstruction algorithm, a more accurate reconstruction will be achieved.

1.3.2 3D segmentation of OPT reconstructions with applications to zebrafish

The OPT is particularly useful for imaging of specimen/objects in the *mm* range; so, zebrafish is a specimen typically suitable for this type of imaging. The zebrafish model

system is very popular in bio-medical research. Zebrafish can be easily embedded in large scale projects as sufficient amounts of samples can be made available. In experimental setup we are interested in the phenotype and the gene expression of the phenotype in a zebrafish sample. Measurements in zebrafish OPT reconstructions require a clean image; noise and debris should be avoided but in the practice of the imaging this is difficult to achieve. Therefore, a segmentation of the specimen is required. An OPT image can contain multiple channels. The bright-field channel typically provides the possibility for generating a whole-mount mask of the zebrafish, or any specimen in general. Noises, debris in the image and transparency of the zebrafish complicate segmentation. These facts limit the success of conventional segmentation algorithms, i.e. adaptive thresholding, mean shift to name a few. In the application of segmentation methods on zebrafish, we observe inaccurate body boundaries and the fainter transparent parts make a distinction between the foreground and the background difficult.

In order to obtain reasonably good segmentation results of zebrafish, a more advanced and intelligent segmentation approach is required. We therefore employ machine learning strategies and, for our studies, a supervised segmentation framework is presented as illustrated in Figure 1.7. The application of supervised machine learning requires a training process from examples and prior to the training procedure images of different specimens are reconstructed and labelled. Because of the information redundancy among adjacent slices in a 3D image, the training data is manually labelled in equidistant intervals. We will investigate the use of a convolutional neural network (CNN) for the training process.
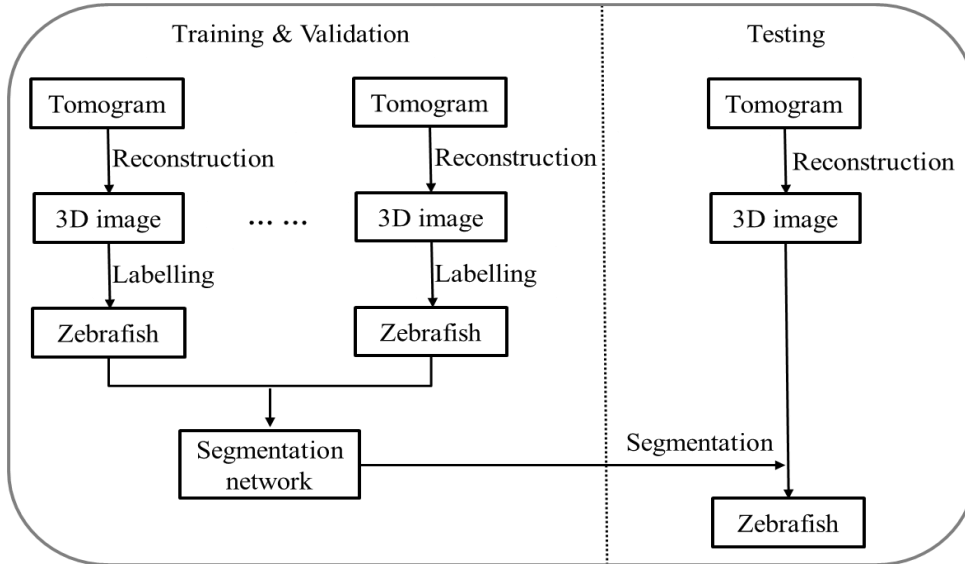
Figure 1.7. The diagram of training a zebrafish model for segmentation. The images used for training segmentation network and testing are from bright-field channel.

The labelling of all slices in the 3D image for training segmentation network is optional; if necessary, it can be obtained by either using interpolation technique or predicting from the segmentation network, such as U-net convolutional networks [29]. We investigate how

such network for 3D image segmentation could be trained either on parts of the 3D image or on the whole 3D image. A typical network that has been successfully employed for segmentation in bio-medical research is the U-net convolutional neural network [29]. The differences between training on manually labelled slices or whole 3D images are the training time and accuracy; this is extensively studied in Chapter 5. Previous results from U-net convolutional neural network have shown its usefulness in bio-medical image segmentation [30]. For training the 3D segmentation network, both 2D U-net and 3D U-net convolutional networks will be employed. The differences in performance on the data between the two networks will be explored. With the trained zebrafish segmentation network, any 3D bright-field OPT image can be semantically segmented as zebrafish or none-zebrafish. Masking the zebrafish in 3D with the corresponding 3D fluorescence channel image accomplishes general fluorescence quantification, providing valuable information for bio-medical research.

1.3.3 Quantification of volumetric fluorescence in zebrafish

We consider zebrafish as a prototypical example for OPT imaging in which we both use the bright-field as well as the fluorescent channels. The rapid development of fluorescent microscope imaging technologies in the past years, enables high-throughput 2D fluorescent imaging platforms now in widely use on both gene expression and proteome scale [31]. High data volumes for protein and/or gene expression benefit the statistical analysis. This is typically the case for such zebrafish. The 2D quantification of the fluorescence provides a relatively rough measurement for analysis. It fails to reconstruct the real, particularly spatial, distribution of fluorescence, thereby losing much useful information.

Therefore, 3D fluorescent imaging techniques, such as CLSM, OPT and MicroCT, play a significant role in obtaining insights in 3D quantification of fluorescence so that real and relatively accurate protein and/or gene expression can be computed. Due to the large amount of data in 3D imaging, high-throughput scale is currently limited by the computing and memory power. To some extent, throughput in 3D imaging and quantification is possible under the condition of efficient sample preparation and reconstruction.

In general, a correct and accurate 3D fluorescent quantification largely depends on sample preparation, reconstruction performance and image processing i.e. segmentation. The protocol of sample preparation varies depending on different imaging techniques but the common demand for a good sample preparation is supposed to suppress noise and highlight fluorescence as much as possible. In this case, prior knowledge is typically required to distinguish between the noise and the fluorescence signal we are interested in. The segmentation process supports in understanding the information about what is useful or significant to the quantification. In order to achieve reliable 3D quantification of fluorescence in a smaller data volume, segmentation can be manually accomplished using prior knowledge. This approach is however, not feasible for larger scale of data, for

instance some-throughput imaging of zebrafish in our OPT system. This is a motivation for us to investigate the utility of a trained 3D zebrafish reference structure for the 3D quantification of fluorescence in zebrafish. In this thesis we present the framework of 3D quantification of fluorescence in zebrafish, but it can be easily transferred to other applications such as 3D quantification of fluorescence in zebrafish/mice brain, liver, kidney, etc.

1.4 Research questions and perspectives

RQ1: To what extent is it possible to increase the processing speed of OPT imaging and reconstruction in an integrated manner?

The 3D reconstruction as a post-imaging process, is therefore separated from OPT imaging system. This separation can be physical, i.e. computations on different computers. The time of one reconstruction varies from minutes to hours depending on the reconstruction algorithm and computation resources. This means that in a worse case it may take multiple hours to generate one 3D OPT image. With the further development of data science in bio-medical research the availability of data becomes increasingly important. To provide more bio-medical data, i.e. 3D OPT images in our case, it is of great importance to decrease the imaging time and increase the efficiency of 3D imaging process. Therefore, we investigate integrating the imaging and reconstruction as a whole and implementing the reconstruction in a parallel fashion. With respect to the reconstruction algorithm, the fast and efficient reconstruction, i.e. FBP is first taken into account. What interests us is how much improvement can be achieved regarding the imaging time and efficiency of computation.

RQ2: To what extent is it possible to reduce the artefacts of 3D image introduced during reconstruction process by misalignment of CoR?

According to Singh *et al.* [19], there are several types of artefacts in OPT reconstruction. One of them is the edge blur artefacts introduced by CoR misalignment. It means that the rotation centre of the imaging system is shifted off the centre of FoV. This shift inevitably exists unless a very accurate mechanical calibration is included. This mechanical calibration process is typically time consuming and requires a lot of operator interaction. Instead of eliminating the shift in the pre-imaging calibration process, we are interested in correcting it in the post-imaging process before reconstruction.

The aforementioned pre-imaging shift will introduce a problem that the CoR is not accordance with the image centre. This means that the reconstructed slices will consequently be corrupted by the edge blur artefacts after application of the reconstruction algorithm. The motivation for this research is to detect and correct the CoR shift in tomogram and further eliminate the corresponding artefacts in the reconstruction process. To this end, a fast and accurate CoR correction algorithm is

needed. How much improvement will be achieved should be qualitatively and quantitatively analysed and explored.

RQ3: Can the PSF of the OPT imaging system can be modelled and applied for deblurring of an OPT reconstruction?

In general, an imaging system is described by the ability of giving a response to a point light source or object, commonly referred to as the point spread function (PSF). A more general term for the PSF is a system's impulse response, being the impulse response of a focused optical system. For OPT imaging and reconstruction, the DoF is expected to be large enough to contain the specimen to be visualized as much as possible. According to the previous studies [32], [33], however, large DoF subsequently results in low in-focus image quality. The trade-off between DoF and image quality should be considered when selecting a lens for the OPT imaging system. A lens with small NA will produce large DoF, allowing imaging of larger specimens but it will result in a relatively low-quality image. Contrarily, a lens with large NA yields relatively high-quality images but cannot image the whole specimen with respect to its size.

One typical way to improve the image quality of 2D imaging system to the best possible resolution, is to apply deconvolution to the images, using a constant theoretical or experimental PSF as the kernel for deconvolution. However, this approach is not strictly suitable for OPT images. Because first, the tomogram normally integrates the information of a specimen at different depths within a wide field, not a fixed depth on the focal plane. Second, the imaging PSF within the field varies at different depths along the optical axis. This means that different PSF is produced when locating the point source at different depths. While the conventional 2D PSF and deconvolution are not feasible for OPT imaging system, we are interested in how a variable PSF could be modelled in OPT imaging system and how much image quality improvement can be achieved when using it in deconvolution of the 3D image.

RQ4: Can the iterative reconstruction eliminate the streak artefacts produced in the fast reconstruction?

With the fast FBP reconstruction, a 3D image can be obtained. For specimens with sparse and large-area signals, it produces satisfactory reconstruction when CoR misalignment is solved. However, when imaging specimens with a dense and concentrated signal such as concentrated fluorescent GFP signal, the FBP algorithm will produce streak artefacts because of the sample limitation. There are several methods such as linear interpolation (LI) and the state-of-the-art normalized metal artefacts reduction (NMAR) to decrease the streak artefacts in CT in an FBP framework [30], which may be useful for OPT 3D reconstruction. But these approaches seem to reduce the artefacts in the reconstruction but do not eliminate the artefacts. In order to eliminate the streak artefacts, iterative reconstruction has proven to be promising in CT reconstruction [34]–[36]. Inspired by the

results in CT imaging, we are interested in exploring if iterative reconstruction can help to eliminate the streak artefacts in OPT reconstruction and how it works. In our work both reconstruction frameworks are available, depending on application of what a certain reconstruction algorithm is chosen.

RQ5: How and to what extent the initialization and the number of iteration steps influence the results in iterative reconstruction?

Iterative reconstruction can eliminate the streak artefacts, thus it could be used as a prospective method to produce an accurate OPT reconstruction at the expense of time compared to fast reconstruction such as FBP. Sometimes these accurate reconstructions are essential. Considering the non-deterministic process of iterative reconstruction, the results are influenced by the parameters of algorithm. In this research we investigate the most customary parameters, i.e. iteration steps and initialization. To explore the effect of iteration steps on reconstruction, we reconstruct the tomogram with different iteration steps. With respect to initialization, we compare the results produced with a setup of no initialization to those with an initial reconstruction that are obtained by fast reconstruction as described in Chapter 2.

With the different reconstructions to compare, an evaluation criterion is required. In bio-medical research it is not always possible to construct a benchmark for both 2D and 3D imaging because of different specimens and experimental setups. For instance, to assess the various reconstructions in our study, there is no theoretically ideal reconstruction that can typically be used as benchmark for evaluation. Therefore, instead of assessing the reconstructions directly, we investigate a framework to assess the segmentation results of the reconstruction indirectly. The benchmarks of the segmentation are easier to obtain by labelling the data. The evaluation of reconstructions of different experimental setups is transferred to the evaluation of the corresponding segmentation results.

RQ6: Is it possible to "learn" a 3D reference structure of zebrafish for 3D fluorescence quantification in zebrafish?

Zebrafish are valuable for studies of a multitude of diseases including cancer, heart disease, obesity, muscular dystrophy and narcolepsy. They are easy to maintain and cost-effective. One key feature is that, following fertilization, zebrafish embryos are transparent and their rapid embryonic development can be observed. Another important reason is that zebrafish as a vertebrate is similar to human, making it a suitable model for many human diseases [37].

Because of the imperfections in sample preparation and experimental imaging environment, a specific fluorescence may exist around a specimen in the background. It is therefore required to segment the specimen from the background. For our study in zebrafish, segmenting and recognizing zebrafish volume from background are of great

importance for fluorescence quantification. For the volume, the fluorescent signal must be counted and quantified just there where zebrafish sample is, otherwise it should not be considered. The segmentation of zebrafish also plays an important role in reference structure detection for high-throughput quantification where relative measurement is considered. This will be elaborated in Chapter 5. To train and learn a theoretically robust segmentation model for 3D reference structure detection, a large number of zebrafish are acquired at different stages and in different experimental environments. By using the trained segmentation model as a classifier, the 3D image reconstructed from a specimen can be identified as reference structure or not. The 3D fluorescent signal, e.g. tumour in Chapter 5, can be quantified and normalized referring to the reference structure that is identified. Compared to 2D, such accurate 3D fluorescence quantification helps to improve the research results in for instance drug discovery.

RQ7: How much 3D information can be achieved and identified from bright-field images of zebrafish and to what extent can the identification of parts be automated?

In the OPT imaging system, bright-field image contains some extent of structures or context information of zebrafish. It also represents the minimum of details of zebrafish structure without any staining techniques. Some structures, e.g. eyes play an important role in accurate positioning analysis of protein and gene expression patterns. In order to explore how much structural detail within zebrafish can be identified for such analysis, we manually label the 3D image of zebrafish in two different very distinct developmental stages, i.e. 5 *dpf* and 25 *dpf*. The results are visualized and the two developmental stages are qualitatively compared to provide an intuitive clue for phenotype analysis of these structures. Furthermore, in view of the trends in order to study the possibility of trending high-throughput analysis in 3D, we are curious about how automated 3D structure detection algorithm can facilitate and accelerate the manual labelling process, as well as how much accuracy can be achieved by using the automated detection. If the accuracy is not satisfactory, how can we propose the possible solutions to improve it for further advanced analysis?

1.5 Thesis structure

This thesis is structured along the research questions presented in the previous paragraph.

In **Chapter 1** "**Introduction**" a brief introduction on the 3D imaging, reconstruction framework in OPT microscopy is given. Besides that, our scope for reconstruction optimisation, 3D segmentation and its application value are elaborated. To clarify these research topics, seven research questions are proposed.

Chapter 2 "**Fast Post-processing Pipeline for Optical Projection Tomography**" presents a fast reconstruction framework to improve the effectiveness and efficiency of OPT 3D reconstruction. It is implemented in a parallel manner and the experiments show

that the average run time for each tomogram is less than 5 minutes in our current setup. In the framework a novel CoR correction method based on interest point detection in sinogram, is proposed by taking the principle of OPT imaging into account. To quantify and compare the reconstructed results of different CoR correction approaches the coefficient of variation (CV) instead of the variance is employed.

Chapter 3 "**Deblurring Images from 3D Optical Projection Tomography Using Point Spread Function Modelling**" focuses on the deblurring of the reconstructed 3D image. When imaging large specimens with OPT imaging system, a large depth of field is required. This normally results in blur, i.e. compromises the image quality. Yet, it is important to obtain the best possible quality 3D image from the OPT, thus deblurring the image is vital. To this end we first model the PSF along optical axis at different depths. Meanwhile the magnification is taken into account in the PSF modelling. Subsequently, deconvolution in the coronal plane based on the modelled PSF is implemented to accomplish deblurring of the OPT image. Experiments with the proposed approach based on 25 3D images including 4 categories of specimens, indicate the effectiveness of quality improvement assessed by image blur measures in both spatial and frequency domain.

Chapter 4 "**Segmentation-driven Optimisation for Iterative Reconstruction in Optical Projection Tomography: An Exploration**" introduces GPU based iterative reconstruction aiming for the best possible, reconstructed from an OPT tomogram. Here possible streak artefacts produced by FBP reconstruction should be eliminated. The reconstruction performance with different initializations and iteration steps is evaluated indirectly based on the segmentation results of the reconstruction, instead of the reconstruction itself. Aiming at producing good segmentation results, a deep learning model is employed. The iteration step and initialization of the iterative reconstruction are considered optimal when evaluation measurement reaches a maximum. The model is trained and tested on three 25 *dpf* 3D zebrafish image from bright-field tomograms.

Chapter 5 "**Automated Detection of Reference Structures for Fluorescent Signals in Zebrafish with a Case Study in Tumour Quantification**" aims at automatically detecting the reference structure to relatively quantify the 3D fluorescence within a zebrafish. We will build and train a segmentation model to automatically detect the zebrafish *Body* and *Eye* reference structure in two different data spaces (i.e. 2D slice and 3D volume, c.f. Chapter 5) and optimize the segmentation model individually. Subsequently, the segmentation performances are compared and evaluated. The approach with the best performance will be considered for the automated detection of reference structure for tumour quantification as a case study for drug research.

Chapter 6 "**Exploration of 3D Structure Annotation and Visualization of Zebrafish Reconstructions from Optical Projection Tomography Imaging**" explores the

possibility of volume region annotation in OPT imaging. The aim of this chapter is to, first give an idea of how much volume region information can be acquired and reconstructed using our OPT imaging system on whole-mount specimen. Second, it explores the possibility of automated annotation of volume region within zebrafish, which are potentially important for the high-throughput research at level of organ systems. In the first case, up to 9 parts or volume regions of a 25 *dpf* zebrafish can be segmented and visualized using OPT. Nevertheless, automated segmentation of such volume regions has proven to be challenging and is still limited by data size and segmentation algorithm.

Chapter 7 "**Conclusions & Discussion**" summarizes our contribution for 3D OPT imaging, reconstruction, structure detection and visualization framework and states some drawbacks of the framework. Future improvements of our framework are discussed guaranteeing the integrity of the framework.

Chapter 2

Fast Post-processing Pipeline for Optical Projection Tomography

This chapter is based on the following publications:

Tang X., Hoff V.M., Hoogenboom J., Guo Y., Cai.F., Lamers G. & Verbeek F. (2016), Fluorescence and bright-field 3D image fusion based on sinogram unification for optical projection tomography. In: Tian T., Jiang Q., Liu Y., Burrage K., Son J., Wang Y., Hu X., Morishita S., Zhu Q., Wang G. (Eds.) *Proceedings 2016 IEEE International Conference on Bioinformatics and Biomedicine Dec 15-18, 2016 Shenzhen, China.* Danvers: *Institute of Electrical and Electronics Engineers* (*IEEE*). 403-410.

Tang X., Zwaan D.M. van der, Zammit A., Rietveld K.F.D. & Verbeek F.J. (2017), Fast Post-Processing Pipeline for Optical Projection Tomography, *IEEE Transactions on NanoBioscience* 16(5): 367-374.

Chapter summary

In order to improve the effectiveness and efficiency of 3D reconstruction for optical projection tomography, we present a fast post-processing pipeline. This pipeline includes image cropping, background subtraction, centre of rotation correction and 3D reconstruction. For OPT imaging, with respect to the centre of rotation correction, a novel algorithm based on interest point detection in the sinogram is proposed. Instead of locating the centre of rotation on single sinogram, we intermittently and evenly select sinograms in the detected full range of a sample to make the located centre of rotation more robust. The post-processing pipeline presented is implemented on a parallel manner and experiments demonstrate that the average runtime for images of size 1036×1360×400 pixel can be less than 1 minute on a computer cluster of which 5 compute nodes are used.

2.1 Introduction

In this section we will state our research question and introduce our perspective on optical projection tomography (OPT) imaging and reconstruction framework. It includes a brief introduction of our contribution to this research as well as related work in this field of research.

2.1.1 Research problem

The aim of an OPT imaging system is to obtain a 3D volume image, so that this volume image can be used for analysis and visualization. This is accomplished by a reconstruction algorithm that is applied on the sinograms derived from the OPT image. With an OPT imaging system a so called tomogram is acquired. The tomogram is a collection of images of a specimen taken at regular angular intervals. For OPT this typically comprised a stepwise acquisition of the images over a full revolution of the sample. The tomogram is transformed to a sinogram in which all projections are represented.

The reconstruction process could, however, introduce various artefacts depending on different imaging setups. This means that for each individual OPT imaging system, exploration and elimination of reconstruction artefacts are necessary. In this chapter, we will focus on the artefacts resulting from the misalignment of centre of rotation (CoR).

Another important issue for OPT imaging system is the speed of reconstruction process. In order to be able to apply OPT in a high-throughput setting as well as to allow quick reconstruction of the imaging, research on fast and efficient reconstruction is important. These two issues represent the general motivation for the work in this chapter.

With the two research questions, we propose a fast post-processing pipeline that is integrated into reconstruction software. In this pipeline, cropping and background subtraction are the first two steps for image pre-processing, followed by a fast and efficient CoR correction and a 3D reconstruction algorithm. This significantly contributes to the innovation in our OPT applications. With the application of cropping and CoR correction, the sample can be placed at any position of the field of view (FoV), decreasing the time for post-processing of tomogram and avoiding the calibration process prior to tomogram acquisition. Originally, a calibration helps to align the CoR to the centre of FoV, which normally takes several minutes. In this pipeline, we implement a parallel computation of both CoR correction and 3D reconstruction to further accelerate the post-processing of OPT tomogram.

2.1.2 Related work

In the recent years, cancer progression [5]–[7], drug discovery [8] and development studies for organ system such as skeleton [9], [10], teeth [38] and blood vessels [12] have benefited from the further development of OPT microscopy. Kumer *et al.* [5] applied OPT to adult

zebrafish to study the synchronous development of cancer and vasculature in adult zebrafish. McGinty *et al*. [8] proposed a fluorescence lifetime optical projection tomography in 2011 for biological research and drug discovery, the time for image acquisition and post-processing including 3D reconstruction were both reported to be ~20 minutes. Later, in 2012, Fieramonti *et al*. [9] extended OPT to optically diffusive samples for studying skeletal and nervous structures in zebrafish, improving the acquisition time to something like ~3 minutes but without considering the artefacts produced when the CoR is inconsistent with the centre of the tomogram. Agarwal *et al*. [7] presented a diagnosis method of early cancer by reconstructing 3D cellular image with OPT. The high resolution of single cells was achieved by using a large NA and scanning the objective focal plan contributed to the extension of DOF, which consequently increased the light dose and acquisition time. More recent, in 2015, Correia *et al*. [12] introduced accelerated OPT by decreasing the number of rotations at tomogram acquisition, aiming to improve the efficiency of OPT system and decrease the light dose the sample is exposed to. In similar fashion, aiming at improving the efficiency of OPT imaging, we present a fast OPT post-processing pipeline which contains pre-processing, CoR correction and 3D reconstruction taking ~1 minute with tomogram size of 1036×1360×400 pixels.

Before applying the inverse radon transform to the sinogram for reconstruction, by definition the position of the CoR should be in the middle of the sinogram, achieved by CoR correction. This was first studied in 1990 [39] in computational CT. Previous studies showed that shifted CoR could introduce severe artefacts or even incorrect results [40]. Furthermore, correcting CoR based on images can bypass the calibration prior to tomogram acquisition, improving the efficiency of the imaging system. In terms of methodologies for CoR correction, there are two mainstream approaches. The first approach is based on signal match for pairs of projection data (180° opposed to each other) [39], [41]–[43]. This is widely used in CT because the intensities from two opposite projected angles are theoretically equivalent in CT imaging. Unfortunately, this method may not be directly suitable for OPT images, as opposite projected data may vary at different sample angles. The differences are caused by the fact that the lens introduces a DoF and only images the front half of the sample [44]. Moreover, feasibility is hampered as the sinogram is often disturbed by fixation artefacts and/ or random noise; both frequently occur in OPT imaging.

The second approach both for CT and OPT is based on iterative reconstruction of the sinogram [22], [45], [46]. The vertical axis that is producing the smallest variance in the reconstructed image is chosen as the CoR, cf. [46]. This approach is however time-consuming and therefore less used in CT. Furthermore, both approaches chose only one sinogram for CoR correction, whereas the CoR fluctuation produced by different sinograms was not taken into account, to some extent resulting in unconvincing CoR.

2.2 Materials and methods

An OPT system that is used for 3D imaging in biomedical research, e.g. embryo or skeleton development, requires the sample, i.e. a zebrafish larvae, first to be prepared for imaging. A clearing of the sample is accomplished with the BABB protocol, cf. § 1.2.4. The data flow of a sample goes from preparation, to OPT image acquisition using our dedicated imaging software, cf. § 1.2.3, to the production of the OPT tomogram. This tomogram is then reconstructed to 3D image by using the OPT reconstruction software which will be further elaborated in this section. The OPT reconstruction software integrates the whole reconstruction pipeline. For the CoR correction and 3D reconstruction tasks, it provides the interface to submit the tasks to our compute cluster, i.e. the Leiden Life Science Cluster (LLSC).

2.2.1 OPT imaging

Our OPT imaging system supports both bright-field and fluorescence illumination. The acquisition time for a bright-field tomogram is less than 3 minutes, and for a fluorescence tomogram it varies depending on the strength of the fluorescence and exposure time; but normally it is less than 10 minutes for a sample in a full revolution, cf. § 1.2.3. Optimisation of sample preparation protocol and image acquisition were implemented as described in chapter 1, cf. § 1.2.4. The tomogram of a single channel from the OPT is a 16-bit image of size 1036×1360×400 pixels, with a file size of 1.05GB. Image of 1036×1360 is acquired over 400 rotation angles in [0°, 360°). For each tomogram, 10 background images of the same size are acquired for the post-processing. The acquisition of the tomogram is separated from the computationally more demanding post-processing. This is accomplished on a cluster computer and communication to the cluster application is realized via a web-service that is available on the acquisition computer.

Application of the post-processing pipeline, implies that the acquired OPT tomograms will be first cropped to the region of interest (RoI). It is followed by a background subtraction of the median of the background images for each channel. Subsequently, CoR correction and 3D reconstruction are applied. The user interface of the post-processing software is shown in Figure 2.1 with an example of a zebrafish in bright-field. Cropping and background subtraction are implemented locally on the left side, and on the right we can upload the data to our computer cluster (LLSC) for CoR correction and reconstruction. Once the task is finished, the resulting 3D image will be automatically stored on the LLSC file server and a file link for downloading is returned. At the same time the reconstructed 3D image is downloaded to the local computer, cf. right panel in the user interface as shown Figure 2.1. Users can also visualize the x-projection and y-projection using the Details button. The file link for users to download the 3D image is provided under the Details button on the right of user interface.

2.2.2 OPT reconstruction software

Figure 2.1. The user interface of the post-processing software; it includes cropping, background subtraction, CoR correction and 3D reconstruction. Once the tomogram is opened, cropping and background subtraction can be done with buttons on the left. With a Start Reconstruction button, the data are automatically uploaded to a dedicated cluster computer. The CoR correction and reconstruction are distributed on the cluster. The reconstructed results and maximum intensity projections are sent to the local computer after completion (right panel).

2.2.3 Cluster computing: the LLSC

The LLSC is a computer cluster for bioinformatics applications. As displayed in Figure 2.2, it consists of three user nodes, 20+ compute nodes and a file server. Each compute node consists of multiple two Intel Xeon dual- or quad-core processors with 16GB RAM, forming a cluster of 108 processors in total. The separate file server has 36TB storage. The nodes and file server are connected using Gigabit Ethernet. The LLSC currently uses the TORQUE job scheduler [47] to allocate the computational tasks, but a future upgrade to the SLURM workload Manager [48] is planned. To ensure consistent performance measurements we will be using the nodes that contain two Intel Xeon 5150 dual-core processors and two Intel Xeon e5430 dual-core processors, given that most of the currently operational nodes in the cluster are of these types.

Figure 2.2. The LLSC cluster with three user nodes, 20 computing nodes (108 processors) and a file server.

2.3 Implementation

In this section we present our specific contributions to the post-processing pipeline, i.e. the CoR correction algorithm and Reconstruction on the LLSC system.

2.3.1 CoR correction

CoR correction involves CoR localization for each of the channels and the CoR alignment of these multiple channels. Considering the artefacts from CoR shift depicted in [22], [49] and the computationally expensive problem of the traditional CoR localization method by using iterative reconstruction [46], a novel CoR localization approach is presented. The CoR localization for each channel is defined as searching for most frequently occurring value from the obtained CoRs of multiple sinograms, which are localized based on interest point detection and CoR optimisation function.

1) Sinogram selection

To make the CoR localization in each channel convincing, multiple sinograms are selected. We select these multiple sinograms by investigating the slice range *NS* from 4 valid sinogram ranges using:
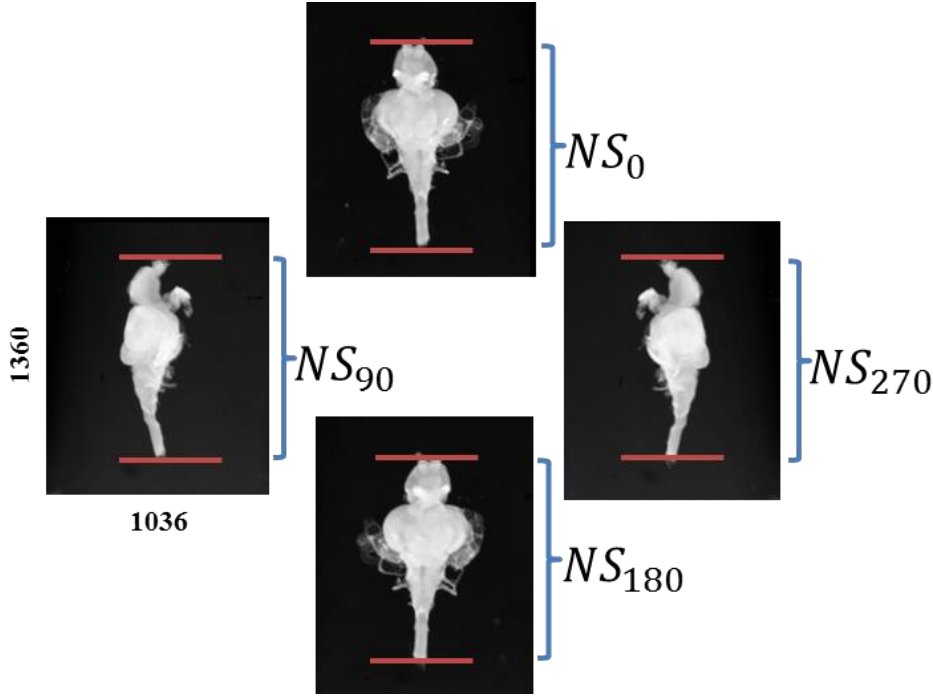
Figure 2.3. The 4 orthogonal sinogram ranges NS_0, NS_{90}, NS_{180} and NS_{270} are from the 4 orthogonal tomogram images ($0°, 90°, 180°, 270°$) of an adult zebrafish brain.

$$NS = NS_0 \cap NS_{90} \cap NS_{180} \cap NS_{270} \tag{1}$$

NS_0, NS_{90}, NS_{180} and NS_{270} respectively represent the 4 slice ranges of 4 orthogonal tomogram images of size 1360×1036, as shown in Figure 2.3. Here, as an example we look at the brain of an adult zebrafish. Within NS, the step for selecting a sinogram $step = ceil(\frac{NS}{\rho})$ is experimentally determined and approximately ρ sinograms from the range NS are evenly selected. The selected set of sinograms is defined as S. In this manner, specimen samples have approximately the same number of selected sinograms for CoR localization regardless of their different sizes.

2) Interest point detection

According to the design of our OPT imaging system, only the front half of the sample is in the DoF, so the projected data from opposite angles may vary differently depending on the rotation angle, specimen size and shape. However, a voxel projected at the left or right boundary of the specimen shares approximately the same image intensity with the opposite projection of the same sample location. This equivalence is shown as a peak and a trough along the sinogram. To illustrate our assumption, a sinogram from the fluorescence tomogram of a zebrafish larva is depicted as an example in Figure 2.4. Point O, A and B are the fluorescence signals of the 6 days post fertilization (*dpf*) zebrafish eye from 3 different angles, and O*, A* and B* are their corresponding opposite projections.
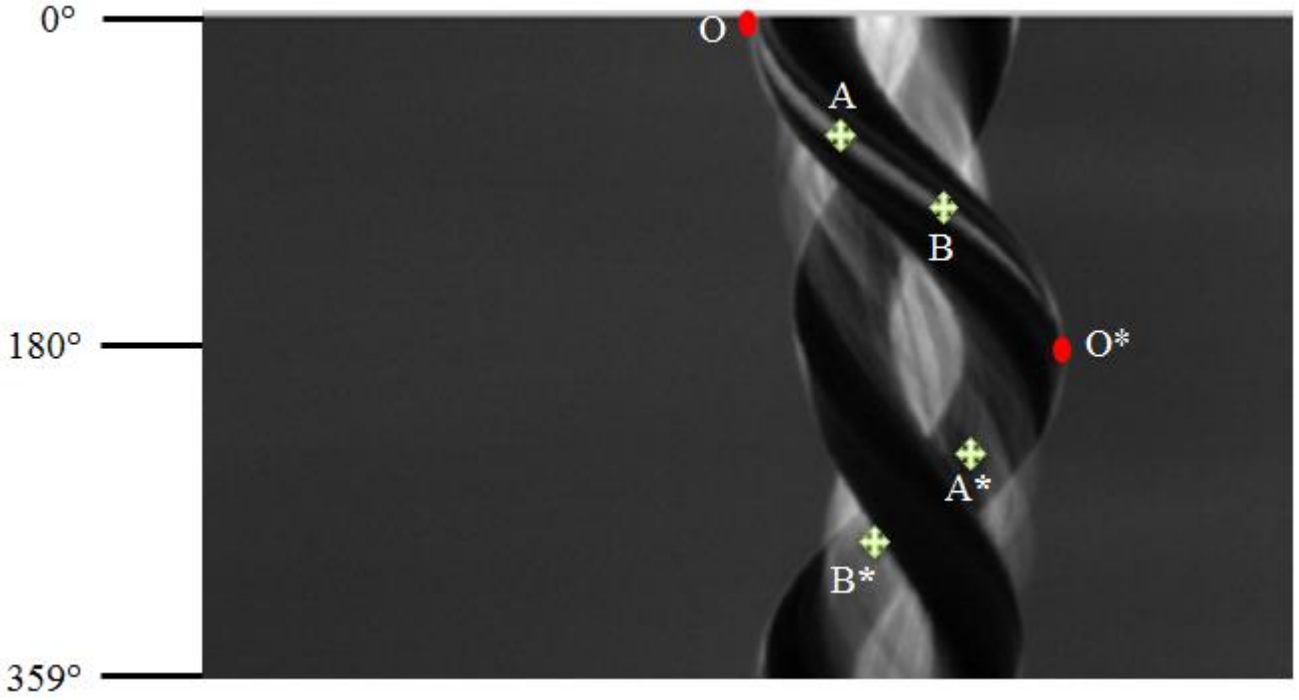
Figure 2.4. A sinogram of a zebrafish embryo showing the differences among pairs of opposite projected data. O and O*, A and A*, B and B* are pairs respectively. O and O* are interest points; while A and A*, B and B* are not.

The projected data for a voxel in the eye should be formed as a sine function passing through O, A, B, O*, A* and B* in CT system, but in OPT only O and O* remain equivalent; while A and A* as well as B and B* differ significantly, being consistent with the assumption above. With this assumption, the CoR should be located with the oppositely projected pairs that are similar to O and O*. The problem of locating the CoR is therefore transformed as search for peaks and troughs on the sinogram edge; in our case defined as interest points.

A sinogram is defined as $S(\xi, \varphi)$ where φ is the rotation angle, and ξ is the phase in each angle. The size of the sinogram is $\phi \times p$ in our case, with ϕ being the number of sample angles and p being the tomogram height after cropping. As depicted in the flowchart in Figure 2.5, the procedure for detecting interest point is based on point selection from initial points $E = \{\xi_k, \varphi_k\}$, $k \in [1, M]$. E is the collection of points using edge detection in a sinogram with M being the number of initial points. $S_b(\xi, \varphi)$ refers to the binary sinogram. After point selection, the detected interest points are $P = \{\xi_j, \varphi_j\}$, $j \in [1, N]$, and $N \leq M$. In Figure 2.5, the detailed algorithm for point selection is presented in the flowchart.

As shown in Figure 2.5, $W0_k$ and W_k are the window patches of an initial point $(\xi_{one}, \varphi_{one}) \in E$, meeting different conditions. ED_k is the image edge detected from W_k. In W_k we define θ as the angle passing through the 0-labeled centre $(\xi_{zero}, \varphi_{zero})$ and the 1-labeled centre $(\xi_{one}, \varphi_{one})$ that are separated by edge ED_k. If ED_k is enclosed, we

set a constraint of $\tan\theta < \frac{1}{3}\sqrt{3}$, where $\arctan\frac{1}{3}\sqrt{3} = 30°$, therefore only points with $\theta < 30°$ will remain. The peak and trough within W_k (indicated with red stars in Figure 2.6) are defined as follows:

$$\text{Peak:}\begin{cases} \xi_{zero} > \xi_{one} \\ |D_\varphi| = (w_1 - 1) \\ D(D_\xi) < 0 \\ 1 \notin sign(d_{\xi 1}) \\ 1 \notin sign(d_{\xi 2}) \end{cases} \qquad \text{Trough:}\begin{cases} \xi_{zero} < \xi_{one} \\ |D_\varphi| = (w_1 - 1) \\ D(D_\xi) > 0 \\ -1 \notin sign(d_{\xi 1}) \\ -1 \notin sign(d_{\xi 2}) \end{cases} \qquad (2)$$

D_φ symbolizes the sum of derivatives of ED_k in the φ direction along the edge curve, while $D(D_\xi)$ is the sum of second derivatives of ED_k in the ξ direction along the edge curve. When $D(D_\xi) < 0$, the function of the ED_k sequence is constrained as being convex, and if $D(D_\xi) > 0$, it is concave, corresponding to the peak and trough, respectively. We break ED_k into upper and lower edges: ED_{k1} and ED_{k2}, both of which are started at the middle of ED_k in the φ direction. Now, $d_{\xi 1}$ and $d_{\xi 2}$ are separately the derivatives of ED_{k1} and ED_{k2} in the ξ direction.

Applying the definition from Eq. (2), false-peak and false-trough (indicated with purple stars in Figure 2.6) are not kept as interest points, as they are not true sine peaks but rather intersections of different sine functions, which should therefore be discarded. Furthermore, when a true trough satisfies R (yellow star in Figure 2.6),

$$R = \begin{cases} \xi_{zero} < \xi_{one} \\ |D_\varphi| = (w_1 - 1) \\ D(D_\xi) = 0 \\ sign(d_{\xi 1}) = 0 \\ sign(d_{\xi 2}) = 0 \end{cases} \qquad (3)$$

or a true peak satisfies Q,

$$Q = \begin{cases} \xi_{zero} > \xi_{one} \\ |D_\varphi| = (w_1 - 1) \\ D(D_\xi) = 0 \\ sign(d_{\xi 1}) = 0 \\ sign(d_{\xi 2}) = 0 \end{cases} \qquad (4)$$

The edge ED_k in W_k of size w_1 is strictly vertical. Then it does not satisfy the definition of peak or trough in Eq. (2). The reason for the inconsistency is caused by the size of W_k. Therefore, to solve this problem, a bigger patch (set as $2w_1+1$ in our experiment) is set to satisfy Eq. (2) in a bigger patch, following the same steps for interest point detection in Figure 2.5.
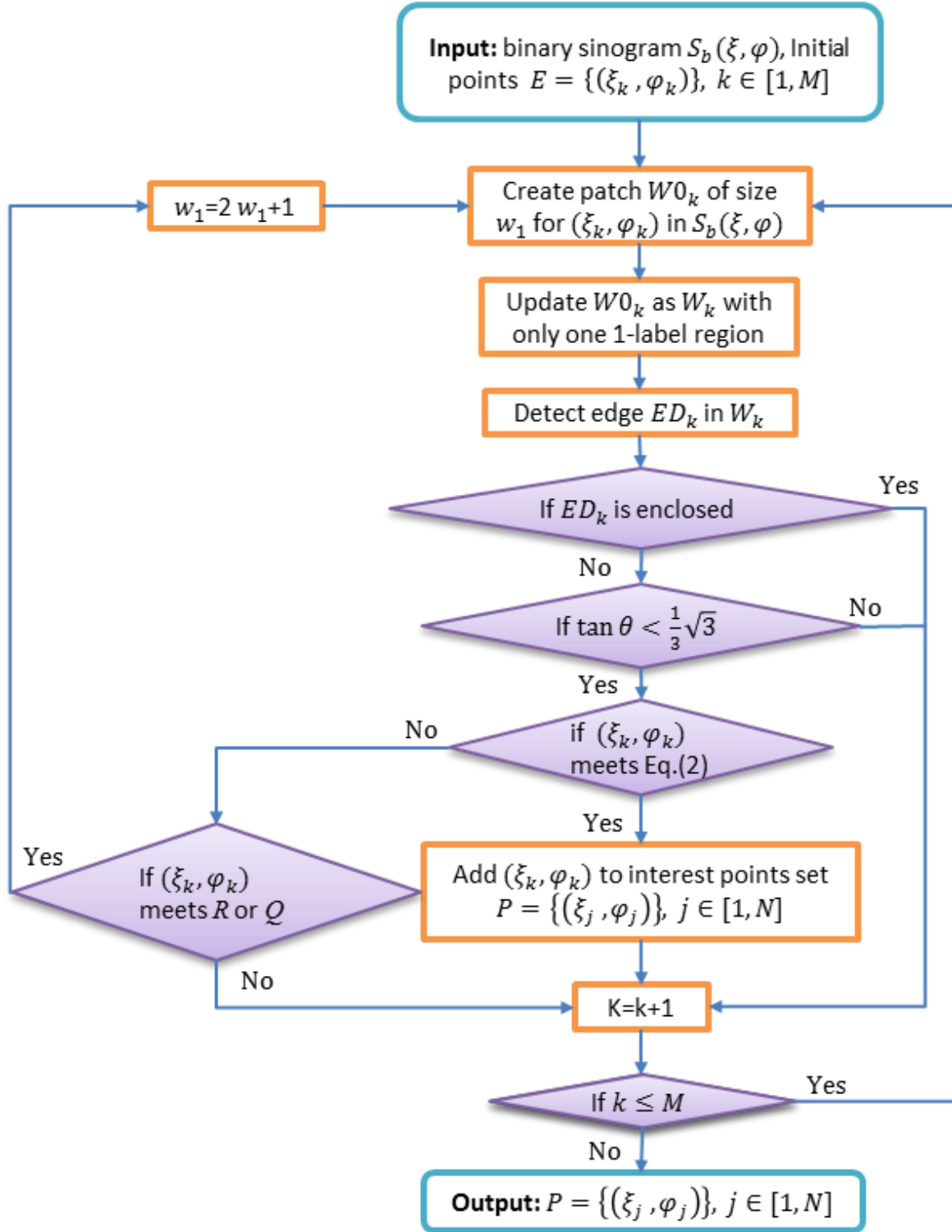
Figure 2.5. Flowchart of the algorithm for interest point detection. The binary sinogram and initial points are obtained through OSTU segmentation [50] and the Sobel edge detector [51] respectively.

3) CoR localization and alignment

According to the definition of CoR localization above, CoR for single sinogram should be first localized. With the interest points $P = \{\xi_j, \varphi_j\}$, $j \in [1, N]$ detected in single sinogram, the CoR range is obtained as $[\xi_{min}, \xi_{max}]$, where ξ_{max} and ξ_{min} are respectively the maximum and minimum of ξ_j in the interest points P. For a specific CoR value c, we locate the corresponding opposite points for P as $P_c' = \{(\xi_j, \varphi_j)_c'\}$, $j \in [1, N]$, which are symmetric by c and have an interval of π in projection. To find a mathematical metric between P and P_c', we define the neighbors of (ξ_j, φ_j) and $(\xi_j, \varphi_j)_c'$ as $r_c(\xi_j, \varphi_j)$ and $r_c'(\xi_j, \varphi_j)$. As shown in Figure 2.4, the projection data between interest point (ξ_j, φ_j) and its opposite point $(\xi_j, \varphi_j)_c'$ should be approximately equivalent, so we localize the optimal CoR in the range of $[\xi_{min}, \xi_{max}]$ for the *i*th sinogram by formulating:

$$C_i^* = \min_c \frac{1}{N} \sum_j^N (r_c(\xi_j, \varphi_j) - r_c'(\xi_j, \varphi_j)) \qquad (5)$$

For the selected ρ sinograms, the localized optimal CoRs are $C = \{C_1^*, C_2^*, \dots, C_\rho^*\}$; so the most frequently occurring value C^* in C is referred to as the CoR for a single channel, either the bright-field or fluorescence channel.
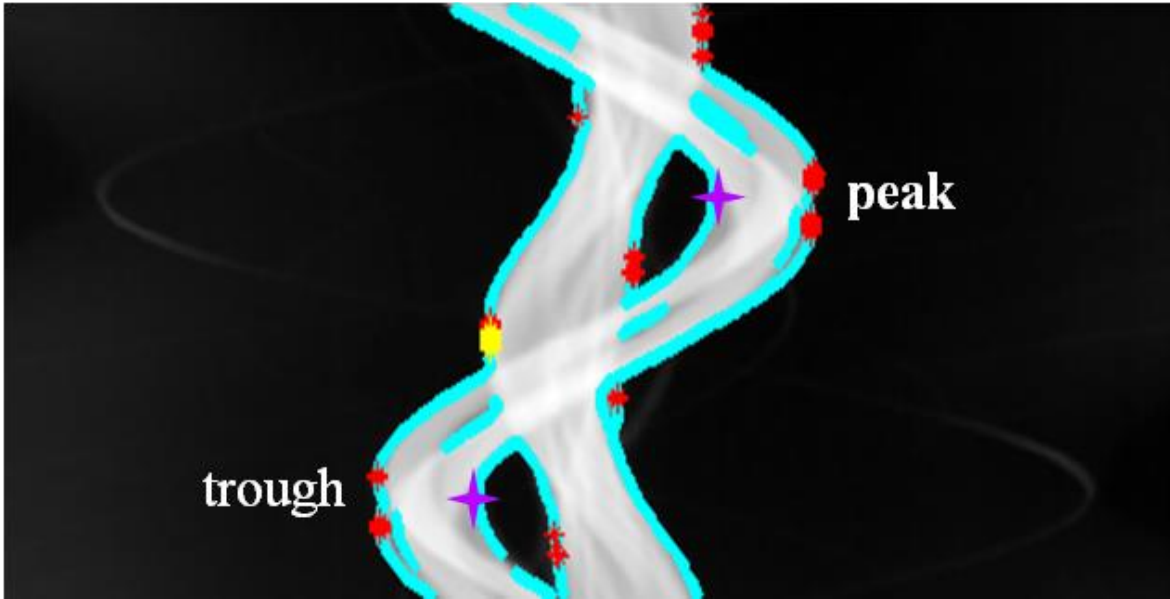
Figure 2.6. A bright-field sinogram (displayed inverted) from a chicken heart images with interest points detected (shown as red and yellow stars). Edge points, i.e. initial points, are shown in cyan; while false-peaks are shown as purple stars.

For multiple channels the sinograms should be aligned to the same size with the same CoR before 3D reconstruction. The disparity between different channels may be a result of a mechanic drift, when images are recorded at different time. We illustrate our alignment scheme in two channels (fluorescence and bright-field), but it is also suitable to multiple channels. The ith sinogram S_i of size $\phi \times p$ in each channel is aligned centered by C^* as S_i' of size $\phi \times q$. This is accomplished by using:

$$q = \begin{cases} 2 \times C^*, & C^* < \frac{p}{2} \\ 2 \times (p - C^*), & C^* \geq \frac{p}{2} \end{cases}, \tag{6}$$

and

$$S_i' = \begin{cases} S_i(1:2C^*, \varphi), & C^* < \frac{p}{2} \\ S_i(2C^* - p:p, \varphi), & C^* \geq \frac{p}{2} \end{cases} \tag{7}$$

As illustrated in Eq. (6), q is calculated to be smaller than p to preserve sufficient sinogram information, as well as to avoid redundant background reconstruction, i.e. S_i is truncated instead of being extended, which consumes more time for reconstructing the background. With Eq. (6) and Eq. (7), the ith sinogram for the fluorescence and bright-field channel are S_{fi}' and S_{bi}' with size of $\phi \times q_f$ and $\phi \times q_b$ respectively. They are aligned to the same CoR with the same size as S_{fi}^* and S_{bi}^* by using:

$$\begin{cases} S_{fi}^* = (z_0, S_{fi}', z_0); \ S_{bi}^* = S_{bi}', & q_f < q_b \\ S_{fi}^* = S_{fi}'; \ S_{bi}^* = (z_0, S_{bi}', z_0), & q_f > q_b \end{cases} \tag{8}$$

where z_0 is a Zero matrix with size of $\phi \times |C_f^* - C_b^*|$, and C_f^* and C_b^* represent the located CoRs for the fluorescence and bright-field channel.

2.3.2 Reconstruction and fusion

Next, from the corrected sinogram, now we have a new sinogram. Applying inverse radon transform or filtered back projection (FBP) [52] to both the bright-field and the fluorescence channel of the obtained samples, the reconstructed images are formed as $R_b^t = \{R_{b1}, \dots, R_{bl}, \dots, R_{bL}\}$ and $R_f^t = \{R_{f1}, \dots, R_{fl}, \dots, R_{fL}\}$ respectively. L is the number of sinogams at imaging time t. The intensities in R_{fl} and R_{bl} refer to the fluorescence and bright-field signals. For transparent specimens, such as the zebrafish larvae from our experiments, the bright-field signals are generally distributed in vertebra and cartilage, providing a description of the silhouette of the zebrafish. Therefore, by fusing R_f^t and R_b^t, fluorescence signals, for instance a pattern of gene expression could be located and analysed within specimens at the specific time t. R_f^t and R_b^t are fused according to their equivalent slice number l and imaging time t. For each voxel in the fused 4D image, $V_{(x,y,l,t)} = (I_f, I_b)$ describes its signals of different channels, and (x, y, l, t) symbolizes

the coordinate in 4D space. (x, y) corresponds to the pixel of the reconstructed image slice, while l and t symbolize the slice number and imaging time. $V_{(x,y,l,t)}$ could be further used in a 3D segmentation procedure and quantification of fluorescence, i.e. gene and/or protein activity, in the specific specimen or organs.

2.3.3 Parallel setting
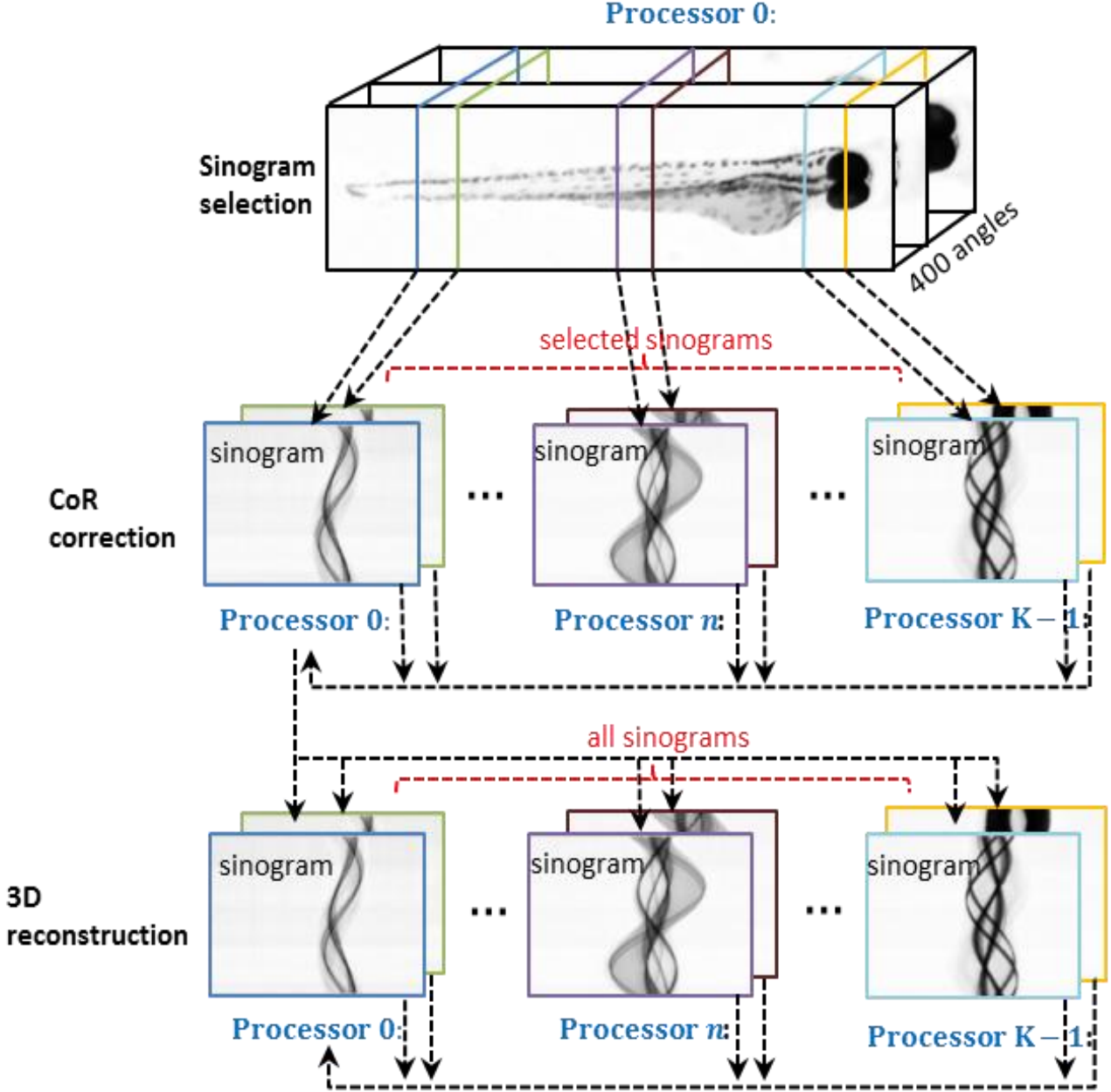
Figure 2.7. Parallel framework for CoR correction and 3D reconstruction. Processor 0 is responsible for sinogram selection, broadcast and collective communication, as well as normalization before image writing.

In order to speed up computations, we implemented a parallel computing scheme on our cluster computers, i.e. LLSC, for both the CoR correction as well as for the 3D reconstruction. This scheme is an essential part of our proposed pipeline. The specific implementation is illustrated in Figure 2.7. K represents the number of available processors. Processor 0 is defined as the Master processor unit responsible for sinogram selection, broadcast and collective communication. The selected sinograms S are then distributed to the K processors for localizing the corresponding CoRs using Eq. (5). All the CoRs from different processors are gathered by the Master processor unit to calculate the C^* of each channel. Subsequently, the CoR alignment of the different channels using Eq. (6), Eq. (7) and Eq. (8) is also processed on the Master processor unit. After CoR correction, the Master processor unit distributes the L slices of aligned sinograms to the K processors for the 3D reconstruction. The reconstructed image slices will be gathered again to the Master processor unit for normalization before image writting.

2.4 Experiments

In this section, we first evaluate the reconstruction pipeline without the CoR correction, qualitatively and quantitatively comparing the results with the pipeline considering the CoR correction. The runtimes of distributed computing for both experimental setups are measured and compared. We present a new CoR correction algorithm in this chapter, therefore further performance comparisons with previous CoR correction algorithm are also included.

2.4.1 Experiments on the fast post-processing pipeline

Instead of calibration of CoR based on adapting the adequate parts in the imaging system, we acquired a tomogram of a zebrafish sample and reconstructed it using the proposed pipeline. The raw 1036×1360×400 OPT tomogram was first cropped, producing a smaller image size of 506×1360×400, followed by a background subtraction from the 10 background images acquired, cf. § 2.2.2. The cropping and background subtraction normally take less than 1 second, the average time of each channel for CoR correction and 3D reconstruction is 15.05s and 18.81s respectively with 5 nodes of 8-core 2.66 GHz CPU+ 16G RAM and 8 nodes of 4-core 2.66 GHz CPU+16G RAM. The number of selected sinograms is set to $\rho = 40$ to balance the effectiveness and computational complexity of CoR localization, and the patch size is $w_1 = 15$. The visualizations of 3D reconstruction without and with the CoR correction are respectively shown in Figure 2.8 (a), (b), (c) and Figure 2.9 (a), (b), (c). It is obvious that Figure 2.9 (a), (b) and (c) contain distinct signals in both channels while Figure 2.8 (a), (b) and (c) show blurred 3D models and signals. Maximum projections of 10 slices of cross-section image selected from the combination model in (c) are magnified in Figure 2.8 (d) and Figure 2.9 (d). Figure 2.9 (d) accurately represented the zebrafish spinal cord and specific GFP fluorescent signals.
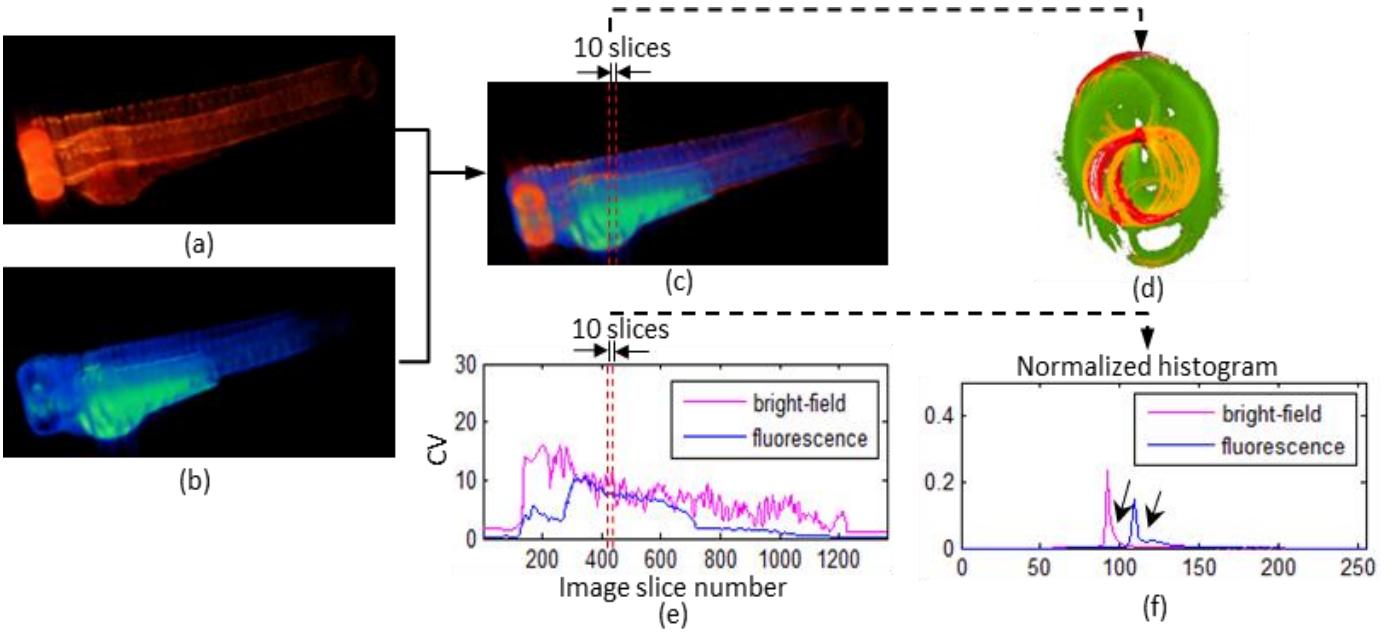
Figure 2.8. An example of 3D reconstruction for a zebrafish without CoR calibration prior to image acquisition . The results are obtained based on the proposed pipeline without CoR correction. (a) , (b) 3D image in bright-field and Fluorescence channel. (c) The combination of (a) and (b). (d) Reconstructed slices between 410 and 419 of (c). The bright-field and fluorescence signals are shown in red and green and the intersections of them are in yellow. (e) The coefficient of variation of all slices corresponding to (c). (f) Normalized histograms for the average of the 10 slices selected from (c) and (e) . black arrows indicate the statistical characteristics of the reconstructed silhouette for the zebrafish.
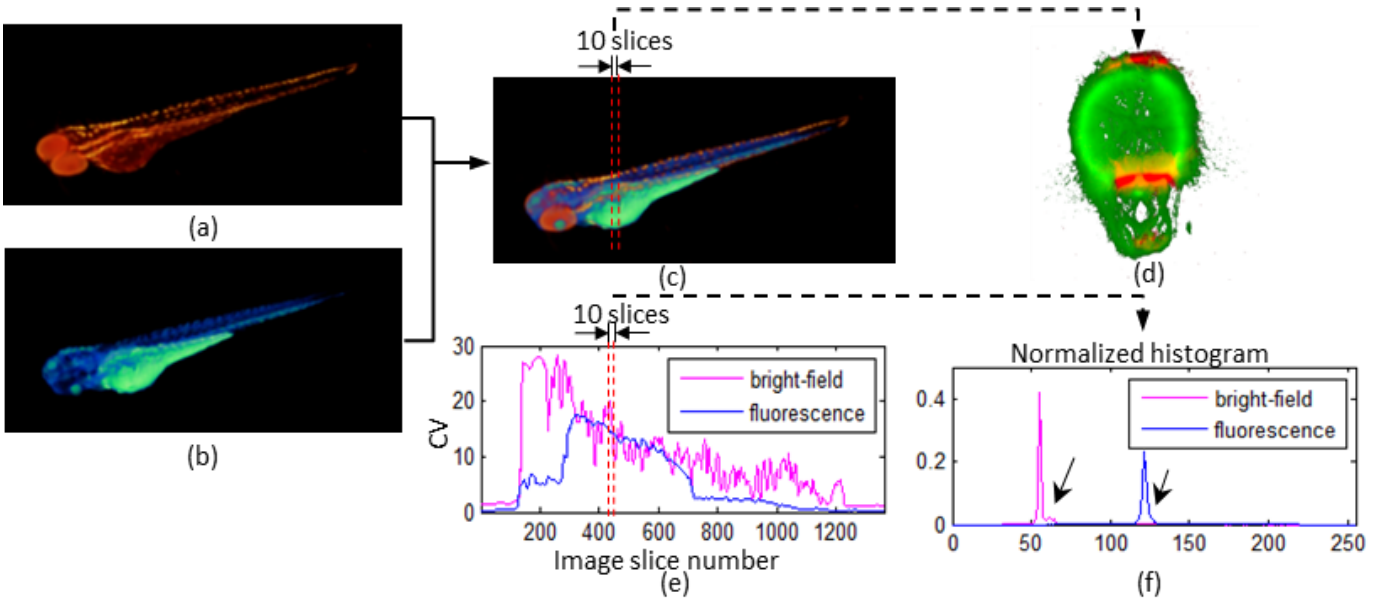
Figure 2.9. 3D reconstruction for the same zebrafish sample as Figure 2.8, without CoR calibration prior to image acquisition , but the omission of the calibration is compensated with the proposed pipeline with CoR correction. (a), (b) 3D image in bright-field and Fluorescence channel. (c) Combination of (a) and (b). (d) Reconstructed slices between 410 and 419 of (c). The bright-field signals (in red) show the outline of the zebrafish. The fluorescence signals (GFP) are

shown in green and the brightness of color indicates the strength of GFP signals, decribing the fluorescent texture. (e) coefficient of variation of all slices corresponding to (c). (f) Normalized histograms for the average of the 10 slices selected from (c) and (e). The peaks (background) are higher and the edges (black arrow) are much sharper compared to those in Figure 2.8 (f).

To quantitatively compare the difference between the reconstructed slices without and with CoR correction, coefficient of variation (CV): $CV = \frac{\sigma}{\mu}$ is calculated for each reconstructed slice as shown Figure 2.8 (e) and Figure 2.9 (e). It should be noted that the CVs calculated in our experiments are based on the raw reconstruction without scaling. In terms of reducing artefacts produced in the reconstruction, we aim to simultaneously maximize the variance and minimize the mean of bright-field and fluorescence signals, presenting the specimen with the least of blur. By comparing Figure 2.8 (e) with Figure 2.9 (e), we can see that after CoR correction CV increases significantly on all slices in both channels.

To observe more details in the reconstructed image slices, the histograms for the average image of the selected 10 slices without and with CoR correction are illustrated in Figure 2.8 (f) and Figure 2.9 (f). The pixel value corresponds to the bright-field or fluorescence signal strength. In practice, the bright-field image is inverted to satisfy the correspondence of pixel value and signal strength. In Figure 2.8 (f) and Figure 2.9 (f), the peaks of the histogram indicate the pixel values of the background, and values for signals are on the right of the peak. We can observe that the background boundary of the histogram (black arrow) in Figure 2.9 (f) is sharper than that in Figure 2.8 (f). The peaks in Figure 2.9 (f) are both higher than peaks in Figure 2.8 (f), indicating that CoR correction clears the background which is smeared by blurred artefacts. This is consistent with the refined and distinct silhouette and texture of reconstructed image with CoR correction.

To illustrate the over-all runtime of the pipeline in our system, we repeat the experiments on 10 more sample specimens including zebrafish, zebra-finch embryo and adult chicken heart. These samples are acquired at different magnifications. The parameters for the CoR correction and the configuration of parallel computing are set as the same as the experiments above. The only difference is that we fixed the image size to the original 1036×1360×400 pixels without cropping to exclude the influence of different image sizes on the runtime. Figure 2.10 describes the runtime of 3D reconstruction (Rec) and that of 3D reconstruction with CoR correction (Rec&CoR). Over all, Rec takes an average time of 26.91s and Rec&CoR takes 54.66s for all the 10 dataset tested.
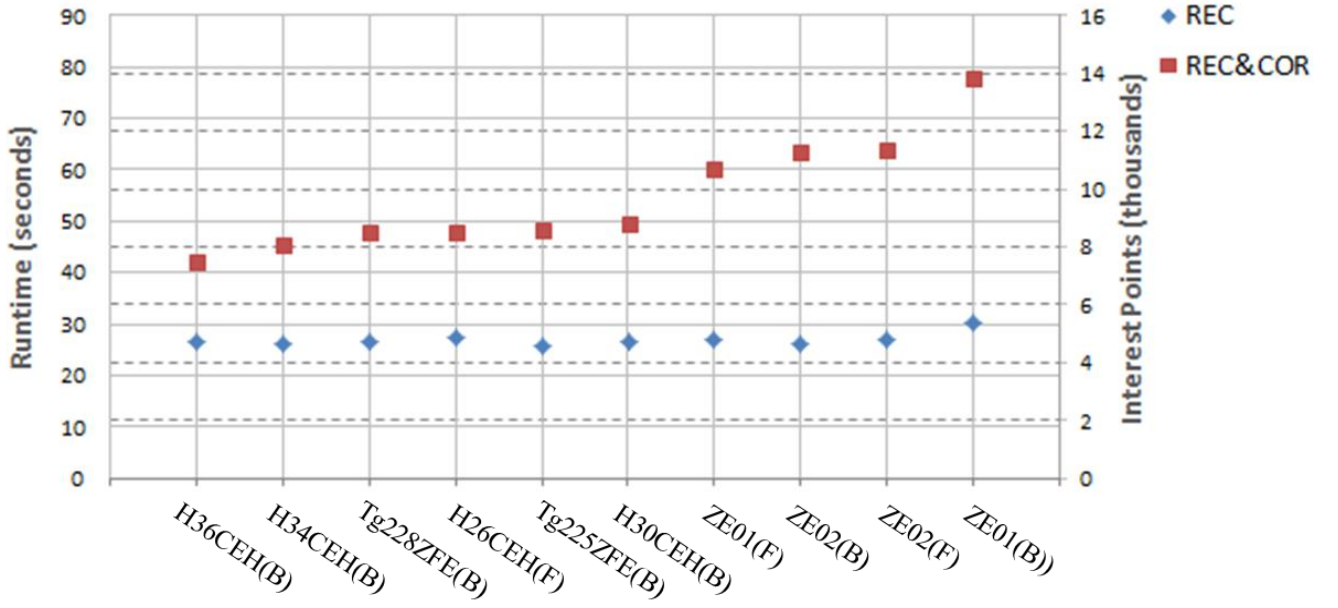
Figure 2.10. The runtime (solid axis) of the pipeline implemented on the cluster. Rec represents the runtime for 3D reconstruction and Rec&CoR indicates the runtime that includes CoR correction as well. The increasing runtime of REC&CoR corresponds to the increasing number of interest points (dashed axis) detected in the CoR correction algorithm on different data.

2.4.2 Comparison of different CoR corrections on different data

Three previous CoR correction approaches are analysed and compared to our method on the 12 OPT images. The Pixel Match method [42] and the Cross Correlation Operation (CCO) method [43] are based on signal match for pairs of projection, both of which are successfully used in CT CoR correction. The most commonly used method for CoR correction in OPT, as described in [46], here referred to as the Automated method. As the results of CoR correction depend on the selection of sinograms, the comparisons of different CoR correction methods are implemented on multiple sinograms selected with the proposed strategy above. For our experiments, the 12 OPT datasets consisting of whole-mount organisms as well as dissected organs. As the developmental stages differ, the size and shape differs. The samples comprise zebrafish embryo (ZE), chicken heart (CEH) and zebra finch embryo (ZFE) using both the bright-field (B) and fluorescence (F) channel. The samples are part of researches in embryo development, skeleton development and heart defection.
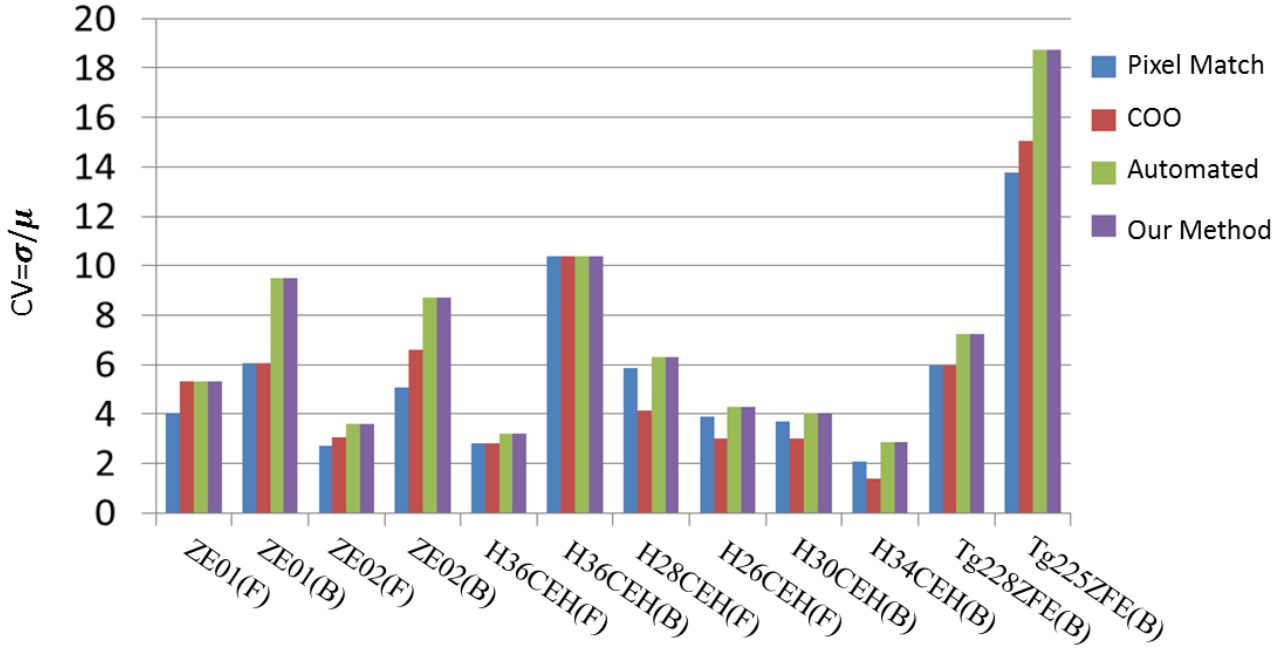
Figure 2.11. The comparison of average coefficient of variation (CV) for reconstruction with 4 different CoR correction methods on 12 datasets. For each dataset, larger CV corresponds to discrimination of information and less artefacts introduced by reconstrution. ZE: zebrafish embryo; CEH : chicken heart; ZFE: zebra finch embryo; Different prefixes refer to different developmental stages of the sample specimens at acquisition. B and F are the bright-field and fluorescence channel.

Table 2.1. Runtimes of different CoR alignment aprroaches in each sinogram on different datasets

Datasets	Pixel Match(s)	CCO(s)	Automated(s)	Ours(s)
ZE01(F)	0.6897	102.8966	1043.1034	10.3448
ZE01(B)	0.6765	102.7941	1042.7941	12.2941
ZE02(F)	0.6897	102.9310	1043.3793	11.7241
ZE02(B)	0.6786	102.8929	1042.8214	11.7241
H36CEH (F)	0.6774	102.9032	1043.2580	2.2258
H36 CEH (B)	0.6667	102.9333	1042.8444	6.2888
H28 CEH (F)	0.6800	103.1200	1043.1200	5.2000
H26 CEH (F)	0.7200	103.2000	1043.2000	4.3200
H30 CEH (B)	0.6774	103.5484	1043.1935	4.1612
H34 CEH (B)	0.6897	103.5517	1043.0344	3.3793
Tg228 ZFE (B)	0.7143	103.1786	1042.8571	2.6785
Tg225 ZFE (B)	0.7857	103.2143	1042.7857	2.7857
Average	**0.6955**	**103.097**	**1043.0326**	**6.4272**

In Figure 2.11 the results are depicted. These indicate the measurement of CV considering both the variance and mean of the reconstructed results, which are more convincing and reliable. The data used for CV calculation is the raw data after reconstruction but before image normalization, excluding the effect of scaling. We

should consider that the CV is only a criterion for evaluating the performance of different reconstruction of the same specimen. It is not suitable for comparing the reconstruction performance across specimens, because the variance and mean value differs in the different specimen structures. However for the same specimen, the CV is very suitable for evaluating reconstruction performance than variance in [46]. In Figure 2.11, the Automated method [46] and our method obtained maximum values for the CV in all the 12 datasets, because both methods achieved the optimal and equivalent CoR in each dataset. The Pixel Match [42] and CCO methods [43] gained different reconstruction and CV performance on the different data. The reason for this variation is that the algorithms in [42] and [43] strongly depend on the symmetry of all opposite projected pixel pairs. In the process of OPT imaging system, however, most of the pairs are not symmetrical.

Achieving competitive performance to the Automated method regarding to reconstruction quality in Figure 2.11, our method performs significantly superior to CCO method and Automated method in terms of its computational complexity; cf. Table 2.1. With the computer configuration of 16Gb RAM and 8-core 3.4GHz CPU, the average runtime of different CoR correction methods for single sinogram are 0.6955s, 103.097s, 1043.0326s and 6.4272s respectively. The Pixel Match method [42] achieves highest runtime performance, but its capability of optimal CoR correction is limited. Overall, our method outperforms the other three by considering the effectiveness and complexity of synchronous computation. It is noteworthy that in our method the runtime of different datasets varies due to the differences in the number of interest points. The other three approaches, however, consume approximately the same runtime for each sinogram because they are considering a fixed number of sinogram pixels.

2.5 Conclusions

In this chapter we presented a fast post-processing pipeline for OPT tomograms including cropping, background subtraction, CoR correction and 3D reconstruction, with focus on parallel computing. For CoR correction, a new automated CoR correction method was proposed, outperforming the other three CoR correction approaches in terms of general efficiency. In terms of 3D reconstruction, we have implemented the inverse radon transform on our cluster computer, i.e. LLSC, to achieve faster reconstruction. A pipeline was implemented using parallel computation and the average runtime based on the 10 datasets with a fixed image size of $1036\times1360\times400$ is 54.66s, using 5 nodes of 8-core 2.66 GHz CPU+16G RAM and 8 nodes of 4-core 2.66 GHz CPU+16G RAM. Furthermore, the proposed CoR correction methodology could suppress random or fixed noise in background, because only peaks and troughs of the sinogram from foreground are detected as interest points. Importantly, the proposed pipeline and CoR correction are also suitable for 3D CT image reconstruction and comparison when used in medical therapy. Currently, the integrated system including imaging, data transfer, pre-processing, CoR correction, 3D reconstruction and visualization is being optimized as a distributed

application. With this integrated system, a profile of the organism/organ enhanced fluorescence probes within it can be imaged, reconstructed and visualized in a very short period of time. In our future work, a quantitative model for locating, calculating and tracking fluorescent signals (gene and/or protein activity) will be established.

2.6 Acknowledgement

Chapter 2

Chapter 3

Deblurring Images from 3D Optical Projection Tomography Using Point Spread Function Modelling

This chapter is based on the following publications:

Tang X., Lamers G.E.M. & Verbeek F.J. (2018), 3D Image Quality Improvement for Optical Projection Tomography via Point Spread Function Modelling. *In Imaging and Applied Optics 2018 (3D, AO, AIO, COSI, DH, IS, LACSEA, LS&C, MATH, pcAOP), OSA Technical Digest (Optical Society of America, 2018)*, paper 3W5G.5.

Tang, X., Lamers, G. and Verbeek, F. (2019). 3D Image Deblur using Point Spread Function Modelling for Optical Projection Tomography. *In Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 2: BIOIMAGING*, ISBN 978-989-758-353-7, pages 67-75. DOI: 10.5220/0007237700670075.

Chapter summary

Optical projection tomography is successfully used in the life-sciences for 3D imaging of specimens of size between $1\,mm$ and $10\,mm$. However, this requires imaging of large specimens at a large depth of field, which normally results in blur in imaging process, i.e. it compromises the image quality or resolution. Yet, it is important to obtain the best possible quality 3D image from the OPT, therefore deblurring of the image is important. The imaging process is modelled through the point spread function: the imaging of a point light source through the lens system. In this chapter we first model the point spread function along optical axis which varies at different depths in the OPT imaging system. Subsequently, the magnification is taken into account in the point spread function modelling. Deconvolution in the coronal plane based on the modelled point spread function is implemented to correct for blurring. Experiments with the proposed approach based on 25 3D images including 4 different categories of samples, indicate the effectiveness of quality improvement assessed by image blur measures from both the spatial and frequency domain.

3.1 Introduction

Here we introduce the origin for imaging blur in optical projection tomography (OPT) and elaborate our motivation of deblurring the 3D image using the point spread function (PSF). The aim of our contribution is to improve the visual resolution by deblurring the 3D OPT image by means of deconvolution based on the modeled PSF of the imaging system. This will, for large samples with a focal plane being at or away from the centre of rotation (CoR), recuperate the imperfections of 3D image resulting from the imaging system. The method for the PSF modelling will be explained in Section 3.2 and the qualitative and quantitative image comparison will be presented in Section 3.3. In Section 3.4 we will give our conclusions.

3.1.1 Background: 3D image deconvolution

As mentioned in Chapter 1 (cf. § 1.1), OPT is an optical 3D imaging technique typically for objects at tissue-, organ- and organism-level in the magnitude range of millimeters, thereby filling a gap between confocal and computational tomography imaging in the resolution range. We use depth of field (DoF) to assess the image quality - DoF is defined as the distance between the nearest and the furthest objects in an image that are in acceptable sharp focus. A point object located within the DoF of the optical system is considered to be in focus, but not necessarily in optimal focus. Beyond the DoF, the object is out of focus [32]. DoF in OPT imaging system is shaped as a double fan symmetric around the focal plane. For OPT imaging and reconstruction, the DoF is expected to be large enough to contain as much of the sample as possible. In this manner the parts of the sample located in the DoF will result in an image more or less in focus. However according to previous studies [32],[33], a large DoF subsequently introduces image blur resulting in low in-focus image quality. The image quality in this chapter is also referred to as image resolution according to some literatures [53], i.e. the extent to which detail can be observed. The trade-off between DoF and image quality should be considered when selecting lens for an OPT imaging system. A lens with low numerical aperture (NA) will produce a large DoF, allowing imaging of larger samples but results in a relatively blurred image.

A 3D image is reconstructed from the OPT tomogram that are obtained by rotating the specimen and acquiring a series of wide-field images at regular angular intervals. This is accomplished over a full revolution of the specimen. The Filtered Back Projection (FBP) algorithm is typically used for 3D image reconstruction [54] in this case, cf. § 2.2. Deconvolution applied to the reconstructed 3D image is defined as 3D image deconvolution in this thesis.

A typical way to improve the image quality to meet the resolution in the best possible way is the application of deconvolution on each reconstructed slice at individual depth, using a constant theoretical or experimental PSF. This is commonly used in 3D scanning microscopy, e.g. confocal laser scanning microscopy (CLSM), where the images

acquired from microscopy are actually a subset of reconstruction slices. However, this approach, in terms of constant PSF deconvolution, is not strictly suitable for OPT imaging. Because the imaging PSF within the DoF in OPT system, varies at different depths along the optical axis. It is, therefore, necessary to model this variation in OPT imaging, which can be subsequently used for 3D image deconvolution. This explains our motivation for PSF modelling in OPT imaging system.

According to Chen *et al.* [55], OPT is typically performed with specimens that extend beyond the Rayleigh length or Rayleigh range of the imaging lens. Therefore, the tangential resolution of the reconstructed 3D image decreases away from the focal plane in a radial manner. When the focal plane coincides with CoR, the tangential resolution, centred at the CoR, decreases in a radial-symmetrical fashion. For an imaging system with a focal plane located away from the CoR the decrease in resolution is more complicated but the highest resolution is still found around the focal plane. In this case the focal plane in the reconstructed slice corresponds to a circle centred at the CoR, rather than a point coincident with the CoR. This subsequently appears as a cylindrical surface in the 3D image centred at the CoR.

The tangential resolution of the OPT 3D image slice decreases radially around the focal plane. Theoretically, the best resolution of the reconstructed 3D image can be achieved by combining all the coronal deconvolutions of different sample angles. The coronal deconvolution means deconvolving the 3D image with the PSF slice by slice in the coronal plane along its depth axis. This depth axis is parallel to optical axis of the modelled PSF. We will only implement the coronal deconvolution in 2 opposite angles, i.e. the reconstructed 3D image and its opposite sample at 180° centred at the CoR, in parallel considering the enormous time consumption of 3D matrix rotation in N angles and the symmetry of the focal plane. When the focal plane is off the CoR during the imaging process, the shift is accounted for by a shift in the model of the PSF.

In this chapter we focus on the presentation of the concept of PSF modelling and coronal deconvolution on 3D OPT data, accompanied by some initial experimental results based on 25 3D images including 4 different categories of samples. Further evaluations on a larger number of data are point of our current research.

3.1.2 Related work

Accounting for the trade-off between large DoF and high resolution, previous studies have proposed several methods to this problem. One possibility is choosing a high NA lens to acquire a high-resolution image and combining multiple focal planes in a simultaneous manner [56] or scanning the focal plane through the sample [57] (Miao *et al.*, 2010) . These multiple focal plane approaches solve the issue of narrow DoF, but the mechanism of multiple measurements and scanning increases the acquisition time and the complexity of the imaging system. Another direction is to use a reasonable NA lens and deblur the image by employing a deconvolution or filter on images before or after reconstruction. Walls *et al.* [32] first applied the frequency-distance relationship (FDR) [58]

in OPT. The corresponding filter was implemented on the sinogram prior to reconstruction. The quality of the 3D image can be further improved with weighted filtered back projection (WFBP) [59]; this is accomplished by considering the intensity distribution of multiple fluorescent spheres of known size along the optical axis. But the implementation of evenly placing each sphere along the optical axis is rather difficult to achieve. Chen *et al.* [55] proposed a way to determine the modulation transfer function (MTF) that contributed to MTF-mask filter and MTF-deconvolution filter in the reconstruction process. The former filter significantly reduced the artifacts produced by sparse projection but the latter filter had limited improvement on tangential image resolution. Additionally, a spatial-invariant experimental PSF was investigated by McErlean *et al.* [60] in order to improve the spatial resolution. However, spatial-invariance of the PSF is not entirely convincing for OPT. Most recently, a new deconvolution approach based on the reconstructed 3D image was proposed by Horst *et al.* [53]. In their approach the PSF was modelled and as such they achieved significant improvement on the reconstructed slice. Nevertheless, they focused on the deconvolution of vertically independent slices and omitted the PSF diffractions along the optical axis that concerns the interaction of different slices.

In this chapter, we present our contribution by modelling an experimental PSF from a single sphere along optical axis, thereby considering the interaction of contiguous slices from the reconstructed volume. At the same time, the magnification, as obtained from a zoom lens, is taken into account in the experiments.

3.2 Materials and methods

In order to model the PSF for our OPT imaging system, we first propose a protocol to prepare for the imaging of a point source, i.e. a fluorescence sphere or bead. The modelling approach will then be introduced and elaborated. This approach will be subsequently used for the deconvolution of 3D image in OPT.

3.2.1 Sample preparation of a single fluorescence sphere

To image the specimens in the range of several millimeters small-valued NA lens is used to obtain the large DoF in our OPT imaging system, i.e. effective NA: 0.0105~0.0705 as part of a Leica Stereo Microscope. The resolution of an optical system is defined as the minimum distance r at which two separate points can be distinguished as individuals. According to the Rayleigh criterion $r = 0.61 * \lambda/NA$ for a circular aperture with $\lambda = 509\ nm$ is the emission wavelength, the minimum size of the experimental fluorescence sphere is supposed to be in the range between $4.40\ \mu$ and $29.57\ \mu$. To make it visible in the image the sphere size is supposed to exceed this range. In our case, we choose the fluorescence sphere of size $43.2\ \mu m$ and diluted it to a concentration of $360\ beads/ml$.

Figure 3.1. The injection protocol with green spots indicating the injection position of fluorescence spheres. The cylinder corresponds to the shape of the agarose block as mounted in the OPT.

To image and model the PSF along optical axis we have developed an injection-based protocol to place the spheres into agarose as follows:

- 1% low melting point (LMP) agarose, cool down to ~37°;
- Drill cylindrical agarose shapes when it is semi-solidified in a petri dish;
- Inject the diluted spheres into the outer wall of the agarose along a line parallel to the central axis, preferably with a small size syringe. We use a 0.5 ml syringe with a needle length of 13 mm and diameter of 0.29 mm, as shown in Figure 3.1;
- Keep the agarose at 4℃ until it is fully solidified (~3 hours);
- Clear the sample with 70%, 80%, 90%, 96%, 100% ethanol, 100% ethanol: BABB (benzyl alcohol: benzyl benzoate = 2: 1) = 1: 1 and BABA.

Our goal is to acquire the images of a single sphere placed at different depths along optical axis. Therefore, randomly sprinkling the spheres into the agarose in a traditional way is not feasible. The main reason is that there may be interactions and overlap between different spheres either at the same or different depths. This makes the selection for imaging of a single sphere image difficult or even impossible. The images of each single sphere at different depths are acquired by means of sample rotation. Each rotation corresponds to a different depth in the OPT imaging system. The sphere injection method in of our protocol significantly reduces the probability of overlapping between different spheres. In this way the images of the same sphere in a full revolution can be easily and efficiently acquired. The OPT imaging system and environment is configured as explained in chapter 1 cf. § 1.2.

3.2.2 PSF modelling concerning different magnifications

For our experiments the acquisition of a single sphere comprises a full revolution in $0.9^{°}$ steps resulting in 400 images. In Figure 3.2 the processes of sphere image acquisition and PSF modelling are depicted. In Figure 3.2(a) and (b), the green dot represents the sphere and the red arrow indicates the sphere rotation. The excitation and emission beams are regarded to be parallel. This is indicated by blue and green arrows in Figure 3.2(a). For

PSF modelling, the focal plane is set at the CoR. The 3D image whose focal plane is shifted from the CoR, requires an equal shift in the PSF. With the protocol (cf. section 3.2.1) the physical rotation radius of the sphere r_b can be easily measured. To this end, we first measure the radius of the cylindrical agarose r_c and image it in the bright-field mode with short exposure time. In the same experimental environment, the sphere is afterwards imaged in the fluorescence mode. r_b is calculated as:

$$r_b = \frac{d_{bi}}{d_{ci}} \cdot r_c \tag{1}$$

with d_{bi} representing the rotation diameter of the sphere in the tomogram, achieved by measuring the distance of two opposite sphere centres that are both in focal plane. d_{ci} is the diameter of the cylindrical agarose in the bright-field image. Dividing r_b by the number of steps required for a rotation of $\pi/2$, i.e. the rotation radius r_b, the depth of each rotation step along optical axis is approximately determined as $r_b/100$ *. In our case the measured $r_c = 4\ mm$, $d_{bi}/d_{ci} = 0.751$, producing $r_b \approx 3\ mm$. Therefore, the physical distance of two adjacent rotations along optical axis is approximately $30\ \mu m$.
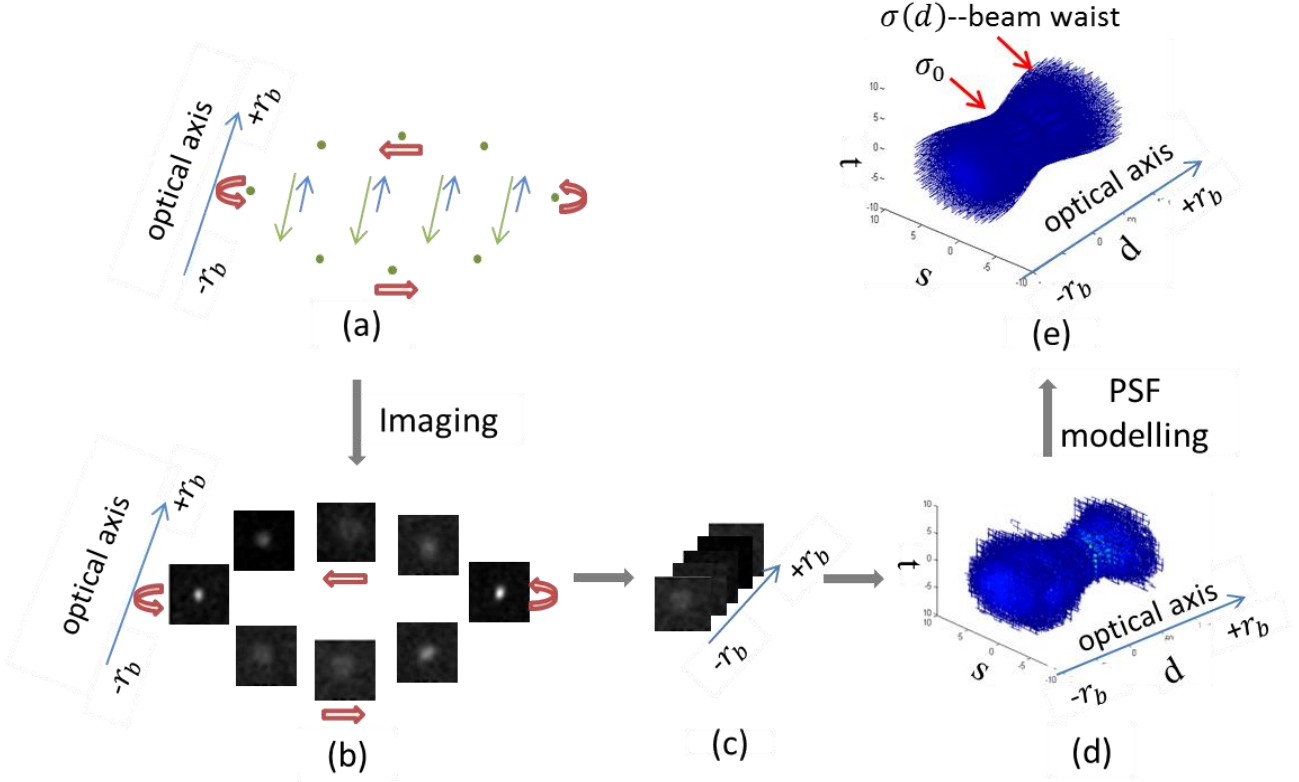
Figure 3.2. Image acquisition and PSF modelling of a single fluorescence sphere. (a) The light path of the OPT imaging system that passes through fluorescence sphere (green dot). The excitation beams and emission beams are separately shown as blue and green arrows. (b) Images of the single sphere acquired at different angles. (c) Images of the single sphere stacked according to the defocus. Half rotation with defocus range $[-r_b, +r_b]$ is required, in our experiment $r_b = 3\ mm$ as calculated from Eq. (1). (d) The experimental and discrete PSF with defocus $[-r_b, +r_b]$. (e) The modelled and continuous PSF with defocus $[-r_b, +r_b]$.
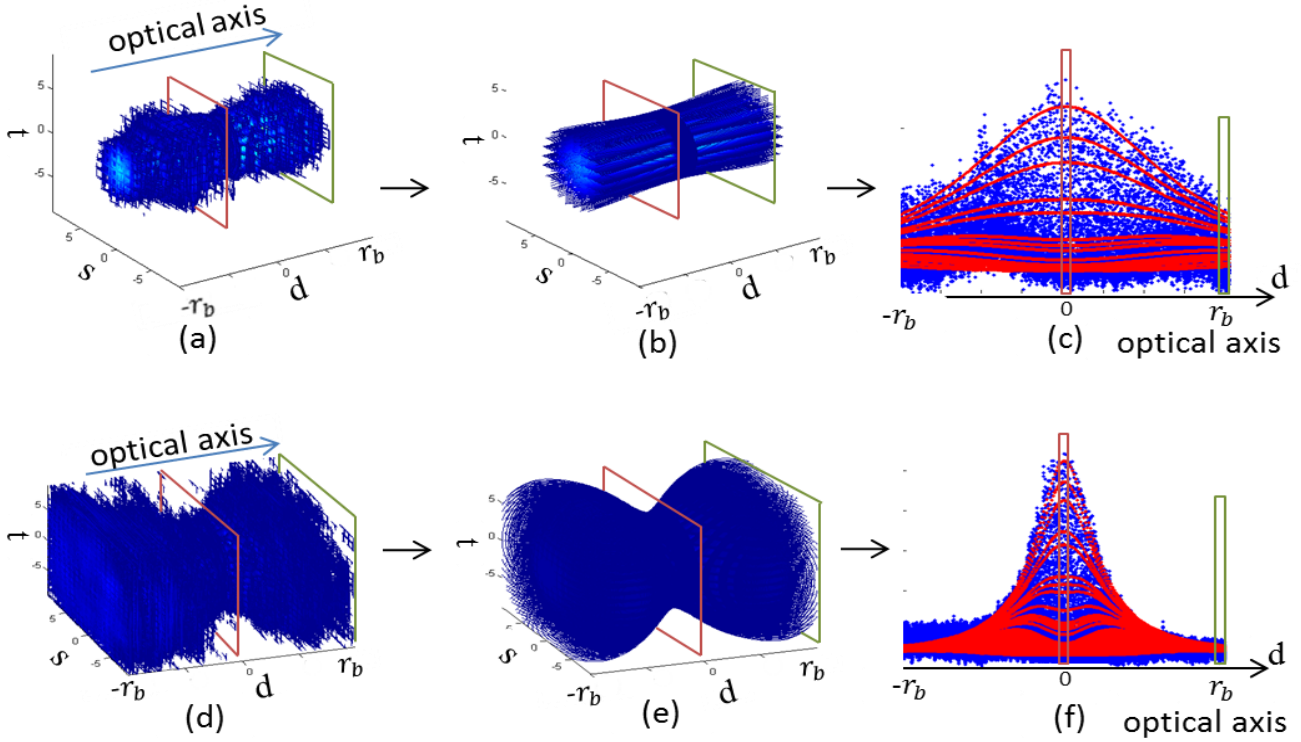
Figure 3.3. PSF modelling along the optical axis. (a), (d) Experimental PSFs acquired from images at magnification of 12.5 × and 25.0 ×. (b), (e) The corresponding modelled PSFs using Eq. (4) and Eq. (5). All the voxels of experimental data in (a) and modelled data in (b) are respectively transformed to blue and red dots in 1D functional in (c) to visualize the modelling performance. The vertical axis in (c) displays the intensity that corresponds to the voxel intensity in (a) and (b). Similarly, voxels in (d) and (e) are transformed to the data in (f).

According to the definition, the optical imaging PSF is assumed as a focused Gaussian-like beam [61], i.e:

$$p(s,t,d) = \frac{1}{2\pi\sigma(d)^2} \cdot exp(-\frac{s^2+t^2}{2\sigma(d)^2}) \tag{2}$$

Where s, t, d are the three axes in 3D space, with d being the optical axis. The $\sigma(d)$ is beam waist (Figure 3.2) given by:

$$\sigma(d) = \sqrt{\sigma_0^2 + (\frac{\lambda d}{\pi\sigma_0})^2} \tag{3}$$

with σ_0 the Gaussian beam waist defined as the $1/e$ value of the field amplitude in focus, λ the emission wave length of fluorescence spheres and d the defocus along optical axis.

For a specific magnification, σ_0 is constant, but it varies when imaging with different magnifications. Additionally, in Eq. (2) and Eq. (3) the beam waist $\sigma(d)$ is typically regarded as the standard deviation of the Gaussian model in previous studies [53]. Different from the Gaussian model [53], we can further generalize the model by employing parameters ρ_1, ρ_2 and ρ_3 as follows:

$$p(s, t, d) = \rho_1 \cdot \frac{1}{2\pi\sigma(d)^2} \cdot \exp(-\frac{s^2+t^2}{2\sigma(d)^2})^{\rho_2} + \rho_3 \tag{4}$$

Instead of equalizing the beam waist and standard deviation as described in [53] and [61], we investigate the relationship between them by multiplying a parameter a with beam waist, thereby considering different magnifications; thus,

$$\sigma(d) = a \cdot \sqrt{\sigma_0^2 + (\frac{\lambda d}{\pi\sigma_0})^2} \tag{5}$$

To relate the beam waist in focus σ_0 and the scale parameter a to the magnification, 6 magnifications i.e. $12.5\times$, $15.0\times$, $17.5\times$, $20.0\times$, $22.5\times$, $25.0\times$, are configured to acquire the images of the same sphere. The magnifications are obtained through zooming. The magnification of $12.5\times$ approximately corresponds to the minimum magnification that renders the sphere visible in our experiment, while $25.0\times$ approximates to the maximum magnification that confirms that a full revolution of the sphere remains in the field of view (FoV). The PSF of each magnification is modelled by creating an optimisation problem and solving it with least square curve fitting. The overall fitting error of the 6 experimental PSFs is 5.00%. The experimental PSFs acquired from images with magnification of $12.5\times$ and $25.0\times$ are shown in Figure 3.3 (a) and (d) respectively. The color of the voxel indicates the intensity of PSF response. (b) and (e) represent the modelled PSFs of the two magnifications. Voxels in 3D space are converted to a 1D space with horizontal axis approximating the optical axis and vertical axis displaying the intensity. The 3D voxels on the slice in (a) and (b) match the 1D points in the box in (c) according to the same color. The experimental PSF differentiation between two magnifications is evident in (a) and (d). By transforming the 3D space to 1D functional, we can intuitively visualize and understand the distribution of the experimental PSF (blue dots) and the modelled PSF (red dots), as well as showing the differences between them.

With the proposed modelling approach on our data, the parameters ρ_1, ρ_2 and ρ_3 have proven to be constant regardless of magnification: $\rho_1 \approx 0.0041$, $\rho_2 \approx 1.0549$ and $\rho_3 \approx 2.9 \times 10^{-5}$. The beam waist σ_0 and parameter a related to the magnification range are estimated as depicted in Figure 3.4. We imply to the model that the fitting errors on the observed data are minimal. We, therefore, employ exponential and a quadratic function respectively. As shown in Eq. (6) and Eq. (7), x represents magnification and p_1 to p_5 are the model parameters.

$$\sigma_0 = p_1 \cdot e^{p_2 x} \tag{6}$$
$$a = p_3 x^2 + p_4 x + p_5 \tag{7}$$

The PSF of any 3D image between $-\infty$ and $+\infty$ along optical axis can be modelled as depicted in this section. The modelling is implemented with the focal plane set at the CoR. However, we acknowledge that in most circumstances of imaging acquisition the focal plane is not in line with the CoR, i.e. with a shift η. Consequently, the modelled PSF will be shifted along the optical axis by η from the focal plane to meet the imaging setup. Besides, the length of the PSF along optical axis is determined by the size of 3D image and the resolution r, because in 3D reconstruction each voxel in the 3D image

corresponds to each pixel in the 2D images. The NA is the effective value achieved from interpolation relating to the magnification. The relationship between effective NA and magnification is determined by the Leica objective lens.
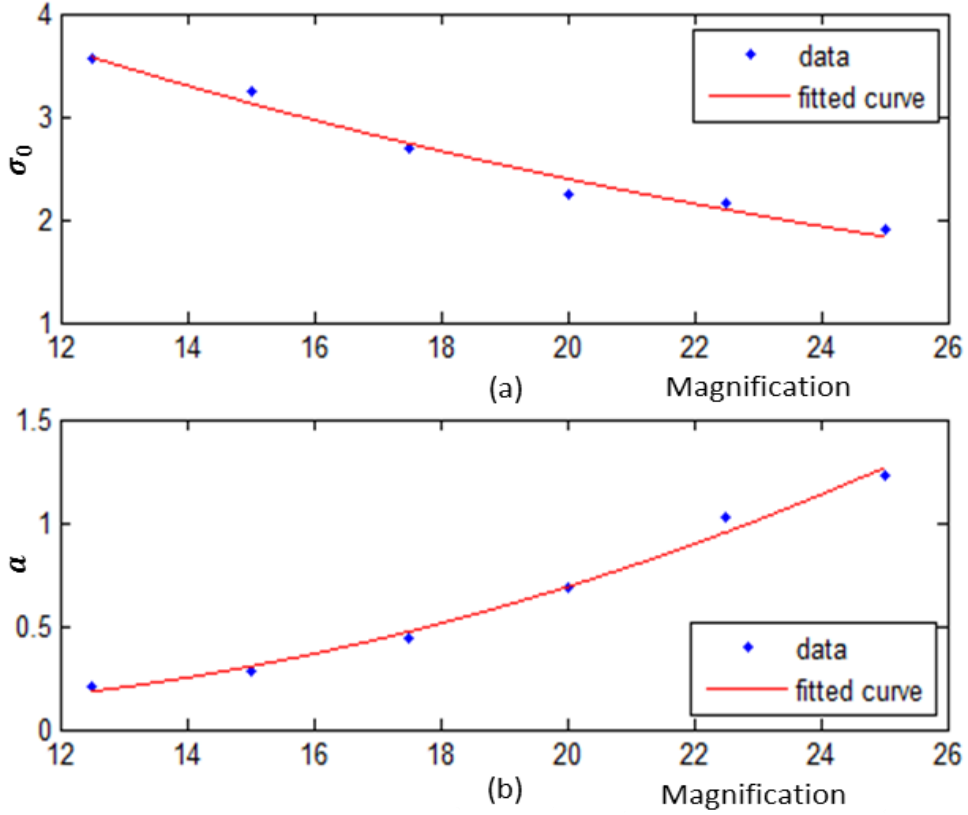
Figure 3.4. Fitting of the parameter σ_0 and a estimated from 6 magnifications. σ_0 can be fitted best by an exponential function as shown in (a) while a can be fitted best by a quadratic function in (b).

3.2.3 Deconvolution of 3D images in coronal plane

The modelled PSF consists of multiple 2D Gaussian patterns along optical axis. Therefore, the 3D image can be deconvolved slice by slice along its depth axis that is parallel to the optical axis. As the slices are coronal sections, the deconvolution is implemented on the 3D image R in the coronal plane as follows:

$$D_{(x,y,d)} = R_{(x,y,d)} */* p_{(s,t,d)} \tag{8}$$

R is the reconstructed 3D image with the depth axis d parallel to the optical axis of the PSF. $*/*$ stands for the operation of deconvolution. Considering the shifted focal plane and the reconstruction symmetry, deconvolution of R', the opposite view of R projected along d, is executed by applying:

$$D'_{(x,y,d)} = R'_{(x,y,d)} */* p_{(s,t,d)} \tag{9}$$

The transform from R to R' is conducted by a matrix rotation of π centred at the CoR. The 3D image with the deconvolution is then achieved by combining D and π back rotation of D'.

3.3 Experiments

The deconvolution performances of the PSF model on 3D OPT images are presented in three different ways. We first give a qualitative comparison between image slices with and without the proposed deconvolution approach. To further quantify the performance differences, we use three image blur metrics to measure the image blur and calculate the performance based on these metrics, in a slice manner. Finally, we present a holistic metric for 3D image deblurring improvement and implement it for performance comparisons. The samples used for the experiments are prepared according to our standard protocol, cf. § 1.2.4.

3.3.1 Image comparison of deconvolution

With respect to the magnifications, the experiments were conducted on images at 2 different magnifications. One is a zebra finch embryo in fluorescence mode with magnification $13.83 \times$ and focal plane shifted by $-0.93\ mm$. Taking the resolution limit and the 3D image size into consideration, the calculated defocus of the PSF along the optical axis ranges from $-6.303\ mm$ to $8.063\ mm$. The deconvolution is performed using the Lucy-Richardson algorithm [62] with a same number of iterations; here 10 is used based on the balance between reconstruction quality and computational time. The result for one coronal slice is shown in Figure 3.5 (c) and for the horizontal slice in Figure 3.5 (d). The corresponding slices prior to deconvolution are displayed in Figure 3.5 (a) and (b). The comparisons of intensity profile along a line with (red) and without (blue) deconvolution are presented in (e) and (f) respectively. The comparisons of the 3D visualization results are shown in Figure 3.6.

In Figure 3.7 another sample is depicted; a specimen of zebrafish larvae. The 3D visualizations of reconstructions with and without deconvolution are displayed in Figure 3.7. Figure 3.8 compares two orthogonal image samples of the zebrafish in detail. The magnification and the shifted focal plane are separately $49.98 \times$ and $-0.5\ mm$, with the computed defocus of the PSF being between $-2.242\ mm$ and $3.246\ mm$. Figure 3.8 (a) and (b) are the slices prior to deconvolution in two orthogonal planes, while (c) and (d) correspond to the deconvolution results. From visual assessment between (c) and (d), we can appreciate that the performance in the horizontal plane is almost as good as it is in coronal plane. This means that deconvolution in the coronal plane simultaneously improves the quality of the image in the horizontal plane to some extent. From a comparison of the quantitative intensity profile in the graph, we state that the proposed deconvolution sharpens and refines the 3D reconstructed images. It enhances the strong signals and makes the intensity profile more distinct.
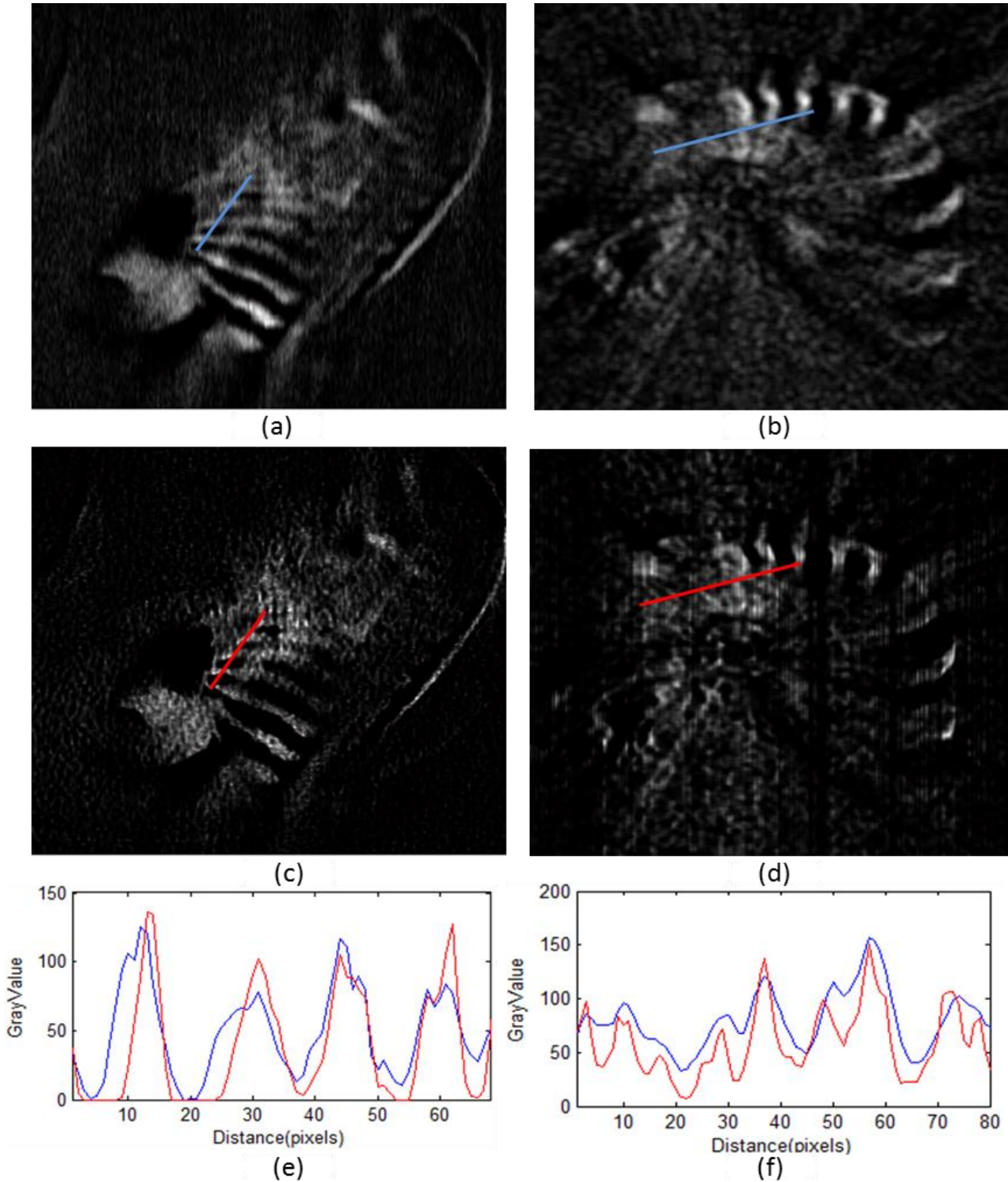
Figure 3.5. Deconvolution results. (a) The coronal slice of the 3D zebra finch with obvious blur around the ribs. (c) Distinct texture appears around the ribs after the deconvolution. (e) The comparison of intensity profiles along a line in (a) and (c). (b) and (d) The horizontal slice comparisons with the line intensity profiles shown in (f). In (c) and (d), more textures are observable in comparison with (a) and (b). In (e) and (f), the red thinner intensity profile, explains the more image sharpness along the red line in (c) and (d), comparing to (a) and (b) separately.

(a)   **(b)**
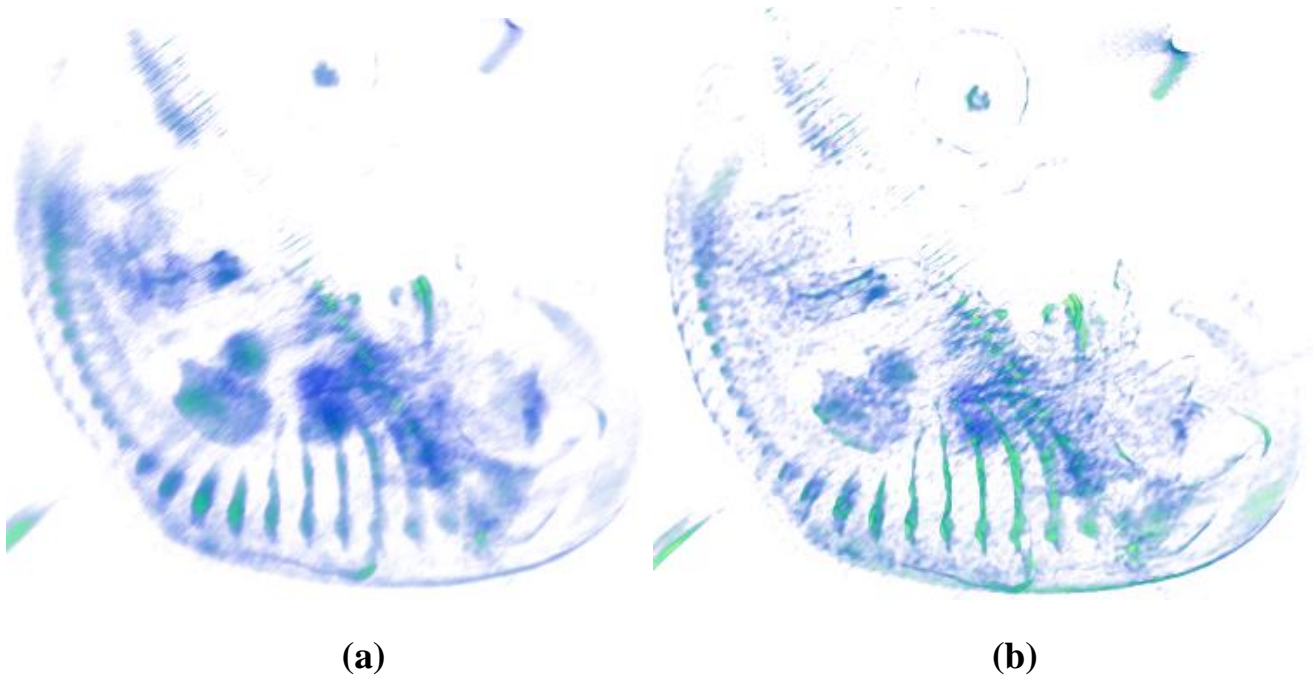
Figure 3.6. Deconvolution results. (a) 3D visualization of the zebra finch embryo without deconvolution. (b) 3D visualization of zebra finch with deconvolution. The visualization is made with Amira software [63]**.**

(a)

(b)
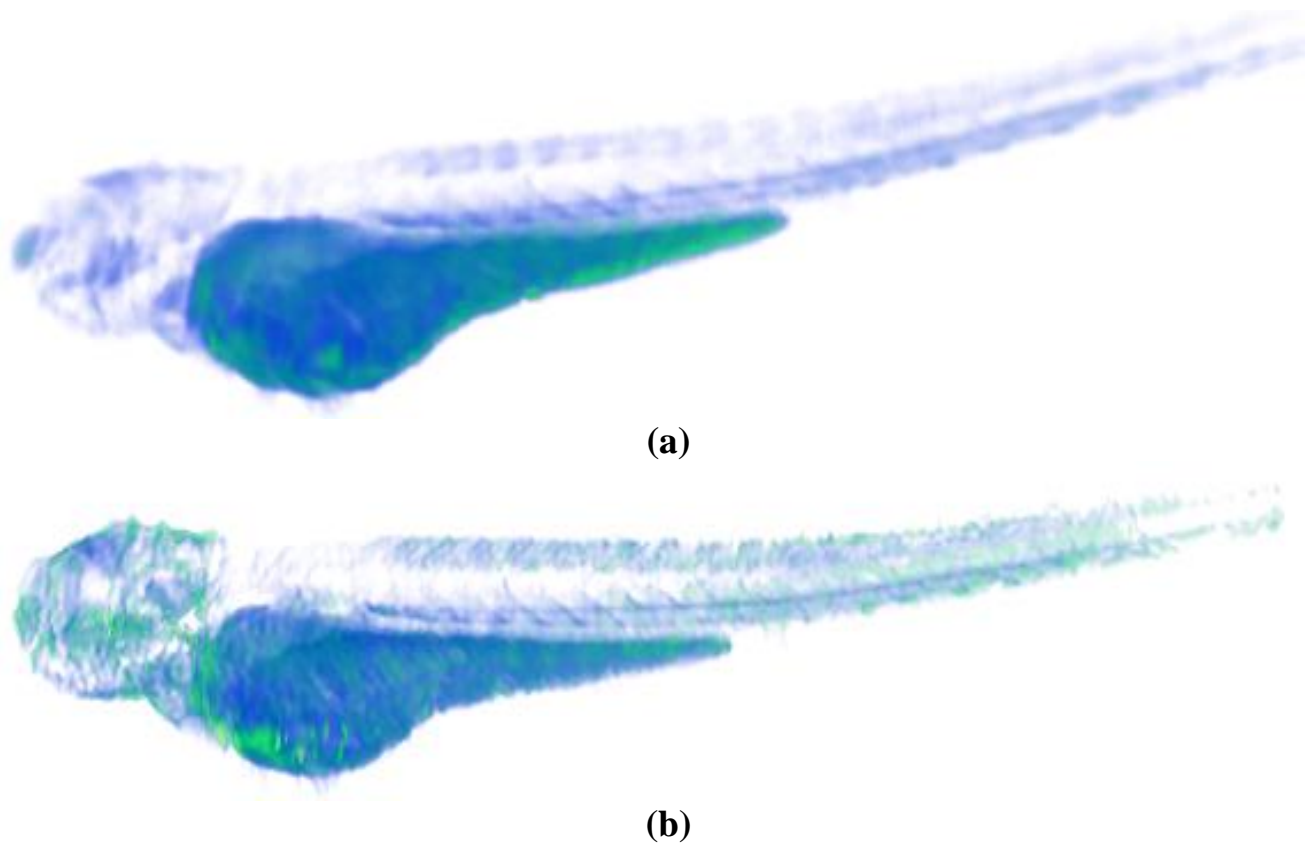
Figure 3.7. Deconvolution results. (a) 3D visualization of the zebrafish without deconvolution. (b) 3D visualization of zebrafish with deconvolution, visualized with Amira software [63]**.**
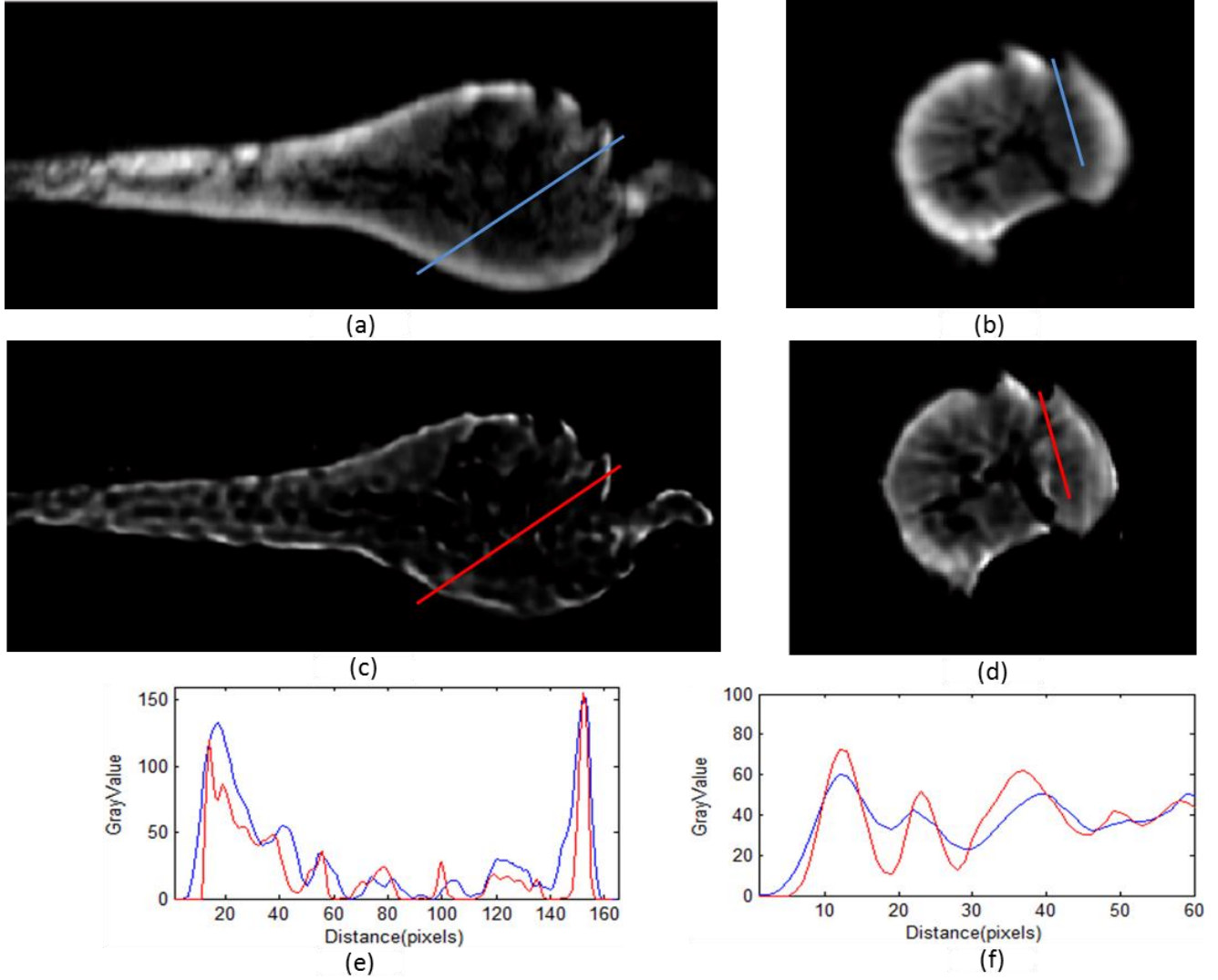
Figure 3.8. Coronal and horizontal slices of 3D zebrafish before ((a) and (b)) and after ((c) and (d)) deconvolution. The deconvolution highlights the strong signals and makes the texture more visible. (e) compares the intensity profile of the same line before (labelled as blue in (a)) and after (labelled as red in (c)) deconvolution, so does (f).

3.3.2 Image blur measurement on slices

To quantify the image blur of each slice, three metrics, known from the literature, were selected; i.e. the just noticeable blur (JNB) [64], the cumulative probability of blur detection (CPBD) [65] and the frequency measure (FM) [66]. These three metrics are employed to evaluate the performance of our method. Both the JNB and CPBD measure represent a sharpness metric by detecting and quantifying the blur in the spatial domain. Different from JNB and CPBD, the FM measure quantifies the sharpness in the frequency domain with an easier and more efficient approach. All the three metrics characterize the sharpness of an image, so the measure increases at improved image quality.
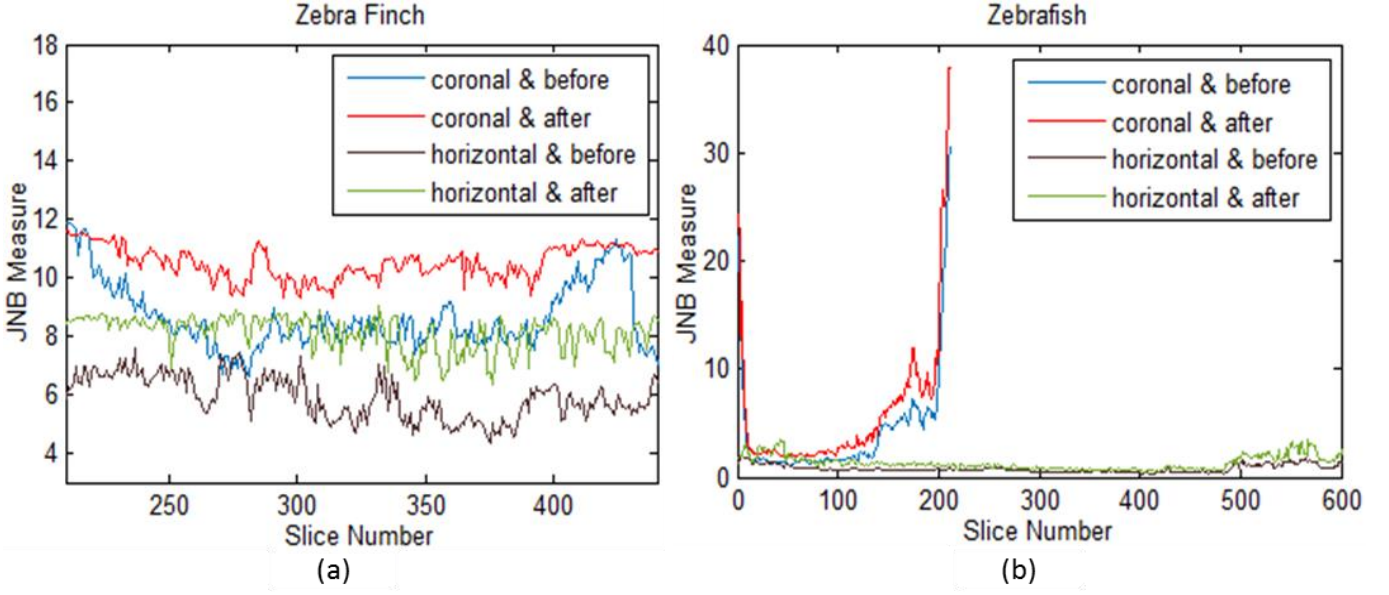
Figure 3.9. (a) JNB measure on the zebra finch data with magnification 13.83 ×. (b) JNB Measure on the zebrafish data with magnification 49.98 ×. Coronal and horizontal are the two orthogonal planes displaying the 3D image.
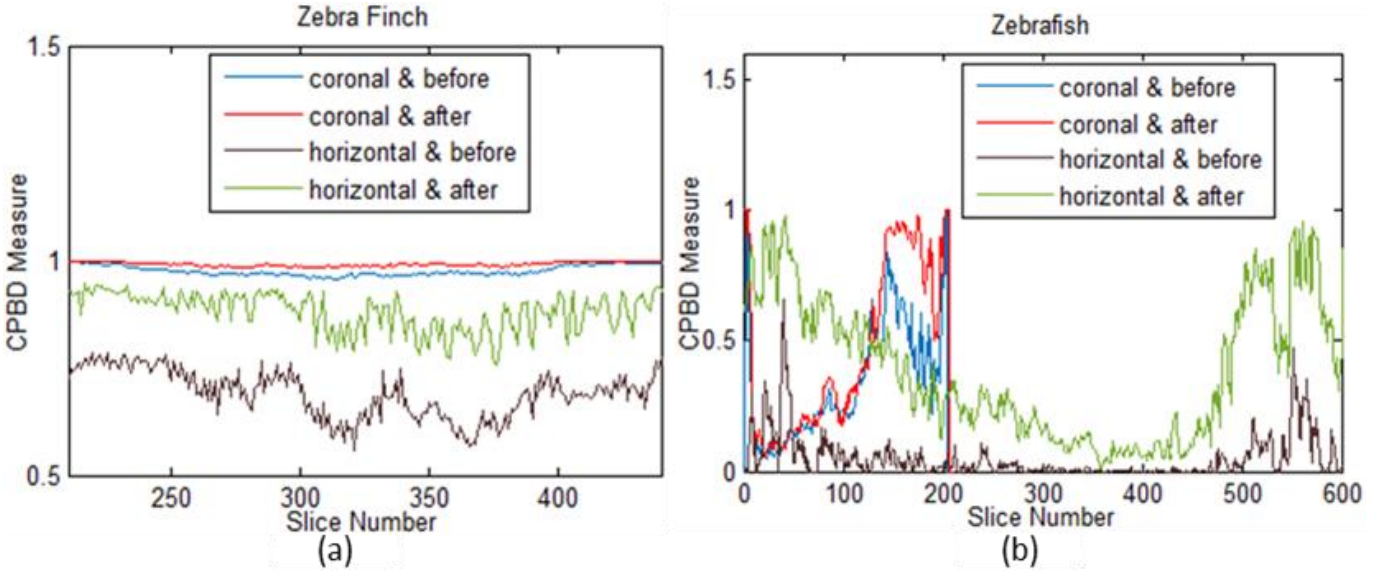
Figure 3.10. (a) CPBD measure on the zebra finch with magnification 13.83 ×. (b) CPBD Measure on the zebrafish with magnification 49.98 ×.

Whilst experiments in section 3.3.1 give us a qualitative comparison between the deconvolved slices and non-deconvolved slices, in this section we quantitatively look into all the slices in different orthogonal planes (coronal and horizontal) with the three image sharpness metrics (i.e. JNB, CPBD and FM). In the graphs depicted in Figure 3.9 - 3.11 we can observe that with all the three metrics the deconvolved slices on both planes show higher measurement values compared to the image slices without deconvolution. This means that for all the slices, regardless of orientation, the deconvolution holistically deblurs the images and thereby significantly improves the image quality.
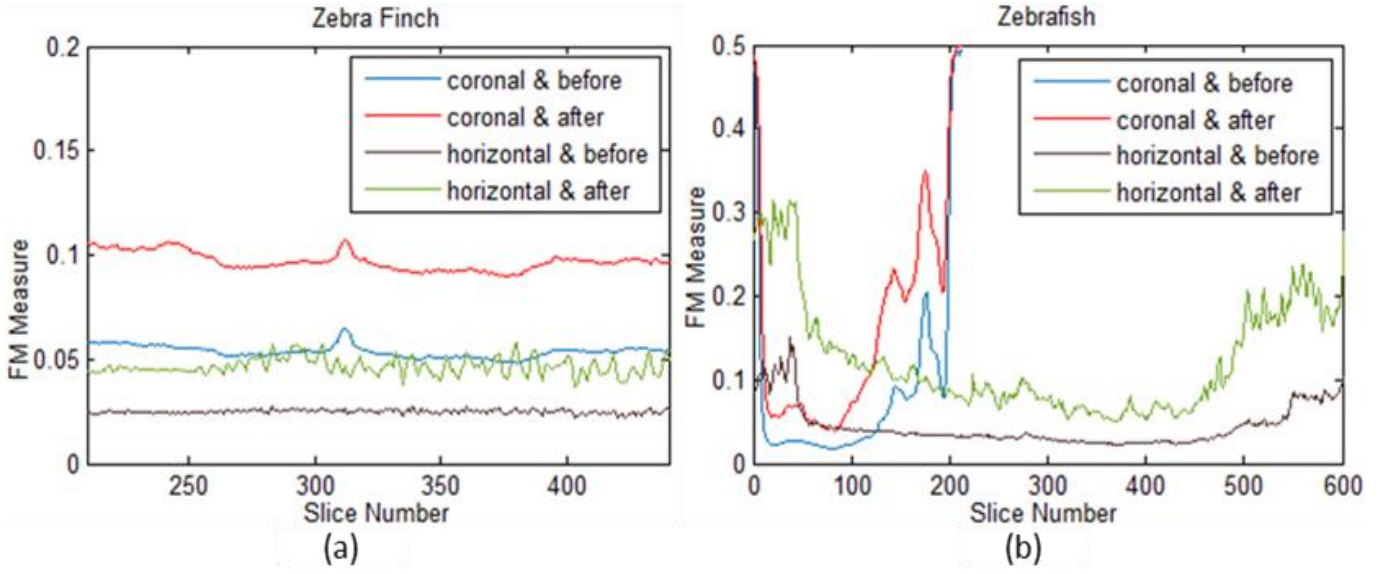
Figure 3.11. The FM before and after deconvolution on the two 3D image data. (a) FM measure on the zebra finch with magnification 13.83 ×. (b) FM Measure on the zebrafish with magnification 49.98 ×.

3.3.3 Quantitative 3D image quality improvement of deblur

To further quantify the deblur of the deconvolution results on the original reconstructed 3D data across the planes, we present the 3D image quality improvement criterion of deblur as I_{3d} in Eq. (10). Improvement in three orthogonal individuals are combined and encoded as a whole and each of them are represented as:

$$I_{3d} = \sqrt{(\frac{1}{Nx}\sum_{ix=1}^{Nx} I_{ix})^2 + (\frac{1}{Ny}\sum_{iy=1}^{Ny} I_{iy})^2 + (\frac{1}{Nz}\sum_{iz=1}^{Nz} I_{iz})^2} \qquad (10)$$

$$I_{ix} = \frac{M_{ix}^d - M_{ix}^r}{M_{ix}^r}, \qquad (11)$$

$$I_{iy} = \frac{M_{iy}^d - M_{iy}^r}{M_{iy}^r}, \qquad (12)$$

$$I_{iz} = \frac{M_{iz}^d - M_{iz}^r}{M_{iz}^r}, \qquad (13)$$

Where i is the slice number and x, y and z are the coordinate axes in 3D space. I_{ix}, I_{iy} and I_{iz} indicate the deblurring performance of slice i on the three different axises. M_{ix}^d and M_{ix}^r are respectively the ith deconvolved and original reconstructed slice on axis x. M_{iy}^d, M_{iy}^r, M_{iz}^d and M_{iz}^r are defined in a similar way on the two different axises. By employing the deblur performance I_{3d}, deconvolution performance of two different methods on the same data should be comparable.

Next, we apply our deconvolution method to 23 more 3D data sets, which contains 3 categories of samples i.e. zebrafish larvae, the adult zebrafish brain and the chicken embryo heart. They are in different stages of development and are acquired at different magnifications. It is important to realize that the metrics in this chapter cannot assess the

image blur across different data, but using the same data they are able to evaluate the performance of different deblurring approaches. Taking advantage of this, we compare the presented deconvolution method with the most commonly used Gaussian-based blind deconvolution [67]. According to our observation of the results for blind deconvolution, the kernel size does not make a visible difference on our 3D images. Therefore, we present our results for blind deconvolution with the kernel size set as 7, taking acceptable computation time into account. From the 3 metrics, i.e. CPBD, JNB and FM, we have selected the most robust metric JNB to measure the image blur of each slice. The results of the 3 categories of samples are presented in Table 3-1 to Table 3-3. For all the 23 data, our deconvolution approach outperforms the Gaussian-based deconvolution, thereby indicating the success of the method.

Table 3.1. 3D image quality improvement of 10 zebrafish embryos based on JNB Measure.

	01	02	03	04	05	06	07	08	09	10
$\mathcal{G}$	0.16	0.21	0.15	0.21	0.29	0.24	0.22	0.16	0.20	0.25
PSF_m	**1.35**	**2.10**	**1.41**	**3.23**	**1.54**	**1.50**	**1.50**	**1.37**	**1.45**	**1.70**

$\star$ $\mathcal{G}$ -- Gaussian-based blind deconvolution. PSF_m -- PSF based modelling deconvolution. 10 zebrafish embryos correspond to 01-10 with age from 3 *dpf* to 7 *dpf*.

Table 3.2. 3D image quality improvement of 6 zebrafish brain based on JNB Measure.

	01	02	03	04	05	06
$\mathcal{G}$	0.25	0.29	0.26	0.24	0.01	0.21
PSF_m	**0.41**	**0.49**	**2.55**	**0.17**	**1.15**	**1.20**

$\star$ 6 adult zebrafish brains correspond to 01-06 with different magnifications.

Table 3.3. 3D image quality improvement of 7 chicken heart based on JNB Measure.

	01	02	03	04	05	06	07
$\mathcal{G}$	0.24	0.16	0.19	0.26	0.27	0.18	0.32
PSF_m	**0.93**	**0.49**	**0.29**	**1.15**	**0.49**	**0.38**	**1.07**

$\star$ 7 chicken embryo hearts at different stages correspond to 01-07.

3.4 Conclusions

In this chapter we have focused on 3D image deblur and quality improvement, under the condition of the limitation of small NA for imaging of large sized samples. We investigated and modeled the PSF along the optical axis, exploring the influence of magnification on PSF. The sample of a single fluorescence sphere is prepared with the protocol in section 3.2.1. The experimental PSF is then modelled to deconvolve the 3D image in a coronal plane. A number of measures for image blur are employed to convincingly evaluate the performance of the deconvolution. They provide quantitative information about how much improvement is achieved. The overall improvement I_{3d}

gives us a criterion to compare image quality improvement regardless of different data. All the experimental results including the image comparisons and quantitative measures sustain the effectiveness of the proposed PSF modelling and deconvolution methodology.

The deconvolution results presented represent a proof of concept. The datasets used in the experiments are composed of 25 samples, i.e. 4 categories: zebrafish embryo, zebra finch embryo, adult zebrafish brain and chicken embryo heart. Regarding the evaluation of performance on a large volume of different datasets, our data are far from perfect in terms of 'large dataset'. However, it presents a clear idea that our model is not constrained by samples of one particular type, it also works on many other types of specimens. This will help to explain its potential capability of improving image quality with similar performance on more 3D data, including those from other OPT imaging systems, which is a part of our current work. In the future we will take further efforts on generalizing the model to other imaging set-ups. In addition, the fluorescent sphere used in the experiments is fixed-size. The effect of sphere size on PSF modelling and deblur performance need to be given more attentions.

3.5 Acknowledgement

Chapter 4

Segmentation-driven Optimisation for Iterative Reconstruction in Optical Projection Tomography: An Exploration

This chapter is based on the following publication:

Chapter summary

The reconstruction of a tomogram to a 3D image has some drawbacks in that there are artefacts introduced in the reconstruction that affect the quality of the image. We have been using the filtered back projection algorithm for the reconstruction; but iterative reconstruction algorithms demonstrate a superior performance as far as artefacts are concerned. In computerized tomography these iterative algorithms are successfully applied. These iterative algorithms will, however, require much more computation time.

In this chapter we study capability of iterative algorithms to remove streak artefacts from reconstructions of optical projection tomograms. Moreover, we explore possible ways to optimize the most customary parameters of the iterative reconstruction algorithms so as to improve its reconstruction performance. Due to the lack of benchmarks for direct reconstruction evaluation in optical projection tomography we consider the assessment according to the performance of segmentation in the reconstruction. We use the zebrafish model, the model system for which our OPT system is used a lot, as we can easily obtain data and build a benchmark. For the segmentation approach we employed the 2D U-net convolutional neural network as it is known for good performance in biomedical image segmentation.

4.1 Introduction

Three dimensional (3D) image reconstruction in OPT can play a crucial role in giving insight into protein distribution and/or gene expression within a research model, e.g. zebrafish, at a tissue and organ level. Given the specific data, a good reconstruction algorithm typically produces a reliable and effective 3D reconstruction, whereas a simple reconstruction technique may introduce computational artefacts that hamper the interpretation of the data. These kinds of artefacts are introduced during the reconstruction process because of imperfection of the data prior to the imaging process. There are two categories of approaches to reduce or eliminate these artefacts. One is considering the perspective of the imaging process, meaning trying to avoid the imaging imperfections that results in artefacts, either from the sample side or from the instrument side. This approach is, however, sometimes quite expensive or even unachievable in some cases. Another category is to computationally improve the reconstruction from an imperfect tomogram. It can be either applying a powerful reconstruction algorithm or employing pre-processing and/or post-processing of a specific algorithm.

In 2005, Walls *et al.* [22] first presented the possible artefacts in the Filtered Back Projection (FBP) reconstruction existing in an OPT imaging system. The main contribution of this work lies in the study of the origins of the reconstruction artefacts from the imaging source, for instance signal decay, CCD imperfection, etc. He studied the reasons why these imperfections in imaging would result in reconstruction artefacts within the FBP [68] framework, but not yet explain if more advanced reconstruction algorithms, such as iterative reconstruction, can reduce or even eliminate these artefacts.

In addition to the work of Walls *et al.*, we explain the so called streak artefacts in the FBP reconstruction and briefly explain the imaging source reason. OPT imaging is characterized by its wide depth of field (DoF) compared to high-resolution microscopes, e.g. confocal microscopy [69], [70]. This means that a point source that is properly focused in one angle of rotation may be blurred or even invisible in its opposite angle. This will result in streak artefacts in the FBP reconstruction because of the severe asymmetry of tomograms.

In CT imaging, the so called metal artefacts can be either reduced with artefacts reduction algorithm within the FBP framework or eliminated by using other reconstruction methods such as iterative reconstruction. With respect to reduction of artefacts, in the last two decades a variety of approaches have been proposed for CT [71]–[75]. These approaches, however, can only decrease the artefacts rather than eliminate them. As for streak artefacts, elimination using other reconstruction methods, the iterative reconstruction stands out. Inspired by the superior performance of metal artefacts elimination when using iterative reconstruction in CT [35], [76]–[78], we are interested in exploring its capability of streak artefacts elimination in OPT.

Iterative reconstruction refers to iterative algorithmic approaches used to reconstruct 2D or 3D images from tomographic imaging techniques. Generally, it starts with an

assumed image, computes projections from the image via a projection function and updates the image according to the difference between the calculated and the actual projections. According to the updating schema for image, it can be categorized into four kinds of approaches, i.e. algebraic reconstruction techniques (ART) [23], iterative sparse asymptotic minimum variance (SAMV) [24], statistical reconstruction [25] and learned iterative reconstruction [26], [79]. They are considered superior when there is a lack of uniform projections or when the projections are sparse, which to some extent fits the character of imaging source for the aforementioned streak artefacts in OPT. Compared to its application in CT, iterative reconstruction was less studied in OPT. Correia *et. al* [12] applied the iterative reconstruction to achieve relatively reasonable results on a sparse collection of projections in 2015. One possible reason for impeding prevalence of iterative reconstruction in OPT is the high computation time. Nevertheless, with the rapid development of computational strategies, e.g. parallel computing and GPU, iterative reconstruction will be more widely used in OPT.

The common approach for evaluating reconstruction performance in tomographic imaging is accomplished by producing and projecting a simulated object, e.g. phantom in CT, and assessing the performance of a reconstruction algorithm based on the projections [80]. Nevertheless, due to the DoF in OPT imaging the projection function is much more complicated than that in CT. In this case, simulating the projections of a phantom as benchmark has, to our best knowledge, never happened in OPT. Therefore, as an alternative we use images from real-life samples for experiments.

Iterative reconstruction has great potential for image reconstruction in OPT. However, there is an intrinsic disadvantage for image reconstruction in OPT in that a good quantitative evaluation on the reconstruction is difficult. This disadvantage mainly results from the lack of benchmarks for samples. In such a situation formulating an alternative evaluation method according to the specific research problem is considered applicable and feasible. With this idea we transfer the evaluation of reconstruction to that of segmentation which can be easily obtained given the reconstruction data; in our experiments we use zebrafish. The inspiration of the transfer originates from the fact that we are expecting better segmentation results from different reconstructions on the same data. Therefore, this transfer can be valid under the assumption that the reconstruction with better segmentation result is preferred, given effect of the same segmentation algorithm. By transferring the evaluation of reconstruction to segmentation, we can approximate the best parameters for reconstruction. Thus, our approach can be interpreted as an optimisation of iterative reconstruction for OPT.

In this chapter we focus on iterative reconstruction in OPT so as to avoid streak artefacts, and furthermore we explore the possibility of segmentation evaluation to optimize the parameters for iterative reconstruction. In section 4.2 the implementation of iterative reconstruction will be briefly introduced. We will focus on the parameter optimisation in section 4.3. The specific scheme and approach for parameter optimisation will be exhaustively discussed there, which will be followed by the experimental results

and discussion. Finally we present our conclusions, as well as raise some limitations and future work.

4.2 Iterative reconstruction for OPT

In general, iterative reconstruction can lead to a more accurate OPT 3D image than that obtained by the FBP [68]. However, a large number of iterations may be required to generate an acceptable result, with each iteration taking about the same amount of time as the FBP. Thus to some extent the effectiveness of iterative reconstruction is achieved at the expense of more computation time. One approach to reduce the number of iterations is to organize the projection data into a series of ordered subsets of evenly spaced projections and update the current estimate of the object after each subset rather than after the complete set of projections. The most commonly used algorithm employing subset is referred to as ordered subset expectation maximization (OSEM) reconstruction [28], [81]. It improves the efficiency of iterative reconstruction with respect to computational time.

According to the expectation maximization (EM) algorithm in iterative reconstruction [82], the intensity of an object projected to the detector follows the Poisson distribution with expected value $\mu = E(I_{(z,\beta,\theta)}) = P_{(z,\beta,\theta)} \times R_{(x,y,z)})$. The object to be reconstructed is assumed as $R_{(x,y,z)}$ and will be updated in an iterative way as follows:

$$R_{(x,y,z)}^{l+1} = R_{(x,y,z)}^{l} \frac{\sum_{\theta_t \in I_t}(I_{(z,\beta,\theta_t)}/(P_{(z,\beta,\theta_t)} \times R_{(x,y,z)}^{l}))}{\sum_{\theta_t \in I_t} P_{(z,\beta,\theta_t)}} \quad (1)$$

with P being the projection function while $I_{(z,\beta,\theta)}$ symbolizes the tomogram at the angle θ in OPT. (z,β,θ) and (x,y,z) are separately the tomogram and 3D image coordinate, with β being the detector axis and θ being the projection angle; whilst z is the slice number and (x,y) indicates the image size of a reconstructed slice. The EM based iterative reconstruction is based on the idea that the reconstruction $R_{(x,y,z)}$ of an object in 3D space can be estimated with the observed or measured data $I_{(z,\beta,\theta)}$ by iteratively updating $R_{(x,y,z)}$ with the EM algorithm, with l being the iteration step. The conventional EM iterative reconstruction updates $R_{(x,y,z)}$ based on a full set of observation I, while the OSEM splits the full set into T ordered subsets $I = \{I_t: t = 1,2,\dots,T\}$, $I_t = I_{(z,\beta,\theta_t)}$ and implements the updating based on I_t in each iteration step [82].

In the OPT imaging system, the light is approximately considered as parallel rather than conical. In this case the projection function $P_{(z,\beta,\theta_t)}$ remains the same at different z positions or on different reconstructed slices. Then the update can be implemented on a slice by slice basis, with each slice updated using:

$$R_{(x,y)}^{l+1} = R_{(x,y)}^{l} \frac{\sum_{\theta_t \in I_t}(I_{(\beta,\theta_t)}/(P_{(\beta,\theta_t)} \times R_{(x,y)}^{l}))}{\sum_{\theta_t \in I_t} P_{(\beta,\theta_t)}} \quad (2)$$
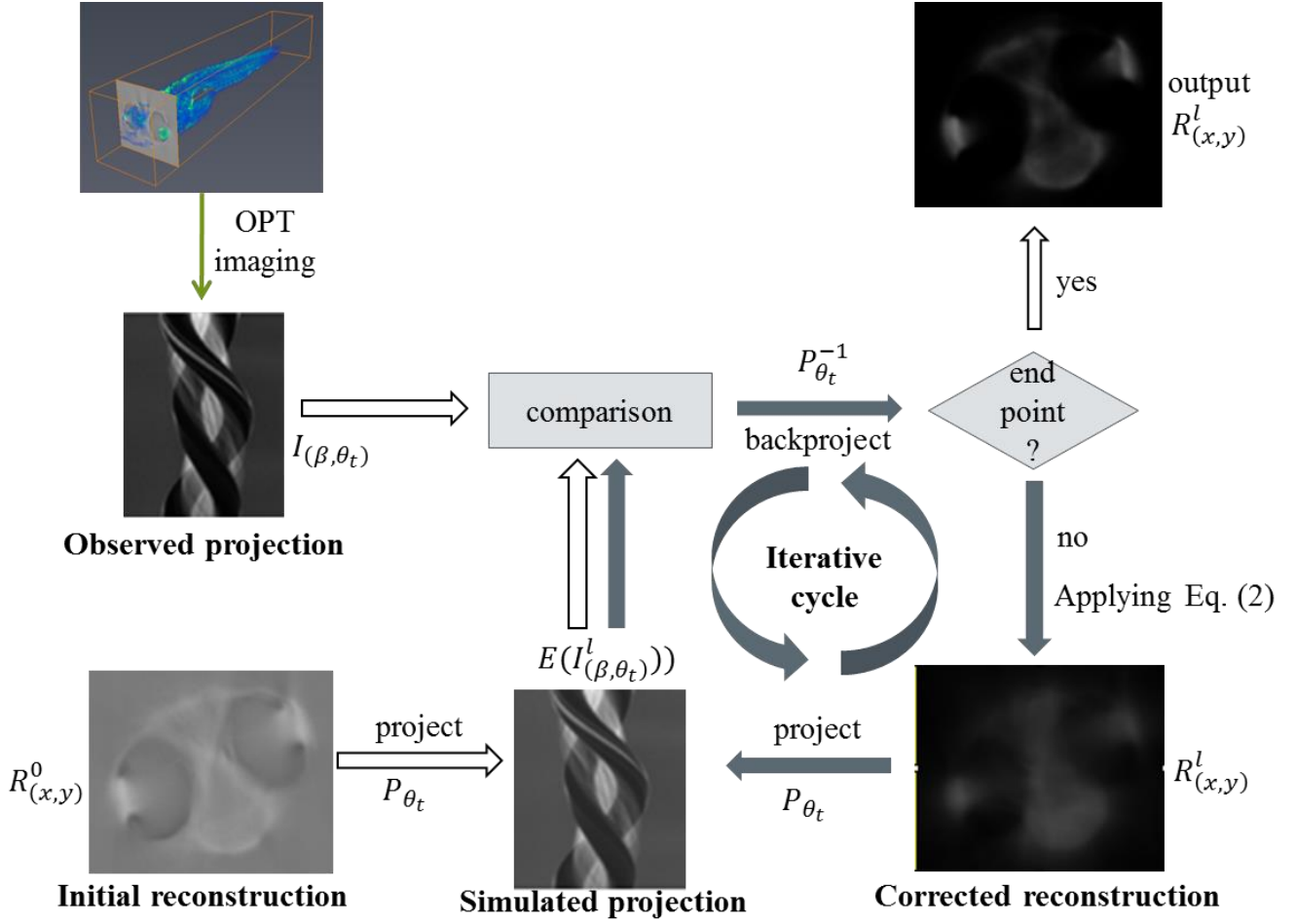
Figure 4.1. Workflow of iterative reconstruction in OPT. Example is given on a zebrafish larvae of 6 days post fertilization (*dpf*).

In Figure 4.1 the workflow of iterative reconstruction is depicted as applied on a zebrafish sample on a slice-by-slice basis. The observed projections, are acquired from the OPT imaging system and are here defined as the sinogram $I_{(\beta,\theta_t)}$. By using the OSEM algorithm, the reconstructed slice $R^l_{(x,y)}$ is updated based on the differences between the observed projection and the simulated projection in reconstruction space using Eq. (2), with a initial reconstruction slice $R^0_{(x,y)}$. The simulated projection is updated as $E(I^l_{(\beta,\theta_t)}) = (P_{\theta_t} \times R^l_{(x,y)})$ and $E(I^0_{(\beta,\theta_t)}) = (P_{\theta_t} \times R^0_{(x,y)})$ for the first update. This difference $I_{(\beta,\theta_t)}/E(I^l_{(\beta,\theta_t)})$ in projection space is reconstructed to 3D image space using back projection $P^{-1}_{\theta_t}$, which is finally used for updating the slice $R^l_{(x,y)}$ as shown in Eq. (2). As we can see from Figure 4.1, in the iterative framework both the end point and the initial reconstruction $R^0_{(x,y)}$ play an important role in the final results, which will be studied in the following sections.

4.3 Parameter optimisation for iterative reconstruction

Parameters for iterative reconstruction can be optimized based on the performance assessed on reconstructed results. Nevertheless, due to the lack of benchmark for direct reconstruction evaluation in optical projection tomography we consider the assessment according to the segmentation performance of the 3D images reconstructed with different parameters. To this end, the framework for this idea will be first elaborated. It is proposed under the assumption that the segmentation method used is reliable and effective. In order to guarantee this, we employ a convolutional neural network (CNN) as it is known for its high performance in image segmentation. In Figure 4.2, an example slice (head part) of 25 *dpf* zebrafish is given, explaining how different reconstruction parameters influence the segmentation performance. Figure 4.2 (a) and (b) are two slices reconstructed from the iterative reconstruction using different parameters. Figure 4.2 (c) and (d) show the corresponding segmentation results based on the reconstructed slices from Figure 4.2 (a) and (b). The benchmark of this slice is displayed in (e), which is obtained from the manual segmentation. When visually comparing the reconstructed slices and the segmentation performance, we observe that different reconstruction parameters could produce different 3D image slices, which will consequently result in distinct segmentation results. For instance the slice in (b) has better segmentation performance shown in (d), comparing to the performance (c) from the slice in (a). With this assumption, we can optimize the parameters based on the segmentation performance of the 3D image slices that are reconstructed with different parameters.
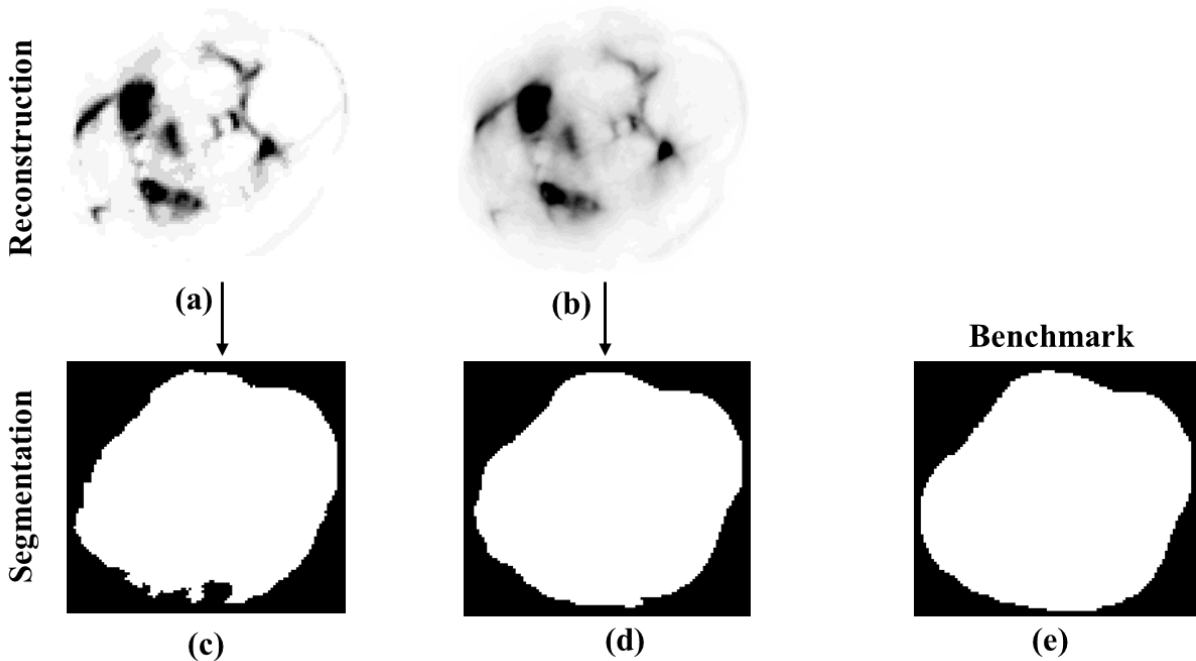
Figure 4.2. An example showing the effect of reconstruction parameters on segmentation performance. (a) and (b) The slices reconstructed from two different parameters. (c) and (d) The corresponding segmentation results. (e) The benchmark of segmentation for comparison.

4.3.1 Framework of parameter optimisation for iterative reconstruction

The framework of parameter optimisation for iterative reconstruction integrates the reconstruction and segmentation process as a whole as shown in Figure 4.3. In OPT a tomogram is referred to as $I_{(z,\beta,\theta)}$, with z and β representing the pixel position of tomogram at the projection angle θ. The 3D image is obtained by implementing the iterative reconstruction algorithm $f(\alpha)$ on the tomograms as $R^{\alpha}_{(x,y,z)} = f(\alpha) I_{(z,\beta,\theta)}$ with α being the parameters required for iterative reconstruction. The reconstructed results can further be segmented according to a specific criterion given the data, achieving the segmentation result $S^{\alpha}_{(x,y,z)}$ depicted in Figure 4.3, where $S^{\alpha}_{(x,y,z)} = g(R^{\alpha}_{(x,y,z)})R^{\alpha}_{(x,y,z)}$. Therefore, the segmentation result can be formulated as:

$$S^{\alpha}_{(x,y,z)} = f(\alpha)I_{(z,\beta,\theta)} \, g\left[f(\alpha) I_{(x_\theta,y_\theta,\theta)}\right] \tag{3}$$
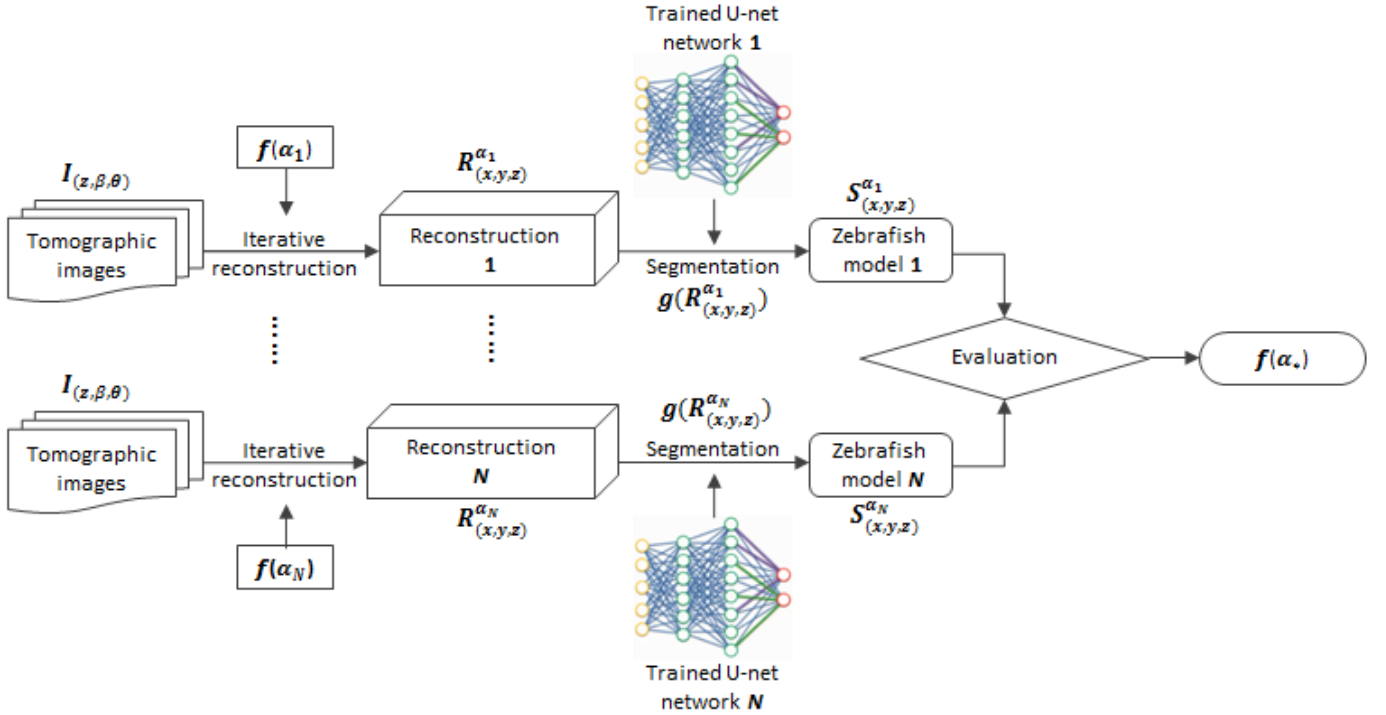
Figure 4.3. The framework for parameter optimisation for iterative reconstruction based on the corresponding segmentation performance. The 2D U-net [29] convolutional neural network is applied to train the segmentation model within each parameter group.

with g indicating the global function term for segmentation. Taking its high performance in biomedical image segmentation into account, we use the 2D U-net [29] convolutional neural network for our segmentation work. This means that, given the same tomogram, training the segmentation network with the 3D image reconstructed from different α will produce different network outputs for segmentation. However these differences originally result from the reconstruction parameter α rather than the network itself given the same configuration. From this notion the optimisation of α can be approximately transferred to

a search for a better α that produces 3D image for a better segmentation, with the idea that the benchmark of the segmentation is known; e.g. the zebrafish model in our work.

The diagram of the pipline for the optimisation is depicted in Figure 4.3. For the same tomogram set, it is assumed that there are N 3D image groups $\{R^{\alpha_1}_{(x,y,z)},\ldots, R^{\alpha_N}_{(x,y,z)}\}$ produced from iterative reconstruction functions $\{ f(\alpha_1),\ldots, f(\alpha_N)\}$ with $\{\alpha_1,\ldots, \alpha_N\}$ representing the variation of parameters. By evaluating the N 3D image groups based on their segmentation performances, we can achieve the best reconstruction parameter ranging from α_1 to α_N. If there are K samples for reconstruction and segmentation, then we have $R^{\alpha}_{(x,y,z)} = \{R^{\alpha}_{1(x,y,z)}, \ldots, R^{\alpha}_{k(x,y,z)}, \ldots, R^{\alpha}_{K(x,y,z)}$ and $S^{\alpha}_{(x,y,z)} = \{S^{\alpha}_{1(x,y,z)}, \ldots, S^{\alpha}_{k(x,y,z)}, \ldots, S^{\alpha}_{K(x,y,z)}\}$. It is worth mentioning that the segmentation network used in each 3D image group is internally trained rather than training it in a global scale across groups. This is because it is considered from an experimental perspective, more practical and valuable to train a segmentation network based on the data from the same reconstruction method, excluding the network preference when training it on the data across reconstruction approaches. With respect to segmentation network we employ the U-net CNN [29] because of its high performance in bio-medical image segmentation, generating a segmentation network group $\{ g(R^{\alpha_1}_{(x,y,z)}),\ldots, g(R^{\alpha_n}_{(x,y,z)}) \ldots, g(R^{\alpha_N}_{(x,y,z)} \}$. With the internal-trained segmenation network, evaluation will be performed to choose the group which has the highest performance, simultaneously finding the optimal iterative reconstrucion parameter $f(\alpha_*)$. There are several parameters required in the OSEM iterative reconstruction algorithm, but we focus on the two most customary ones, i.e. iteration number and initial reconstruction, as η and γ respectively. The different combinations of η and γ comprise $\{ \alpha_1 ,\ldots, \alpha_n ,\ldots, \alpha_N \}$, meaning that if $\eta = \{\eta_1, \ldots, \eta_i, \ldots, \eta_p\}$ and $\gamma = \{\gamma_1, \ldots, \gamma_j, \ldots, \gamma_q\}$, then $\alpha_n = (\eta_i, \gamma_j)$ and $N = p \times q$.

4.3.2 Segmentation approach

In order to provide a reliable and effective segmentation approach for reconstructed slices, CNN is employed as a result of its recent promising performance in medical image segmentation. As for our data, the zebrafish samples are transparent. It is challenging for traditional segmentation approaches to segment the foreground from the background when their intensities are very similar. A CNN can learn the structural and context information at different scales of resolution and, in that, it differs from traditional segmentation methods. For this reason, a CNN is very suitable for our research questions. We define our segmentation task as a binary segmentation on transparent samples in intensity image space. This means a small network such as 2D U-net rather than a complex network, is more desirable for our problem.

1) Network structure

The 2D U-net segmentation network [29] feeds 2D images as input layer, i.e. the reconstructed slices from the 3D image. Overall, the network contains encoder, decoder and a *Merge* layer between them as shown in Figure 4.4. In both the encoder and the decoder, there are 4 *Convolutional* layers and within each individual layer two 3x3 *Conv-*
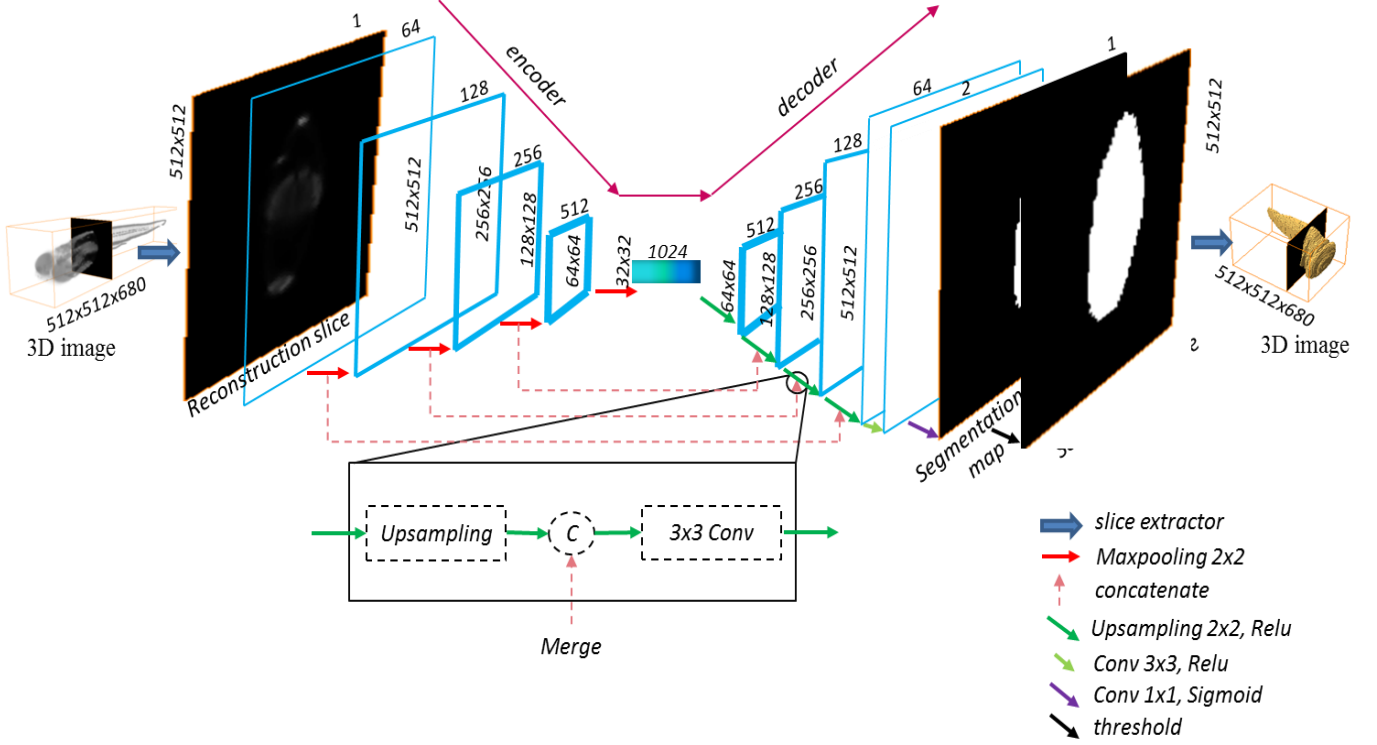
Figure 4.4. The structure of the 2D U-net network for zebrafish slices. Each slice is fed into the network as a single sample. The network contains both encoder (down layers) and decoder (up layers) for the deep CNN. The encoder is accomplished by a convolution and a max pooling operation while the decoder consists of a convolution and an upsampling operation. The *Maxpooling* and *Upsampling* layers are separately represented by red and green arrows. The *Convolution* layers are implemented with a 3x3 kernel at each depth of the encoder and decoder. Between similar layers on both sides there is a merge operation, allowing deep convolution layers to be merged with more shadow convolutional layers. *c* is the concatenation operation to merge layers from different depths.

olution layers and one *Dropout* layer are integrated. Different depths in the encoder and the decoder are separately connected by *Maxpooling* and *Upsampling* layers. The bridge-like *Merge* layer between the symmetric layers combines deep *convolution* layers and shadow ones. This typically improves the performance in segmentation problem. As the *Output* layer, the map activated by Sigmoid function [83], reflects the response to the zebrafish segmentation ranging from 0 to 1. The segmentation mask is generated when implementing a threshold on the map.

2) Network training

A. Training scheme

The segmentation network for zebrafish in each α_n-specified 3D image group is independently trained. The number of segmentation networks for the parameter optimisation framework is dependent on the parameter combination, as $N = p \times q$ for two parameters in our case. For a specific group $R_{(x,y,z)}^{\alpha_n}$, a certain ratio, typically small, of reconstructed slices from the 3D image are sampled at even intervals for training while the rest, a large ratio, are used for testing. The rationale for training on a small rather than large ratio of data is due to the fact of information redundancy among adjacent slices in 3D. Moreover, this can, to a large extent, reduce the workload of manual labelling. With a smaller ratio of data for training, the labelling for a single zebrafish could still be hundreds of slices. The image size of a typical OPT image is 512x512x1360. Here 1360 indicates the number of slices in the image of 512x512 per slice. To further decrease the labelling workload and make it more efficient, an interpolation approach is applied. The tool for interpolation labelling is available in software [63]. The results of interpolation labelling can be further verified manually and efficiently. The 3D images from the FBP reconstruction are used for labelling because of its capacity of offering comprehensive information.

Within each group the training slices and the corresponding labelled maps are used to train the segmentation network that is constructed referring to the work in [29]. Different from their network, the size of feature map after each convolutional layer maintains the same. The *Re*LU activation is used for all convolutional layers except for the last one which employs the sigmoid activation such that the output of the network ranges from 0 to 1. Another difference between the U-net network and ours is that we employ the simplest binary cross entropy as the loss function without considering the weight of each pixel in the image. The loss energy function would be:

$$E = \sum_{X \in \Omega} y(X)\log(\hat{y}(X)) + (1 - y(X))\log(1 - \hat{y}(X)) \tag{4}$$

where $y(X)$ is the labelled value at the pixel position $X \in \Omega$ with $\Omega \subset Z^2$, i.e. $y(X) \in \{0,1\}$. $\hat{y}(X)$ represents the predicted value at the same pixel position. The network is trained with the Adam optimizer [84] implementing Keras [85].

B. Learning rate

In the framework of stochastic gradient descent (SGD) [86],[87] optimisation technique for machine learning problem, the learning rate is considered to be carefully chosen to guarantee the convergence of the loss function. In order to set an effective learning rate to train the segmentation network, we first need to observe the intrinsic correlation between the data and the learning rate. Therefore, the different-fixed-learn-rate scheme is first applied to the data to investigate the different loss decay profiles. The profile of the loss function with an increasing epoch, using different fixed learning rates, is presented in
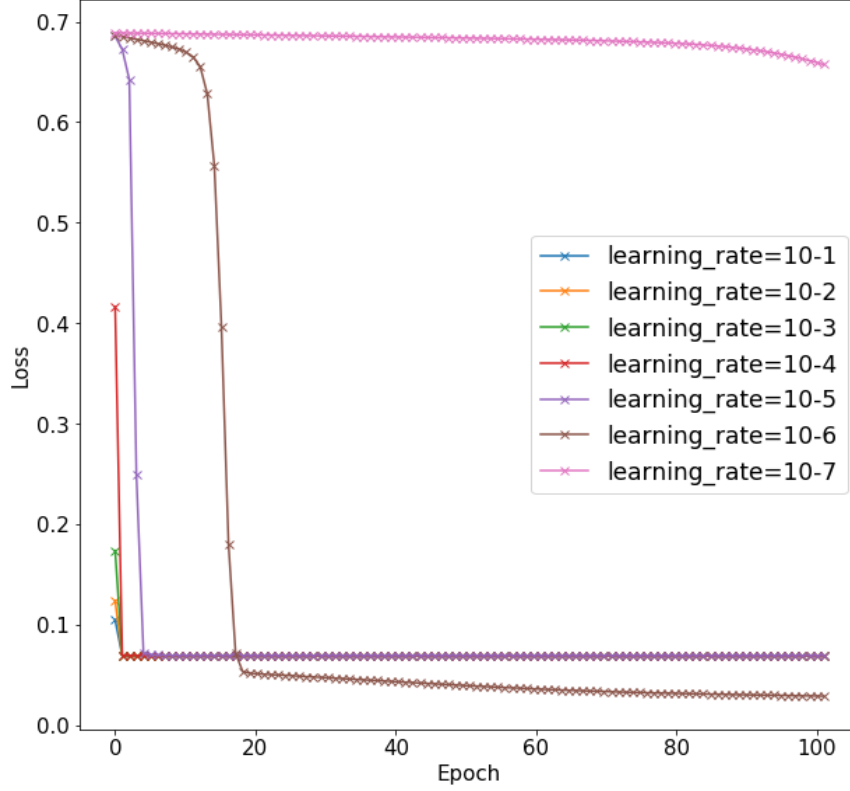
Figure 4.5. The profiles of the decay loss for different learning rates in 100 epochs. The learning rates observed are from 10^{-1} to 10^{-7}.

Figure 4.5. It easily and quickly falls into local minimum when the learning rate is set to or larger than 1e-5. To avoid this, the learning rate is supposed to be smaller than 1e-5. But when the learning rate is equal or smaller than 1e-7, the loss decreases extremely slow. 1e-6 is therefor considered to be a good initial learning rate for the data. However, as the epoch increases and the loss decreases, the loss decreases very slowly with the small learning rate 1e-6.

We observe a correlation between our data and learning rate. This can be seen in Figure 4.5, where loss curves for decreasing learning rate are plotted. As the learning rate increases, the learning process quickly drops to the local minimum. To avoid the learning process from falling into the local minimum, a small learning rate is required in the beginning of the training process while a relatively larger learning rate is lately needed to avoid the cost of extremely slow convergence. Moreover, the learning rate increase is supposed to be flat in the beginning to avoid the local minimum, while much steeper in the end to decrease the convergence cost. To meet these requirements, the learning rate scheme is designed as:

$$lr^{(i+1)e} = lr^{ie} * \lambda^{1+N_{ie}} \tag{5}$$

where $lr^{(i+1)e}$ and lr^{ie} represent the learning rate at the ith and $(i+1)$th effective epoch respectively. An epoch is considered to be effective only when it decreases the loss function. For an effective epoch the learning rate will be recorded and used for the next update, while the learning rate of ineffective epoch will be propagated to the next epoch.

This means that if the current epoch fails to decrease the loss function, the same learning rate will be used for the next epochs until an effective epoch emerges. In Eq. (5) N_{i_e} counts the effective epochs and λ indicates the base of the exponential function which is much close to 1 but greater than 1, which is set as 1.001 for our experiments. We use the effective epoch count rather than epoch as the exponential power, to prevent the loss function from falling into the local minimum at the early stage and to limit the growth rate of learning rate at the late stage around the global optimal solution, resulting from a number of ineffective epochs. The initial learning rate is set as 1e-6 and we set the upper limit as 1e-4 to avoid divergence as the learning rate increases.

In Figure 4.6 we compare the loss decrease profiles between the fixed learning rate (1e-6) and the proposed learning rate in Eq. (5). It is observed that the proposed scheme starts from a point in the network which has a larger loss, but as the learning rate tardily increases the loss decreases slowly avoiding the local minimum. However, as the epoch increases, the loss rapidly falls down and reaches a smaller level that a fixed learning rate 10^{-6} fails to obtain.
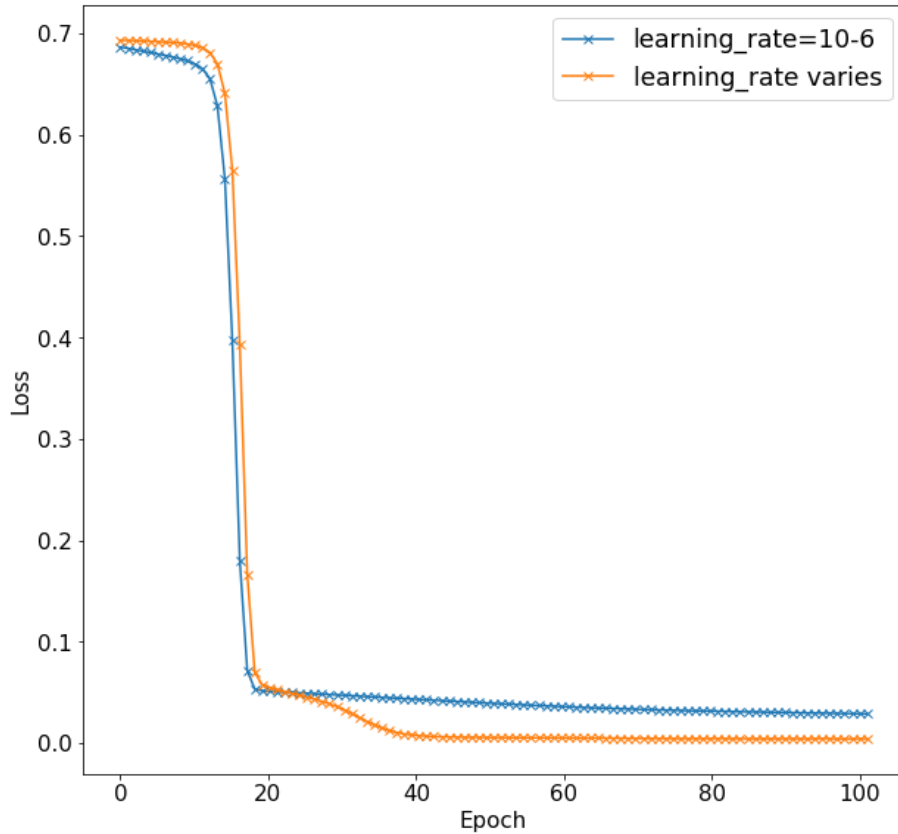
Figure 4.6. The comparison of the decay loss between learning rate 10^{-6} and the proposed learning rate scheme in 100 epochs. The proposed learning rate increases from 10^{-6} to 10^{-4}.

3) Segmentation

The segmentation describes the process of implementing the trained network on the kth test voxel $R^{\alpha_n}_{k(x,y,z)}$. We have a 2D kernel for training the segmentation network. It shows limited ability to consider the relationship between slices. It is assumed that the segmentation from network $g(R^{\alpha_n}_{(x,y,z)})$ is indicated as $M^{\alpha_n}_{(x,y,z)} \in \{0,1\}$, a binary map corresponding to the testing voxel. Considering the correlation between adjacent slice, a refining post-processing is introduced as follows. If $M^{\alpha_n}_{(x,y,z)} \oplus M^{\alpha_n}_{(x,y,z-1)} = 1$ and $M^{\alpha_n}_{(x,y,z)} \oplus M^{\alpha_n}_{(x,y,z+1)} = 1$, then $M^{\alpha_n}_{(x,y,z)}$ is set to equal to $M^{\alpha_n}_{(x,y,z-1)}$ and $M^{\alpha_n}_{(x,y,z+1)}$. This process particularly works for correcting the isolated segmentation error in Z direction. The segmentation result after the refined process refers as $S^{\alpha_n}_{(x,y,z)}$.

4.3.3 Evaluation criterion

The evaluation component as depicted in Figure 4.3 is accomplished by figuring out the group which has the best segmentation performance using the internally trained network, given the 3D zebrafish data and N different parameter groups $\{ \alpha_1, \ldots, \alpha_N \}$. Different from 2D image segmentation, evaluation of image segmentation in 3D should be implemented based on a unit of 3D image because the slices are well ordered. This means that each 3D sample should be independently evaluated, rather than evaluating all the slices in 2D across samples. Because what interests us is not only the overall performance of segmentation on image slices but also the profile of performance change along the ordered slices, such as how the performance varies from the tail to head for zebrafish. Taking this into account the objective function for optimisation is formulated as follows:

$$\alpha^* = \min_{\alpha} \frac{1}{K} \sum_{k=1}^{K} \frac{\sigma_-(F1(S^{\alpha}_{k(x,y,z)}))}{F1(S^{\alpha}_{k(x,y,z)})} \tag{6}$$

K is the number of samples used for evaluation. For each sample, i.e. a zebrafish 3D image excluding the training slices in our experiment, both the overall performance in 3D and negative deviation σ_- of all slices to the overall performance are considered. The criterion in Eq. (6) for evaluation is defined as the modified coefficient of variation of F1-score (MCVF1). The negative deviation $\sigma_-(F1(S^{\alpha}_{k(x,y,z)}))$ is defined as the standard deviation of F1 score [88] for the slices that perform worse than that of the overall 3D volume. $F1(S^{\alpha}_{k(x,y,z)})$ gives the overall segmentation performance of the 3D volume for sample k with reconstruction parameter α. In our work this is accomplished by calculating the $F1$ scores for the whole segmentation volume $S^{\alpha}_{k(x,y,z)}$. The aim of the parameter optimisation is to find the optimal α which produces 3D image group that has a maximum overall performance for segmentation as well as minimum negative deviation for performance. The details of experimental implement regarding to the segmentation evaluation of 3D image will be presented in the experimental section.

4.4 Experiments

In this section, we will first give an example to show the streak artefacts produced in the FBP reconstruction, for comparison followed by a result with artefacts eliminated, using iterative reconstruction method. The second part of the experiments is the parameter optimisation for iterative reconstruction, implemented on two zebrafish datasets cleared with different protocols. In the third experiment, we investigate the effects of different reconstruction methods, i.e. the FBP and iterative reconstruction, on segmentation performance, taking the streak artefacts into account.

4.4.1 Streak artefacts and elimination

In this section an intuitive comparison between the FBP [68] and iterative reconstruction [28] on a zebrafish for our OPT imaging system is given. There are quite dense but small-sized GFP [89] signals inside the zebrafish, which provides an extreme case regarding to how the streak artefacts are produced in OPT imaging and reconstruction. Figure 4.7 (a) and 4.8 (a) show the same tomogram which is one of the 400 ones evenly acquired in a full revolution. The corresponding reconstructed volumes are presented in Figure 4.7 (b) and 4.8 (b) from which a large difference can be observed. To further zoom in and visualize the difference we use both maximum projection along Z direction (Figure 4.7 (c) and 4.8 (c)) and 3D volume visualization (Figure 4.7 (d) and 4.8 (d)).

With the results in Figure 4.7 and 4.8, for both algorithms the pre-processing on the tomograms includes background subtraction and centre of rotation (CoR) correction in Chapter 2. The FBP reconstruction is implemented as elaborated in cf. § 2.3.2. For iterative reconstruction the iteration number η is set as 10 and the 3D image from the FBP is used as the initial reconstruction γ in this example. From the observation of different reconstructed results, we can see that the iterative reconstruction, in a superior manner, outperforms the FBP in terms of artefact suppression on the given data.

4.4.2 Parameter optimisation

To implement the experiments for the parameter optimisation, first the dataset and experimental settings are introduced. The iteration number and initial reconstruction are studied and optimized for reconstruction process. In this section, the details of experiments for optimizing these two parameters are explained.

1) Dataset and experimental settings

In order to implement the experiments for parameter optimisation, 6 zebrafishes are used for OPT imaging. They are split into two groups and are prepared with different protocols that result in different contrast and intensity distributions. Figure 4.9 shows two example images of zebrafish in bright-field mode of which each sample corresponds to a
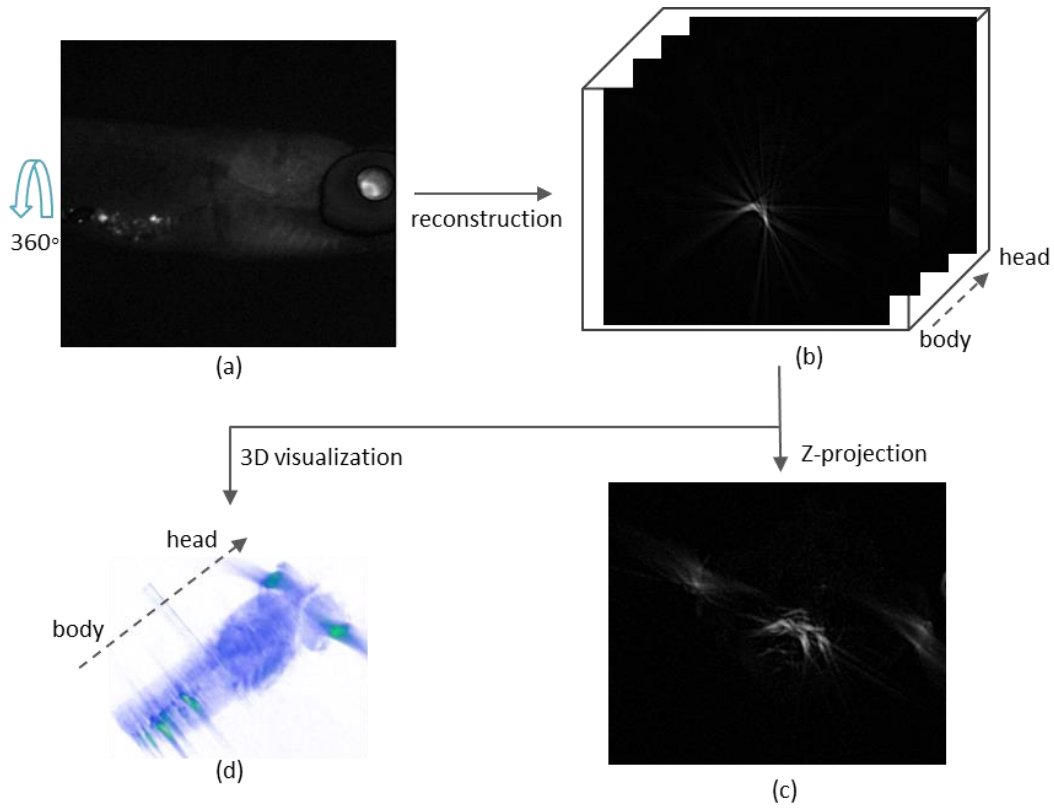
Figure 4.7. Streak artefacts existing in OPT reconstruction with the FBP reconstruction. (a) The tomogram of zebrafish in fluorescence mode. (b) Streak artefacts existing in the reconstructed slices using the FBP algorithm. (c) Z-projection of all reconstructed slices along with streak artefacts. (d) 3D volume rendering with streak artefacts.
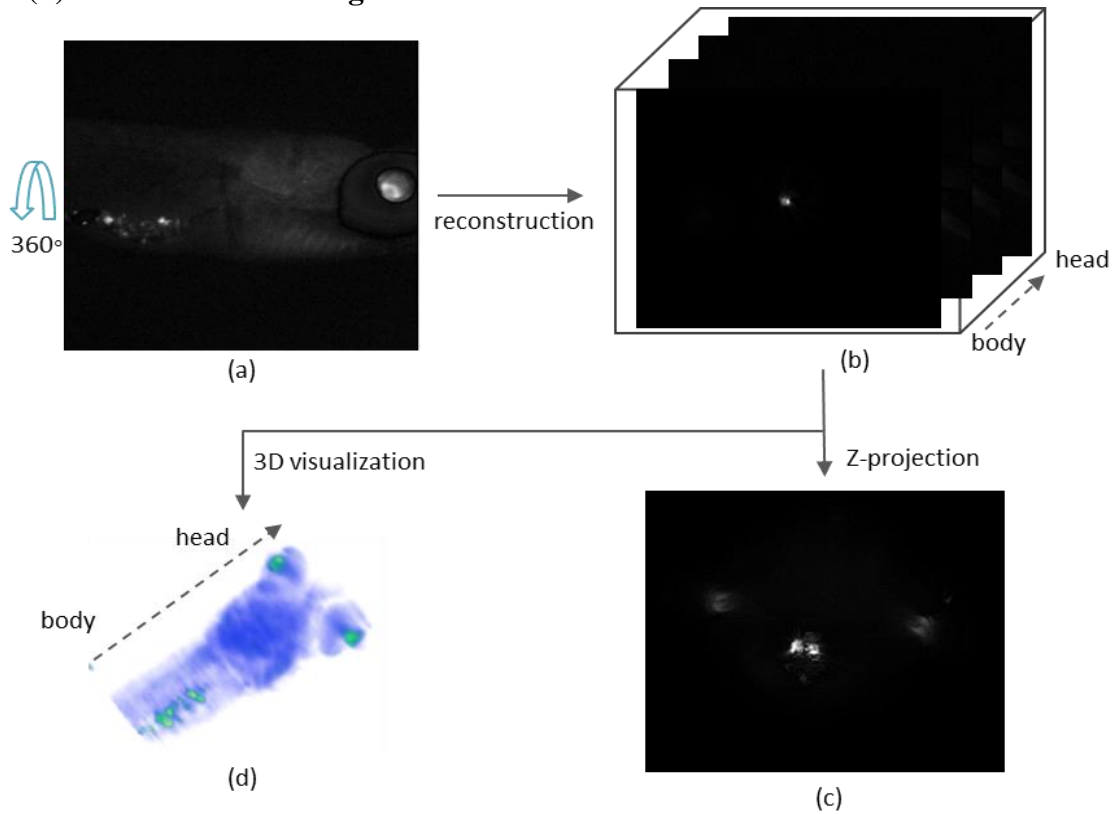
Figure 4.8. Streak artefacts eliminated with OSEM reconstruction in OPT. (a) The tomogram of

zebrafish in fluorescence mode. (b) Streak artefacts removed in the reconstructed slices using OSEM algorithm. (c) Z-projection of all reconstructed slices without streak artefacts. (d) 3D volume rendering with no streak artefacts.

different protocol for sample preparation. The motivation for using bright-field rather than fluorescence images is that, in our imaging system, bright-field images provide the outline information of a zebrafish. This can provide more prior knowledge to benchmark for segmentation. Figure 4.9 (a) corresponds to the tomogram of 5 *dpf* cleared zebrafish with BABB protocol [20], while (b) displays the 25 *dpf* zebrafish with CUBIC protocol [90],[91]. In the bright-field mode, we can easily observe the difference of intensity distribution between them. For each preparation protocol or dataset, there are three zebrafishes studied for parameter optimisation. Each reconstructed sample, i.e. zebrafish 3D image, consists of 1360 slices produced from a 400 tomogram image set using the reconstruction algorithm, compromising 4080 slices for each group or experimental implement.

We study the two most problematic parameters, i.e. iteration number η and initial reconstruction γ. First, for each dataset the effect of iteration number on segmentation performance is investigated. To this end, 5 groups of 3D images reconstructed from different iteration numbers are generated, combing the cost of computation and effectiveness of the reconstruction. Each group corresponds to one of five iteration numbers, i.e. $\eta = \{5, 10, 15, 20, 25\}$. We consider 5 to be a reasonable step size for the assessment of reconstruction performances based on different numbers of iteration, concerning the computational expense and experimental requirements.

(a)   (b)
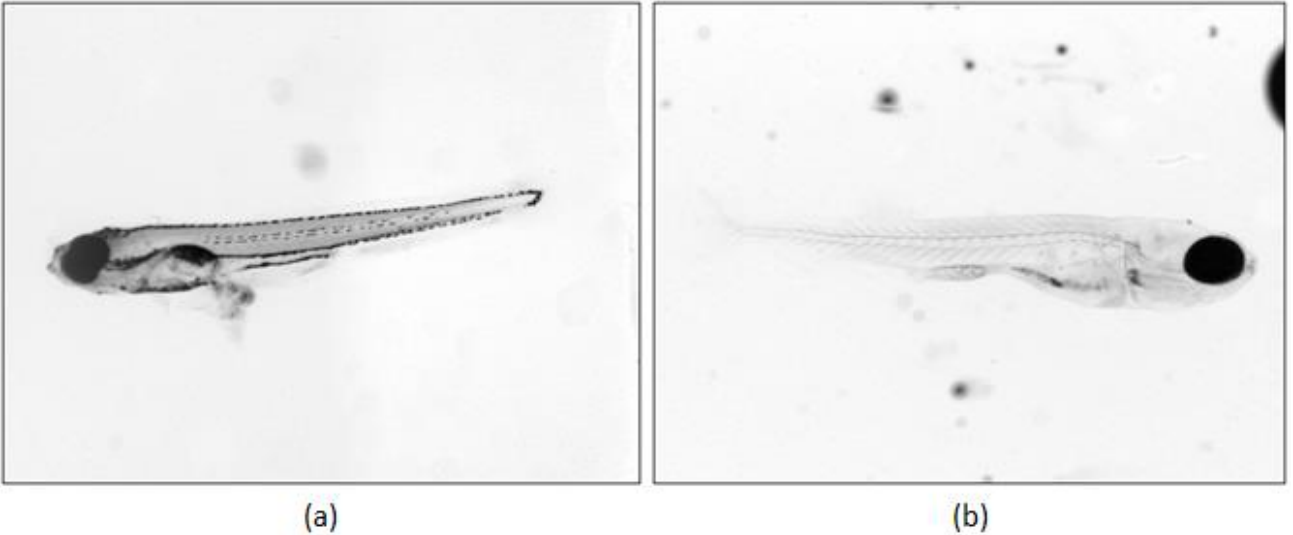
Figure 4.9. Samples of zebrafish with different preparation protocols in the bright-field mode. (a) One example of a tomogram for the 5 *dpf* zebrafish clearing with BABB [20] protocol. (b) One example of tomogram for the 25 *dpf* zebrafish cleared with CUBIC [90], [91] protocol.

The iteration number η is optimized by comparing the segmentation performance of each group. For initial reconstruction γ we compare the performance with an initial 3D

image from the FBP results (FBP-initial) and zeros (No-initial). To obtain the segmentation network all the zebrafish slices within each group are evenly distributed into training set (20%) and test set (80%). The segmentation network within each group is trained with the validation rate as 10%. The segmentation ground truths are labelled based on the FBP results. The segmentation performance MCVF1 of each group consists of overall F1 scores in 3D scale and the corresponding negative deviation on the test data as depicted in cf. § 4.3.2.

2) *Iteration number and initial reconstruction*

A. Experiments on 5 *dpf* zebrafish

The first data used in the experiments are three 5 *dpf* zebrafish are prepared with the BABB protocol [20]. The MCVF1 performance of each group defined in cf. § 4.3.2 and the performance of individual samples are presented in Table 4.1 and Table 4.2. The two tables show the performance differences of different initial reconstruction settings, with Table 4.1 presenting the results with the No-initial reconstruction setting while Table 4.2 corresponding to the FBP-initial results. Results are obtained based on the same training configuration of segmentation network.

For each sample with a particular iteration number, the performance consists of overall 3D F1 score $F1$ and negative deviation $\sigma_-(F1)$ of all slices. Such as for the Fish1 with 5-iterations, 98.23% represents $F1$ and 2.78% is $\sigma_-(F1)$. The smaller the MCVF1 is, i.e. higher $F1$ and smaller $\sigma_-(F1)$, the higher segmentation performance is. The results of the highest segmentation performance for each sample across iteration numbers are in bold. From the results obtained on the 3 zebrafishes, 10-iterations achieves the best overall performance for both No-initial and FBP-initial reconstruction. Regarding the different initial schemes, the FBP-initial method outperforms the No-initial approach in terms of three segmentation performances yielded with the same iteration number.

Table 4.1 and Table 4.2 present quantitative measures for segmentation performance on the 3D images. It, however, does not provide performance details of each slice inside of the 3D volume. In order to achieve these segmentation performances of all slices within each sample, each individual 3D image needs to be investigated. Figure 4.10 and 4.9 provides an example with that on the first zebrafish. Slices from the left to the right correspond to the zebrafish from the head to the tail. The F1_score of each slice ranges from 0 to 1. Each point in Figure 4.10 and 4.9 stands for a F1 score of one slice. The value of 1.0 on the left and right side of the figures implies a background slice in the reconstructed 3D data. Close to the background slices are the critical slices which are indistinguishable for segmentation. That is why they have quite low F1 score performances on both sides of the 3D data, i.e. the critical slices between the head or tail and background.

Table 4.1. Segmentation performance of iterative reconstruction implemented with various iteration numbers and No-initial reconstruction on the three 5 *dpf* zebrafishes.

No-initial	Fish1 (%)	Fish2 (%)	Fish3 (%)	MCVF1 (%)
5-iterations	98.23 (**-2.78**)	97.94 (-2.77)	97.88 (**-3.35**)	3.03
10-iterations	**98.24** (-2.82)	**98.14** (**-2.69**)	**97.94** (-3.37)	**3.02**
15-iterations	97.54 (-3.19)	97.38 (-3.63)	97.25 (-3.98)	3.70
20-iterations	97.34 (-3.10)	97.11 (-3.40)	96.91 (-3.84)	3.57
25-iterations	97.74 (-3.84)	97.70 (-3.10)	97.69 (-3.50)	3.56

Table 4.2. Segmentation performance of iterative reconstruction implemented with various iteration numbers and FBP-initial reconstruction on the three 5 *dpf* zebrafishes.

FBP-initial	Fish1 (%)	Fish2 (%)	Fish3 (%)	MCVF1 (%)
5-iterations	98.29 (-2.79)	98.01 (**-2.56**)	97.95 (-3.18)	3.55
10-iterations	**98.39** (**-2.63**)	**98.27** (-2.61)	**98.17** (**-3.16**)	**2.85**
15-iterations	98.29 (-2.82)	98.14 (-2.68)	98.06 (-3.18)	2.95
20-iterations	98.36 (-2.77)	98.04 (-2.77)	97.95 (-3.30)	3.00
25-iterations	98.31 (-2.76)	98.01 (-2.88)	98.02 (-3.57)	3.13
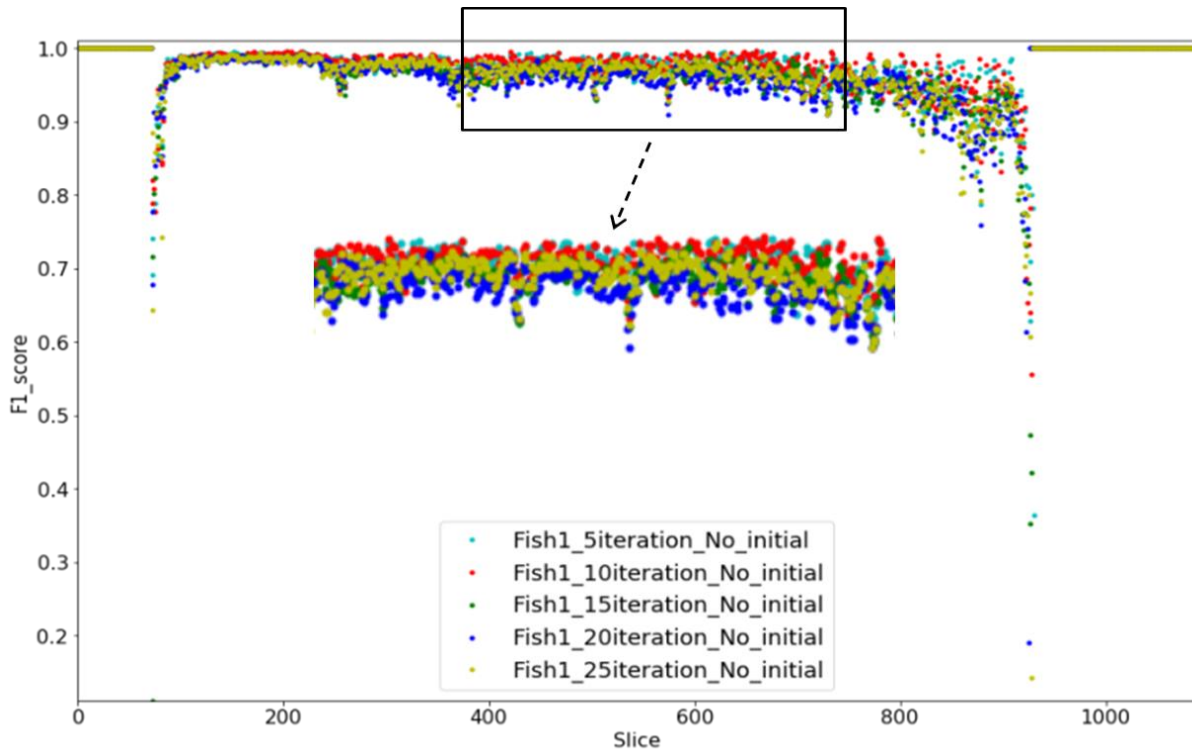
Figure 4.10. Segmentation performance of reconstructed slices with different iteration numbers in a No-initial setting on the 5 *dpf* Fish1. Each point represents the F1 score of each slice for segmentation. The rectangular area is zoomed.
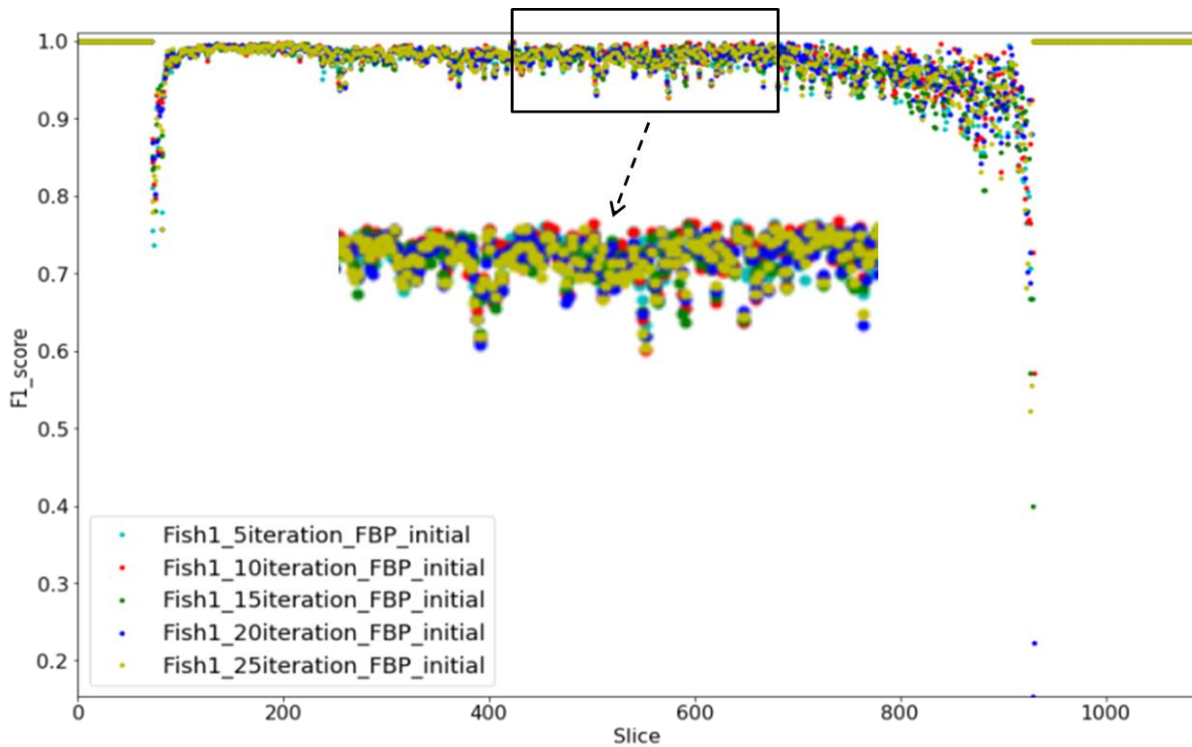
Figure 4.11. Segmentation performance of reconstructed slices with different iteration numbers in a FBP-initial setting on the 5 *dpf* Fish1. Each point represents the F1 score of each slice for segmentation. The rectangular area is zoomed.

By comparing the distribution of F1_score on all slices in Figure 4.10, we can also see that the 3D image group with 10-iterations has more points with larger F1_score and less with smaller F1_score, compared to other groups. This is in correspondence with the results of Fish1 in Table 4.1, i.e. larger $F1$ and smaller $\sigma_-(F1)$ value. Similar results can be observed in Figure 4.11 for FBP-initial reconstruction. The comparisons of segmentation performance on the 5 *dpf* zebrafishes reconstructed with different iterations, both for No-initial and FBP-initial, indicate that the reconstruction with the 10-iterations produces the most desirable results.

B. Experiments on 25 *dpf* zebrafish

As far as the variation of intensity distribution and experimental environment are concerned, three 25 *dpf* zebrafishes cleared with CUBIC protocol [90],[91] are also used for experiments. The performances of each sample and each 3D image group are demonstrated in Table 4.3 and 4.4, with MCVF1 corresponding to the performance of each group while $F1$ and smaller $\sigma_-(F1)$ showing the performance of each sample. As with Table 4.1 and 4.2, the best performance of individual sample and group is given in bold. Overall, the segmentation performance indicates that the FBP-initial reconstruction method outperforms the No-initial approach. Furthermore, in both cases, 10 iterations outperform other number of iterations, according to the MCVF1, $F1$ and $\sigma_-(F1)$ values. This is consistent with the observations regarding the 5 *dpf* zebrafishes.

However, compared to the 5 *dpf* zebrafishes, segmentation performances on the 25 *dpf* zebrafishes are generally lower and more deviated across 3D image groups. For example if we compare the Fish1 of both dataset in the No-initial case (in Table 4.1 and 4.3), the Fish1 of 5 *dpf* zebrafish as shown in Figure 4.9 (a), has a $F1$ range of [97.34, 98.24] that outperforms the range of [97.37, 97.97] for the Fish1 of 25 *dpf* zebrafish in Figure 4.9 (b). $\sigma_-(F1)$ of the 5 *dpf* Fish1 in Table 4.1 ranges from 2.78 to 3.84, smaller than that of 25 *dpf* Fish1 in Table 4.3 with the range from 3.51 to 4.83. Similar comparisons between the two datasets, corresponding to the two protocols, can be made for the other samples. In general, we conclude that the CUBIC clearing protocol makes the zebrafishes more transparent for OPT imaging compared to the BABB protocol. We need to acknowledge that a higher transparency of sample results in a higher similarity or a lower contrast between the foreground (zebrafish) and background tomograms and the 3D images, particularly at the object edge. Such 3D image will consequently have a lower segmentation performance and will be more sensitive to the number of iterations in the process of reconstruction.

Table 4.3. Segmentation performance of iterative reconstruction implemented with various iteration numbers and No-initial reconstruction on the three 25 *dpf* zebrafishes.

No-initial	Fish1 (%)	Fish2 (%)	Fish3 (%)	MCVF1 (%)
5-iterations	97.81 (-3.72)	96.51 (-9.24)	97.73 (-3.69)	5.72
10-iterations	**97.97** (**-3.51**)	**96.72** (**-5.78**)	97.84 (**-2.23**)	**3.95**
15-iterations	97.87 (-3.86)	95.18 (10.56)	**97.89** (-2.29)	5.79
20-iterations	97.86 (-4.40)	85.83 (-30.4)	97.82 (-3.32)	14.44
25-iterations	97.37 (-4.83)	76.08 (-40.2)	96.99 (-4.47)	20.79

Table 4.4. Segmentation performance of iterative reconstruction implemented with various iteration numbers and FBP-initial reconstruction on the three 25 *dpf* zebrafishes.

| FBP-initial	Fish1 (%)	Fish2 (%)	Fish3 (%)	MCVF1 (%)
5-iterations	98.08 (-3.45)	**97.21** (-6.07)	98.17 (-3.08)	4.30
10-iterations	**98.29** (**-3.11**)	97.18 (**-5.75**)	**98.38** (**-2.58**)	**3.90**
15-iterations	97.71 (-3.99)	94.15 (-10.7)	97.48 (-3.76)	6.4
20-iterations	97.90 (-4.21)	91.56 (-25.4)	98.02 (-2.77)	11.60
25-iterations	97.90 (-3.73)	83.30 (-32.7)	97.84 (-3.28)	15.47
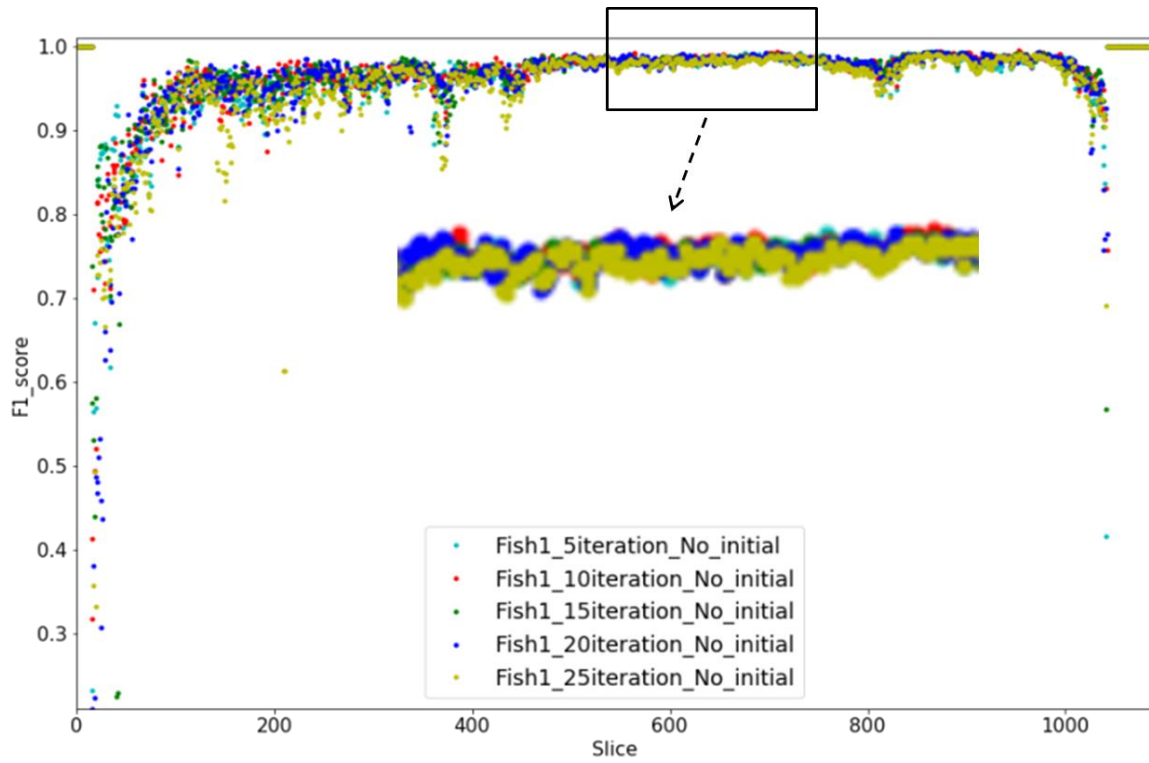
Figure 4.13. Segmentation performance of reconstructed slices with different iteration numbers in a No-initial setting on the 25 *dpf* Fish1. Performance is represented by F1 score of each slice. The rectangular area is zoomed.
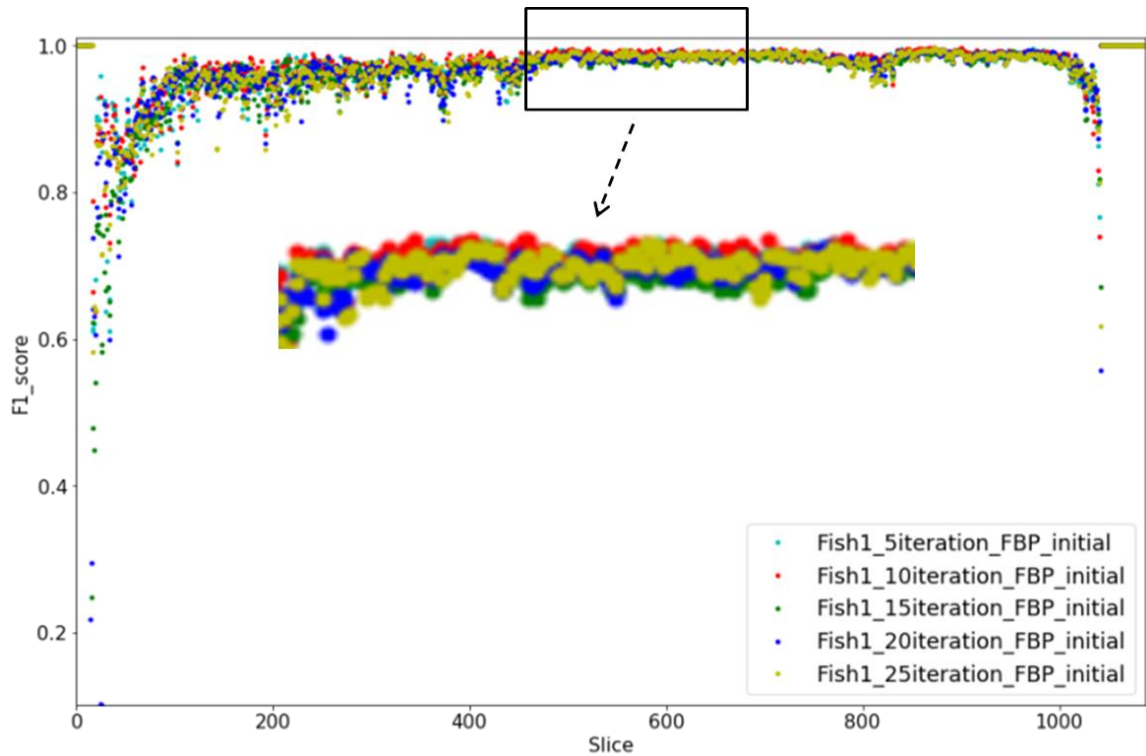
Figure 4.14. Segmentation performance of reconstructed slices with different iteration numbers in a FBP-initial setting on the 25 *dpf* Fish1. Performance is represented by F1 score of each slice. The rectangular area is zoomed.
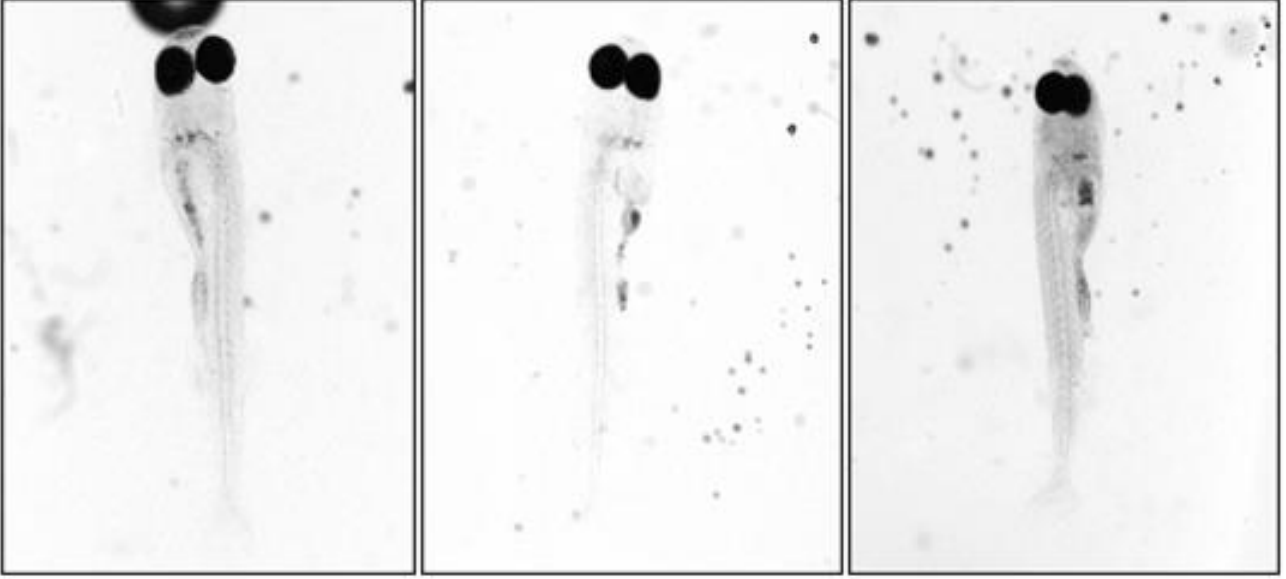
Figure 4.12. The tomogram of the 25 *dpf* Fish1, Fish2 and Fish3. They are shown in mounting orientation in the OPT imaging system. Fish2 has a lower contrast than the other two, with comparable background noise.

In Table 4.3 and 4.4 we can readily see that as iteration number increases the performance of sample Fish2 decreases dramatically whilst sample Fish1 and Fish3 decrease much less. This is because Fish2 has a lower contrast or a higher transparency as we can see from Figure 4.12. When the iteration number increases the information loss of the zebrafish body in the reconstruction is much more obvious. This makes it more difficult to be segmented correctly.

Regarding the performance of individual slices within the 25 *dpf* Fish1 for both No-initial and FBP-initial configuration, they are separately shown in Figure 4.13 and 4.14. For both cases, 10-iterations reconstruction outperforms other groups from either 2D slice or 3D volume scale.

4.4.3 Comparison of segmentation performance between OSEM and FBP

In this section the segmentation performance from the FBP reconstruction and the iterative reconstruction (OSEM) will be compared, experimented on the same datasets as for parameter optimisation, cf. § 4.4.2. In order to eliminate the contingency of gaining a good performance based a segmentation network that is trained on a fixed training set, different training sets are taken into account. They are evenly sampled from the 3D reconstruction slices with a different step size from 2 to 20, corresponding to the training ratio from 50% to 5%. 50% is considered to be a reasonably high training ratio for training a segmentation network in 3D OPT image, considering both accuracy and training cost. Higher training ratio might produce a bit higher accuracy, yet the training cost increases dramatically. As the aim of the comparisons among different reconstruction groups is to find the best-performed reconstruction parameters, rather than search for the global highest accuracy. Therefore, it is not necessary to train each network

with the highest training cost, e.g. a training ratio of 90%. The comparisons for both data are shown in Table 4.5 and 4.6. The performances are obtained with the same configuration for training segmentation networks.

In this experiment, the iteration number is set to 10 for the iterative reconstruction, because of its good performance, cf. § 4.4.2. In the training process of segmentation network, the number of epochs is set as 500 with batch size of 12 and each training process starts from a similar point in the network which can be seen from the loss and accuracy value. The MCVF1 of each 3D image group is calculated from individual samples, cf. § 4.3.2. The reconstruction method that has the smallest MCVF1 for each training ratio (bold) performs the best. By comparing the average MCVF1 we conclude that on both zebrafish datasets the FBP-initial iterative reconstruction achieves the best segmentation performance and are most desirable.

We further investigate the effect of streak artefacts on the segmentation performance. The average MCVF1 of different ratios of training set, produced with a specific reconstruction method, is seen as the criterion for assessing the overall performance of the reconstruction method. For the 5 *dpf* zebrafishes, the overall performances of the three reconstruction methods are 6.00, 3.90, and 3.83. Because the 5 *dpf* zebrafishes are cleared with the BABB protocol that maintains high specimen contrast; the pigments on the zebrafish skin are more concentrated, cf. Figure 4.9 (a). This introduces more streak artefacts in the FBP reconstructed slices. This explains the reason why iterative reconstruction methods (both No-initial and FBP-initial) highly outperform FBP method in Table 4.5. Different from the BABB protocol, the 25 *dpf* zebrafishes are cleared with CUBIC protocol. In this case, we obtain tomograms with lower contrast and the pigments are also less concentrated, introducing less streak artefacts in the FBP reconstruction. This allows the FBP reconstruction to produce 3D images achieving comparable overall performance (5.51) for segmentation, comparing to the iterative reconstructions (5.81 and 5.32).

Excluding the effect of streak artefacts on segmentation, we only look at the overall performance of OSEM reconstruction on the two different datasets in Table 4.5 and 4.6. The 5 *dpf* less transparent zebrafishes outperform the 25 *dpf* zebrafishes which are more transparent. When we compare the performances between the OSEM and FBP, it can be seen that the results for the 25 *dpf* zebrafishes are comparable, but the OSEM achieves better performance than the FBP for the 5 *dpf* zebrafishes. The reason for the difference is that the FBP reconstructions for the 5 *dpf* zebrafishes have much more streak artefacts than the 25 *dpf* zebrafishes do, which deteriorates the segmentation performance.

Table 4.5. Comparisons of segmentation performance using different reconstruction algorithms on the three 5dpf zebrafishes.

5dpf	FBP				No-initial OSEM				FBP-initial OSEM			
Training Ratio	Fish1 (%)	Fish2 (%)	Fish3 (%)	MCVF1 (%)	Fish1 (%)	Fish2 (%)	Fish3 (%)	MCVF1 (%)	Fish1 (%)	Fish2 (%)	Fish3 (%)	MCVF1 (%)
2 (50.0%)	98.02 (-2.57)	97.83 (-3.23)	98.01 (-3.76)	3.26	98.90 (-2.30)	98.92 (-2.18)	98.90 (-2.81)	2.46	98.87 (-2.17)	98.92 (-2.02)	98.86 (-2.67)	**2.32**
5 (20.0%)	97.10 (-3.17)	96.15 (-4.57)	96.61 (-5.53)	4.57	98.24 (-2.82)	98.14 (-2.69)	97.94 (-3.37)	3.02	98.39 (-2.63)	98.27 (-2.61)	98.17 (-3.16)	**2.85**
8 (12.5%)	97.02 (-3.29)	96.61 (-4.67)	97.02 (-5.50)	5.05	97.81 (-3.21)	97.89 (-3.17)	97.89 (-3.62)	3.40	97.65 (-3.17)	97.66 (-3.25)	97.63 (-3.21)	**3.29**
10 (10.0%)	94.86 (-5.54)	93.80 (-5.66)	95.08 (-7.40)	6.42	97.39 (-3.39)	97.28 (-3.64)	97.14 (-3.89)	3.74	97.50 (-3.29)	97.50 (-3.25)	97.16 (-4.19)	**3.68**
12 (8.30%)	94.99 (-5.77)	93.51 (-6.14)	95.04 (-6.32)	6.43	96.81 (-3.78)	96.53 (-4.15)	96.23 (-4.38)	4.25	97.22 (-3.26)	97.32 (-4.24)	97.04 (-4.58)	**4.15**
15 (6.70%)	95.26 (-5.29)	94.86 (-6.56)	95.76 (-7.32)	6.71	97.20 (-3.69)	97.04 (-3.89)	96.90 (-4.46)	**4.13**	97.22 (-4.26)	97.14 (-3.34)	97.05 (-4.87)	4.27
18 (5.60%)	94.28 (-5.91)	92.06 (-7.23)	93.60 (-7.06)	7.21	96.78 (-4.34)	96.80 (-4.36)	96.71 (-5.19)	4.79	96.80 (-4.60)	96.83 (-4.13)	96.67 (-4.54)	**4.57**
20 (5.00%)	93.61 (-6.89)	91.2 (-7.47)	93.31 (-8.90)	8.36	96.58 (-4.54)	96.23 (-5.63)	96.11 (-5.25)	**5.34**	96.30 (-4.80)	96.05 (-5.18)	95.65 (-5.90)	5.51
Average				6.00				3.90				**3.83**

Table 4.6. Comparisons of segmentation performance using different reconstruction algorithms on the three 25dpf zebrafishes.

5dpf	FBP				No-initial OSEM				FBP-initial OSEM			
Training Ratio	Fish1 (%)	Fish2 (%)	Fish3 (%)	MCVF1 (%)	Fish1 (%)	Fish2 (%)	Fish3 (%)	MCVF1 (%)	Fish1 (%)	Fish2 (%)	Fish3 (%)	MCVF1 (%)
2 (50.0%)	98.04 (-4.34)	98.11 (-3.51)	98.44 (-1.98)	3.34	98.85 (-2.42)	98.45 (-3.48)	98.91 (-1.95)	**2.65**	98.85 (-2.66)	98.47 (-3.76)	98.96 (-1.86)	2.79
5 (20.0%)	96.93 (-5.51)	96.70 (-5.41)	98.07 (-3.03)	4.78	97.97 (-3.51)	96.72 (-5.78)	97.84 (-2.23)	3.94	98.29 (-3.11)	97.18 (-5.75)	98.96 (-1.86)	**3.89**
8 (12.5%)	97.25 (-5.23)	96.67 (-5.65)	97.69 (-3.06)	4.78	97.95 (-3.85)	96.06 (-7.92)	98.15 (-3.07)	**5.08**	97.89 (-4.02)	95.76 (-7.77)	98.12 (-3.31)	5.17
10 (10.0%)	96.62 (-5.71)	96.55 (-5.42)	97.83 (-3.03)	4.87	97.67 (-3.99)	94.90 (-8.94)	97.91 (-3.02)	5.49	97.70 (-3.73)	95.73 (-7.18)	97.87 (-3.11)	**4.81**
12 (8.30%)	95.73 (-9.91)	95.78 (-6.40)	97.53 (-2.93)	6.65	97.49 (-4.18)	94.49 (-9.52)	97.74 (-3.63)	**5.99**	97.52 (-4.38)	94.58 (-10.5)	97.75 (-2.73)	6.07
15 (6.70%)	95.59 (-7.27)	95.96 (-5.21)	97.18 (-3.53)	**5.55**	97.14 (-4.82)	93.97 (-11.2)	97.33 (-3.74)	6.86	97.19 (-4.55)	94.47 (-8.38)	97.40 (-4.27)	**5.95**
18 (5.60%)	94.13 (-10.3)	95.70 (-5.82)	96.73 (-4.06)	7.05	97.07 (-5.35)	91.84 (-13.1)	97.12 (-4.17)	7.91	97.18 (-4.70)	93.54 (-11.7)	97.38 (-3.34)	**6.85**
20 (5.00%)	95.62 (-7.08)	94.30 (-9.99)	97.24 (-3.28)	7.08	96.81 (-6.07)	91.82 (-14.8)	96.96 (-3.45)	8.52	96.76 (-5.43)	92.08 (-10.7)	97.07 (-4.01)	**7.04**
Average				5.51				5.81				5.32

4.4.4 Discussion

In this section an overall experimental discussion will be given. In § 4.4.1 the qualitative comparison between the FBP and iterative (OSEM) reconstruction is presented, in terms of the elimination of artefacts in one specific zebrafish. This inspired and motivated us to go one step further and study the parameter optimisation for the iterative reconstruction with respect to its potential performance in OPT reconstruction. This means that the effect of different reconstruction parameters needs to be investigated. We accomplished this in § 4.4.2 by studying the two most problematic parameters, i.e. the number of iterations and the initial reconstruction, on two zebrafish datasets with different image intensity distributions. According to these experiments, the combination of 10-iterations and FBP-initial is proved to produce the most desirable and preferable 3D image, compared to the other combinations. Additionally, 10-iterations are also acceptable for reconstruction implementation concerning the computation cost. When comparing the segmentation performances across datasets, we find that the contrast rich dataset performs better, which is reasonable from theoretical perspective. Because a highly transparent sample produces a more indistinguishable intensity distribution in both tomogram and 3D image in OPT, therefore bringing more complications for the segmentation.

Followed by parameter optimisation, comparisons of segmentation performance between the FBP and iterative reconstruction are displayed in a quantitative manner. It is known that sample preparation plays a crucial role in the OPT imaging, as well as the 3D reconstruction process. One should realize that the BABB protocol provides higher specimen contrast that generally leads to a better 3D image from the iterative reconstruction and consequently achieves a higher segmentation performance, however introduces streak artefacts when using the FBP reconstruction. Moreover, with the BABB clearing protocol, the iterative reconstruction can eliminate the streak artefacts during the reconstruction process, so the segmentation result outperforms that of the 3D images from the FBP algorithm. With the CUBIC clearing protocol, both reconstruction methods are able to avoid the streak artefacts in the 3D image, therefore obtaining a comparable segmentation performance between different approaches. In this work the experiments are implemented on finite zebrafish samples. With the trend of high-throughput analysis on zebrafish, in the near future more samples will be considered to confirm this tendency.

4.5 Conclusions

The research presented in this chapter is the development and implementation of an iterative reconstruction, specifically OSEM, and the further exploration of approaches to optimize the reconstruction parameters. The OSEM algorithm produces superior 3D image in comparison with FBP algorithm in terms of streak artefact elimination when the signals in the tomograms are very concentrated. The method used for further improvement of the reconstruction is realized by optimizing the parameters for the OSEM algorithm,

which requires the evaluation of reconstructed 3D image. Restricted by the lack of benchmarks for reconstruction in real-life imaging, we have used an alternative approach that is inspired by the segmentation evaluation. Notably, the way of integrating segmentation evaluation into parameter optimisation for iterative reconstruction, may not result in the achievement of the globally optimal parameters. But to our best knowledge, it provides a good and reasonable way for guaranteeing an optimized and efficient reconstruction result, considering both the reconstruction quality and computational cost. It is worth point out that, even though the OSEM produces promising results for highly transparent samples, e.g. the 25 *dpf* Fish2 with 10 iterations, it could also be possible that a sample is too transparent to produce any significant reconstruction results with the same OSEM parameters. In such an extreme case, decreasing the iteration number might help recover more information in compromise with image blur.

4.6 Acknowledgement

Chapter 4

Chapter 5

Automated Detection of Reference Structures for Fluorescent Signals in Zebrafish with a Case Study in Tumour Quantification

This chapter is based on the following publication:

Chapter summary

Zebrafish as a vertebrate model plays an important role in biomedical researches such as development, disease, toxicological and drug discovery studies. In this chapter we assume that fluorescent markers represent a specific signal of interest. We aim to quantify these signals in zebrafish, to provide accurate experimental information, for e.g. drug discovery, in an automated and efficient way. We first define the quantification approaches with a case study in tumour growth. Based on the definition, the reference structures for the quantification, obtained from bright-field images, are studied.

In order to automatically detect the reference structures from 3D bright-field images with a high performance, we use the deep learning approach to obtain a segmentation of the reference structures for each sample. The 3D images are obtained and reconstructed from the optical projection tomography imaging. According to our experiments, the automated approach for detecting reference structures is a promising method for the relative quantification of fluorescent signals in zebrafish.

5.1 Introduction

In this chapter we will focus on the application of imaging of zebrafish in OPT for disease modelling. In biomedical research, zebrafish has become a widely used model organism in the last decade because of its fecundity, its physiological and genomic similarity to mammals, the existence of many genomic tools and the ease with which large, phenotype-based screens can be performed [92]. Mammalian models of absorption, distribution, metabolism and excretion (ADME) or pharmacokinetics and efficacy, are considered expensive and laborious and consume quantities of precious compounds. Compared to this, zebrafish is more cost-effective and can therefore be a useful alternative to mammalian models.

5.1.1 Research questions

Drug discovery involves a complex iterative process of biochemical and cellular assays, working up to *in vivo* validation in animal models and ultimately in humans. Zebrafish is considered promising in accelerating the process of drug discovery with a comprehensive advantage of scale, high-throughput screening and physiological complexity. In disease modelling and treatment, e.g. drug discovery for tumour treatment, zebrafish has revealed its effectiveness and advantages [93]. Using zebrafish as disease model for tumour means exposures to different levels of drug treatment. The performance of this drug treatment can be expressed in qualitative terms, i.e. using the visual signal, as well as in quantitative terms, i.e. measurements of the intensities and extend of the signal.

To this end a specific fluorescent signal, i.e. the expression of a representative gene is used. However, compared to visual qualitative assays, quantitative assessment is more comparable and transferable and hence more convincing. The quantification of the expression of specific disease within zebrafish such as tumour [93] can give a direct and accurate insight of tumour size and shape, as well as make the precise comparison of different treatment groups.

Depending on different research demands and available facilities, the disease phenotype can be represented as 2D or 3D microscope images. For a whole-mount zebrafish, 2D images from a stereo-fluorescence microscope can provide fast information on the structure of a tumour in a single specimen; de facto this is a projection of 3D information. For measurement and phenotypical description, this should be considered with caution. Projecting and imaging a zebrafish from a different angle will result in a different 2D image. This difference can result in inaccurate and biased quantification. In contrast, 3D imaging of disease phenotype reconstructs the 3D structure, therefore potentially more accurate for disease phenotype quantification. Therefore, we take it as a research question: given a 3D image, what are the possible solutions for the quantification of disease phenotype and to what extent the quantification process can be automated.

5.1.2 OPT as a solution for whole-mount imaging

As mentioned, zebrafish is a good model system for the disease studies. For overview of disease and disease progression in zebrafish, whole-mount imaging is indispensable. With confocal laser scanning microscopy (CLSM), light microscopy (LM) and, to a lesser extent, scanning electron microscopy (SEM) the size of the specimen limits the application of whole-mount imaging [15], [94]. With MRI [16] the strength of the magnetic field determines the resolution that can be obtained for whole-mount imaging and a *mm*-scale object requires quite a strong magnet. With a right choice of optics, optical projection tomography (OPT) [95] can conveniently operate with *mm*-scale objects. It can display gene expression or a specific staining in the bright-field or fluorescence channel, while the specimen as a whole can be visualized. In this manner OPT adds an important range of scale that can be investigated. It allows for the acquisition of whole-mount images of animal/plant tissues as well as organs/organisms [18], [19]. OPT has also been studied for its capability of imaging with excellent spatial resolution and contrast and minimal shadowing artefacts produced from back-projection reconstruction after tomogram acquisition. Therefore, we take OPT into account to assess the whole-mount 3D imaging of zebrafish.

5.1.3 Multi-channel analysis of whole-mount zebrafish

In order to exclude biological variation and individual differences, large scale analysis of zebrafish for disease treatment is necessary. This means that multiple zebrafishes will need to be quantified and averaged to describe the disease progression or the performance of drug exposure at different time-points in disease.

To our best knowledge, two methodologies are used for the quantifying of a disease model. One is measuring a read-out in absolute sizes of the disease marker in either 2D or 3D, named as absolute quantification. This involves a segmentation of the signal that is representing the disease, i.e. the marker, and pixel/voxel size calibration. For absolute quantification, only the fluorescence channel, with fluorescent disease marker, is required. In this particular case, the pixel/voxel size calibration for the imaging system is required so as to make measurements comparable and transferrable between different systems. Another approach is calculating the relative ratio of disease phenotype, i.e. fluorescent signal, referring to a specific structure such as *Body* or *Eye*. This is mostly depending on detection of the reference structure (RS) and has generality across imaging systems and between specimens. In this case often both modalities, i.e. fluorescence and bright-field, are required; with one for the fluorescent marker and the other for the RS. Taking advantage from computational techniques and resources, in this chapter we construct and detect RS for relative quantification of disease phenotype. To that end we introduce two definitions so as to support our approach for relative quantification.

Definition 1 Phenotype: The total appearance of an organism determined by interaction during development between its genetic constitution (genotype) and the environment [96].

Definition 2 Reference Structure (RS): a structure that is a part of the organism under study that is used for a relative comparison over scale and/or time. Usually, the RS is used in a normalization to establish an effect in a phenotype.

Relative quantification of 2D disease phenotype for each zebrafish can be achieved by standard image analysis tools [97]. When considering throughput of the data, a more automated approach is preferred. However, quantification of 3D disease phenotypes for a zebrafish specimen is much more difficult to obtain than it is in 2D. Therefore, in order to prevent manual labor and enable application of analysis on a larger scale, automated image analysis is necessary. We focus on the automation of obtaining RS to accelerate the throughput of the 3D analysis of zebrafish.

The research presented in this chapter concentrates on relative quantification of disease phenotyping, specifically constructing 3D reference structures in zebrafish. We have chosen to work with two RSs that are always visible: the *Body* and the *Eye*. The *Body* represents the overall phenotype of a sample, providing a normalization standard for all the samples. The *Eye* is a local RS with less deviation among specimens of the same stage, it is also easier to detect because of its clear texture. The goal of this research is to automatically detect both RSs and using them for the calculation of the relative quantification of fluorescent signals. Here we focus on the 3D quantification of signals from a disease. The automated RS detection can be generalized to other 3D fluorescent signals in zebrafish.

The RSs we need for the 3D relative quantification are automatically detected from the bright-field 3D image by using segmentation techniques. Concerning the large-scale requirement, we are into exploring an approach for automated detection of the RSs. The transparency and inhomogeneity of the zebrafish make segmentation performance cumbersome, particularly in 3D, when using traditional segmentation algorithms; i.e. threshold-based, region growth, graph cut and traditional machine learning techniques. The challenge must be seen in the high similarity of intensity between voxels inside and outside zebrafish, as well as the edge/surface discontinuities [98]. Fortunately compared to *Body*, *Eye* has a more dense tissue, therefore more discriminative. To meet our requirements for automated RS detection, advanced segmentation approaches will have to be explored.

5.1.4 Related work

During the last twenty years, to our best knowledge there were just a few research topics on automated analysis in zebrafish. Mikut *et al.* [99] contributed a survey on automated processing of zebrafish-related data and generalized the workflow for analysis of biomedical research on zebrafish model. They showed some examples of automated

image analysis, including cell tracking during embryogenesis, heartbeat detection, anatomical landmarks, dead embryo detection, recognition of tissues, and quantification of behavioral patterns. In general, the microscopy images could be classified into two categories as mentioned: bright-field images and fluorescence images. Analysis of the fluorescence images is relatively easy compared to bright-field because fluorescent intensity typically reflects specific information of interest.

Previous work on fluorescent imaging varies with the specific research topic. Understanding bacterial infection was accomplished with a template-based segmentation method through which the shape of a zebrafish larva was detected. The bacterial load was obtained from the fluorescent channel and normalized to the size of the larva, or specific parts thereof [100]–[103]. An automated segmentation was utilized to zebrafish heart based on 2D light-sheet fluorescent images, accompanied by 3D reconstruction in the second stage [104]. They followed the pipeline that 3D volume is reconstructed based on the segmentation results of tomograms. Segmentation of the axial skeleton and spine of the zebrafish are also common in developmental research [105],[106]. Their segmentation was implemented on images of fluorescent maker in notochord sheath cells and with conventional segmentation techniques acceptable results have been achieved [105], [106]. More lately segmentation of developing zebrafish vasculature was proposed from light sheet fluorescence microscopy imaging [107]. They used the open source software Fiji [97] for segmentation of the fluorescent marker and achieved satisfied results [107]. Different from the aforementioned non-learning algorithms, Zhang et al. [108] first brought the deep learning technique to vessel segmentation on images from 3D confocal imaging in 2019. They obtained promising results on accurate segmentation of challenging vessel data that were labelled with green fluorescent protein (GFP). The 3D image data were acquired with confocal microscopy and the segmentation was implemented on the reconstructed slices [108].

In addition, whole-mount specimen segmentation can be achieved using the bright-field image of the sample. It is typically related to phenotype and behavior analysis in the research of development and drug discovery [98], [109]–[111],. Wu et al. [109] proposed a hybrid method which integrates region and boundary information into an active contour model considering the ambiguity of edges for 2D image segmentation. Later, Xiong et al. [110] presented the level-set model to segment zebrafish on image slices from confocal microscopy and achieved promising results on 3D images. Inspired by the good performance of level-set model, Guo et al. [98] integrated mean shift to level-set model for accurate 2D zebrafish image segmentation in bright-field channel, and then used the segmented masks for 3D reconstruction of zebrafish surface. Recently, instead of phenotype analysis, Ishaq et al. [111] classified zebrafish deformation, i.e. normality or deformation, based on 2D bright-field images for drug discovery using a deep neural network.

5.1.5 Structure of this chapter

The research presented in this chapter is to explore a segmentation approach, aiming at automated detection of the RSs for fluorescent signals in zebrafish. The detected RS is able to help with relative quantification of fluorescent signals in zebrafish. In section 5.2 we present the materials and methods used for our research. In section 5.3 we further elaborate the design and implementation of the segmentation approach for automated detection of the two 3D RSs that we use here. In section 5.4 the experiments and results will be presented, with a case study in tumour quantification, followed by conclusions and discussion in section 5.5.

5.2 Materials and methods

Here, we will first explain the specimen and sample preparation of zebrafish used for automated RS detection. The 3D imaging and reconstruction framework will follow afterwards. Based on the reconstructed 3D image, the approach to the relative quantification of fluorescent signals will be formed.

5.2.1 Zebrafish

Both for the *Body* and *Eye* reference structures, 38 zebrafish samples are used to learn parameters of the segmentation approach for automated detection. These zebrafishes are from three different stages including 5 *dpf*, 6 *dpf* and 7 *dpf* and all of them are cleared with the BABB protocol (cf. [20]). As we are only interested in the bright-field image for the detection, the samples are not necessarily stained or marked with fluorescent markers. For our experiments we have eight 5 *dpf* zebrafishes, fifteen 6 *dpf* and fifteen 7 *dpf* zebrafish without staining.

5.2.2 OPT imaging and reconstruction

3D imaging with OPT is suitable for zebrafish imaging as it can deal with the size range to produce whole-mount images. With OPT both bright-field and fluorescence modalities can be accomplished and the acquisition of these channels is done in a sequential manner. In confocal microscopy the details are inspected at cellular level but it compromises information at the global level [112],[113]. The strength of OPT is that it enables observation of the whole specimen, i.e. the zebrafish, at a tissue level with depth ranging from millimeter to centimeter. This character of a large depth and field of view (FoV) enables the analysis the zebrafish as a whole on volumetric level. This also explains the conditions of our research work for detecting the RSs for fluorescent signals in zebrafish.

The bright-field images from the 38 samples are acquired using the OPT imaging system as depicted in chapter 1, cf. § 1.2, producing a 3D tomogram of $1360 \times 1036 \times 400$ for each sample; 1360×1036 per image over 400 angles in full revolution. The 3D bright-field image of each zebrafish is obtained from the corrected tomogram set by

using the iterative reconstruction algorithm as described in chapter 4, cf. § 4.2. The transform from original OPT tomogram set to corrected ones are accomplished by applying the centre of rotation (CoR) correction approach in chapter 2, cf. § 2.3. The application of the iterative reconstruction [28] results in a 3D image sized $R \times R \times 1360$, where R is the determined by the CoR value and 1360 refers to the number of slices in 3D. The 3D images of all the 38 zebrafishes are subsequently used to learn the parameters of segmentation approach for RS detection. This design of the automated RS detection will be elaborated in the section on design and implementation.

5.2.3 Relative quantification

The relative quantification of fluorescent signals is based on the reconstructed 3D images from both the bright-field and fluorescence channel. Specifically, the fluorescent signals from the 3D image of the fluorescence channel are first segmented through a threshold-based algorithm. In this manner, a sub-volume is produced from which the fluorescent signal is quantified. For RS detection, we need first to segment or identify the RS from the 3D bright-field image, producing the sub-volume of the reference structure. The ratio of the two volumes obtained from the segmentation results is defined as the relative quantification of the fluorescent signal.

5.3 Design and implementation

In this section, we focus on the design and implementation of automated RS detection for 3D quantification. With the 3D bright-field image of zebrafish, cf. § 5.2.2, the RS, i.e. *Body* or *Eye* will be identified by the supervised segmentation approach. We investigate how this can be accomplished by a convolutional neural network. For each of the RSs binary ground truth is realized with annotation software, such as TDR or Amira. In order to train a high-performance segmentation network for each RS, we employ U-net segmentation network [29], [114] implemented in both 2D and 3D image space.

5.3.1 Segmentation of reference structures

In order to reduce the computational load, we rescale the 3D image to $512 \times 512 \times 680$ for the 2D U-net segmentation and to $128 \times 128 \times 340$ for the 3D U-net segmentation, thereby compromising resolution. To our best knowledge, 3D U-net has the highest performance when each 3D image is directly fed into the network. However, constrained by the memory, this is not feasible for our 3D image due to its large size. Alternatively, we resized and cropped the 3D image as $128 \times 128 \times 340$ and feed the 3D U-net in a patch way. From the 38 samples as introduced in§ 5.2.1, 35 samples are used for training and validating the segmentation network, whilst 3 samples for testing or evaluating the performance of the segmentation approach. So, there are 23800 slice samples in total for training and validating the 2D U-net, but much less volume patches for 3D U-net training. In order to increase the sample size for 3D U-net, we decrease the patch size to $64 \times$

64×64 with an overlap of 16 to cover 3D context well. This results in 1113 cubic patch samples for training and validation. Concerning the high imbalance of voxels between RS and background, in particular for *Eye*, we exclude the patches that have no object in the ground truth data. For the remaining patch samples, data augmentation including distortion and flip is employed before being fed into the network.

1) 2D Unet

The 2D U-net network [29] feeds 2D images in the *Input* layer. As our starting point is 3D images, we need a slice extractor to provide the 2D images to the net. In this manner, the slice extractor also contributes in stacking all slices back to 3D volumetric images after *Output* layer. The network structure has been elaborated in chapter 4, cf. § 4.3.2.

2) 3D Unet

In the 3D U-net [114], different from the 2D U-net, 3D image are used as *Input* layer. This is schematically depicted in Figure 5.1. As we now work in 3D, a slice extractor is no longer required, and instead a cubic patch is used. The equation operation fits the cubic patch into the *Input* or *Output* layer. The *Output* layer of the network is named as the segmentation map. This will, to some extent simplify the description of network layers including 3D Input and number of kernels. Besides the features on each reconstruction slice, the encoder and decoder in 3D U-net also considers the correlation between adjacent slices by using 3D *Maxpooling, Convolution* and *Upsampling* layers. Compared to a 2D U-net approach, this produces smoother segmentation results. Similar to 2D U-net there is one more *Convolution* layer before each *Maxpooling* or *Upsampling* layer (not shown in Figure 5.1). The *Merge* layer after each *Upsampling* layer integrates shadow layer into a deeper one, this yielding even more informative layers. The segmentation result of the volumetric patch is achieved based on the thresholding of 3D segmentation map, the *Output* layer of network.

5.3.2 Learning scheme

A CNN can only be successful when it is properly designed and parameterized. For this, the proper loss functions, optimizers and learning rate will be employed. In this section, we elaborate on these schemes and how they should be applied.

1) Loss & Metrics

Loss and metrics functions play an important role in training networks, because they provide a criterion for measuring the similarities between prediction and truth, and determine the level of convergence for the training process. With the Sigmoid activation function at the *Output* layer and the binary segmentation problem for both RSs, we first
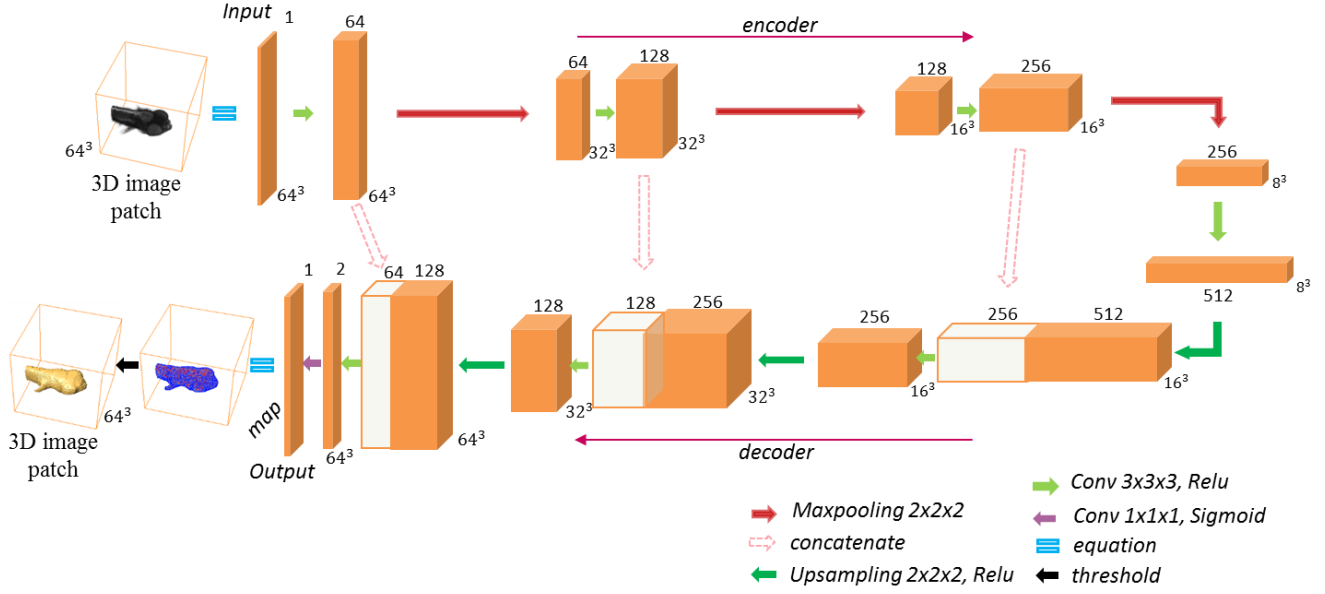
Figure 5.1. The 3D U-net framework for volumetric segmentation of the *Body* RS in zebrafish. A volumetric patch of reconstructed 3D image, i.e. 64x64x64, is fed into the network as *Input* layer. It inherits the characteristics of *Max pooling, Upsampling* and *Convolution* layers; moreover, it includes the concatenation operation to merge similar layers of different depths. In this structure, every operation is implemented in 3D. The equations operation at *Input* and *Output* layer help to overcome the visualization gap between 3D image and 4D network layers. The values of the segmentation map at the *Output* layer are in the range of [0, 1], representing the network response to the *Body* RS. By applying threshold=0.5 to the segmentation map, the binary 3D *Body* RS is obtained as shown in the 3D image patch.

use the binary cross entropy loss [115],[116] and accuracy metrics. In this case, the segmentation performance of both background and RS, i.e. foreground, are taken into account for updating the network weights according to:

$$EL = \sum_{i=0}^{N} g_i^r \log(p_i^r) + (1 - g_i^r) \log(1 - p_i^r) \qquad (1)$$

where EL is the loss, whilst p_i^r represents the probability of voxel i being predicted as the RS r, and g_i^r symbolizes the corresponding ground truth. This normally results in very small losses and high accuracies during training process when the classes are highly imbalanced. In classification problems, class imbalance exists if the number of samples for each class varies. Such imbalance may have a different impact on the classification results depending on the level of imbalance. For the 3D images in our work, the number of voxels for RS class is generally smaller compared to the background class.

With respect to segmentation map in both networks, it only records the response of the positive class, i.e the RS. This means that the network can be updated according to the assessment of the segmentation performance for the positive class. To this end, the Dice coefficient and Dice loss [117] are applied for the training of the network. With the Sigmoid activation, the *Output* layer consists of one plane for the RS. Now P and G are

the set of predicted and ground truth binary labels. The Dice similarity coefficient between two volumes is defined as [117], [118]:

$$D(P,G) = \frac{2|P \cap G|}{|P| + |G|} \tag{2}$$

It weighs FP and FN (precision and recall) equally. The Dice coefficient loss is then defined as follows:

$$DL = 1 - \frac{2 \sum_{i=1}^{N} p_i^r g_i^r + \epsilon}{\sum_{i=1}^{N} p_i^r + \sum_{i=1}^{N} g_i^r + \epsilon} \tag{3}$$

where p_i^r represents the probability of voxel i being predicted as the RS and g_i^r symbolizes that of the ground truth. ϵ is a secondary functional term which helps the loss function converge more effectively.

Considering the high imbalance of 3D image for RS detection, in addition the Tversky loss function [119] is also applied. This metric was formulated based on the Tversky index [120], which gives FP higher weights than FN in the training of the network. The Tversky index between prediction and ground truth volume is formulated as:

$$T(P,G; \alpha, \beta) = \frac{|P \cap G|}{|P \cap G| + \alpha|P - G| + \beta|G - P|} \tag{4}$$

where α and β control the magnitude of penalties for FP and FN, respectively. Accordingly, the Tversky loss function is formulated as:

$$TL = 1 - \frac{\sum_{i=1}^{N} p_i^r g_i^r + \epsilon}{\sum_{i=1}^{N} p_i^r g_i^r + \alpha \sum_{i=1}^{N} p_i^r g_i^0 + \beta \sum_{i=1}^{N} p_i^0 g_i^r + \epsilon} \tag{5}$$

Here p_i^r and g_i^r have the same meaning with Dice coefficient loss, and p_i^0 and g_i^0 are separately the probability of voxel i belonging the background (label = 0) in prediction and ground truth. According to the literature [119], for $\alpha = 0.3$ and β=0.7 the Tversky loss function has the best performance in managing highly imbalanced data. Therefore, we adopt these values to the CNN for our segmentation.

The aforementioned three state-of-the-art metrics and loss functions are widely used in biomedical image segmentation [121] because of their stability and robustness. In the results section we will apply the three metrics and loss functions to our data and segmentation network, and explore how they perform on the results.

2) Optimizer & Learning rate

An Optimizer is an optimisation algorithm that regulates and determines the route of converging for the loss function. Our volumetric data exhibit sparseness, i.e. compared to the object voxels the ratio of the background voxels is high. The Adam optimizer [122] is typically useful for such sparse data. Adam was designed to combine the advantages of Adagrad [123],[124] and RMSprop [125],[126] with momentum [122] as an improved version of stochastic gradient decent (SGD) [87] for training deep learning models. This makes it suitable to work with sparse gradients on noisy data. Another advantage of Adam is that

the rule for step size updating, is invariant to the magnitude of gradient. This helps to go through areas with low gradients such as saddle points and ravines [122].

In order to accelerate training process and to some extent improve the performance of the deep network [127], we employ the stochastic gradient descent with warm restarts (SGDR) [128] in a cyclical learning rate scheme. The idea of this strategy is to decay the learning rate from maximum l_{max}^i to minimum l_{min}^i in a cyclic fashion, using:

$$l_t = l_{min}^i + \frac{1}{2}\left(l_{max}^i - l_{min}^i\right)\left(1 + \cos\left(\frac{T_t}{T_i}\pi\right)\right) \tag{6}$$

where i is the current cycle for learning decay. The beginning of each cycle is referred to as a restart. T_i is the length of the ith cycle, deciding the number of epochs in this cycle. Experimentally, it is preferred to increase T_i as i increases during training. The increasing step for each restart cycle is controlled by T_{mult}. T_t accounts for the number of epochs that have been performed since the last restart or in the current cycle i. This means that the learning rate will decay for each epoch within each cycle, and the decay speed will decrease as the cycle progresses. Experimentally we set $l_{max}^i = 10^{-4}$, $l_{min}^i = 10^{-6}$, $T_{mult} = 1.5$ and number of epochs as 500. In this manner we are achieving reasonable results. For more information about the learning rate, we refer to [128].

5.4 Experiments and results

The experiments are first applied to the 3D bright-field images of the 38 zebrafish samples to train and evaluate the RS segmentation network (cf. § 5.3.1). Both networks and different metrics are used for comparison to achieve the highest performance of the segmentation network. This network is then employed for the case study in tumour quantification (cf. § 5.2.4) as a test for automated RS detection and phenotype quantification.

5.4.1 Detection of 3D reference structures

In order to use the segmentation network for automated detection of the RSs from the 3D bright-field image, we first need to train and optimize the segmentation network with the 35 training samples. This means that the performance of the segmentation network needs to be evaluated first on the 3 testing samples with evaluation metrics, before it can be used further for 3D quantification. To this end, we introduce the evaluation metrics, followed by optimisation and evaluation experiments implemented on both RSs for this study.

1) Evaluation metrics

To evaluate the performance of different loss functions on both networks for each RS, we spilt the 38 samples into three sets: 28 samples for training the networks, 7 samples for validation and 3 samples for testing. The performance is compared by

applying five different evaluation metrics to the 3 test samples from individual prediction. The evaluation metrics we employ include the Dice similarity coefficient (DSC) [117] or *F*1 score, sensitivity, specificity, *F*2 score and area under the Precision-Recall curve, i.e. APR score [129]–[131]:

$$F1 = \frac{2TP}{2TP + FP + FN} \tag{7}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{8}$$

$$Specificity = \frac{TN}{TN + FP} \tag{9}$$

$$F2 = \frac{5TP}{5TP + FP + 4FN} \tag{10}$$

where *TP*, *TN*, *FP* and *FN* are the true positive, true negative, false positive and false negative rates, respectively. Sensitivity or recall, measures the proportion of actual positives that are correctly identified as such and it also quantifies the ability to avoid false negatives. Specificity or precision, quantifies the ability to avoid false positive. The *F*2 score is an effective measure for cases where recall is more important than precision compared to *F*1 that equally measures the recall and precision. To critically evaluate the segmentation performance of different networks for highly unbalanced data, in our case in particular for *Eye*, we use the APR score.

2) Detection of 3D Body reference structure

Compared to 2D, the 3D segmentation of the *Body* RS is more complicated. Specifically, 3D data contains both image information of each slice and correlation between adjacent slices. By using 3D based segmentation techniques such as 3D U-net, we can collect and extract, to some extent, both kinds of information. To this respect it offers more features compared to the 2D *Body* RS. However, because of the transparency of the specimen in the bright-field channel, the difficulty of segmentation in zebrafish, especially the surface, also increases; even more so from 2D to 3D. Figure 5.2 (a) gives us an intuitive idea of what transparency means in a single reconstructed slice of a 3D image. Due to the intensity similarity between background and transparent tissue, the difficulty of segmentation with transparencies in the specimen on a single slice can be assessed by comparing Figure 5.2 (a) and (b).

To investigate the ability of the U-net based segmentation network on transparent specimens, we evaluate both 2D U-net network excluding correlations between adjacent slices and 3D U-net network including correlations between adjacent slices. Both are trained based on the three different loss functions (cf. § 5.3.2). Figure 5.3 shows the results of a test sample achieved from different methods, i.e. two networks with three loss

functions, and the corresponding errors (yellow for *FP* and red for *FN*), compared to the ground truth. From these qualitative results, it is difficult to conclude which methodology performs best on the data. However, we observe that the loss function has an impact on the segmentation errors. From binary cross-entropy loss [116] to Tversky loss [119], regardless of different networks used, the *FN* error decreases whilst the *FP* error increases. This means that both segmentation networks have the highest ability to avoid *FN* errors when using the aforementioned Tversky loss, and the highest capability to avoid *FP* errors when using binary cross-entropy loss.

(a) Reconstructed slice   (b) Groundtruth of segmentation
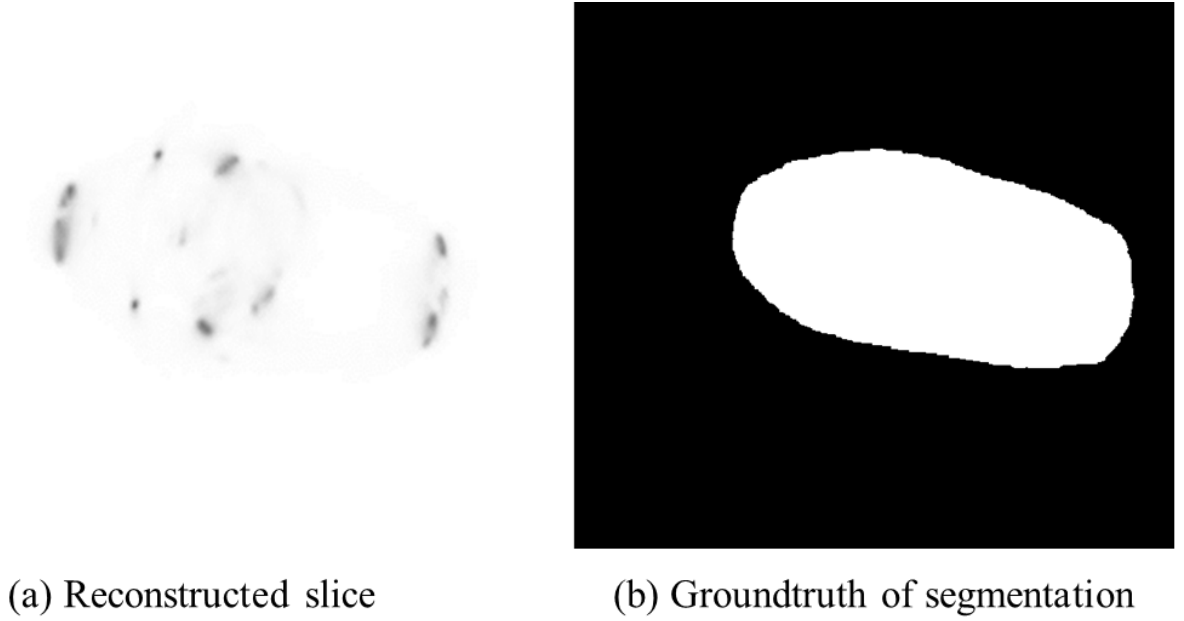
Figure 5.2. An Example for the reconstructed slice of 3D zebrafish image and the corresponding segmentation result for the slice of the *Body*.

To further quantify the performances of different methods and loss functions, five evaluation metrics are reported in Table 5.1 based on the average performance of the three tested samples. The best result of each individual metric across different segmentation networks and loss functions is presented in bold. Since in our work avoiding FP errors is equally important as avoiding FP errors, the combination of network and loss function that has the most bold results in the DSC and APR metrics, is regarded as the overall best performing method. By evaluating and assessing the results in Table 5.1, we concluded that 3D U-net with Dice loss has the best performance on the current dataset. It achieves the highest DSC/F1 and APR score as 93.9% and 88.4% separately.
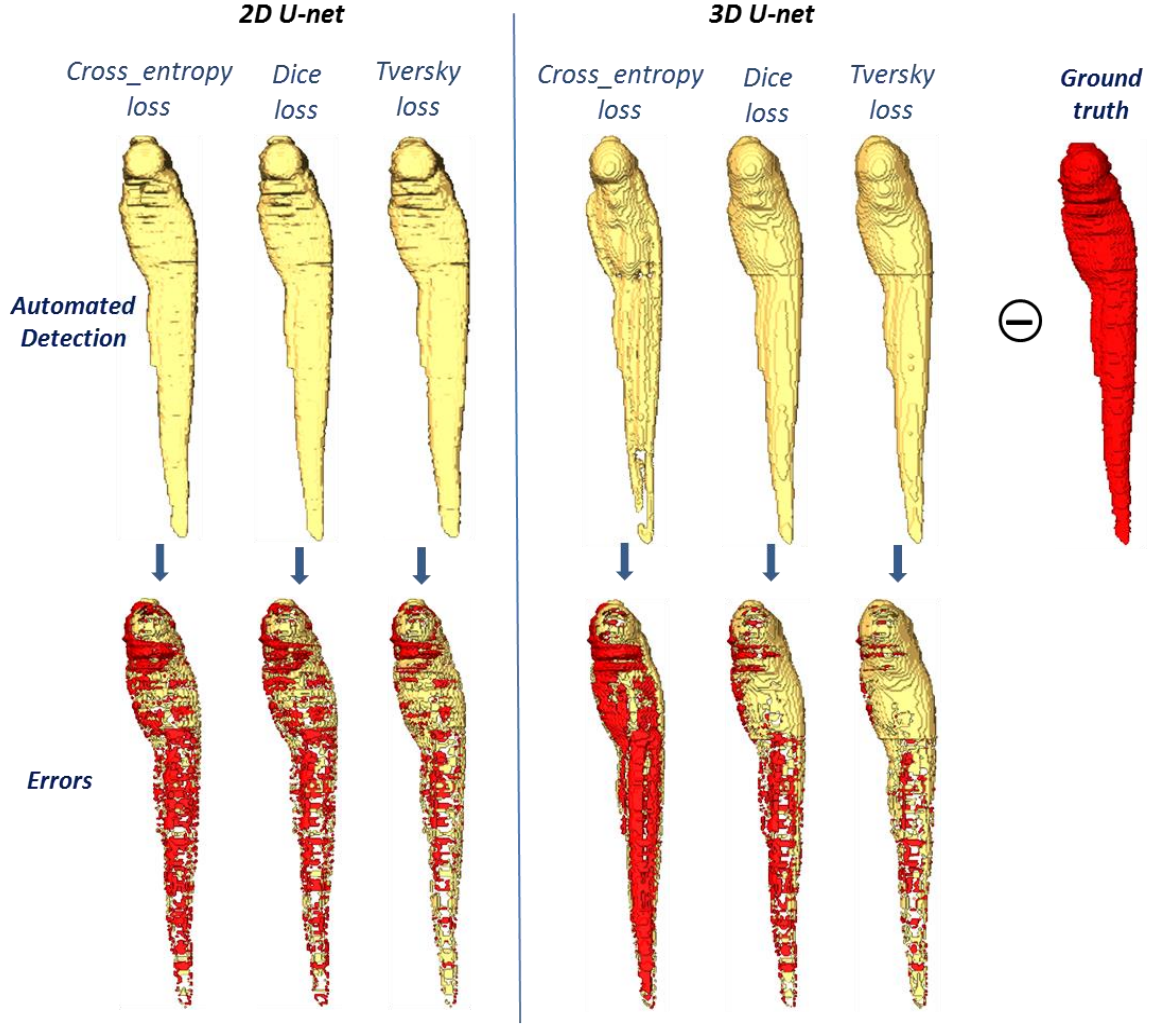
Figure 5.3. Comparison of the detected _Body_ (volume rendering), as a RS on a zebrafish sample. The results from different segmentation networks and loss functions are displayed on the top row, as well as the ground truth model manually labelled on the top right. On the bottom there are six errors (volume rendering), corresponding to different segmentation methods or metrics, with yellow showing _FP_ and red for _FN_.

3) Detection of 3D Eye reference structure

The tissues in _Eye_ are relatively dense, which guarantees a sufficiently distinct range of image intensities for _Eye_. One should, however, be aware that similar intensity patterns also exist in other parts of the zebrafish. A good segmentation method is supposed to discriminate _Eye_ from other tissues considering the 3D specific configuration of that shape. The experiments for _Eye_ segmentation are implemented according to the description in § 5.3. Ground truth labels for _Eye_ are obtained through manual labelling using annotation software; i.e. TDR [132] and Amira [63].

Figure 5.4 presents the comparisons of segmentation results on the same sample, when applying the 2D/3D networks and various loss functions. When assessing the segmentation results, we conclude that 2D U-net with Dice loss and 3D U-net with binary cross entropy loss fail to identify or segment the volume of _Eye_ correctly. 2D U-

net with binary cross entropy loss is able to identify parts of *Eye*, whilst with Tversky loss there is an over-segmentation of the volume. Similar to the *Body*, the best performing methods for the *Eye* are also 3D U-net with Dice loss, with much less errors both in the volume and on the surface. In order to identify the best segmentation performance for *Eye*, a quantified evaluation on the dataset is required.

Table 5.1. The average evaluation results of the *Body* detection on three test samples. We compare 6 methods including 2D U-net and 3D U-net across three different loss functions. In order to evaluate the results in a comprehensive and critical way, five evaluation metrics are employed. Referring to the visible errors in Figure 5.3, it is easier to understand the differences. The bold number represents the highest accuracy among 6 approaches for each individual evaluation metric.

Network	Metrics	Loss_function		
		Cross_entropy[116]	Dice [117]	Tversky[119]
2D U-net	DSC	92.80%	93.03%	92.47%
	Sensitivity	92.37%	94.20%	95.70%
	Specificity	**99.88%**	99.84%	99.79%
	F2	92.53%	93.70%	94.40%
	APR	86.37%	86.73%	85.83%
3D U-net	DSC	90.60%	**93.90%**	92.37%
	Sensitivity	90.53%	94.47%	**97.47%**
	Specificity	99.83%	99.87%	99.74%
	F2	90.50%	94.27%	**95.40%**
	APR	82.60%	**88.40%**	85.63%

Table 5.2. The evaluation of *Eye* detection on three test samples. Five evaluations with 6 approaches for segmentation are presented. The bold number in represents the highest accuracy among 6 approaches for each individual evaluation metric.

Network	Metrics	Loss_function		
		Cross_entropy [116]	Dice[117]	Tversky[119]
2D U-net	DSC	54.70%	0%	35.00%
	Sensitivity	43.00%	0%	**99.70%**
	Specificity	99.98%	100%	99.21%
	F2	46.87%	0%	55.47%
	APR	34.43%	0.20%	22.03%
3D U-net	DSC	0%	**91.80%**	90.64%
	Sensitivity	0%	90.78%	93.62%
	Specificity	100%	**99.99%**	99.98%
	F2	0%	91.17%	**92.40%**
	APR	0.20%	**84.53%**	82.40%
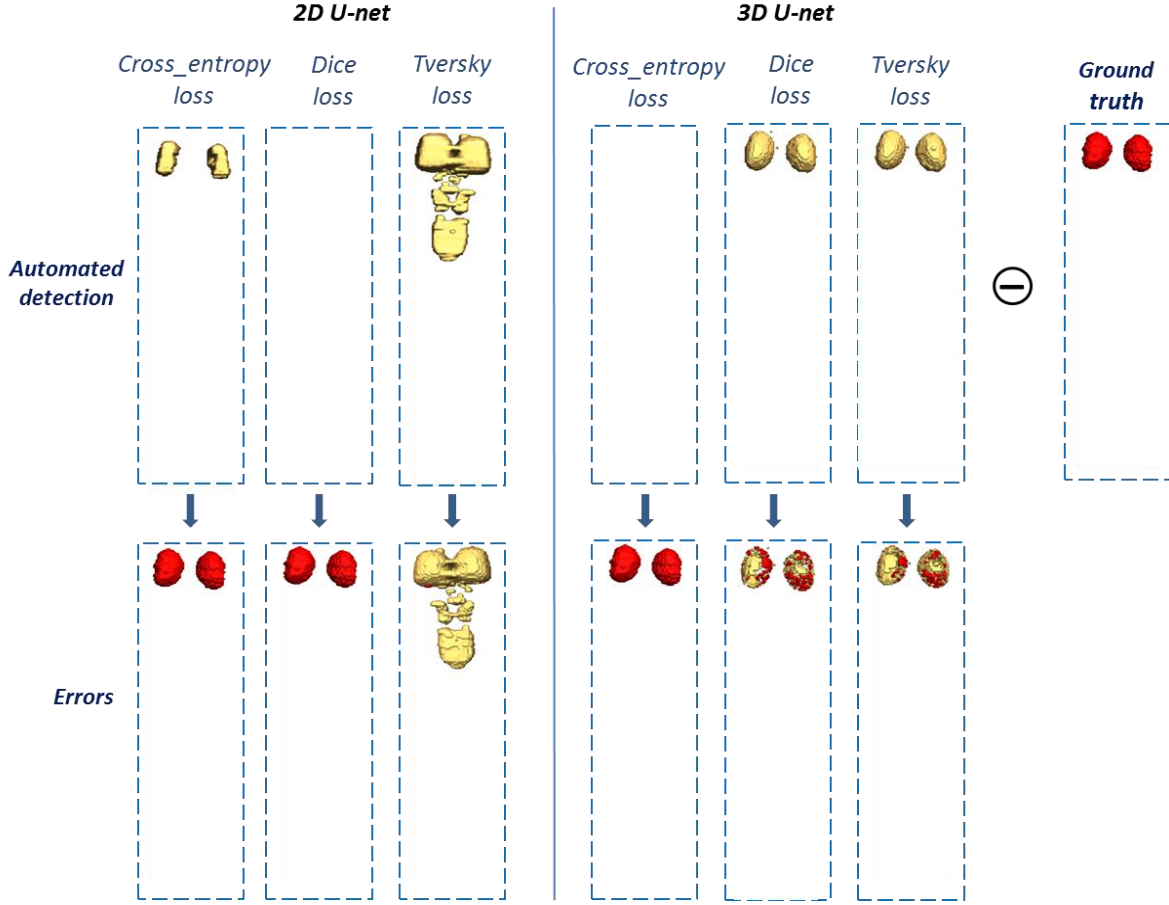
Figure 5.4. Comparison of the detected *Eye* (volume rendering), as a RS on the same zebrafish sample. Similar to Figure 5.3, the detected results and ground truth are presented on the top, while the segmentation errors are showing on the bottom with yellow showing *FP* and red for *FN*.

A qualitative report on the evaluation results on the three test samples is given in Table 5.2. Because the 2D U-net with Dice loss and 3D U-net with binary cross entropy loss fail to detect *Eye*, the accuracies of DSC, sensitivity, F2 and APR are 0%. But specificities are reported 100% because they are able to avoid FP errors. 2D U-net with Tversky loss performs best for sensitivity at 99.7%, but not best for specificity at 99.21%, meaning it has the highest ability to prevent FN errors but lowest ability to avoid FP errors. 3D U-net with Tversky loss has best performance for F2 score at 92.40%. However, overall we conclude that the 3D U-net with Dice loss performs best, achieving a DSC score of 91.80%, specificity of 99.99%, and APR of 84.53%.

5.4.2 Case study in tumour

Tumour growth and remission of neuroblastoma in zebrafish can be observed in a controlled experiment with and without a tumour inhibitor. The experimental condition includes a longitudinal exposure to the tumour inhibitor isotretinoin so that the relative quantification can be used for determining the tumour size at a specific stage of treatment. In this manner the measurements are independent of the variation in the individual

samples and imaging environment (e.g. exposure time and magnification). The performance of the treatment at that stage will be statistically averaged based on the quantification of multiple samples. In this section we describe the 3D relative quantification of tumour as a case study. Prior that, 2D relative quantification is explained for comparison.

1) 2D relative quantification with manual labelling

2D quantification provides a fast and intuitive view of fluorescent signals (tumour) and the RS, i.e. *Body* or *Eye*, in zebrafish in terms of a projection from the sample in both channels. We wish to obtain a measurement with which we can compare samples. Therefore, in 2D we normalize a tumour for a sample n by dividing the tumour area t_n by area of a RS f_n, achieving the area ratio $r_n = t_n/f_n$, which is depicted in Figure 5.5 (A) and (B) for a different RS. In the example of Fish 1, the area ratios from the two different RSs are separately 0.3694% and 3.5006%, with threshold-based segmentation for the tumour and manual labelling for the RSs. The performance of a treatment at stage i is determined by the average ratio of the N samples, $R_i = \frac{1}{N}\sum_{n=1}^{N} r_n$. A satisfactory segmentation of the tumour in the fluorescence channel is basically easy to achieve by using traditional threshold-based algorithms. However, for the segmentation of the RSs, more advanced methods [109],[98],[111] are needed.

When averaging the relative ratios of N samples, the treatment performance of all samples R_i should be calculated at the same projection angle, which is difficult to achieve. Another drawback of 2D quantification is that it fails to provide information about the shape of the tumour and the RSs. This limits its capability of representation in true 3D space. Therefore, we change to 3D quantification.

2) 3D relative quantification with manual labelling

Before 3D quantification, the OPT tomograms from individual channel are reconstructed to a 3D image, cf. § 5.2.2. This 3D reconstruction is required for the volumetric analysis. The pipeline of 3D relative quantification for each sample is similar to that of the 2D quantification, so is tumour segmentation in the fluorescence channel. The challenge of the 3D quantification lies in obtaining the volumetric RS. Specifically, the large number of reconstructed slices in 3D image makes manual labelling of a 3D RS impractical. Thus there is a demand for automated RS detection. Figure 5.6 presents an example of workflow for 3D tumour quantification at treatment stage i using the individual RS. In the next section, the state-of-the-art automated segmentation strategy based on convolutional neural network (CNN), cf. § 5.4.1, will be presented and implemented. In the same example of Fish 1, the 3D relative quantification results of tumour from the two different RSs are separately 0.1865% and 2.6962%.
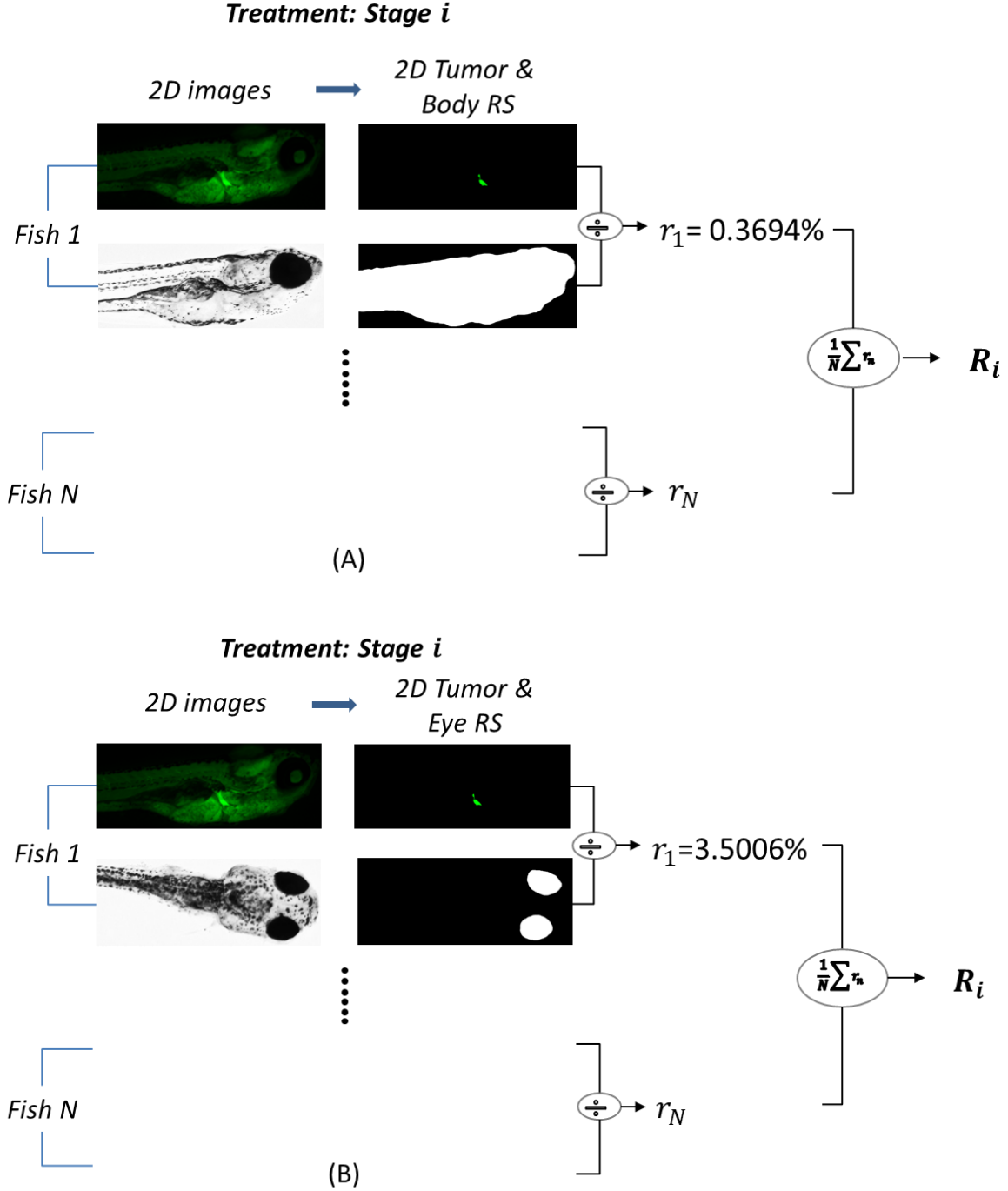
Figure 5.5. 2D relative quantification of tumour at a specific treatment stage *i* using *Body* (A) and *Eye* (B) RS labelled from the OPT tomogram. Each zebrafish sample is represented as a 2D image in the two channels (tumour in fluorescence channel and zebrafish structure in bright-field channel). The segmentation of tumour and the manual labelling of RS (*Body* or *Eye*) are used for the quantification. In the example of Fish 1 with $r_n = t_n/f_n$, the 2D relative quantification results r_n for the *Body* and *Eye* are separately $r_1 = 0.3694\%$ and $r_1 = 3.5006\%$. By averaging the 2D relative quantification r_n of N samples, the performance of treatment at time-point *i*, i.e. R_i is achieved.
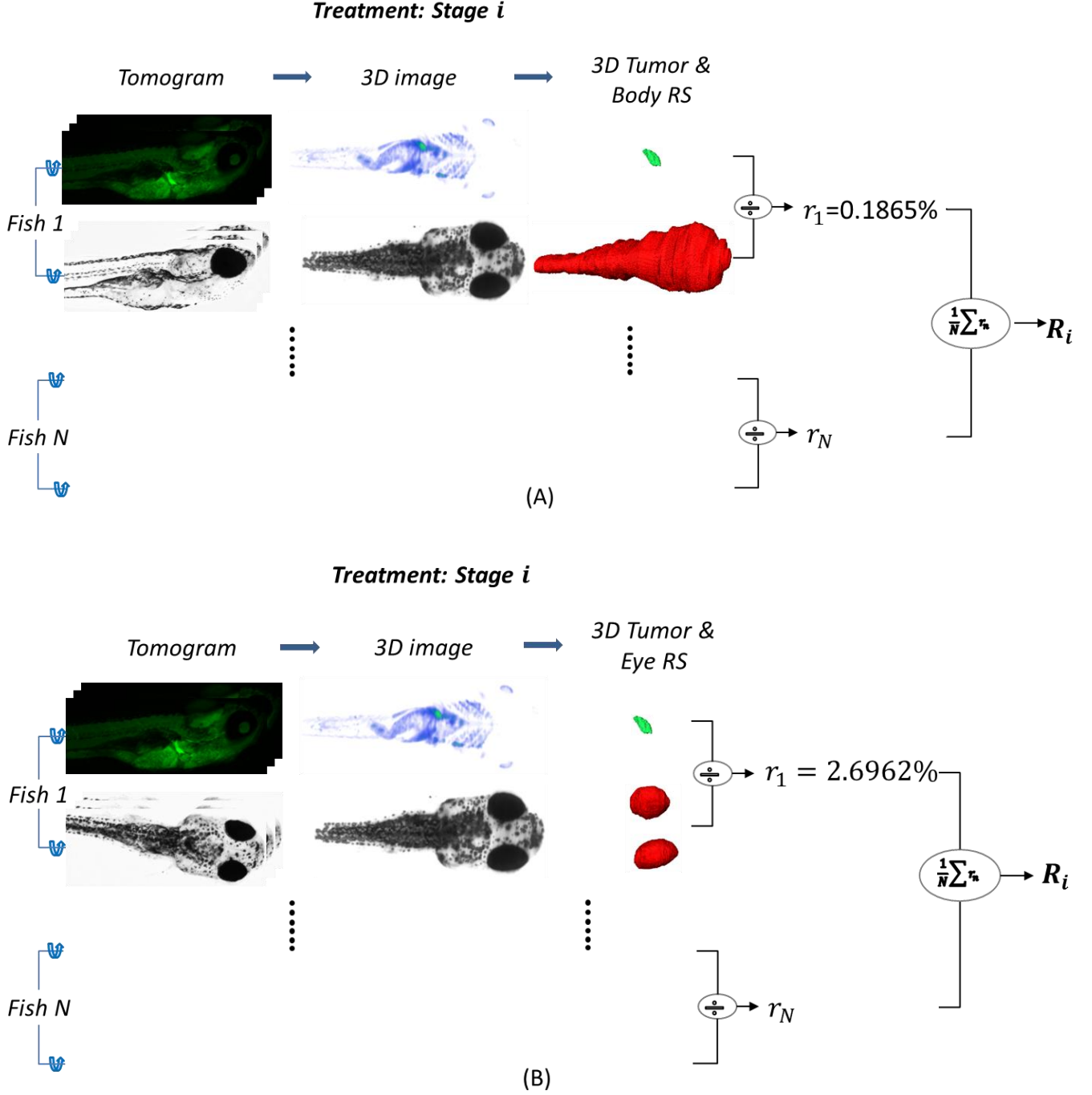
Figure 5.6. 3D relative quantification of tumour using two different 3D RSs. Each zebrafish sample is represented in two different channels, with the fluorescence channel for the tumour and the bright-field channel for the RS. (A) The 3D quantification based on the _Body_. (B) The 3D quantification based on the _Eye_. The 3D quantification is obtained by calculating the volume ratio between the tumour and RS. Specifically, in this example the 3D quantification results for Fish 1 are $r_1 = 0.1865\%$ and $r_1 = 2.6962\%$, calculated based on the two different RSs. As with 2D quantification the performance of treatment at time-point _i_, i.e. R_i is achievable by averaging the r_n of N different samples.

3) *Comparisons of automated detection and manual labelling of RS for 3D quantification*

Studies of tumour growth in zebrafish require measurements of the size and shape of the tumour. In our work the relative quantification of tumour is based on images in the fluorescent and bright-field channel (cf. § 5.2.4). Here we use the output of experiments on the automated detection method to automatically detect the RSs, i.e. *Body* or *Eye*, for a case study in tumour growth. Manual labelling a volume such as the *Body* normally takes at least 1-2 hours for one sample, while automated detection takes a few seconds. Figure 5.7 and 5.8 provide the comparisons of tumour quantification performance based on the RSs from automated detection and manual labelling. Because of its good performance for both RSs, we adopt 3D U-net to the tumour quantification for this case study. The sample compared in this section is a 25 *dpf* zebrafish and it is not included in the 38 samples mentioned before. The tail was not included in the imaging process because as a higher magnification is used, this results in an incomplete zebrafish. However, the segmentation method succeeds to detect the incomplete zebrafish even though it is trained and validated on the 35 complete zebrafishes.

For the *Body* RS, the automated detection using 3D U-net with Dice loss, achieves best overall performance when referring to the manual labelling result as the ground truth. Therefore, we use this segmentation network to automatically detect the *Body* in this case study. The relative ratios of the tumour referring to the *Body* from automated detection r and manual labelling r_f (§ 5.2.4), are separately 0.1883% and 0.1865%, with a quantification error of 0.9651%. The quantification error E_r explains the relationship between the FP and the FN error. A positive error in this case, i.e. the volume from automated detection is larger than the ground truth, means that the FN error is more than the FP error. With respect to *Eye*, the automated method has the best performance of 95.04% for DSC score, 93.77% for sensitivity, 99.98% for specificity, 94.27% for F2 score and 90.37% for APR, using Dice loss function. Dividing the number of tumour voxels by *Eye* volume, we obtain the relative ratio as $r = 2.7679\%$, larger than $r_f = 2.6962\%$ from manual labelling as the ground truth, resulting in the positive ratio error $E_r = 2.7261\%$. This is consistent with the fact that the FN error exceeds the FP error shown on the right of Figure 5.8.
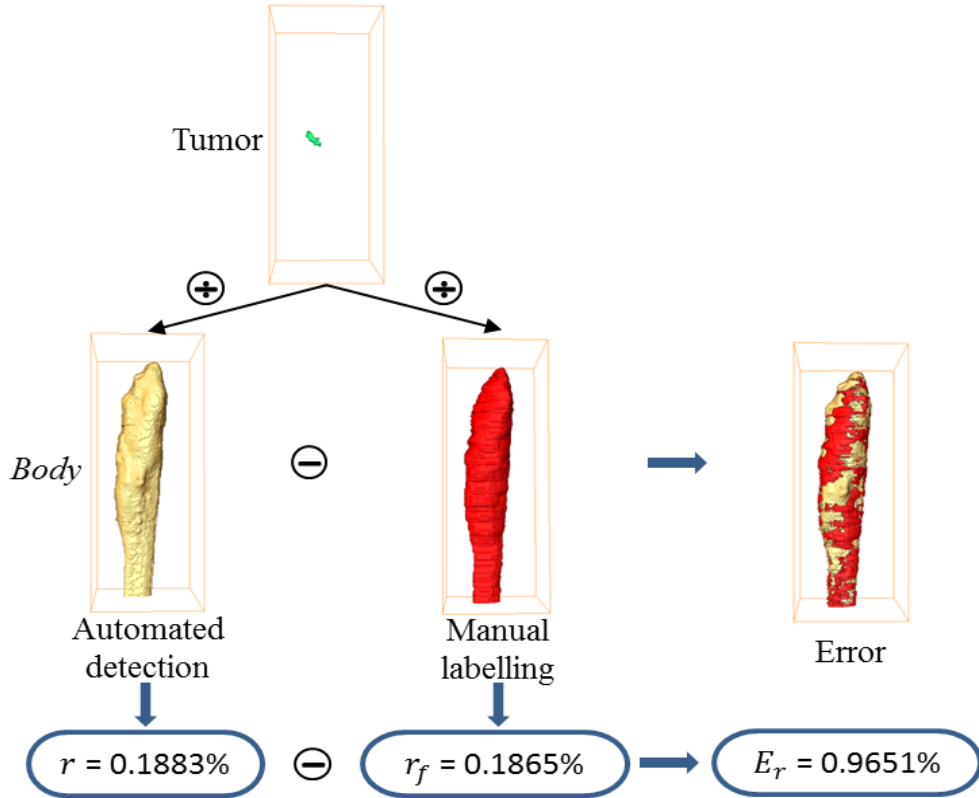
Figure 5.7. Comparison of tumour quantification based on the volumetric *Body* RS obtained from automated detection and manual labelling.
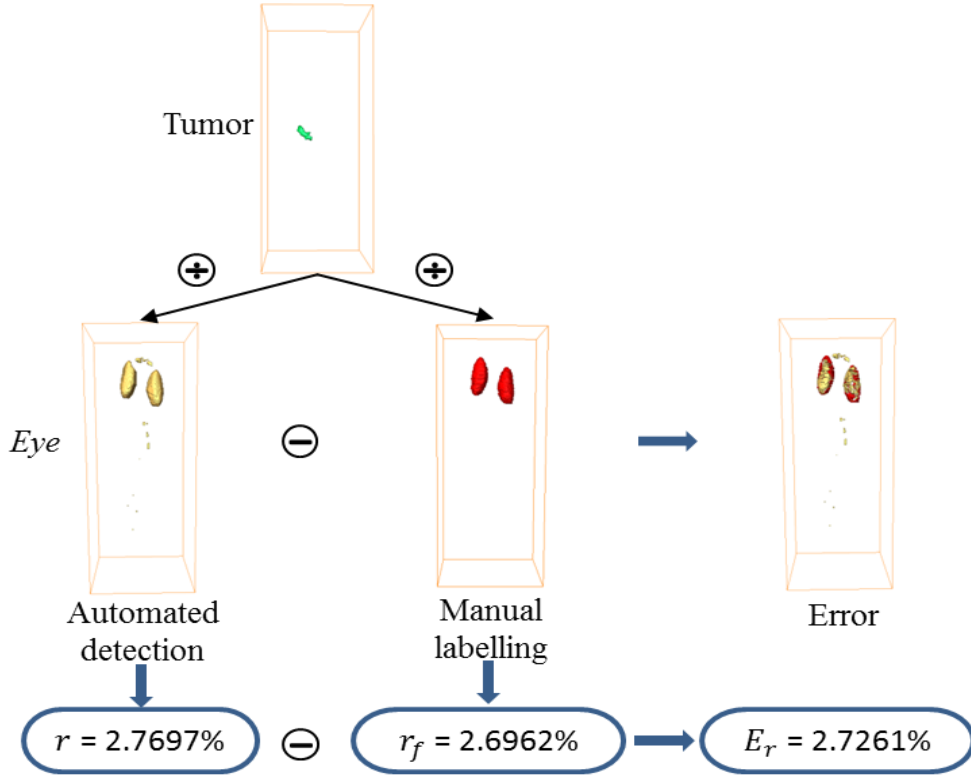

Figure 5.8. Example of tumour quantification based on the volumetric *Eye* RS obtained from automated detection and manual labelling.

5.5 Conclusions and discussion

In the previous sections, we introduced the concept of 2D and 3D relative quantification for fluorescent signals in zebrafish and focused on the technical solution of automated reference structure detection for the 3D quantification. In § 5.4, we compared the detection performances of different segmentation methods for both RSs and conclude that the 3D U-net segmentation network with Dice loss function performs best for automated detection of both RSs on the 38 samples. An overall promising accuracy of over 90% is achieved with respect to five evaluation metrics for both RSs. Subsequently, we compared the relative quantification of tumour between the automatically detected RSs and manually labelled ones. We further investigated how the segmentation errors influence the relative ratio r, compared to the ground truth ratio r_f. From our experiments it is shown that when FN exceeds FP this results in positive quantification error. Whereas, an FN smaller compared to FP results in a negative quantification error. The overall quantification error that we have established is 0.9651% for the zebrafish body RS and 2.7261% for the zebrafish eye RS. Given the experimental setting this is acceptable and reasonable. Nevertheless, given these acceptable outcomes we still can do the effort of further automation of the laborious manual labelling. Moreover, based on the results of the automated detection, further improvement can be accomplished by a careful manual error correction. In this case, the hybrid of artificial intelligence (AI) and human intelligence (HI) gains the best performance.

In this research project we focused on quantification of tumour growth, but the approach can be generalized to the quantification of fluorescent signals in zebrafish. Ideally, they are labelled with fluorescent markers for OPT imaging and reconstruction. With the promising results of automated detection on the limited dataset, better results can be achieved when training the network with a larger dataset. This way we can further improve the accuracy of automated detection of the RSs. Additionally, another contribution of this research is the introduction of a pipeline for relative quantification using automatically detected RSs. This pipeline can be transferred to high-throughput analysis of zebrafish. In the case study for tumour quantification, we presented and evaluated the pipeline for just one sample. Once more samples are available, we would continue with a statistical analysis of the performance of treatment using the proposed pipeline. This is motivated by the fact that statistical analysis of samples at either the same stage or different stages is getting increasingly important for drug discovery.

5.6 Acknowledgment

Chapter 5

Chapter 6

Exploration of 3D Structure Annotation and Visualization of Zebrafish Reconstructions from Optical Projection Tomography Imaging

Chapter summary

In this chapter we investigate the last parts of 3D optical projection tomography imaging pipeline: annotation and visualization. Our results focus on zebrafish samples without specific staining from the bright-field channel of the OPT. With respect to annotation, we first present the manual method from different software packages including our own annotation software. This provides a clue about how much 3D information within zebrafish can be obtained from the optical projection tomography 3D imaging system. Beside manual annotation, an automated 3D annotation method is explored to give an insight in the extent to which deep learning can automate 3D annotation process. With the results the limitations are analysed and new perspectives are presented.

6.1 Introduction

The zebrafish is an important vertebrate model organism that is widely used in biomedical research, e.g. developmental biology, disease biology, toxicology and behaviour. In this research, the identification and quantification of marker signals from genes and proteins have been providing basic evidence for gene analysis [133].

Detailed analysis of genes includes the identification of certain proteins in organs rather than looking at a whole-mount study of zebrafish [134]. Moreover, these organs can be physically dissected for an individual proteomic analysis [134]–[137]. Such techniques aim to answer the questions, e.g. is a gene X is expressed in an organ and how much of the corresponding proteins can be identified in this organ. From imaging, and multi-channel OPT in particular, expression patterns can be visualized *in situ* so that location and distribution of genes and proteins can be established.

Spatial analysis of patterns of gene expression from whole-mount zebrafish has been applied in the last decades [11], [138]–[141]. This has often been done using *in situ* hybridization (ISH). In this manner the localization of gene-expression has been accomplished. However, research on detailed analysis within organs is limited by the techniques available for the imaging of organs and subsequent annotation. To overcome these limitations, we explore annotation with experiments to detect zebrafish organs or volume regions from optical projection tomography (OPT).

In the application of imaging with OPT, we assume that patterns of gene expression can be visualized in the fluorescence channel through fluorescent *in situ* hybridization (FISH) or using a fluorescent reporter gene (e.g. GFP) . In general, the bright-field channel provides structural information for zebrafish and this includes some of the organs. Our research questions therefore are: (1) to what extent the organs or volume regions can be annotated from OPT 3D images of whole-mount zebrafish; (2) to what extent this annotation process can be automated.

In order to answer these questions, we have manually annotated 5 *dpf* and 25 *dpf* zebrafish from 3D images obtained from the OPT imaging system. In addition, automated annotation is explored on the 5 *dpf* zebrafish for the annotation of 4 labels. The results of the annotation are presented with different visualization software packages.

In order to set the terminology right for the remainder of this chapter, three definitions are introduced. First, we define (1) *The anatomical domain* as a demarcation of an anatomical structure and this structure is being generally accepted as a named structure in anatomy of a species. We consider (2) named structures to follow unambiguous labels that are organised in *controlled vocabularies* or *curated ontologies*. An annotation label for an anatomical domain is extracted from either of these curated namespaces. In addition to the anatomical domains we distinguish (3) *the volume region* as volumetric part of an organism presenting a dedicated region that is not coinciding with a recognized anatomical domain but it can be addressed for its named location.

With these definitions in hand we can make annotation in zebrafish with the intention to detect volume regions and anatomical domains that can be found in the bright-field 3D image without any specific contrast techniques being employed.

This chapter is structured as follows. In section 6.2 we will introduce the sample preparation, the OPT imaging system, the annotation methods and the visualization software. In section 6.3 we explain the annotation and experiment with visualization using different software packages. The results include both manual and automated annotation methods. Conclusions and discussion are summarized in section 6.4.

6.2 Methods and materials

In order to visualize and quantify a region of interest within a zebrafish, samples need to be first prepared and imaged. With the tomograms acquired from the OPT imaging system, 3D reconstruction is applied to obtain a 3D image. This 3D image is used for segmentation to find the structures for annotation, e.g. zebrafish *Eye* that can subsequently be visualized as a label. The workflow is shown in Figure 6.1.
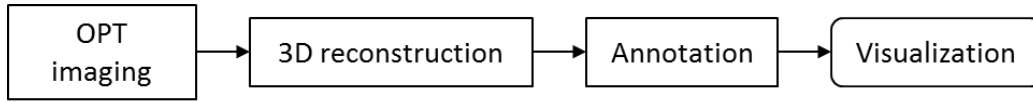
Figure 6.1. The workflow summary of OPT 3D imaging system.

6.2.1 Zebrafish and OPT 3D imaging system

During zebrafish development from larva to juvenile, structures and organs become more distinct and can be visualised using the OPT imaging system. In order to visualize and compare them, we use a 5 *dpf* and a 25 *dpf* zebrafish for our experiments. Each sample is prepared with the BABB or the CUBIC protocol (cf. § 4.4.2). The samples are prepared and imaged only with the clearing solution for the sole purpose of making the specimen transparent. No additional staining is used, meaning that the 3D structures of the zebrafish that we can obtain from our OPT imaging system represent the minimum information. With a counterstaining or a specific staining, the visualization of more structures can be accomplished. The OPT imaging system and reconstruction configurations as described in § 5.2.2 are used.

6.2.2 Annotation method

With the 3D image obtained from the reconstruction, manual annotation is first considered to provide a clue about how much 3D information within zebrafish can be identified without staining, using the OPT imaging on both zebrafishes. A manual annotation of the 3D image is accomplished with annotation software, e.g. Amira [63], TDR-3Dbase [12]–[14], Vaa3D [144], etc. In our work, Amira and TDR-3Dbase software are used for manual annotation. Aside from the manual annotation, we also explore the possibility of automated annotation for the visualization. The approach we used for the

automated annotation is semantic segmentation using the 3D U-net, which is set up according to chapter 5 (cf. § 5.3). Differently, multi-label segmentation rather than binary segmentation (cf. § 5.3) is required.

6.2.3 Visualization software

With respect to the descriptive parameters for phenotype characterisation of a volume region within zebrafish, visualization supports the qualitative information of the shape, size and context. Quantitative information is obtained from image analysis software. For visualization we use Amira [63], TDR-3Dbass [132], [142], [143] and MeshLab [145] software to visualize the individual regions in the 3D image, using several annotation methods.

6.3 Experiments

The first experiment is a manual annotation of zebrafish of different developmental stages and subsequent visualization of the annotated volume regions from the 3D OPT images. This aims to answer the question on the extent to which organs and/or volume regions can be annotated from a whole-mount zebrafish 3D image. The second part of the experiments focuses on the exploration of automated 3D annotation in 5 *dpf* zebrafish. As the 25 *dpf* zebrafish has more visible structures but fewer samples, the automated annotation of all parts is more challenging. Therefore, we start a simple with the 5dpf zebrafish that has much less distinct anatomical domains.

6.3.1 Manual annotation and visualization

1) Amira

In order to visualize a *5dpf* zebrafish in 3D or further do quantitative analysis on it, as a point of departure, the bright-field tomogram from the OPT imaging system is used. This tomogram consists of 400 images in our OPT setup. In Figure 6.2 (a) the tomogram is depicted that composes the input for the reconstruction algorithm. By using an effective reconstruction algorithm (cf. § 4.2), the tomogram is transposed to a raw 3D volumetric image of the zebrafish with a reduced amount of artefacts/noise. The raw 3D image is manually annotated and visualized using Amira [63] software (Figure 6.2 (b)). For volume region quantification in zebrafish annotation is required. Figure 6.2 (c) expresses an intuitive concept of 3D volume region segmentation or detection on a 5 *dpf* zebrafish. Obvious annotation labels are the zebrafish *Eye*, *Head*, *Muscle* and *Belly*; the latter three are location domains whereas Eye is an anatomical domain. The volume regions in this example are obtained based on manual segmentation of raw 3D image and subsequently visualized using Amira. Because of young age, transparency, and preparation method, most internal organs are not very well visible for zebrafish at this stage.
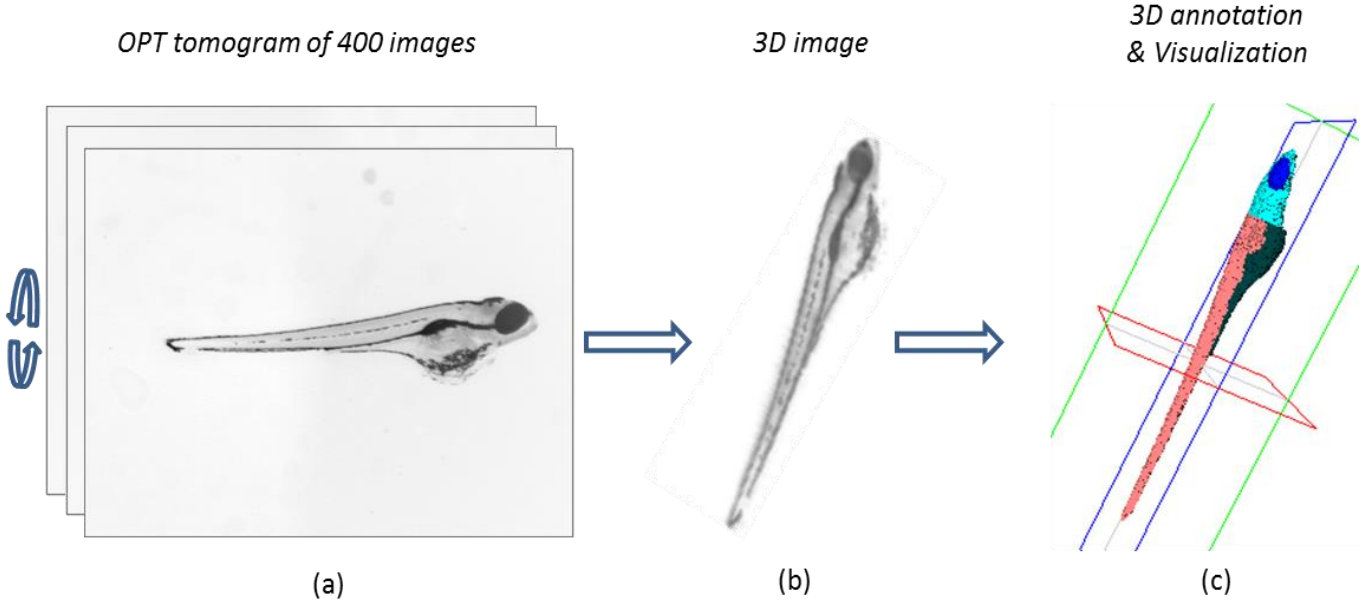
Figure 6.2. Workflow for 3D OPT imaging, reconstruction and segmentation on a 5 *dpf* zebrafish. (a) OPT tomogram of 400 images over 360° in OPT used for reconstruction. (b) Reconstructed 3D image from the 3D reconstruction algorithm. (c) Manually annotated 4 volumetric parts: *Eye* (blue), *Head* (light blue), *Muscle* (pink) and *Belly* (dark green).

Different from the 5 *dpf* zebrafish embryo in Figure 6.2, in Figure 6.3 the result of the imaging of a 25 *dpf* zebrafish with more internal structures is depicted. After reconstruction and semantic segmentation, the visualization is realized. This sample, i.e. tomogram, contains quite some noise, resulting in more background noise in the reconstruction as we can see from Figure 6.3 (b). The visualization results of the reconstruction are achieved by using a threshold in Amira as for 5 *dpf* zebrafish in Figure 6.2. The volume regions and anatomical domains we can distinguish from this sample are 9 different parts: *Eye* (blue), *Head* (light blue), *Muscle* (pink) and *Fin* (purple) in Figure 6.3 (c); *Blood vessel* (red) and *Belly* (dark green) in Figure 6.3 (d); *Notochord & Spinal cord* (yellow), *Swim bladder* (green) and *Intestine* (brown) in Figure 6.3 (e). As we can see, the *Swim bladder* has shrunk as a result of deflation. The volume regions and anatomical domains can provide significant information such as location or reference for fluorescent markers. It helps to understand volumetric structures of zebrafish in the OPT imaging system. Each individual part of the 25 *dpf* zebrafish is shown in Figure 6.4, using a surface rendering technique for visualization.

OPT tomogram of 400 images

3D image

3D
reconstruction

(a)

(b)

3D annotation v1

(c)

3D annotation v2

(d)

3D annotation v3

(e)
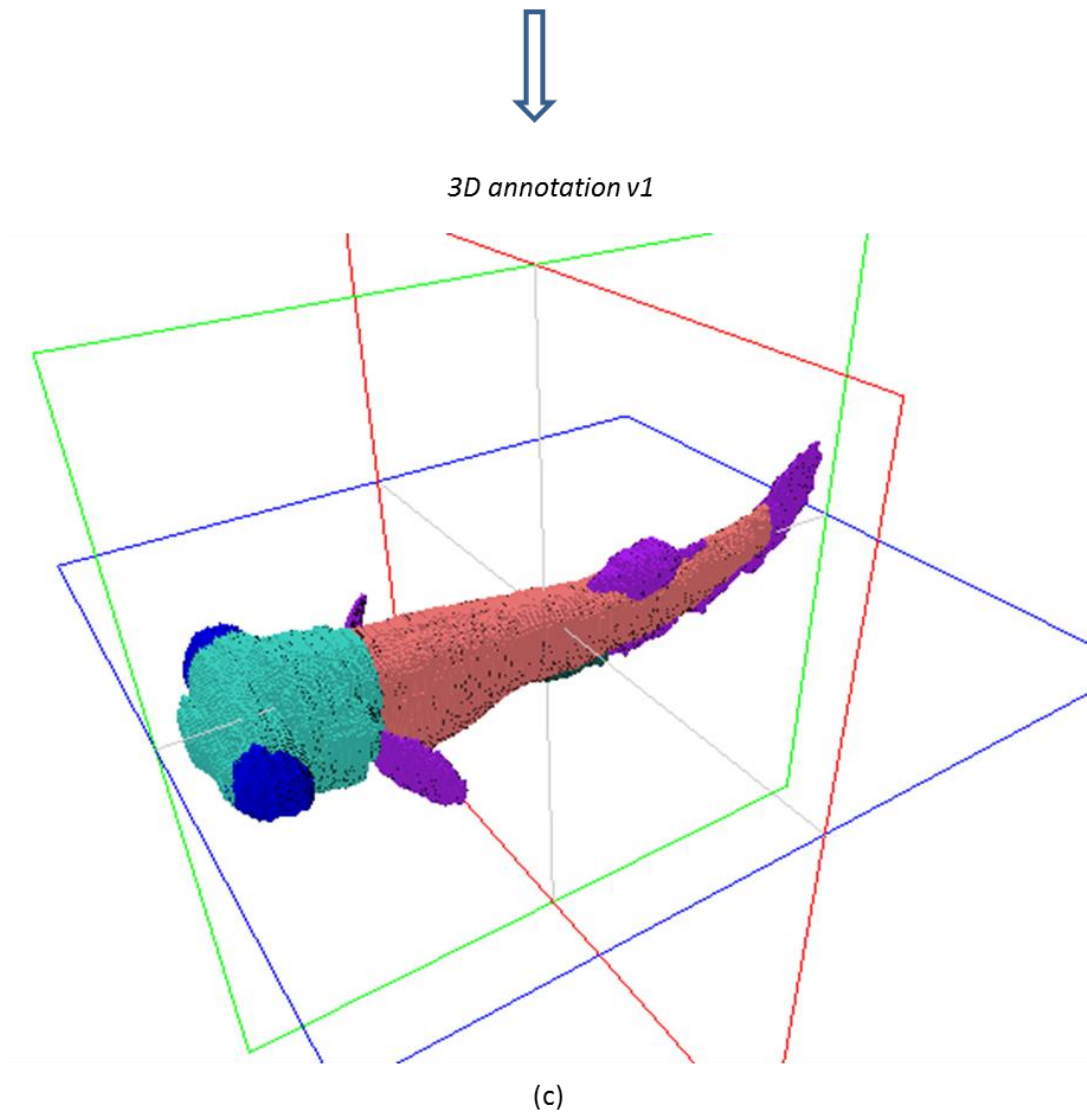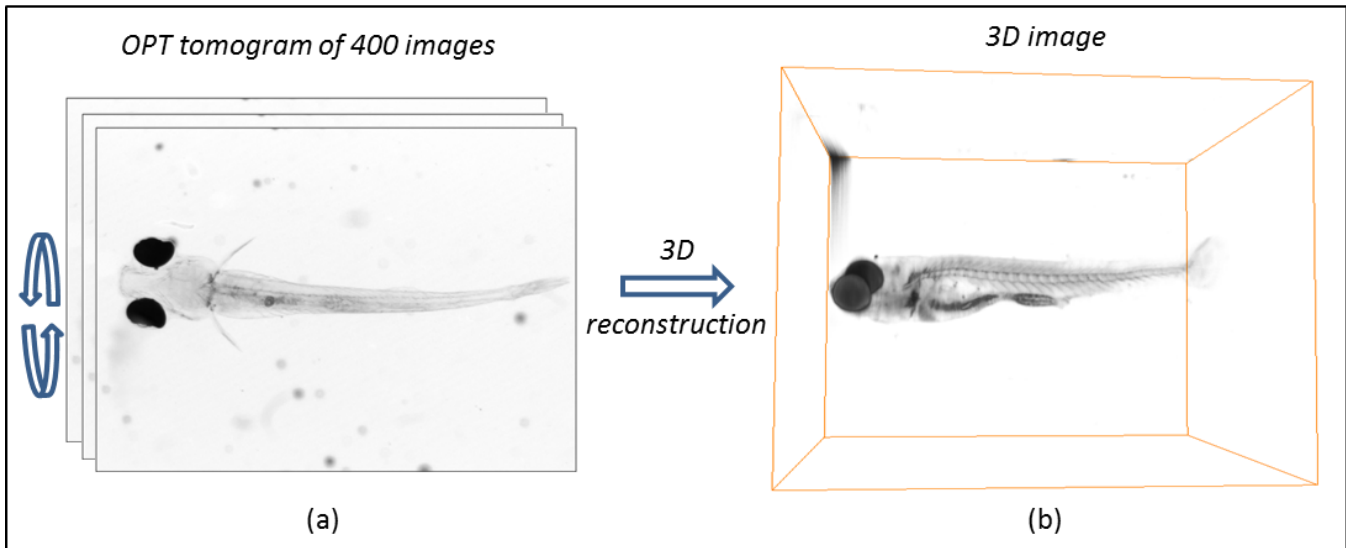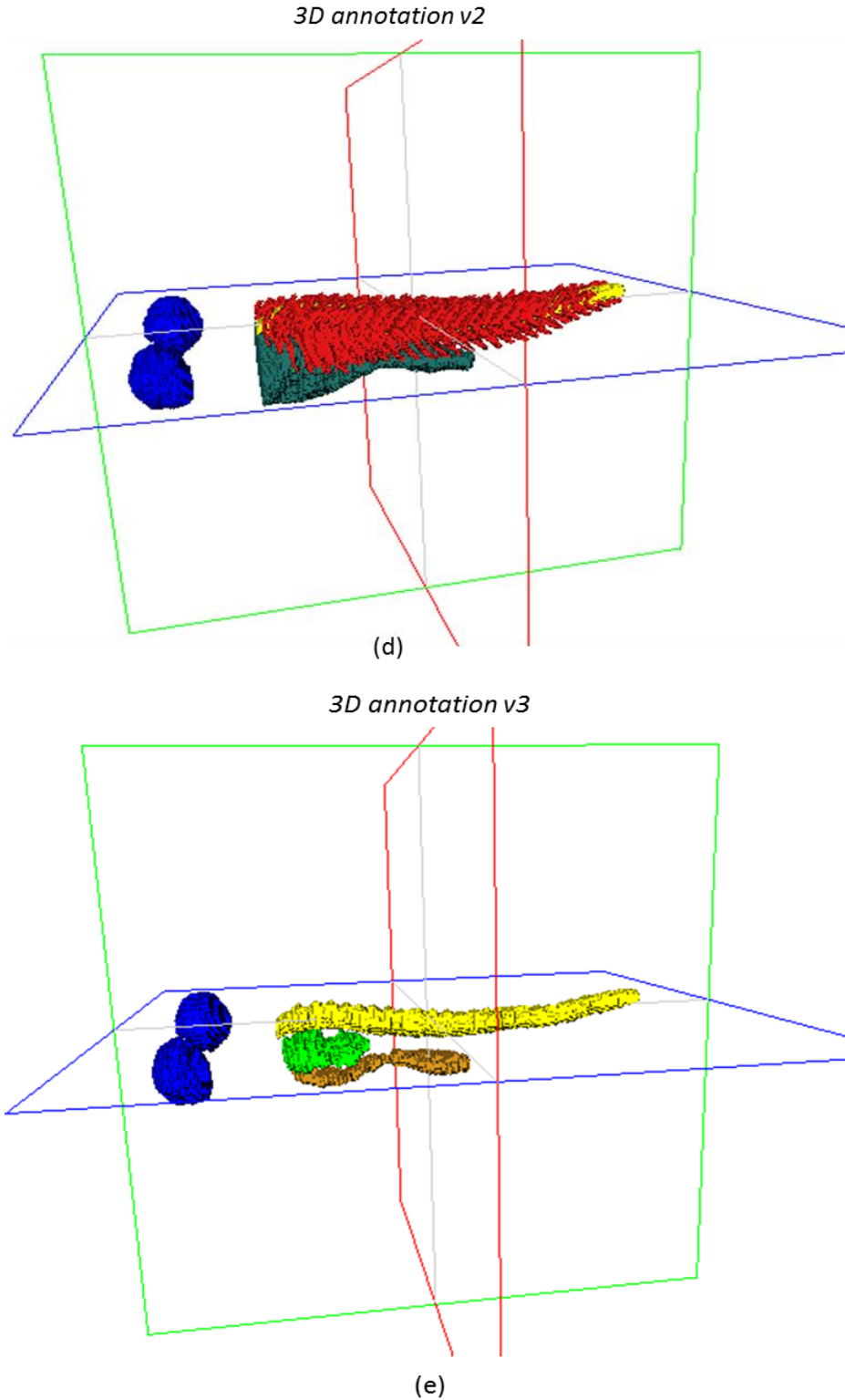
Figure 6.3. 3D OPT imaging, reconstruction and annotation on a 25 *dpf* zebrafish. (a) and (b) showing the projections and reconstruction respectively. (c), (d) and (e) visualizing 9 parts of zebrafish: *Eye*, *Head*, *Fin*, *Muscle*, *Blood vessels*, *Belly*, *Notochord & Spinal cord* (*N&SC*), *Swim bladder* (*SB*) **and** *Intestine*. **V1, v2 and v3 represent three visualizations with increasing depth from outside to inside, with v1 showing the** *Eye*, *Head*, *Fin* **and** *Muscle*, **v2 showing the** *Blood vessels* **and** *Belly*, **and v3 showing the** *N&SC*, *SB* **and** *Intestine*.

(a)

(b)
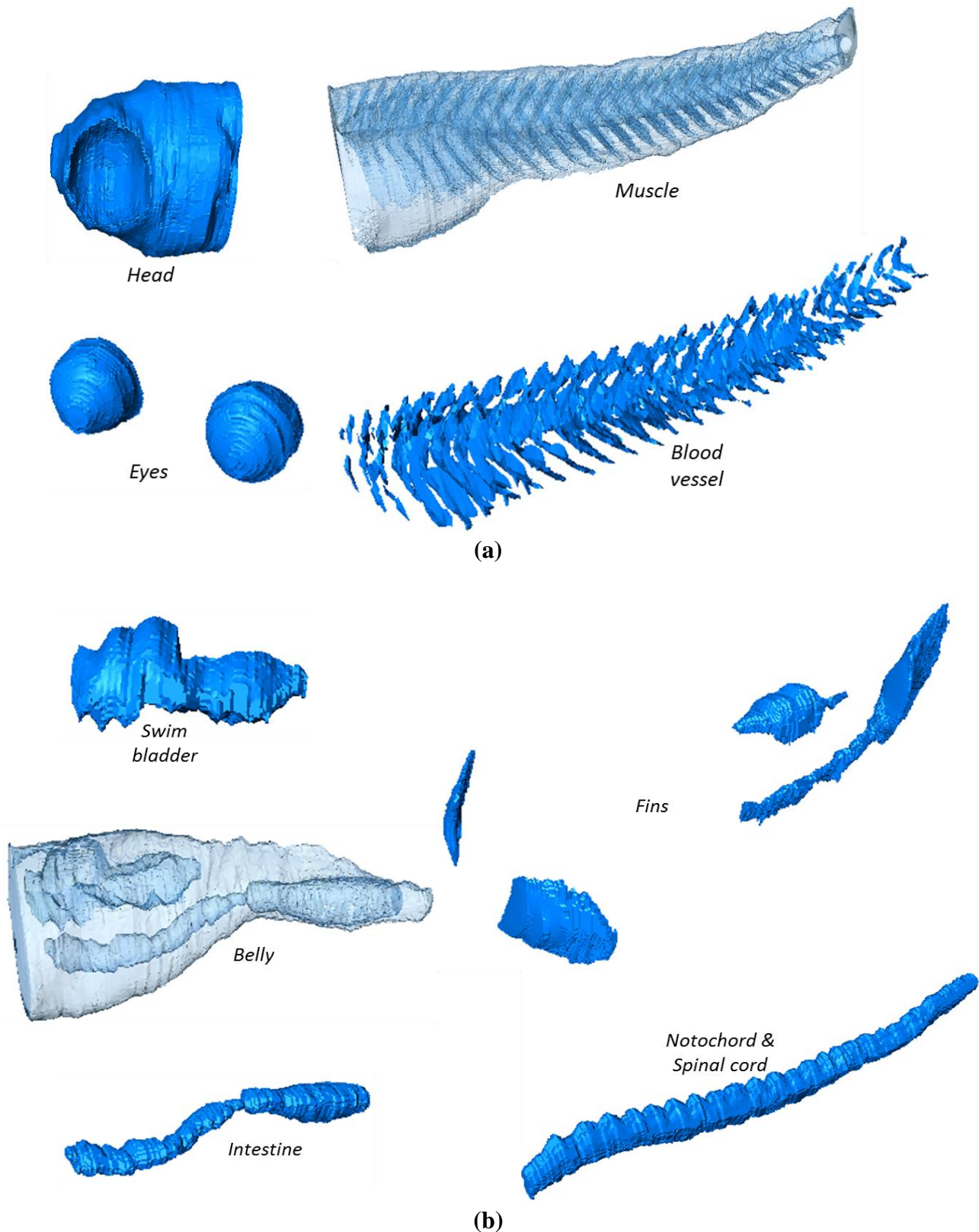
Figure 6.4. The individual part of the 25 *dpf* zebrafish. (a) The parts of *Head*, *Eye*, *Muscle* and *Blood vessel*, with *Muscle* being visualized transparently. (b) The parts of *Swim bladder*, *Belly*, *Intestine*, *Fin* and *Notochord & Spinal cord*, with *Belly* being visualized transparently.
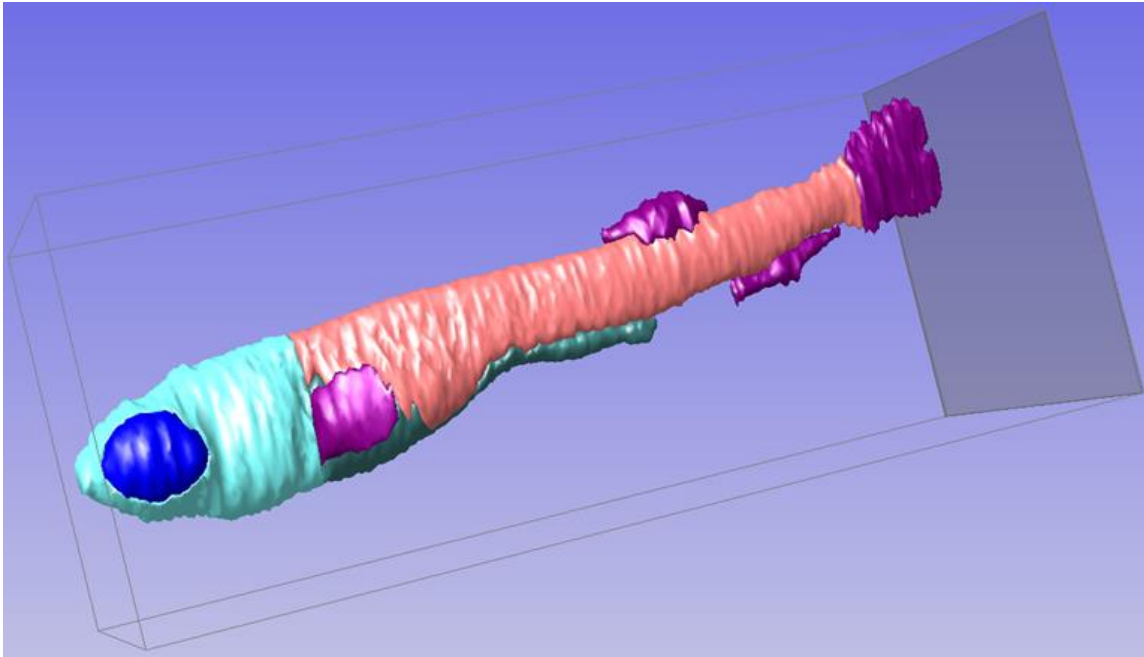
(a)
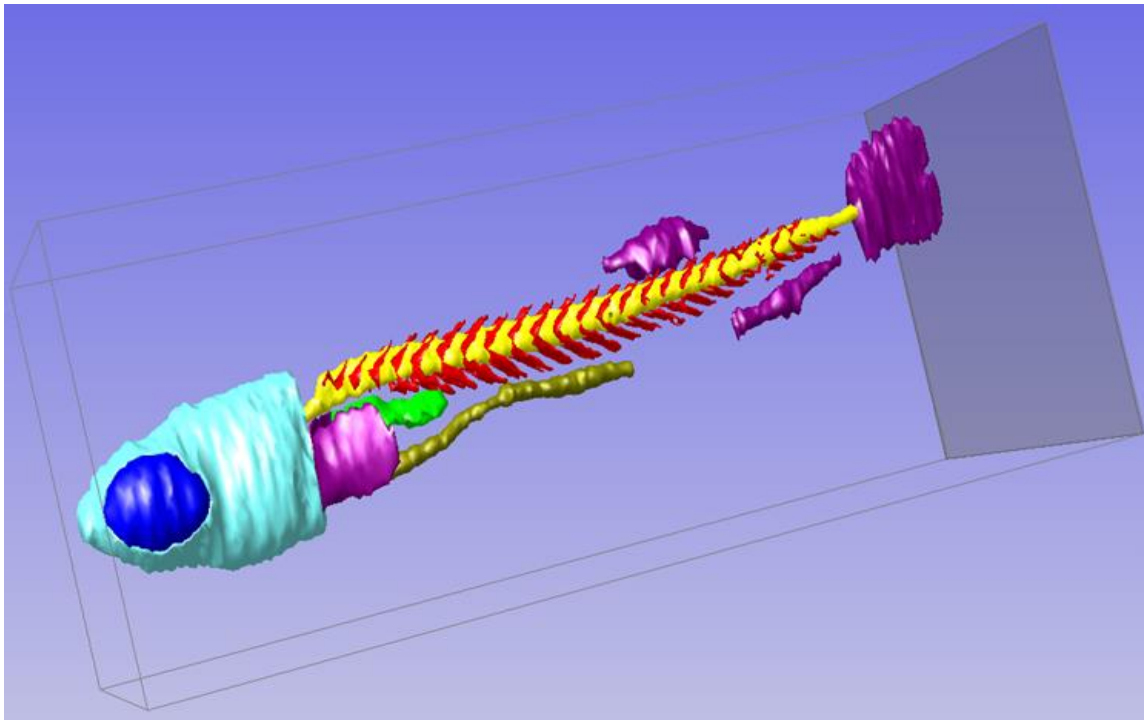
(b)

Figure 6.5. Annotation of another 25 *dpf* zebrafish from OPT 3D image using TDR-3Dbase software. In this whole-mount OPT imaging of the zebrafish, 9 parts are identified. (a) The parts of *Head*, *Eye*, *Muscle*, *Belly* and *Fin*. (b) *Blood vessel*, *Swim bladder*, *Intestine* and *Notochord & Spinal cord* are annotated within the zebrafish.

2) *TDR-3Dbase and MeshLab*

In Figure 6.5, another example of the annotation on a 25 *dpf* zebrafish from OPT 3D image is depicted. The annotations are done with the TDR-3Dbase [132], [142], [143] software. With TDR-3Dbase, manual annotation is well incorporated through the use of an LCD-

tablet with a pen (Wacom Cintiq). Similar to the previous results, nine labels are identified including *Eye* (blue), *Head* (light blue), *Muscle* (pink), *Belly* (dark green), *Fin* (purple), *Blood vessel* (red), *Swim bladder* (green), *Intestine* (brown) and *Notochord & Spinal cord* (yellow). As we can see from the example in Figure 6.5, TDR-3Dbase provides an informative and smooth visualization from a triangulation employing a surface rendering technique. Further refinement can be accomplished in other software. The annotation result of TDR-3Dbase is then transferred to MeshLab [145]. Now, the whole-mount zebrafish can be visualized with the volume regions (*Head*, *Muscle*, *Belly* and *Fin*) being transparent. This is depicted in Figure 6.6.
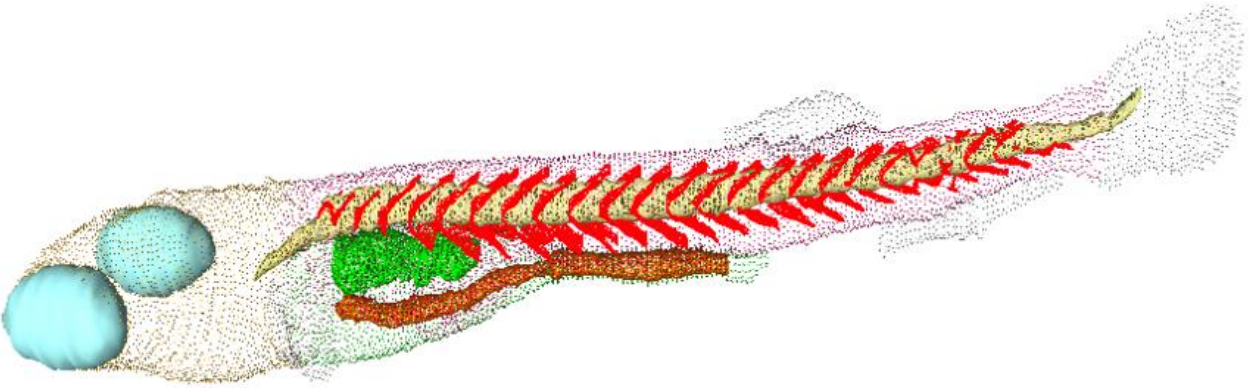
Figure 6.6. Visualization of the 25 *dpf* zebrafish using MeshLab software. The *Head*, *Muscle*, *Belly* and *Fin* are visualized as transparency in 4 different colours, whilst the *Blood vessel*, *Swim bladder*, *Intestine* and *Notochord & Spinal cord* are visualized with iso-surface technique. The mesh is obtained from TDR-3Dbase.

6.3.2 Automated 3D annotation of 5 *dpf* zebrafish

From the motivation of accelerating image analysis for zebrafish from biomedical experiments, segmentation of volume regions is explored. This has not been described in earlier literature. Inspired by the promising results achieved with 3D U-net segmentation network for binary tasks as applied in Chapter 5; i.e. on the zebrafish *Muscle* and *Eye*, we are interested in the performance for multi-label segmentation.

Different from the segmentations in Chapter 5, for a multi-label segmentation, i.e. the four volume regions we are using, the segmentation outputs four competing segmentation maps or classes excluding background. Voxels are labelled as one of the four classes that produce the highest prediction value from the segmentation network (cf. § 5.3). If this highest value is smaller than a specified threshold, the voxel will be recognized as background. The four prediction maps from the segmentation network, relevant to four volume regions accordingly, are trained from the manually labelled zebrafish in a

supervised way. The high class imbalance, e.g. *Eye* and *Muscle*, however typically results in a preferred segmentation for class with more positive samples, i.e. voxels. The learning schemas for this four-label segmentation are set similar to the approach of Chapter 5. Based on the results of Chapter 5, we use Dice loss [117] to train and validate the network on the same 35 zebrafish and test it with 3 zebrafish. The results are shown in Figure 6.7 and Table 6.1. In Figure 6.7 the qualitative performance between manual and automated segmentation on the three test zebrafish is compared. In Table 6.1 the quantitative evaluation with the five metrics similar to those used in Chapter 5 (cf. § 5.4.1) is presented. In each of the test samples, the performance of each volume region is quantified. By comparing the performance of different volume regions, it shows that the volume region *Muscle* has the highest segmentation performance with an average accuracy of 70.39% on the 3 test samples. Limited by the high class imbalance, i.e. very different numbers of voxels for the different parts, the network fails to detect the volume regions for *Eye* because it has much less voxels for the classifier.
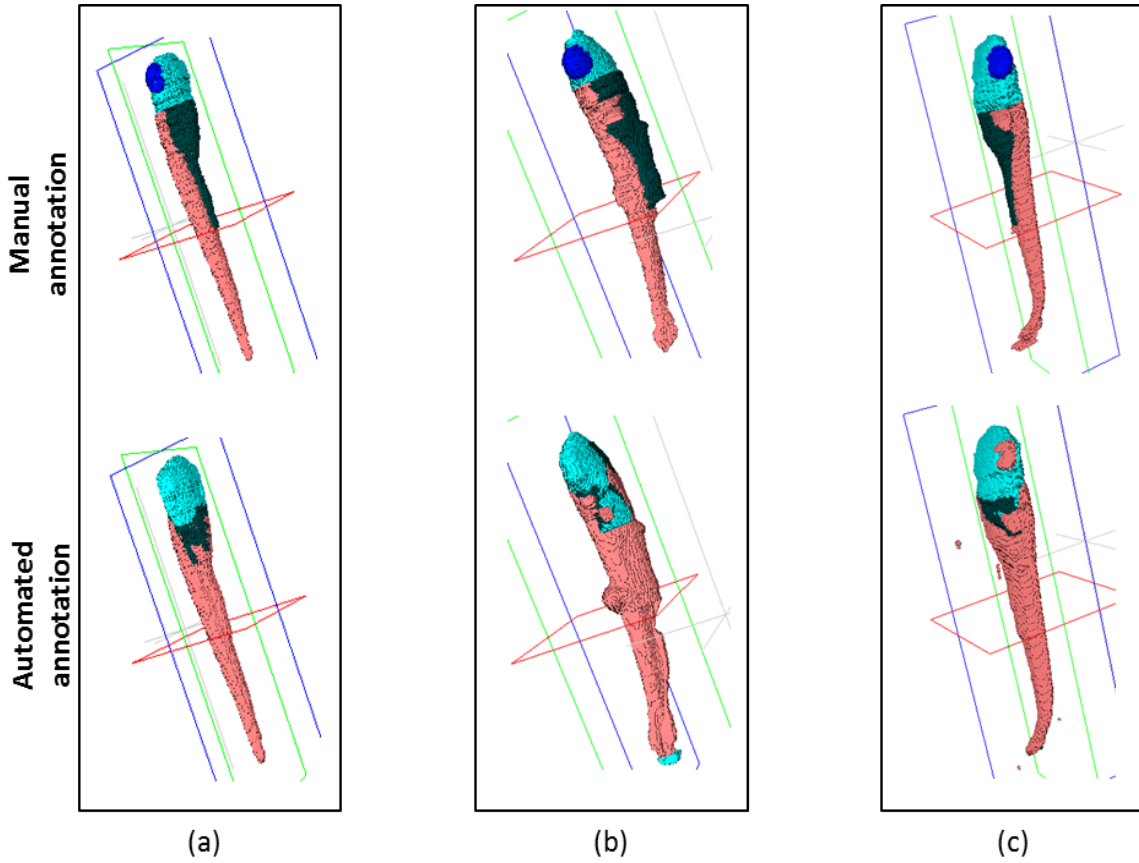

Figure 6.7. Comparisons between the performance of manual and automated segmentation on 4 parts of the 5 *dpf* zebrafish: *Eye*, *Head*, *Muscle* and *Belly*, where (a), (b) and (c) correspond to three different test samples respectively. The manually labelled models are displayed in the top row and the automated labels are in the bottom row.

Table 6.1. Quantitative evaluation of the three test samples of zebrafish. The results are obtained using the multi-label 3D U-net segmentation network. The evaluation metrics are the same as used in Chapter 5. The best performed volume region for each test sample across different metrics is printed in bold.

	Fish1(%)				Fish2(%)				Fish3(%)			
	Muscle	*Belly*	*Head*	*Eye*	*Muscle*	*Belly*	*Head*	*Eye*	*Muscle*	*Belly*	*Head*	*Eye*
DSC	**78.43**	47.22	70.82	0	**63.58**	40.15	29.64	0	**69.15**	30.12	64.41	0
Sensitivity	**91.34**	36.90	87.96	0	**85.10**	41.14	25.01	0	**96.36**	28.73	57.76	0
Specificity	99.69	**99.91**	99.66	1	99.46	99.58	**99.83**	1	99.49	99.76	**99.85**	1
F2	**85.70**	40.44	80.20	0	**74.95**	40.74	26.68	0	**83.26**	29.27	60.25	0
APR	**62.84**	24.46	52.20	0.18	**43.29**	16.51	9.39	0.17	**51.98**	9.36	42.34	0.17

6.4 Conclusions and discussion

In this chapter, we explored the last parts of our OPT imaging pipeline, the annotation and the visualization of 3D images. The outcome can be further used for phenotype characterisation. We presented examples to show to what extent the organs or regions can be annotated from a whole-mount zebrafish based on both manual and automated segmentation. With the annotation software packages, e.g. TDR-3D base [132], [142], [143], Amira [63] or Vaa3D [144], the advantage of manual annotation is that it can provide accurate segmentation including the expert-knowledge. The process is, however, quite laborious. Achieving comparable segmentation accuracies with automated annotation is challenging. Nevertheless, it can save a huge amount of labour and time. For instance, for a 3D OPT bright-field image of one single zebrafish, manual annotation normally takes hours to days depending on the level of detail. The automated annotation takes several seconds and sufficiently reasonable results are achieved. In practice, the combination of automated and manual method is recommended for an organ or volume region annotation.

With the best parameters and configurations achieved from the segmentation method described in Chapter 5, the performance for automated multi-label annotation task is not yet satisfactory. In future work further improvement need be accomplished by improving segmentation method and using larger set of images.

Based on the results, we analyse and discuss the reasons from the perspective of the data, algorithms and resources, and give further insights into performance improvement for each individual aspect. Furthermore, possible directions for further improvement are given.

1) **Data limitation**: We categorize the data limitation into three types: data characteristics, data quality and data size. Data characteristics represent the inherent structure of zebrafish as represented in a computer and it is further determined by using a specific 3D imaging technique. The data characteristic is, in proactive, determined by individual samples and imaging techniques. It defines an upper limit of the performance of segmentation of a given algorithm. Because of

the transparency of zebrafish, 3D segmentation of boundary between some parts, e.g. *Head* and *Muscle*, *Muscle* and *Belly*, in itself is challenging. The image quality related to noise in the background is segmented as foreground, with an example shown in Figure 6.7. (c). Noise is often introduced during sample preparation. Any remains, dust or bubbles during the sample preparation could result in artefacts in the 3D image. Because of the high intensity similarity between noise and some parts of zebrafish in this example, 3D U-net segmentation network fails to distinguish them when feeding the network with volume patches. In order to decrease the segmentation error resulting from noise, a cautious sample preparation process is required. With respect to data size, the performance is achieved with 35 training and validation samples. In our view this is still insufficient for training a good 3D segmentation network. To improve the segmentation performance, more zebrafish samples will be added in the next steps.

2) **Algorithm limitation**: The huge difference of performance between different parts/classes is raised by the class imbalance. The segmentation can be treated as a multi-label classifier in 3D U-net, with each voxel representing a classification sample. The segmentation task is interpreted as classifying all the voxels of a 3D image into 4 classes with a probability assigned to each class. The network fails to segment or detect the *Eye*, but it performs reasonably for *Muscle* segmentation or detection. This is probably because in all the training voxels for classification, the number of voxels as *Muscle* is much larger than that as *Eye*. This gives the *Muscle* class an unfair advantage over *Eye*. Further research and exploration that aim to solve this problem include data resampling and loss rebalancing. In data resampling, voxels of lower frequency class can be repeatedly sampled until all the classes reach to an approximate balance. Another possible solution for class imbalance would be assigning each voxel with a class weight (i.e. reciprocal of class frequency) and integrating it into prediction when implementing loss function. For both possible solutions the improvement performance need to be further verified because of the high imbalance between some classes. In 3D and multi-label classification, this imbalance research problem has become a hot and challenging topic in the recent years in medical imaging field for MRI/CT images.

3) **Resource limitation**: Constrained by the limitation of the GPU memory, the training and validation batches are fed into the network. In this way the structural and spatial information among patches are not taken into account for training. This means the smaller the patch size is, the less structures will be trained and learnt, resulting in lower performance of segmentation. In our segmentation task of four labels, the patch size is set as 64x64x64, while the image size is 128x128x340 with GPU memory size of 12GB. An increase of the GPU memory can be another improvement to the segmentation accuracy. This generally means a change of the GPU setup and it comes with a cost that need be assessed. However, given the number of different images this might be a good investment.

All considered, the research described in this chapter gives insights in future development of automation of the annotation process. In some cases, for very specific structures, a manual annotation would still be the most efficient manner of obtaining an accurate result. From the visualization examples it can be assessed that each of the software packages has its own advantages. But with a dedicated annotation software package like TDR-3Dbase [132], [142], [143], good results can be accomplished while refinements in the visualization can be accomplished with other software packages.

We have also given our considerations the respect to limitations that are imposed on the data. With time these might change, however, when method improvement is considered it is good to be aware of these limitations in an assessment for further developments.

6.5 Acknowledgement

Chapter 6

Chapter 7

Conclusions & Discussion

Chapter summary

In this thesis, we have addressed seven research questions regarding four themes: 3D imaging, 3D reconstruction, 3D segmentation and applications in the biomedical domain. In this chapter we concisely answer and summarize these seven research questions from the research presented in the previous chapters. Subsequently, limitations that we have found are addressed and possible solutions with respect to data, algorithms and theory are discussed. Lastly, ideas on future research are presented.

7.1 Main contributions

The main contributions of the work presented in this thesis can be summarized by answering the seven research questions as follows:

RQ1: To what extent is it possible to increase the processing speed of OPT imaging and reconstruction in an integrated manner?

We have made an attempt to make the process of tomogram reconstruction convenient for users by offering an efficient and reliable way of 3D imaging for OPT. Our OPT system enables to acquire a tomogram and reconstruct a sample in the millimetre scale, e.g. zebrafish larvae, in a few minutes. By using the OPT reconstruction software as presented in Chapter 2 (cf. § 2.2.2), an OPT tomogram is uploaded to our computer cluster for reconstruction. The computations for reconstruction from tomogram to 3D image are parallelized over the cluster through a smart scheduling schema. With the current image size, the time for reconstruction is around one minute using 5 compute nodes of 8-core 2.66 GHz CPU+16G RAM and 8 nodes of 4-core 2.66 GHz CPU+16G RAM. The exact time used for a specific sample is determined by the CPU resources available in the cluster as well as the image intensity distribution of a sample. OPT users will receive the reconstructed data through a web interface (link) provided by the software after the completion of the computation. By the increased sample throughput, more samples can be processed and thereby the integrated system brings facilitates the statistical analysis of the biological samples.

RQ2: To what extent is it possible to reduce the artefacts of 3D image introduced during reconstruction process by misalignment of Centre of Rotation (CoR)?

In the exploration of OPT imaging and reconstruction, we found that the misalignment between the CoR and the image centre normally introduces ring artefacts around the object edges of the reconstructed slices. This is independent of the reconstruction algorithm used. This means that the prerequisite for a good quality reconstruction, i.e. the 3D image, is to correct for this misalignment. We investigated how to apply for this correction and we have presented a fast and accurate CoR correction algorithm, cf. § 2.3.1. The algorithm is implemented on the OPT tomograms and it corrects the sinograms in a straightforward way. Similar to the 3D reconstruction, the CoR correction is also implemented in a parallel manner. By using the proposed CoR correction before reconstruction, the ring artefacts are effectively eliminated from the reconstructed slices. Besides the qualitative comparison of the ring artefacts, quantitative improvement is analysed and evaluated cf. § 2.4.1 and § 2.4.2.

RQ3: Can the Point Spread Function (PSF) of the OPT imaging system be modelled and applied for deblurring of an OPT reconstruction?

This question is answered in Chapter 3. It is of interest to know the relationship between projection blur and imaging depth, therefore we propose a protocol to acquire a tomogram image set of a single fluorescent bead. Our protocol accommodates for a decrease in the probability of overlap between different point sources in a full 3D revolution. We model the PSF using a generalized 3D Gaussian model. We have simplified this model in a workable manner and relate the model to magnification. The model can be easily used for 3D image deconvolution and deblur from just the value of magnification. In Chapter 3 both qualitative and quantitative comparisons are given based on different magnifications and different specimens including zebrafish larvae, zebra finch embryo, chicken heart etc. Moreover, we found that the performance of the proposed deconvolution and deblur approach increases on samples imaged with larger magnifications. This is because smaller magnifications correspond to flatter 3D deblur models whilst larger magnifications relate to steeper, i.e. sharper, models. The results are shown in the modelling section of Chapter 3.

RQ4: Can the iterative reconstruction eliminate the streak artefacts produced in the fast reconstruction?

Iterative reconstruction is implemented in a way that takes the observed projections, i.e. tomogram images, into account and uses it as a reference for updating the current reconstruction. If there are streak artefacts in the current reconstruction, through an iterative reconstruction workflow these will be propagated and reflected on simulated sinogram and be further compared to the observed projection. Aiming at minimizing the error between simulated and observed projection, the algorithm guarantees that the new reconstruction will converge in the correct direction. In Chapter 4, an example of streak artefacts in zebrafish is given and we present the effectiveness of iterative reconstruction on streak artefact elimination based on the results of multiple samples we have tested.

RQ5: How and to what extent the initialization and the number of iteration steps influence the results in iterative reconstruction?

We explore the effects of iteration steps and initial reconstruction having on the reconstructions. Thereby, we focus in particular on the zebrafish specimen. By using the reconstructions for segmentations, we measure the segmentation performance in zebrafish. By evaluating the segmentation performance, we quantify the influences of iteration steps and initial reconstruction and further optimise the parameters for the iterative reconstruction. To achieve this, a highly reliable segmentation algorithm is required, that takes the transparency of the samples into account. Based on the experimental results, we find that the combination of 10-iteration and FBP-initial

reconstruction has the best performance with the current preparation protocol and data. We further demonstrated that without a phantom for reference of quality, and empirical approach provides sufficient information on quality of the reconstruction.

RQ6: Is it possible to "learn" a 3D reference structure of zebrafish for 3D fluorescence quantification in zebrafish?

This research question deals with the application domain of OPT; here we have focussed on our typical model system, i.e. the zebrafish. In order to avoid influence of the variation in individual samples and imaging environment, i.e. exposure time and magnification, on tumour quantification for drug discovery, relative quantification is proposed and defined in Chapter 5. This quantification approach can be further generalized to other fluorescent signals in zebrafish. In terms of 3D relative quantification of fluorescence in zebrafish, we focus on the automated detection of reference structures. This detection is defined as "learn" as we aim to avoid the laborious manual labelling in a volume image. By using the current state-of-the-art volumetric segmentation approaches in biomedical imaging, we trained a robust segmentation approach to detect the two reference structures, i.e, *Body* and *Eye*. Based on the 38 training samples we have achieved promising results. For both reference structures we achieve an accuracy of over 90%. We think that the accuracy can be further improved by adding more data in the training procedure.

RQ7: How much 3D information can be achieved and identified from bright-field zebrafish OPT imagery and to what extent such identification can be automated?

From an unstained zebrafish 3D OPT bright-field image we are able to distinguish a number of well-defined regions in the volume. For a 5*dpf* zebrafish we can observer four volume regions: i.e. *Eye*, *Head*, *Muscle* and *Belly*. Whilst in a 25*dpf* zebrafish, a more advanced developmental stage, more comprehensive volume regions can be observed: i.e. *Eye*, *Head*, *Belly*, *Fin*, *Muscle*, *Blood vessel*, *Notochord & Spinal cord*, *Swim bladder* and *Intestine*, cf. § 6.3.1. In order to explore the automated detection of 3D structures, we trained 35 zebrafish samples aged from 5*dpf* to 7*dpf* and independently tested the automated detection method on 3 samples. The average accuracy for *Head*, *Muscle*, *Belly and Eye* is 54.96%, 70.39%, 39.16% and 0% respectively. Constrained by the transparency characteristics, data availability, benchmark labelling, computational resources and algorithm intelligence, automated 3D detection on zebrafish is still in its infancy and remains a challenge.

7.2 Achievements of research presented in this thesis.

All considered, we set out to study the use of OPT and pondered on how to maximize the information that can be obtained from an OPT image. In order to do so, we reviewed the quality of the images that are obtained from OPT.

The quality is influenced by the reconstruction algorithm, the artifacts that are introduced from the reconstruction and the artifacts that are introduced from the imaging. We have shown from our research questions that we have addressed, in addition, the speed of operation as we consider this an important asset of OPT imaging in a research workflow.

We have shown that the artifacts from the reconstruction, i.e. rings and streaks can be corrected in an efficient manner. The imperfections in the imaging causing deblur can be restored by a specifically designed process of deconvolution. It is clear from the algorithms that are designed and probed that OPT imaging is a typical form of computational imaging. It requires sufficient computational resources and smart algorithmic approaches in order to be efficient and valuable.

In the chapters on optimisation of the images we have used several different samples typical for the range of magnitude common to OPT. In the chapters on application of OPT we use zebrafish and worked on typical manners to support the analysis of OPT images from zebrafish. In that we have invoked machine learning approaches that are state-of-the-art. The rationale behind the use of the machine learning in segmentation is to be able to automate these processes. That is, now that we can obtain good quality images from the OPT, we must develop the processing of these images. We have shown that is approach can be successful.

7.3 Limitations and possible solutions

Further to the presentation of the results we here consider some of the limitations as well as manners how to overcome these limitations. To that end we take for perspectives in the next paragraphs, i.e. data, hardware, algorithms and theory.

7.3.1 Data perspective

(1) The amount of biomedical image datasets, especially 3D images are, in general, relatively insufficient for existing computational architectures. Large image resources such as ImageNet [146] do have a better performance for this matter. In this data-limited context, typical for scientific research, we accelerate and facilitate the process of producing 3D OPT microscopy images. By employing the use of cluster and parallel computing with a fast reconstruction algorithm, i.e. FBP, we created a fast reconstruction system. It enables users to obtain 3D images fast with a reasonable quality for most cases,

yet it fails to avoid streak artefacts at locations where small and concentrated signals appear. To completely eliminate the possible streak artefacts, iterative reconstruction is applied on top of the fast reconstruction. Currently, the iterative reconstruction is, however, explored on GPU without parallel optimisation. This means that in terms of speed of 3D image data acquisition, the iterative reconstruction work is far from optimal. The combined optimisation of implementation on GPU and parallel computing for iterative reconstruction will be considered in the further research now that we know how to combine fast and precise reconstruction methods.

(2) In Chapter 3, the deblur experiments are implemented and reported on 25 3D images. The size of the data we have is far from sufficient with respect to requirements for statistical analysis and big data. The effectiveness of the proposed methods needed to be further verified as more data become available. In Chapter 4, parameters of iterative reconstruction are optimized based on experiments of zebrafish with two different clearing protocols. By using each protocol we image three zebrafish, which are utilized for reconstruction with different parameters. Even though the test performance of single zebrafish reaches up to more than 98%, with the training ratio of 20% and test ratio of 80%, it might be more convincing if more samples can be used. In Chapter 5 we train the two reference models based on 35 zebrafishes and achieve promising results on 3 test data, but there is room for improvement to achieve a higher accuracy. We will therefore consider keeping adding training samples as they become available and then regularly retrain the model, so as for the automated annotation in Chapter 6.

7.3.2 Hardware perspective

Imaging resources: The 3D imaging process is accomplished in a full revolution and samples are manually mounted in the FoV. Because of the manual operation, a perfect mount of the sample to meet the requirements for reconstruction cannot be guaranteed. To address this problem, three imaging parameters are introduced. The camera rotation and prism tilt in the microscopy synergistically determine the direction of CoR in the image space. Ideally, the CoR is supposed to be parallel with one of the image axes for reconstruction, but, in practice of imaging this is difficult to accomplish. The operator can, however, decrease the differences between them by adjusting two screws, a laborious operation. Another parameter is the prism rotation which determines the distance between CoR and image centre in the parallel axis. It is also difficult and time consuming to adjust this distance to an ideal value of 0. We solve this problem by presenting the CoR correction in Chapter 2.

From our perspective, there are two possible solutions for these problems. One is keeping the manual placement and fixing the cameral rotation and prism tilt screw after a satisfactory CoR direction is reached. The other solution would be replacing the manual mounting of the sample with a fixed automated mounting of the sample. In such way the three imaging parameters and sample mounting are fixed and optimized. This is

promising in terms of both imaging quality and efficiency. It requires, however, accurate mechanical motorized parts, for example how to guarantee the accurate and same position when placing each sample. This is absolutely challenging now but might be feasible in the near future.

7.3.3 Algorithmic perspective

(1) CoR correction: The CoR correction algorithm is dependent on the signal intensity of sample from the imaging system. This means that if signals are weak the tomogram cannot provide the algorithm with sufficiently strong signals to calculate the CoR. This is the main drawback of our CoR correction approach and currently we cannot establish a generic solution to avoid this drawback. However, the combination of bright-field and fluorescence channels can give sufficient information.

(2) 3D PSF modelling: The 3D PSF in Chapter 3 is modelled based on a fluorescently labelled bead of fixed size, which is larger than but close to the resolution of the imaging system. In such case we approximate the model built from the sphere as the PSF model at the specific resolution, but a theoretical PSF is considered more powerful. In this work we assume that PSF model is linearly related to sphere size. The effect of the bead size on the modelling and deblur is not taken into account. But the theoretical relation between them needs be further investigated. With respect to the magnification effect, the model is constructed based on 6 different magnifications. Subsequently, we estimated the PSF model on the 3D image according to magnification consistency, which is theoretically regarded as reasonable. However, if we critically think about this rationale, we need prove the optimality of implementation based on magnification consistency. This means performance comparison of deblur with different magnifications implemented on the same 3D image needs to be considered, for experimental evidence of the abovementioned optimality.

7.3.4 Theoretical perspective

(1) 3D PSF modelling: In Chapter 3 we generalize the 3D Gaussian model with more parameters for the PSF modelling and deblur tasks. In practice, it works well for both tasks. The generalized 3D Gaussian model we used for deblur is supported by previous work and experimental observations, which is based on statistical assumption rather than imaging theory. This means that there is a lack of theoretical proof showing the expediency and effectiveness of PSF modelling from the aspect of imaging process. This question motivates us to think is if there is another model working for OPT image deblur, benefiting from the inference of the imaging process. To answer this question, knowledge of optical physics and inferential mathematics are required. This might be an interesting and challenging research to continue with.

(2) 3D segmentation: The challenge of zebrafish segmentation in 3D image is a good estimation of the location of the surface based on the limited surface information such as pigments. The ground truth of zebrafish model for training are estimated based on human knowledge, i.e. manual labelling. There is no benchmark model for a zebrafish, so the manually labelled model is regarded as the closest to a benchmark. Under this assumption, we trained the zebrafish model based on the approximate benchmark and used it for evaluation. From practice in machine learning we know that when sufficient data is available the learnt model can achieve a very high accuracy but never reach to 100%, compared to manually labelled benchmark. One possible addition is to combine information from phenotypes of zebrafish such as smoothness, connectivity, etc., to evaluate the estimation of the rather transparent surface.

7.4 Outlook

Some 3D imaging techniques are non-invasive and at the same time provide sufficient interior information of the biological sample. This makes 2D imaging of physical sections less necessary. Stepping up one dimension, i.e. from 2D to 3D, with large image sizes introduces a big challenge for both hardware and software.

For image acquisition in OPT, the automated mounting of samples is a promising direction of research. Given automated mounting, a calibration becomes obsolete with the assumption that samples can be ideally placed for obtaining an approximately perfect reconstruction. Another point is that automated mounting will shorten the overall time for imaging each sample. Both aspects contribute a large decrease in imaging time and accelerate the imaging process. Taking advantage of this automation process, the system will be suitable for high throughput screening. But the challenge is that the progress of automation is determined by the collaborative development of mechanical engineering and computer technology.

The fast reconstruction algorithms in the OPT system have yielded reconstructions for hundreds of samples. From these samples we observed two kinds of artefacts in our work. But there might be more artefacts for OPT imaging. The scientific literature in this field is still limited. A systematic analysis of artefacts in OPT would be an interesting research topic. To this end an example can be taken from models that have been used for artefacts reduction in CT imaging.

The research presented in this thesis has addressed several research areas, from microscopy, algorithms, image processing and analysis, distributed and parallel computing to machine learning and AI. This shows that research can progress if state-of-the-art methodology is merged and applied. In the future progress for the field of OPT, this will also be the case. Merging the different aspects of OPT in a smart manner will allow application of OPT in biomedical research to flourish and grow.

In the workflow of imaging the OPT takes a niche for its ability to produce good images of samples in the range of magnitudes of millimetres. Efforts to further optimize

this will be beneficial in understanding of biological phenomena as well as in imaging workflows in general.

References

[1] L. Quintana and J. Sharpe, "Optical projection tomography of vertebrate embryo development," *Cold Spring Harb. Protoc.*, vol. 6, no. 6, pp. 586–594, 2011.

[2] S. Z. M. Muji *et al.*, "Optical tomography: A review on sensor array, projection arrangement and image reconstruction algorithm," *International Journal of Innovative Computing, Information and Control*, vol. 7, no. 7 A. pp. 3839–3856, 2011.

[3] J. Sakurai, A. Ahamed, M. Murai, M. Maeshima, and M. Uemura, "Tissue and cell-specific localization of rice aquaporins and their water transport activities," *Plant Cell Physiol.*, vol. 49, no. 1, pp. 30–39, 2008.

[4] E. B. Brown *et al.*, "In vivo measurement of gene expression, angiogenesis and physiological function in tumors using multiphoton laser scanning microscopy.," *Nat. Med.*, vol. 7, no. 7, pp. 864–868, 2001.

[5] S. Kumar *et al.*, "Quantitative in vivo optical tomography of cancer progression & vasculature development in adult zebrafish," *Oncotarget*, vol. 5, no. 28, pp. 2–11, 2016.

[6] Q. Miao *et al.*, "Dual-mode optical projection tomography microscope using gold nanorods and hematoxylin-stained cancer cells.," *Opt. Lett.*, vol. 35, no. 7, pp. 1037–9, 2010.

[7] N. Agarwal, A. M. Biancardi, F. W. Patten, A. P. Reeves, and E. J. Seibel, "Three-dimensional DNA image cytometry by optical projection tomographic microscopy for early cancer diagnosis," *J. Med. Imaging*, vol. 1, no. 1, p. 017501, 2014.

[8] J. McGinty *et al.*, "In vivo fluorescence lifetime optical projection tomography," *Biomed Opt Express*, vol. 2, no. 5, pp. 1340–1350, 2011.

[9] L. Fieramonti *et al.*, "Time-Gated Optical Projection Tomography Allows Visualization of Adult Zebrafish Internal Structures," *PLoS One*, vol. 7, no. 11, 2012.

[10] A. Bassi, L. Fieramonti, C. D'Andrea, M. Mione, and G. Valentini, "In vivo label-free three-dimensional imaging of zebrafish vasculature with optical projection tomography.," *J. Biomed. Opt.*, vol. 16, no. 10, p. 100502, 2011.

[11] R. A. Baldock, F. J. Verbeek, and J. L. Vonesch, "3-D reconstructions for graphical databases of gene expression," *Semin. Cell Dev. Biol.*, 1997.

[12] T. Correia *et al.*, "Accelerated optical projection tomography applied to in vivo imaging of zebrafish," *PLoS One*, vol. 10, no. 8, 2015.

[13] L. A. Shepp and B. F. Logan, "FOURIER RECONSTRUCTION OF A HEAD SECTION.," *IEEE Trans. Nucl. Sci.*, 1974.

[14] D. M. Alanis, D. R. Chang, H. Akiyama, M. A. Krasnow, and J. Chen, "Two nested developmental waves demarcate a compartment boundary in the mouse lung," *Nat. Commun.*, 2014.

[15] W. Denk, J. H. Strickler, and W. W. Webb, "Two-photon laser scanning fluorescence microscopy.," *Science*, vol. 248, no. 4951, pp. 73–6, 1990.

[16] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank, "Brain magnetic resonance imaging with contrast dependent on blood oxygenation.," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 87, no. 24, pp. 9868–72, 1990.

[17] J. Sharpe *et al.*, "Optical projection tomography as a tool for 3D microscopy and gene expression studies," *Science (80-.).*, vol. 296, no. 5567, pp. 541–545, 2002.

[18] J. Michálek and M. Čapek, "Artifact-free 3D Reconstruction for Optical Projection Tomography.," *Microscopy and Microanalysis*, vol. 20, no. S3. pp. 1362–1363, 2014.

[19] M. Singh *et al.*, "Comparison of optical projection tomography and optical coherence tomography for assessment of murine embryonic development," *SPIE BiOS*. p. 93340J, 2015.

[20] Merel van't Hoff, "Optical Projection Tomography Optimizing Processes from Specimen to Images," Leiden, 2015.

[21] J. Radon, "Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser

Mannigfaltigkeiten," 1983.

[22] J. R. Walls, J. G. Sled, J. Sharpe, and R. M. Henkelman, "Correction of artefacts in optical projection tomography.," *Phys. Med. Biol.*, vol. 50, no. 19, pp. 4645–4665, 2005.

[23] R. Gordon, R. Bender, and G. T. Herman, "Algebraic Reconstruction Techniques (ART) for three-dimensional electron microscopy and X-ray photography," *J. Theor. Biol.*, 1970.

[24] H. Abeida, Q. Zhang, J. Li, and N. Merabtine, "Iterative sparse asymptotic minimum variance based approaches for array processing," *IEEE Trans. Signal Process.*, 2013.

[25] J. A. Fessler, "Penalized Weighted Least-Squares image Reconstruction for Positron Emission Tomography," *IEEE Trans. Med. Imaging*, 1994.

[26] J. Adler and O. Öktem, "Learned Primal-Dual Reconstruction," *IEEE Trans. Med. Imaging*, 2018.

[27] K. Hammernik *et al.*, "Learning a variational network for reconstruction of accelerated MRI data," *Magn. Reson. Med.*, 2018.

[28] H. M. Hudson and R. S. Larkin, "Accelerated Image Reconstruction Using Ordered Subsets of Projection Data," *IEEE Trans. Med. Imaging*, 1994.

[29] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2015.

[30] Y. Zhang and H. Yu, "Convolutional Neural Network Based Metal Artifact Reduction in X-Ray Computed Tomography," *IEEE Trans. Med. Imaging*, 2018.

[31] N. Hamilton, "Quantification and its applications in fluorescent microscopy imaging," *Traffic*. 2009.

[32] J. R. Walls, J. G. Sled, J. Sharpe, and R. M. Henkelman, "Resolution improvement in emission optical projection tomography," *Phys. Med. Biol.*, vol. 52, no. 10, p. 2775, 2007.

[33] J. G. McNally, T. Karpova, J. Cooper, and J. A. Conchello, "Three-dimensional imaging by deconvolution microscopy," *Methods*, vol. 19, no. 3, pp. 373–385, 1999.

[34] G. L. Zeng, "Image reconstruction - A tutorial," *Comput. Med. Imaging Graph.*, 2001.

[35] M. Beister, D. Kolditz, and W. A. Kalender, "Iterative reconstruction methods in X-ray CT," *Physica Medica*. 2012.

[36] A. K. Hara, R. G. Paden, A. C. Silva, J. L. Kujak, H. J. Lawder, and W. Pavlicek, "Iterative reconstruction technique for reducing body radiation dose at CT: Feasibility study," *Am. J. Roentgenol.*, 2009.

[37] O. Ishaq, S. K. Sadanandan, and C. Wählby, "Deep Fish," *SLAS Discov. Adv. life Sci. R D*, 2017.

[38] P. S. and S. Pacheco, "3-D Reconstruction of Tooth Development and Gene Expression Using Optical Projection Tomography," *Journal of Molecular Imaging & Dynamics*, vol. 2, no. 1. OMICS International., 2012.

[39] S. G. Azevedo, D. J. Schneberk, J. P. Fitch, and H. E. Martz, "Calculation of the Rotational Centers in Computed Tomography Sinograms," *IEEE Trans. Nucl. Sci.*, vol. 37, no. 4, pp. 1525–1540, 1990.

[40] Barbara Zitová and Jan Flusser, "Image registration methods: A survey," *Image and Vision Computing*, vol. 21, no. 11. pp. 977–1000, 2003.

[41] B. Olander, "Centre of Rotation Determination Using Projection Data in X-ray Micro Computed Tomography," Linköping University Electronic Press, Linköping University, Linköping University, Radio Physics, 1994.

[42] K. Bc Jan, "3D computed tomography," Czech Technical University, 2004.

[43] Y. Min, G. Haidong, L. Xingdong, M. Fanyong, and W. Dongbo, "A new method to determine the center of rotation shift in 2D-CT scanning system using image cross correlation," *NDT E Int.*, vol. 46, pp. 48–54, Mar. 2012.

[44] J. Sharpe *et al.*, "Optical projection tomography as a tool for 3D microscopy and gene expression studies," *Science (80-.).*, vol. 296, no. 5567, pp. 541–545, 2002.

[45] A. Brunetti and F. De Carlo, "A robust procedure for determination of center of rotation in

tomography," in *Optical Science and Technology, the SPIE 49th Annual Meeting*, 2004, pp. 652–659.

[46] T. Donath, F. Beckmann, and A. Schreyer, "Automated determination of the center of rotation in tomography data.," *J. Opt. Soc. Am. A. Opt. Image Sci. Vis.*, vol. 23, no. 5, pp. 1048–57, 2006.

[47] G. Staples, "TORQUE---TORQUE resource manager," in *Proceedings of the 2006 ACM/IEEE conference on Supercomputing - SC '06*, 2006.

[48] L. Tedersoo *et al.*, "Global diversity and geography of soil fungi," *Science (80-.).*, 2014.

[49] X. Tang *et al.*, "Fluorescence and bright-field 3D image fusion based on sinogram unification for optical projection tomography," in *Bioinformatics and Biomedicine (BIBM), 2016 IEEE International Conference on*, 2016, pp. 403–410.

[50] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man. Cybern.*, vol. 9, no. 1, pp. 62–66, 1979.

[51] I. Sobel and G. Feldman, "A 3x3 isotropic gradient operator for image processing," *a talk Stanford Artif. Proj.*, pp. 271–272, 1968.

[52] A. Kak and M. Slaney, "Principles of Computerized Tomographic Imaging," *Engineering*, p. 344, 1988.

[53] J. van der Horst and J. Kalkman, "Image resolution and deconvolution in optical tomography," *Opt. Express*, vol. 24, no. 21, pp. 24460–24472, 2016.

[54] A. C. Kak and M. Slaney, *Principles of computerized tomographic imaging*. SIAM, 2001.

[55] L. Chen *et al.*, "Incorporation of an experimentally determined MTF for spatial frequency filtering and deconvolution during optical projection tomography reconstruction.," *Opt. Express*, 2012.

[56] L. Chen, N. Andrews, S. Kumar, P. Frankel, J. McGinty, and P. M. W. French, "Simultaneous angular multiplexing optical projection tomography at shifted focal planes.," *Opt. Lett.*, vol. 38, no. 6, pp. 851–3, 2013.

[57] Q. Miao, J. Hayenga, M. G. Meyer, T. Neumann, A. C. Nelson, and E. J. Seibel, "Resolution improvement in optical projection tomography by the focal scanning method.," *Opt. Lett.*, vol. 35, no. 20, pp. 3363–5, 2010.

[58] W. Xia, R. M. Lewitt, and P. R. Edholm, "Fourier correction for spatially variant collimator blurring in SPECT," *IEEE Trans. Med. Imaging*, vol. 14, no. 1, pp. 100–115, 1995.

[59] a Darrell, H. Meyer, K. Marias, M. Brady, and J. Ripoll, "Weighted filtered backprojection for quantitative fluorescence optical projection tomography.," *Phys. Med. Biol.*, vol. 53, no. 14, pp. 3863–81, 2008.

[60] C. M. McErlean, E. Bräuer-Krisch, J. Adamovics, and S. J. Doran, "Assessment of optical CT as a future QA tool for synchrotron x-ray microbeam therapy.," *Phys. Med. Biol.*, vol. 61, no. 1, pp. 320–37, 2016.

[61] H. Kogelnik and T. Li, "Laser Beams and Resonators," *Proc. IEEE*, vol. 54, no. 10, pp. 1312–1329, 1966.

[62] W. H. Richardson, "Bayesian-Based Iterative Method of Image Restoration*," *J. Opt. Soc. Am.*, 1972.

[63] D. Stalling, M. Westerhoff, and H. C. Hege, "Amira: A highly interactive system for visual data analysis," in *Visualization Handbook*, 2005.

[64] R. Ferzli and L. J. Karam, "A no-reference objective image sharpness metric based on the notion of Just Noticeable Blur (JNB)," *IEEE Trans. Image Process.*, vol. 18, no. 4, pp. 717–728, 2009.

[65] N. D. Narvekar and L. J. Jaram, "An improved no-reference sharpness metric based on the probability of blur detection," *IEEE Int. Work. Qual. Multimed. Exp.*, vol. 20, no. 1, pp. 1–5, 2009.

[66] K. De and V. Masilamani, "Image Sharpness Measure for Blurred Images in Frequency Domain," *Procedia Eng.*, vol. 64, no. Complete, pp. 149–158, 2013.

[67] T. F. Chan and C. K. Wong, "Total variation blind deconvolution," *IEEE Trans. Image Process.*,

1998.

[68]   D. E. Dudgeon and R. M. Mersereau, "Multidimensional Digital Signal Processing," *IEEE Communications Magazine*. 1985.

[69]   J. B. Pawley, *Handbook of biological confocal microscopy: Third edition*. 2006.

[70]   A. Nwaneshiudu, C. Kuschal, F. H. Sakamoto, R. Rox Anderson, K. Schwarzenberger, and R. C. Young, "Introduction to confocal microscopy," *J. Invest. Dermatol.*, 2012.

[71]   Y. S. Li, Y. Chen, J. H. Ma, L. M. Luo, and W. F. Chen, "Metal artifact reduction in CT based on adaptive steering filter and nonlocal sinogram inpainting," *Chinese J. Biomed. Eng.*, 2011.

[72]   J. Wang, S. Wang, Y. Chen, J. Wu, J. L. Coatrieux, and L. Luo, "Metal artifact reduction in CT using fusion based prior image," *Med. Phys.*, 2013.

[73]   K. Y. Jeong and J. B. Ra, "Metal artifact reduction based on sinogram correction in CT," in *IEEE Nuclear Science Symposium Conference Record*, 2009.

[74]   L. Gjesteby *et al.*, "Metal Artifact Reduction in CT: Where Are We After Four Decades?," *IEEE Access*, 2016.

[75]   D. Giantsoudi *et al.*, "Metal artifacts in computed tomography for radiation therapy planning: Dosimetric effects and impact of metal artifact reduction," *Physics in Medicine and Biology*. 2017.

[76]   L. L. Geyer *et al.*, "State of the Art: Iterative CT Reconstruction Techniques," *Radiology*, 2015.

[77]   D. Hahn *et al.*, "Statistical iterative reconstruction algorithm for X-ray phase-contrast CT," *Sci. Rep.*, 2015.

[78]   G. Wang, D. L. Snyder, J. A. O'Sullivan, and M. W. Vannier, "Iterative deblurring for CT metal artifact reduction," *IEEE Trans. Med. Imaging*, 1996.

[79]   H. Chen *et al.*, "LEARN: Learned Experts' Assessment-Based Reconstruction Network for Sparse-Data CT," *IEEE Trans. Med. Imaging*, 2018.

[80]   C. Ghetti, O. Ortenzia, and G. Serreli, "CT iterative reconstruction in image space: A phantom study," *Phys. Medica*, 2012.

[81]   H. M. Hudson and R. S. Larkin, "Ordered Subsets of Projection Data," *IEEE Trans. Med. Imaging*, 1994.

[82]   L. A. Shepp and Y. Vardi, "Maximum Likelihood Reconstruction for Emission Tomography," *IEEE Trans. Med. Imaging*, 1982.

[83]   E. W. Weisstein, "Sigmoid Function," *MathWorld - A Wolfram Web Resour.*, 2006.

[84]   D. P. Kingma and J. L. Ba, "Adam optimizer," *arXiv Prepr. arXiv1412.6980*, 2014.

[85]   RStudio, "Keras," *R Cheat Sheet*, 2017.

[86]   N. K. Sinha and M. P. Griscik, "\A Stochastic Approximation Method," *IEEE Trans. Syst. Man Cybern.*, 1971.

[87]   L. Bottou, "Stochastic gradient descent tricks," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, 2012.

[88]   N. Chinchor, "MUC-4 evaluation metrics," in *Proceedings of the 4th conference on Message understanding - MUC4 '92*, 1992.

[89]   M. Chalfie, Y. Tu, G. Euskirchen, W. W. Ward, and D. C. Prasher, "Green fluorescent protein as a marker for gene expression," *Science (80-.).*, 1994.

[90]   Hermes Spaink, "Comparison of Clearing Protocols for OPT Imaging," Leiden, 2018.

[91]   K. Tainaka *et al.*, "Whole-body imaging with single-cell resolution by tissue decolorization," *Cell*, 2014.

[92]   L. I. Zon and R. T. Peterson, "In vivo drug discovery in the zebrafish," *Nature Reviews Drug Discovery*. 2005.

[93]   S. He *et al.*, "Synergy between loss of NF1 and overexpression of MYCN in neuroblastoma is mediated by the GAP-related domain," *Elife*, 2016.

[94]   D. B. Murphy and M. W. Davidson, *Fundamentals of Light Microscopy and Electronic Imaging: Second Edition*. 2012.

[95] J. Sharpe, "Optical projection tomography as a new tool for studying embryo anatomy," *Journal of Anatomy*, vol. 202, no. 2. pp. 175–181, 2003.

[96] M. Thain and M. HICKMAN, *Dictionary of Biology*, 9th Editio. London, UK: Penguin, 1994.

[97] J. Schindelin *et al.*, "Fiji: an open-source platform for biological-image analysis.," *Nat. Methods*, 2012.

[98] Y. Guo, Z. Xiong, and F. J. Verbeek, "An efficient and robust hybrid method for segmentation of zebrafish objects from bright-field microscope images," in *Machine Vision and Applications*, 2018.

[99] R. Mikut *et al.*, "Automated Processing of Zebrafish Imaging Data: A Survey," *Zebrafish*, 2013.

[100] and F. J. V. A. E. Nezhinsky, "Efficient and Robust Shape Retrieval from Deformable Templates," in *ISoLA 2012, Part II, LNCS 7610*, 2012, pp. 42–55.

[101] A. A. E. Nezhinsky, E. Stoop, A. van der Sar and F. J. Verbeek, "Numerical Analysis of Image Based High-Throughput Zebrafish Infection Screens: Matching Meaning with Data," in *Proceedings of the Int. Conf. on Bioinformatics Models, Methods and Algorithms*, 2012, pp. 257–262.

[102] and A. M. van der S. E. J. Stoop, T. Schipper, S. K. Rosendahl Huber, A. E. Nezhinsky, F. J. Verbeek, S. S Gurcha, G. S. Besra, M. Vandenbroucke-Grauls, W. Bitter, "Zebrafish embryo screen for mycobacterial genes involved in the initiation of granuloma formaion reveals a newly identified ESX-1 component," *Dis. Model Mech.*, pp. 526–536, 2011.

[103] A. E. Nezhinsky and F. J. Verbeek, "Pattern recognition for high throughput zebrafish imaging using genetic algorithm optimization," in *5th IAPR Conference on Pattern Recognition in BioInformatics*, 2010, pp. 301–312.

[104] R. R. S. Packard *et al.*, "Automated Segmentation of Light-Sheet Fluorescent Imaging to Characterize Experimental Doxorubicin-Induced Cardiac Injury and Repair," *Sci. Rep.*, 2017.

[105] L. Lleras Forero *et al.*, "Segmentation of the zebrafish axial skeleton relies on notochord sheath cells and not on the segmentation clock," *Elife*, 2018.

[106] S. Wopat *et al.*, "Spine Patterning Is Guided by Segmentation of the Notochord Sheath," *Cell Rep.*, 2018.

[107] E. Kugler, K. Plant, T. Chico, and P. Armitage, "Enhancement and Segmentation Workflow for the Developing Zebrafish Vasculature," *J. Imaging*, 2019.

[108] K. Zhang *et al.*, "Zebrafish Embryo Vessel Segmentation Using a Novel Dual ResUNet Model," *Comput. Intell. Neurosci.*, 2019.

[109] T. Wu, J. Lu, Y. Lu, T. Liu, and J. Yang, "Embryo zebrafish segmentation using an improved hybrid method," *J. Microsc.*, 2013.

[110] Z. Xiong and F. Verbeek, "Segmentation of Zebrafish Larva Inhomogeneous 3D Images Using the Level-Set Method," *Electron. Imaging*, 2017.

[111] O. Ishaq, S. K. Sadanandan, and C. Wählby, "Deep Fish: Deep Learning-Based Classification of Zebrafish Deformation for High-Throughput Screening," *J. Biomol. Screen.*, 2017.

[112] C. Cullander, "Fluorescent probes for confocal microscopy," *Methods Mol. Biol.*, 1998.

[113] Z. Földes-Papp, U. Demel, and G. P. Tilz, "Laser scanning confocal fluorescence microscopy: An overview," *International Immunopharmacology*. 2003.

[114] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-net: Learning dense volumetric segmentation from sparse annotation," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2016.

[115] P. T. De Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, "A tutorial on the cross-entropy method," *Ann. Oper. Res.*, 2005.

[116] C. Robert, "   Machine Learning, a Probabilistic Perspective   ," *CHANCE*, 2015.

[117] F. Milletari, N. Navab, and S. A. Ahmadi, "V-Net: Fully convolutional neural networks for

volumetric medical image segmentation," in *Proceedings - 2016 4th International Conference on 3D Vision, 3DV 2016*, 2016.

[118] F. J. Verbeek, "Deformation correction using euclidean contour distance maps," in *Proceedings - International Conference on Pattern Recognition*, 1992.

[119] S. S. M. Salehi, D. Erdogmus, and A. Gholipour, "Tversky loss function for image segmentation using 3D fully convolutional deep networks," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017.

[120] A. Tversky, "Features of similarity," *Psychol. Rev.*, 1977.

[121] G. Litjens *et al.*, "A survey on deep learning in medical image analysis.," *Med. Image Anal.*, 2017.

[122] D. P. Kingma and J. L. Ba, "Adam: A method for stochastic gradient descent," *ICLR Int. Conf. Learn. Represent.*, 2015.

[123] J. Duchi, E. Hazan, and Y. Singer, "Adaptive Subgradient Methods for Online Learning and Stochastic Optimization," *JMLR*, 2011.

[124] J. Dean *et al.*, "Large Scale Distributed Deep Networks," in *Advances in Neural Information Processing Systems 25*, 2012.

[125] G. E. Hinton, N. Srivastava, and K. Swersky, "Overview of mini-batch gradient descent," *COURSERA: Neural Networks for Machine Learning*. 2012.

[126] T. Tieleman, G. E. Hinton, N. Srivastava, and K. Swersky, "Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude," *COURSERA Neural Networks Mach. Learn.*, 2012.

[127] L. N. Smith, "Cyclical learning rates for training neural networks," in *Proceedings - 2017 IEEE Winter Conference on Applications of Computer Vision, WACV 2017*, 2017.

[128] I. Loshchilov and F. Hutter, "Fixing Weight Decay Regularization in Adam," in *Proceedings of the 2019 International Conference on Learning Representations (ICLR'19)*, 2019.

[129] Y. Sasaki, "The truth of the F-measure," *Am. Rev. Respir. Dis.*, 1982.

[130] J. Davis and M. Goadrich, "The relationship between Precision-Recall and ROC curves," 2006.

[131] K. Boyd, K. H. Eng, and C. D. Page, "Area under the precision-recall curve: Point estimates and confidence intervals," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2013.

[132] F. J. Verbeek, M. M. de Groot, D. P. Huijsmans, W. H. Lamers, and I. T. Young, "3D base: A geometrical data base system for the analysis and visualisation of 3D-shapes obtained from parallel serial sections including three different geometrical representations," *Comput. Med. Imaging Graph.*, 1993.

[133] M. Corredor-Adámez *et al.*, "Genomic annotation and transcriptome analysis of the zebrafish (Danio rerio) hox complex with description of a novel member, hoxb13a," *Evol. Dev.*, 2005.

[134] D. S. Kelkar *et al.*, "Annotation of the zebrafish genome through an integrated transcriptomic and proteomic analysis," *Mol. Cell. Proteomics*, 2014.

[135] F. J. Verbeek, "Theory & Practice of 3D-reconstructions from serial sections," in *A Practical Approach*, R. A. B. and J. Graham, Ed. (Oxford: Oxford University Press), 1999, pp. 153–195.

[136] F. J. Verbeek and D. P. Huijsmans, "A Graphical Database for 3D Reconstruction Supporting (4) Different Geometrical Representations," in *Medical Image Databases*, 1998.

[137] S.van der Plas-Duijvestein, "Advancing Zebrafish Models in Proteomics," Leiden University, 2018.

[138] M. C. M. Welten *et al.*, "ZebraFISH: Fluorescent in situ hybridization protocol and three-dimensional imaging of gene expression patterns," *Zebrafish*, 2006.

[139] F. J. F. Laroche *et al.*, "The embryonic expression patterns of zebrafish genes encoding LysM-domains," *Gene Expr. Patterns*, 2013.

[140] J. P. Junker *et al.*, "Genome-wide RNA Tomography in the Zebrafish Embryo," *Cell*, 2014.

[141] F. J. Verbeek, K. A. Lawson, and J. B. L. Bard, "Developmental bioinformatics: Linking genetic data to virtual embryos," *International Journal of Developmental Biology*. 1999.

[142] F. J. Verbeek *et al.*, "A Standard 3D Digital Atlas of Zebrafish Embryonic Development for Projection of Experimental Data," in *Internet Imaging*, 1999.

[143] D. Potikanond, M. Belmamoune, M. C. M. Welten, and F. J. Verbeek, "Visual Integration of 3D Digital Atlas and 3D Patterns of Gene Expression in Zebrafish," *Int. Symp. Integr. Bioinforma.*, 2011.

[144] H. Peng, A. Bria, Z. Zhou, G. Iannello, and F. Long, "Extensible visualization and analysis for multidimensional images using Vaa3D," *Nat. Protoc.*, 2014.

[145] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, and G. Ranzuglia, "MeshLab: An open-source mesh processing tool," in *6th Eurographics Italian Chapter Conference 2008 - Proceedings*, 2008.

[146] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2010.

Summary

Optical projection tomography (OPT) is a tomographic 3D imaging technique used for specimens in the millimetre scale. 3D images are computed from a tomogram and therefore OPT is considered as computational imaging. In order to provide imaging and image analysis solutions for large scale biomedical research, optimisation of the OPT reconstruction is required. The aim of the optimisation presented in this thesis includes: (1) accelerate the reconstruction process; (2) reduce the reconstruction artefacts; (3) improve the image quality of 3D image; (4) Find optimal parameters for the iterative reconstruction.

Starting from the optimisations that we have elaborated and implemented in the OPT imaging workflow, we have worked on case studies in zebrafish imaging. In this thesis we present one such particular case study (5) as it falls nicely in the order of magnitude for specimens in OPT imaging. The case study is on quantification of tumours in zebrafish and it is explored with image segmentation and object detection using artificial intelligence (AI) techniques.

The acceleration of the reconstruction process (ad. 1) aims to reduce the time of imaging process from tomogram to reconstruction. This supports biomedical research in high-throughput or large-scale imaging. In our work, we accelerate the OPT 3D reconstruction by implementing filtered back projection (FBP) in a parallel manner. In this thesis, we refer to this approach as fast reconstruction. With our (current) computational resources users are able to acquire a millimetre scale whole-mount 3D image in several minutes.

The reduction of the artefacts in OPT reconstruction (ad. 2) aims to provide a 3D solution of reconstruction with less artefacts that are introduced during the reconstruction process as a result of limitations of the reconstruction algorithm and the imaging setup. In this thesis, two different types of artefacts are covered, i.e. ring artefacts and streak artefacts. With respect to these artefacts, both the cause and the solution are addressed in this thesis. In the FBP algorithm, the ring artefacts are introduced by the misalignment of centre of rotation (CoR), whilst the streak artefacts are results of absence of sufficient signal in a tomogram. As a solution, a CoR correction algorithm is proposed in the FBP reconstruction framework. In order to eliminate the streak artefacts, iterative reconstruction is explored in OPT imaging.

The improvement of the image quality of an OPT 3D image (ad. 3) in this thesis aims to find a method for deblurring of the reconstructed images. This is done by deconvolving the 3D image with an empirically derived point spread function (PSF). The qualitative and quantitative comparisons of the results that are obtained with our deconvolution method illustrate the effectiveness of the improvements that are accomplished.

The Optimisation of the parameters for iterative reconstruction (ad. 4) aims at gaining the best possible 3D images suitable for segmentation and detection of structures within the sample. This is essential for the signal quantification, e.g. tumour quantification in the later developmental stages. The most customary parameters including iteration number and initial reconstruction are explored and optimised based on the segmentation performance of corresponding reconstructions.

The case study on tumour quantification (ad. 5) demonstrates an application of the optimised OPT imaging system. Compared to 2D quantification, we expect to provide some new insights to the tumour quantification in 3D for drug research. Based on the samples we have available for tumour quantification, we conclude that the relative ratio of tumour to a specific organism or organ in 3D, is much smaller than that in 2D. This to a large extent justifies the reason why the OPT 3D imaging is necessary for specimens in the millimetre scale. Supported by more computational resources and advanced analytical techniques, we are confident about its prospect and popularity of OPT in the area of biomedical research.

Nederlandse samenvatting

Optische projectietomografie (OPT) is een tomografische 3D-beeldvormende techniek die wordt gebruikt voor objecten ter grootte van millimeters. Op basis van een tomogram wordt een 3D-beeld berekend; daarom wordt OPT dan ook beschouwd als zogenaamde computationele beeldvorming. Om oplossingen te bieden voor het maken van beelden en het doen van beeldanalyse op die beelden voor biomedisch onderzoek op grote schaal, is optimalisatie van de OPT-reconstructie vereist. Het doel van de optimalisatie zoals beschreven in dit proefschrift omvat: (1) het versnellen van het proces van reconstructie; (2) het verminderen van artefacten uit de reconstructie; (3) het verbeteren van de beeldkwaliteit van het 3D-beeld; (4) het vinden van optimale parameters voor iteratieve reconstructie.

Met optimalisaties die we hebben uitgewerkt en geïmplementeerd in de OPT-beeldvormingsworkflow als uitgangspunt, hebben we gewerkt aan casestudy's in (3D) beeldvorming van zebravissen. In dit proefschrift presenteren we één specifieke case-study (5), aangezien deze goed past in de orde van grootte van objecten geschikt voor toepassing in OPT. Deze case-study gaat over de kwantificering van tumoren in zebravissen. Er wordt specifiek gebruik gemaakt van beeldsegmentatie en objectdetectie met behulp van technieken uit de kunstmatige intelligentie (AI).

De versnelling van het reconstructieproces (ad. 1) heeft tot doel de tijd van het beeldacquisitie- en beeldbewerkingsproces van tomogram tot reconstructie te verkorten. Dit ondersteunt met name biomedisch onderzoek in high-throughput beeldbewerking. In ons werk versnellen we de OPT 3D-reconstructie door het *Filtered Backprojection* algoritme (FBP) parallel te implementeren. In dit proefschrift wordt deze aanpak de *snelle reconstructie* genoemd. Met onze (huidige) computationele middelen kunnen gebruikers in enkele minuten een 3D-beeld verkrijgen van het complete object.

De vermindering van de artefacten in OPT-reconstructie (ad. 2) heeft tot doel een 3D-oplossing voor reconstructie te bieden met minder artefacten zoals die tijdens het reconstructieproces worden geïntroduceerd als gevolg van beperkingen van het reconstructie-algoritme en de OPT-opstelling. In dit proefschrift komen twee verschillende soorten artefacten aan bod, namelijk ring-artefacten en streak-artefacten. In dit proefschrift wordt aandacht besteed aan zowel de oorzaak als de oplossing met betrekking tot deze artefacten. In het FBP-algoritme worden de ring-artefacten geïntroduceerd door een verkeerde uitlijning van het centrum van rotatie (CoR), terwijl de streak-artefacten het gevolg zijn van het ontbreken van voldoende signaal in een tomogram. Als oplossing wordt een CoR-correctiealgoritme gepresenteerd binnen de FBP reconstructie. Om de streak-artefacten te elimineren, zijn iteratieve reconstructiemethoden voor OPT-reconstruction onderzocht.

De verbetering van de beeldkwaliteit van een OPT 3D-beeld (ad. 3) wordt gerealiseerd met het formuleren van een methode voor het verscherpen van de

gereconstrueerde beelden. Dit wordt gedaan door het 3D-beeld te deconvolueren met een empirisch afgeleide point-spread functie (PSF). De kwalitatieve en kwantitatieve vergelijkingen van resultaten die met deze methode zijn verkregen illustreren de effectiviteit van de bereikte verbeteringen.

De optimalisatie van de parameters voor iteratieve reconstructie (ad. 4) is gericht op het verkrijgen van de best mogelijke 3D-beelden die geschikt zijn voor segmentatie en detectie van structuren binnen het object van onderzoek. Dit is essentieel voor de kwantificering van het specifieke signaal, b.v. tumor kwantificering in late stadia van ontwikkeling, De meest gebruikelijke parameters, waaronder iteratie aantal en de zogenaamde initiële reconstructie, worden verkend en geoptimaliseerd op basis van de prestaties van segmentatie van vergelijkbare reconstructies.

De case-study over kwantificering van tumoren (ad. 5) laat een toepassing van geoptimaliseerde bewerkingen van beelden uit het OPT-systeem zien. In vergelijking met de gebruikelijk 2D-kwantificering verwachten we een aantal nieuwe inzichten te kunnen verschaffen door de kwantificeting van de tumor in 3D; voor geneesmiddelenonderzoek is dit van groot belang. Op basis van de zebravisjes die we hebben kunnen gebruiken voor tumorkwantificatie, concluderen we dat de relatieve verhouding van tumor tot een specifiek organisme of orgaan in 3D veel kleiner is dan hetgeen gemeten wordt in 2D. Dit rechtvaardigt in grote mate de reden waarom de 3D OPT techniek nodig is voor objecten in de orde grootte van millimeters. Met verdere ondersteuning door meer computationele middelen en geavanceerde analytische technieken, hebben we veel vertrouwen in de OPT techniek qua toepasbaarheid in het biomedische onderzoeksgebied.

Curriculum Vitae

Xiaoqin Tang was born on 30th 1988 in Sichuan, China, where she grew up and attended primary, middle and high school. In 2008, she continued her study at Southwest University in Automation (control) in Chongqing, the city she likes a lot and sees as her hometown in the sense of language and culture. At that university she got her bachelor degree there in 2012 with the thesis titled "Application of Internet of Things in Supermarket". Followed by that, she started her master in Computer Application Technology at the same university. In line with her research interest, she joined the Multimedia Lab to do research in machine learning and video classification, supervised by Prof. Guoqiang Xiao. In 2015 she obtained her master degree with the thesis titled "Human Action Recognition Based on Global and Local Movement Features". During her master study she obtained the first-grade scholarship. Supported by the Chinese Scholarship for a 4-year PhD study in the Netherlands, she started her career in Leiden University in 2015. During the PhD, she works on 3D imaging and image analysis in the Computational Bio-imaging group supervised by Prof. Fons Verbeek, providing computational and analytical solutions to researchers in drug discovery and biology. The techniques she has used for her research include image reconstruction, algorithm design, deep learning, statistical modelling and data visualization. She likes thinking ahead and providing predictive and analytical solutions for projects in different fields.

Acknowledgements

The journey towards an academic title in the Netherlands started with excitement and full passion. It was immediately followed by the language and culture barriers which were tough but interesting. The most enjoyable part for me to understand how people talk is trying to figure out the differences between English and Chinglish (Chinese English) at both language and culture level, which I am still learning and will continue doing in the future. The path to academic achievements is often combined with doubt and uncertainty, however exploration and confidence can always help to give a satisfying answer. I am glad that I made it so far and I am happy to share this moment. However, I would like to first express my sincere gratitude to my colleagues, friends and families.

First of all, I would like to thank my promoter Fons for his guidance, support, discussions and comments in microscopy imaging and image analysis. He is like a father who can always make you feel the freedom to explore, and at the same time encourage and support you to overcome the difficulties when you feel tired. As my supervisor, he also taught me how to be open-minded to people from different backgrounds and with different cultures. I would also like to thank my supervisor Guoqiang Xiao for supporting and guiding me to step into the researches in image analysis and machine learning. As my first academic tutor and mentor, he continuously inspires me to do the right thing with big efforts.

Many thanks go to Merel, Jerry, Alan, Dennis, Yuanhao and Fuyu for their contributions to OPT 3D reconstruction project on which we worked together. I would also like to thank Prof Spaink, Rob, Shuning, Hermes and Dani for their contributions and collaboration on the drug research in zebrafish. Especially thank Gerda and Merijn for their help and support on data acquisition and experimental protocols in IBL.

In some groups, there are always a few colleagues who are proficient in managing servers and clusters and also like to help others. At this point, I would like to thank Kristian and Leon for their time and contribution to support me and our group. Furthermore, I would like to thank Marloes and Marcello for their help and care about my study at Leiden University.

In addition, I would like to thank all the colleagues and friends for your companionship including Katy, Lu, Sacha, Zhan, Junling, Song, Fuyu, Yuanhao, Enrique, Mohamed, Alan, Mehrdad, Yunpeng, Feibo, Chen, Shima, Irene, Lisa, Leon, Danyi, Mariam, Yi, Xue, Erick, Rohola, Erwin, Ziyu, Quanchi, Yufei, Jing, Kelvin, Punan, Yuzhi and Wei. The daily life with you composes the memorable moments in my life. Special thanks go to my friend Will for improving my English and inspiring me to be more open to the

world. Last but not least, I would like to thank my oldest friends Nianbin, Jing and Jia for always being there.

The great gratitude goes to my parents and mother-in-law for their unconditional love and financial support. Last but not least, I would like to show the greatest appreciation to my husband for his care, love and respect. It was the luckiest thing in my life to have met you in the crowds. The time for us being together was short, but every moment was a beautiful and memorable story.

<div align="right">
Xiaoqin Tang

唐小琴

January 2020
</div>