Optical Coherence Tomography - Histology Registration

An automated framework towards facilitation of intra-operative pathology

Smruti Rekha

Technische Universiteit Delft

TUDelft
Delft
University of
Technology

Challenge the future

Optical Coherence Tomography - Histology Registration

An automated framework towards facilitation of intra-operative pathology

by

Smruti Rekha

in partial fulfillment of the requirements for the degree of

Master of Science

in Computer Science

Track: Media and Knowledge Engineering

at the Delft University of Technology,
to be defended publicly on Tuesday January 29, 2016 at 12:00 PM.

LU
MC

LEIDEN UNIVERSITY MEDICAL CENTER

TUDelft Delft University of Technology

PREFACE

This dissertation is an original and independent work by Smruti Rekha (Student ID - 4324889). All work presented henceforth was conducted in the Laboratorium voor Klinische en Experimentele Beeldverwerking (LKEB) at the Leiden University Medical Center under the supervision of Prof. dr. B. Lelieveldt and Dr. ir. J. Dijkstra. The presented work is in partial fulfillment of the requirements for the degree of Master of Science in Computer Science with a specialization in Media and Knowledge Engineering (MKE) at the Faculty of Electrical Engineering,Mathematics and Computer Science at Delft University of Technology.

The main contributions in this work are the following:

- Establishment of a multi-stage registration framework for aligning histology and Full Field Optical Coherence Tomography (FF-OCT) images, which would go a long way in the facilitation of better intra-operative pathology.

- Development of an edge-preserving modality transformation algorithm which helps in creating more realistic pseudo(same modality) images. This helps in the solving inter-modality registration problem using techniques such as Demon registration or other optical-flow based algorithms, which are traditionally applicable only to intra-modality registration scenarios.

- Development of an optical-flow based registration framework for inter-modality images. This framework is based on a recently published optical flow algorithm called Deep Flow, which is known to handle large deformations.

A subset of the work has been accepted in the International Symposium of Biomedical Imaging (ISBI) - 2016 conference. The work in its entirety will be submitted in the 19th international conference on Medical Image Computing and Computer Assisted Intervention (MICCAI) - 2016.

Smruti Rekha
January 29, 2016
Delft

ABSTRACT

Registration between histology and *ex-vivo* Full Field Optical Coherence Tomography (FF-OCT) can help in many clinical applications including identifying surgical margin in a tumour tissue during intra-operative pathological diagnosis, quantifying features in FF-OCT for diagnosis and reduction of tissue processing time. In this work, we present a framework for non-rigid registration between Histology and *ex-vivo* FF-OCT images. The proposed framework consists of a two-stage registration process. The first step consists of large-scale misalignment correction while also establishing the match between FF-OCT and one of the several histology tissue samples based on iterative closest point of prominent edge points. The second step starts with an area-based affine registration algorithm and culminates with a Deep Flow based registration algorithm. To facilitate the use of optical-flow based algorithm for inter-modality registration, mutual transform based modality transformation has been used here. Traditionally used mutual transform does not preserve the edges well. Since edges are an important component of the objective function used for minimization, this affects the accuracy of the registration algorithm. To address this problem, an edge-preserving mutual transform has been introduced in this work. This newly proposed variant of mutual transform as well as the use of Deep Flow algorithm for registration is seen to improve the accuracy significantly.

CONTENTS

LIST OF FIGURES

1

INTRODUCTION

Today, surgeons have far more information in their arsenal than before. This is heartening because when a patient is open under anesthesia on the operating table, every second longer increases the chances of infection as well as the time that would be needed post-operation to make a complete recovery. The role of intra-operative pathology is crucial in such a situation, especially to provide the surgeon with immediate information regarding whether there is still some cancerous tissue present, to determine precisely the surgical margins of a tumour or to assess the adequacy of the sample.

Currently, histopathology is the gold standard for intra-operative consultations. But it involves offline processing of the biopsy samples collected and it cannot be relied upon for on the spot decisions during surgery. Full-Field Optical Coherence Tomography is a promising new technology which can help provide the necessary evidence in near real time *in-situ*, for taking such surgical decisions. The following sections in this chapter gives an introduction to these technologies and establishes the need for an automatic registration algorithm between traditional histology images and FF-OCT images.

1.1. HISTOPATHOLOGY

Histopathology refers to the microscopic examination of tissue in order to study the manifestations of disease. Specifically, in clinical medicine, histopathology refers to the examination of a biopsy or surgical specimen by a pathologist, after the specimen has been processed and histological sections have been placed onto glass slides.

A histopathological examination usually starts with surgery, biopsy or autopsy. After obtaining the 'specimen', the standard tissue processing process is applied to it, moving it through a series of steps so that the soft tissue is supported in a medium that allows sectioning. The five main steps in the preparation of histology slides are Fixing, Processing, Embedding, Sectioning and Staining [3]. Figure 1.1 shows an overview of the steps involved in histological slide preparation and the order in which it is done.
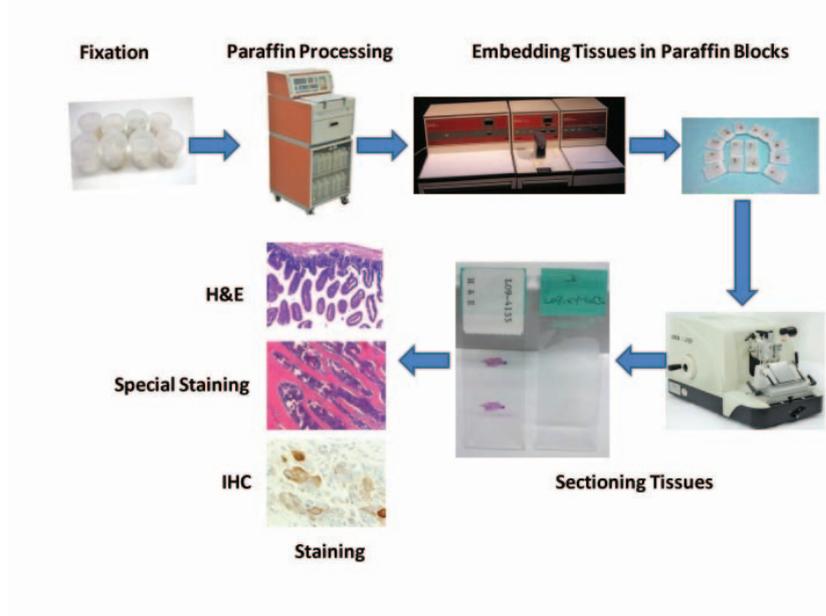
Figure 1.1: Histology Procedure Overview

Fixing is done to preserve the tissue in as natural a state as possible. It can be done either with the use of chemicals such as paraffin/formalin or by freezing the specimen. Tissue processing is done to remove water from the tissues so that it can be replaced with something that solidifies and provides the necessary rigidity to the tissue for fine slicing. Depending on the method used for fixing, the steps involved in processing differ. Frozen section processing is rapid and requires less equipment. But the quality of the final slide is poor and lacks morphological detail. As such, it is widely used in intra-operative pathology for coarse determinations that help decide the next step in surgery. On the other hand, chemical section processing is slow, but the minute details of the tissue are preserved.

Embedding is the process in which the water in the tissue is replaced with an alternative medium that solidifies and supports as well as fixes the specimen. This is followed by sectioning. This can be done in 3 ways, vertical sectioning, horizontal sectioning and Mohs surgery technique [4]. Vertical sectioning is generally used for light microscopic applications. The thickness of each section is usually of the order of 5 μm. The sliced sections are stained using appropriate histology stains. Some commonly used stains for histopathological examination of the tissue are hematoxylin-eosin (H&E) or Periodic acid–Schiff (PAS) [5]. Figure 1.2 shows a sample H&E stained slide digitized using a whole slide scanner.
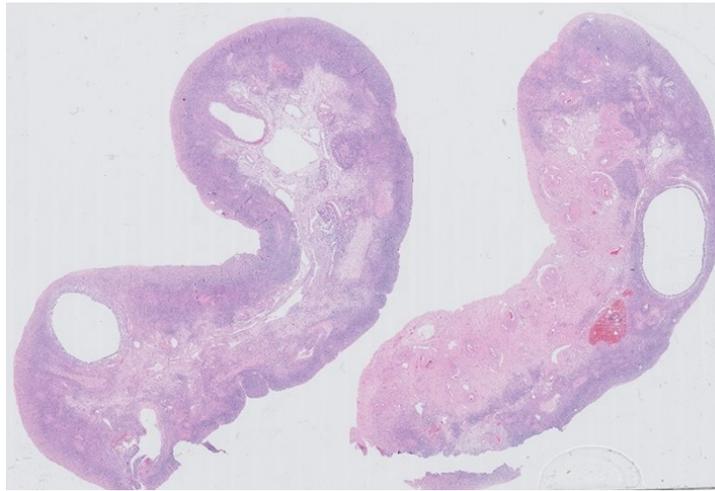
Figure 1.2: Sample H&E stained Histology Image

Although this is used as a gold standard for disease diagnosis, there are several disadvantages in choosing histological examination. First of all, inappropriate processing or some fundamental mistake in any of the aforementioned steps might result in unusable tissue specimens that will not yield any useful diagnostic information. Then there is the time element - 6 to 24 hours for fixation, 15 to 45 minutes for dehydration, 20 to 45 minutes for clearing, 30 to 45 minutes for embedding. Added to this is the danger of tissue shrinkage and damage. For an offline diagnosis using the tissue sample, the time needed for examination does not pose much of a problem. However for deciding the course of a surgery, diagnosis based on histology leaves much to be desired.

1.2. OPTICAL COHERENCE TOMOGRAPHY

In order to overcome the limitations mentioned in section 1.1, it is worth exploring Optical Coherence Tomography (OCT) as a means of (a) providing a fast, reliable and morphologically detailed microscopic image of the tissue that can be used in intra-operative pathology, and (b) decreasing the processing time that comes with histological examination while retaining the clarity of the slide.

OCT is an established medical imaging technique that uses light to capture micrometer-resolution, three-dimensional images from within optical scattering media (e.g., biological tissue). It is based on low-coherence interferometry, typically employing near-infrared light. The use of relatively long wavelength light allows it to penetrate into the scattering medium. While tissue preparation for OCT imaging is simple and consists of placing the tissue sample as is, in water in the tissue holder and viewing it through a gel that minimises refraction, there is much less tissue deformation in the process. These advantages have led to a growing interest in the use of Optical Coherence Tomography (OCT) in intra-operative pathology [6] [7].

Two general approaches for OCT based microscopy have been reported to date. The first approach is

based on the principle of point-scanning imaging developed in confocal microscopy [8]. The second approach involves full-field illumination and detection, which is sometimes termed Full-Field Optical Coherence Tomography (FF-OCT). This is an alternative technique to scanning-based approach based on white-light interference microscopy [9].

FF-OCT produces tomographic images in the en face orientation by arithmetic combination of interferometric images acquired with an area camera and by illuminating the whole field of view using a low-coherence light source [10]. Since the large depth-of-field needed in other OCT approaches is not required here, the setup provides 2D or 3D tomographic images with a 3D resolution of close to 1μm. This is typically one order of magnitude higher than for standard commercially available OCT systems, and matches the cellular resolution required for tissue examination. Figure 1.3 shows a sample FF-OCT image which corresponds to the sample histology image shown in figure 1.2.
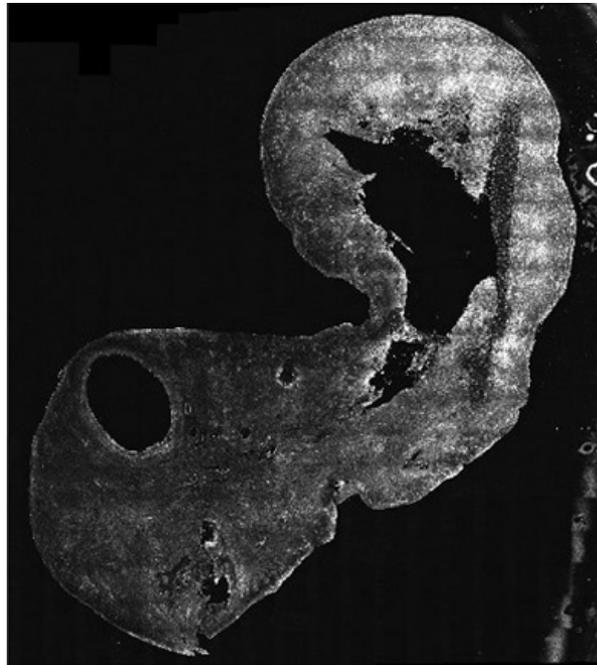
Figure 1.3: Sample FF-OCT Image

The FF-OCT typically comprises a Michelson interferometer with identical microscope objectives in both arms, known as a Linnik interferometer (see Figure 1.4). The setup appears simple in design, but can be difficult to align. Getting the zero path difference means balancing centimeter paths to less than one micrometer.

Comparing morphological features revealed by FF-OCT with histological or frozen sections, pathologists have achieved high levels of sensitivity in their results. Two pathologists could distinguish between in situ normal and benign tissue and invasive carcinomas, with a sensitivity of 97% and 90%, respectively [11]. For this reason, we anticipate that FF-OCT should be a valuable tool for intra-operative diagnosis as well as tissue selection. Indeed, we expect FF-OCT to help reduce the rates of repeat operation and multiple biopsies, as

well as saving time and space in refrigeration for tissue conservation.
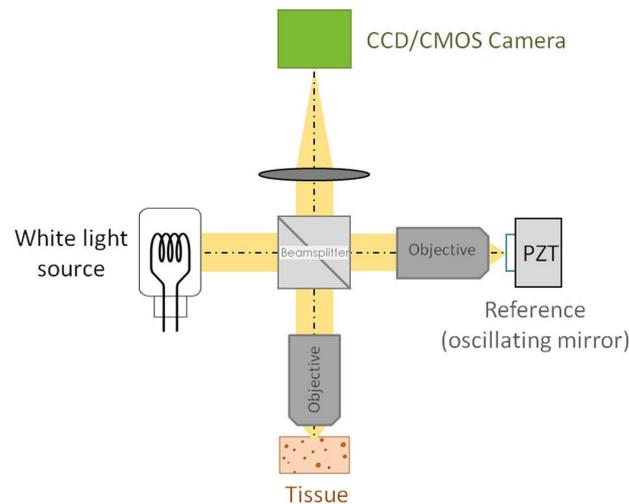
Figure 1.4: Experimental setup for Full-Field Optical Coherence Tomography

In-vivo imaging is difficult with FF-OCT because of the appearance of motion artifacts and the standard bulk FF-OCT set up. There has been some recent improvement in FF-OCT imaging [12] to make it more useful for *in vivo* and *in situ* imaging applications such as intra-operative pathology. While some of them have addressed the inconvenience caused due to bulkiness of the set-up, motion artifacts still remain a problem. With constant advances in the technology of cameras and light sources, one can reasonably predict that the acquisition speed of FF-OCT will be considerably improved in the near future. FF-OCT would then become a powerful tool for ultrahigh-resolution *in vivo* imaging without any contrast agent, making *in situ* examination possible without the need for histological processing of tissues. This makes FF-OCT an ideal candidate for intra-operative pathology in the future.

1.3. NEED FOR AN AUTOMATED FRAMEWORK

Establishing a dense correspondence between histology and *ex-vivo* FF-OCT can help in identifying surgical margin in tissue using FF-OCT. Identifying descriptors for diagnosis using FF-OCT is also very much at its nascent stage and often involves observing corresponding sections from histology and FF-OCT [11]. Registration between histology and FF-OCT for establishing this correspondence can help speed up the aforementioned clinical procedures. Any shift to a new technology requires an acclimatization period. Extensive use of FF-OCT in pathological diagnosis would require training of personnel to extract the same visual information from FF-OCT images as they would from histological slides. Histological slides are examined by trained pathologists who usually undergo a lengthy (5 to 5.5 yrs in the UK) training program. The aim of the work presented here is to perform registration of FF-OCT and histology images so that the aforementioned deformations in tissues are compensated and the transition from using pathological slides for diagnosis to using

OCT images is made as seamless as possible.

Moreover, with the advent of fast whole-slide image scanners, pathology labs have moved towards a fully digital workflow, which aims to replace viewing of glass slides under a classical microscope with viewing of digitized slides on a computer monitor. A completely automated framework for registration between an incoming FF-OCT image and a histological digitized slide would fit right into such a digital workflow for making the transition to using FF-OCT in intra-operative pathology in the future.

This thesis provides a framework for the problem of inter-modality non-rigid registration between whole-slide scanner histology and *ex-vivo* FF-OCT images. Since deformations are unavoidable at the time of tissue preparation (since tissue is squashed against the glass of tissue holder) we assume that the corresponding sections could be deformed and hence there is a need for non-rigid registration. A slide prepared for histological examination could contain multiple tissue samples from the same subject. We consider establishing a match between an incoming *ex-vivo* FF-OCT image and one of the several tissue samples present of the same subject as an integral part of the problem as it is in the spirit of automatizing the clinical procedure as discussed before.

1.4. ORGANIZATION OF THE THESIS

As a starting point, we present a bird's eye view of the approach followed by delving into the details of each aspect of it. The proposed framework consists of a two-stage registration process. The first step consists of large-scale misalignment correction while also establishing the match between FF-OCT and one of the several histology tissue samples based on iterative closest point [13] of prominent edge points. The second step starts with an area based affine registration algorithm and culminates with a Deep Flow based registration algorithm. Optical flow based algorithms are generally not used for inter-modality registration problems. This is becuse the brightness constancy assumption (in optical flow algorithms) between the two images in different modalities does not hold true. This is solved by introduction of an edge preserving gray-level based modality transformation. This is motivated by [14] but is different from it in the sense that it uses edge-orientation information as a means for establishing confidence of joint occurrence of gray-levels used in modality transformation. Figure 1.5 presents the overview of the proposed algorithm. The incoming arrows to an algorithmic block represents the required inputs for the block and the outgoing arrows from a block represents the output of the block. The details of the proposed framework are organized in the remainder of the thesis as follows:

Chapter 2 presents a brief literature survey for inter-modality registration in general. This is followed by prior art in various algorithms needed for point-cloud generation and point-cloud based affine registration. This is followed by a discussion on different techniques used for non-rigid registration and various advance-
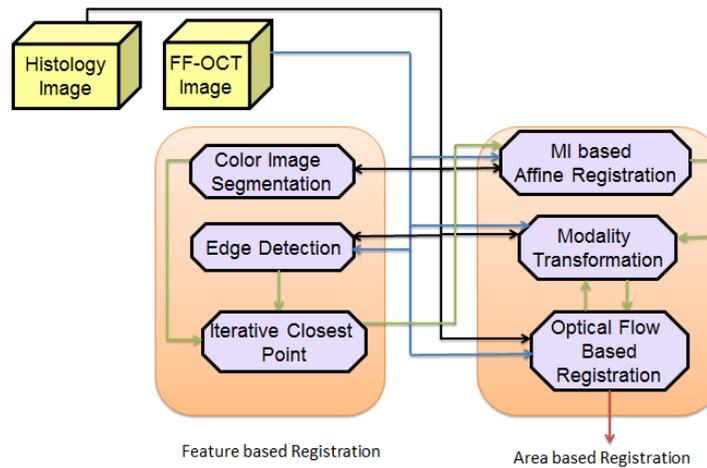
Figure 1.5: Overview of the proposed framework

ments made in Demon registration. Lastly, prior art in OCT-Histology registration is discussed towards the end of this chapter.

In **Chapter 3**, Section 3.1 describes the various algorithms needed for point-cloud based affine registration. This includes the supervised edge detection algorithm used for the purpose of point-cloud generation and iterative closest point algorithm for rigid registration. Section 3.2 elaborates the multi-stage MI based affine registration algorithm followed by methodology used for performing modality transformation and Deep flow based registration to achieve our final goal of non-rigid registration between histology and *ex-vivo* FF-OCT images.

Chapter 4 starts off with materials and methods used for acquiring the images. It is then followed by the results of the proposed algorithms and evaluates the results obtained in comparison with some of the existing registration algorithms.

Chapter 5 concludes the thesis with a discussion of results and its implications to the future work in this field.

2

LITERATURE SURVEY

This chapter introduces the relevant prior art for different techniques used in this work. Since registration between FF-OCT and Histology is our end goal, we first look at techniques used for inter-modality registration in section 2.1.

As described in figure 1.5, the first step in the proposed framework is feature based affine registration using edge points as features. Robust edge detection is an important part of the feature based affine registration algorithm. Hence section 2.2.1 is dedicated to literature survey of various supervised edge detection algorithms. Color image segmentation is needed for separating the different tissues in the histology image. Section 2.2.2 covers the necessary prior-art for color based image segmentation and section 2.2.3 is dedicated to different variants of iterative closest point algorithm which is needed for point-based registration.

Lastly, to account for deformations between FF-OCT and Histology, non-rigid registration algorithm is used. Section 2.3 covers the literature survey for this part of the work.

2.1. INTER-MODALITY REGISTRATION

Image registration is the method by which correspondence or alignment between two images is established by estimating a spatial transformation from one image to the other. Figure 2.1 shows the spatial transformation (T) that has to be estimated for the image on the left to be aligned or registered with the image on the right.

The most basic components for an image registration algorithm are a fixed image, a moving image, com-
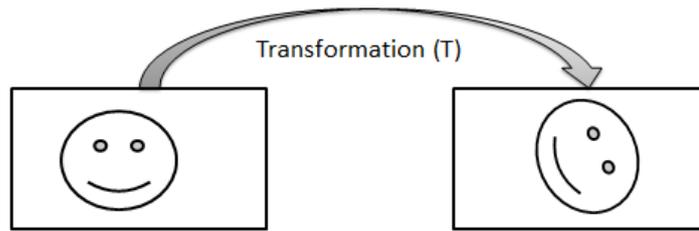
Figure 2.1: Spatial transformation

parison metric, transform estimator and an interpolator. Fixed image is the image to be registered with and the moving image is the image which needs to be registered. Comparison metric is a way of comparing the similarity or dissimilarity of different corresponding feature points or areas in the two images. Transform estimator is the functional block which estimates an optimal transformation such that the moving image is best registered with the fixed image. This estimation is often done iteratively until the algorithm converges. Such an iterative algorithm often uses an optimzation algorithm such as gradient descent algorithm. An interpolator is needed so as to resample the moving image after transformations to obtain pixel intensity values in non-integer locations. All these functional blocks of registration is illustrated in figure 2.2.
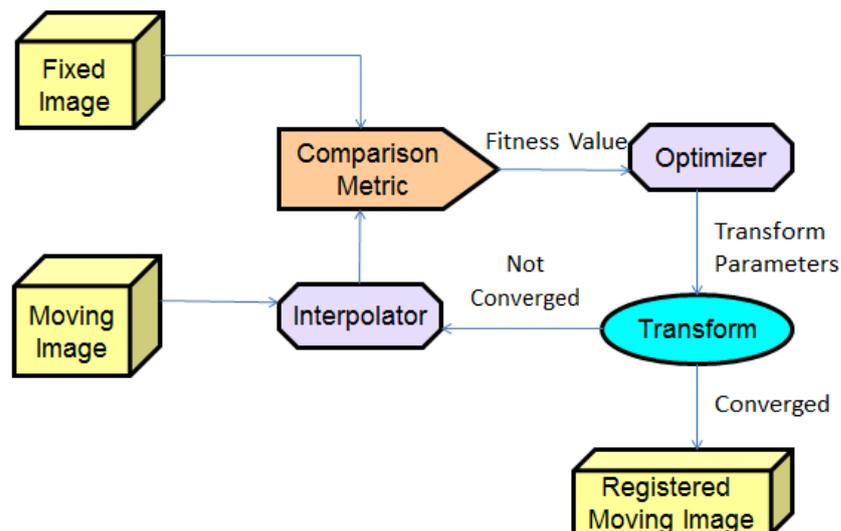
Figure 2.2: Functional blocks in a registration algorithm

Registration between images of two different modalities is a well researched topic. The two widely used techniques for registration are feature-based and area-based registration. A complete overview of area-based and feature-based registration techniques is given in Zitova *et.al.* [15].

Feature-based registration methods typically consists of four main steps. They are, feature detection, feature matching, transform parameter estimation, and image transformation accompanied with interpolation. Feature detection and feature matching gives the fitness value for the optimizer as seen in figure 2.2. Features

used in feature-based registration could be one of edges, contours, corners or salient points. The choice of features is essentially application specific. For e.g., salient points in a registration between CT angiographic images and MR angiographic images could mean invariant points which are visible across modalities [16]. Feature matching is the step in which correspondence between the two images are found. In feature based techniques, correspondence between feature points can be established using similarity measures alone [17] or in combination with spatial relationships [18]. The correspondences thus established could sometimes be conflicting each other and it might not be possible to accommodate a transformation where-in all the established correspondences coincide. Hence a transformation model is defined and parameters of the transformation is established such that similar feature points are matched in an optimal way within the constraints of the transformation model.

Metrics for area-based registration could either be based on Correlation or Mutual Information(MI). MI based metric is often preferred in inter-modality registration. Pluim *et.al.* [19] provides a complete review of MI based registration in medical image analysis.

Mutual information [20] is a information theoretic similarity metric and is defined in equation 2.1.

$$MI(A, B) = H(A) + H(B) - H(A, B) \tag{2.1}$$

where, A and B are the images to be registered. H is the entropy measure and MI is the mutual information measure. It can be inferred from the equation that maximizing the similarity measure $MI(A, B)$ is related to minimizing the the joint entropy measure $H(A, B)$, which is in fact a dissimilarity measure. The most common entropy measure used for MI calculation is Shannon entropy [21]. However, there are a few instances of using Jumarie entropy (such as [22]) for registration as well. Other variations of MI includes Normalized Mutual Information (NMI) [23] and Entropy Correlation Coefficient (ECC) [24]. These variations were suggested to make the fitness value independent of the overlap between the fixed image and the moving image. Equations 2.2 and 2.3 give their corresponding definitions.

$$NMI(A, B) = \frac{H(A) + H(B)}{H(A, B)} \tag{2.2}$$

$$ECC(A, B) = 2 - \frac{2}{NMI} \tag{2.3}$$

Since images from different modalities often have complimentary imformation, correlation does not serve the purpose of providing us with a good fitness value. However there have been a few instances of using gray level transformations to create pseudo similar modality images and treat the problem as a intra-modality registration problem. A global histogram based transformation of images from different modalities

was introduced in [25]. Following the transformation, registration was performed using correlation as the fitness value. One of the major drawbacks of this approach is that the transformation between gray levels was a global one. Images from multiple modalities in general do not have a global one-to-one correspondence between the different gray levels in the two images. A gray level g_1 might correspond to a gray level g_2 in one part of the image and g_2' in a different part of the image. This local mapping based transformation was proposed by Kroon *et.al.* [14]. Chapter 3.2 discusses this in further detail.

The next step in a registration problem is the definition of 'allowed' spatial transformations between the two images. The set of 'allowed' spatial transformations defines the search space for the problem. The different classes of transformation that can be considered are: Rigid, Similarity, Affine and Non-rigid.

Rigid transformation is a transformation where only rotation and translation of the moving image is allowed. The search space is the smallest in this type of transformation. Similarity transformation is the one in which the allowed degrees of freedom are rotation, translation and uniform scaling along each dimension. Affine transformation is a transformation where rotation, translation and non-uniform scaling along different dimensions are considered. Non-rigid transformation is where deformations in space are considered. Based on the model used for deformation, the search space for these transformations can vary a lot. These types of transformations are covered in detail in section 2.3.

Once the search space and fitness metric is defined, the registration problem effectively boils down to finding an optimum point in the search space which optimizes the fitness metric under consideration. If the search space is small enough, this could be found out by evaluating the fitness metric for all possible transformations. But this is not practical for two reasons:

1. The number of dimensions in the search space even in $2D$ similarity transform is quite high. For affine and non-rigid transformations evaluating the fitness value for all possible transformations is next to impossible.

2. Transformations parameters can take real values. Evaluating 'all' possible transformations would involve quantization of transformation parameters. The true optimum could lie between the quantized levels.

Hence to find an optimum transformation, iterative optimization algorithms are used. Suppose the initial set of parameters is represented as μ_0 and our aim is to arrive at an optimal set of parameters ($\hat{\mu}$) iteratively. The iterative updation of parameters is given in equation 2.4.

$$\mu_{k+1} = \mu_k + a_k \times d_k \tag{2.4}$$

where, k represents the iteration number, d_k is the search direction and a_k is the gain factor. Popular optimization techniques which follows the above equation include Gradient Descent [26], Quasi-Newton [27],

Levenberg-Marquardt [28], Stochastic Gradient Descent [29] [30] and Evolutionary strategy [31]. A comparative study between several such optimization algorithms can be seen in [32].

2.2. Point Cloud based Affine Registration

The three major algorithmic blocks leading to an affine registration are robust edge detection, color based image segmentation and iterative closest point based registration. A brief literature survey for each of these blocks is discussed in this section.

2.2.1. Supervised Edge Detection

Edge detection is a critical low level computer vision task which has been a subject of research ever since Roberts introduced a filter to detect edges in his thesis [33]. This was followed by the introduction of an isotropic edge detector by Sobel[1]. The ability of an edge detector to detect actual edges in the image depends on the scale at which image is being viewed. There has been several attempts at designing multi-scale edge detectors. One of the earliest papers on multi-scale edge detectors is proposed by Ziou *et. al.* [35]. An overview of several such detectors is given by Ziou and Tabbone [36].

However, visually salient edges need not necessarily correspond to dominant gray scale or color gradients in multiple scales as observed by Martin *et.al.* [37]. Figure 2.3 gives an illustration of prominent visual edges(ground truth) in a few images from the BSDS500 database [38]. It can be seen that the texture in the cloth on the floor (top left image), hay stack (top center image), ground (top right image) would have all led to detection of several edge points using unsupervised edge detection algorithms and the resulting edge maps would have been significantly different from the ground truth seen in the bottom row. This has led to a recent revival in the field of edge detection using supervised techniques.

One of the earliest edge detection methods purely based on learning algorith was suggested by Dollar *et.al.* [39]. Probabilistic boosting tree classification algorithm was used to classify the edges. The features used for classification includes multi-scale oriented gradients, Haar wavelet responses and histogram of filtered images obtained from filters such as multiscale Difference of Gaussian. The classifier was trained for both positive and negative examples for obtaining better discrimination. Zheng *et.al.* [40] extend this work further by developing a hierarchical edge classifier where the result of [39] is used as the result of the first stage. This is refined by considering context information and high-level shape information to arrive at the final classification.

Mairal *et.al.* [41] used discriminative sparse coding techniques to solve the edge detection problem. Normally in a sparse coding problem, learning consists of constructing a dictionary ($D = [d_1, d_2, ..., d_k] \in \mathbb{R}^{n \times k}$)

---

[1]It is interesting to note that Sobel never actually published this work. It was first reported by R.Duda and P.Hart [34] in which the author had credited Sobel for this filter.
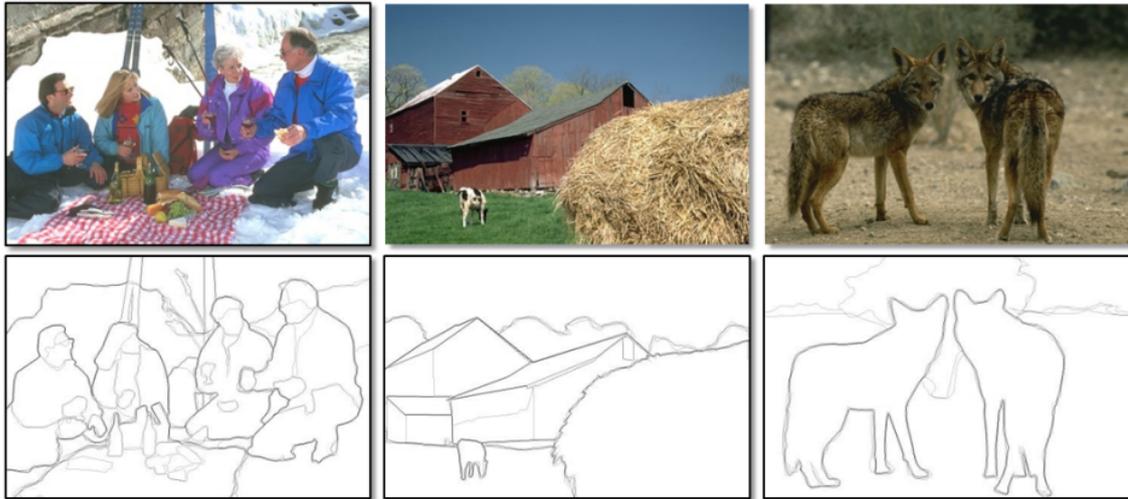
Figure 2.3: Illustration of Visual Edges from BSDS500 database

based on training signals $y \in \mathbb{R}^n$. In the testing phase, incoming signal is then represented as a linear combination of the dictionary elements $y = \alpha \times D$, where α is a k dimensional sparse vector. In discriminative sparse coding, dictionaries for each class is constructed independently based on training data for each class. In case of edge detection problem, the two classes represents edges and non-edges in the images. This was later extended to the problem of contour detection in [42].

As seen before, low level representation such as color gradients do not by themselves give a robust representation of the edges as percieved by human vision. Several aforementioned learning based techniques can be seen as a way to bridge this gap between low level representation and high level perceived features. Lim *et.al.* [43] introduce an efficient mid-level features/representation to bridge the same gap. These features, termed as sketch tokens, are learned based on the ground-truth and are seen to be quite effective in contour detection.

In the hierarchical edge cassifier [40] the misclassification in the initial stages of classification is passed on further to the next stages of classification. Instead of ignoring the context information in the first stage and then using context information later to modify the results, Dollar *et.al.* [44] propose structured forests for edge detection where classifiers are trained based on structured labels (based on topological distribution of edges and non-edges in the ground-truth image) and not just on pixel based labels. Orientation of the edge is not taken into consideration for this work. Hallman *et.al.* [45] use edge orientation across multiple scales along with other color based features for training a random forest to detect edges. Contextual information which comes while using structured labels is not used in this work.

Kivinen *et.al.* [46] present a deep neural network based method for edge detection. This is one of the earliest papers using deep neural architectures for the purpose of edge detection. This was later extended in [47] for contour detection.

2.2.2. COLOR IMAGE SEGMENTATION

Image segmentation consists of achieving a compact region-based description of the image scene by decomposing it into meaningful regions sharing similar attributes. It is a process of dividing an image into a set of homogeneous regions such that the union of any two adjacent regions is non-homogeneous. The features used for segmentation may include one or more of gray-level intensities, color, texture and edge based features. Using these features, segmentation can be approached as an optimum threshold selection problem [48] or a clustering problem [49]. It may also be posed as a region growing [50], merging or splitting based approach [51]. Markov Random Fields [52] and graph-theoretic approaches [53] have also been widely used for segmentation. Since the prior art in image segmentation algorithm is huge, in this section we only cover some of the most popular color image segmentation algorithms. An illustration of the result of a typical color image segmentation algorithm is shown in figure 2.4.

Figure 2.4: Illustration of Color Image Segmentation

In [54], Cheng and Li proposed a color image segmentation algorithm based on fuzzy homogeneity histogram. They defined homogeneity as a composition of two components: standard deviation and discontinuity of the intensities, and used the homogeneity histogram to detect uniform regions. The scale space filter was used to compute the optimum thresholds.

Normalized Cut [53] is an algorithm where the segmentation task is posed a graph partitioning problem. For this purpose, an incoming image is represented as a graph where each node is a pixel and each edge between any two nodes is representative of the pixel affinities for them to remain in the same group. The affinity could be a function of on any of the aforementioned features. The algorithm proposed by them is different from the traditional min-cut algorithm in the sense that it penalizes a cut which results in a very small partition.

Modeling the affinity metric which would define a proper segmentation for the application under consid-

eration can be quite tricky sometimes and might lead to over-segmentation. Several affinty measures or in a broader sense, different segmentation algorithms might lead to different segmented images. Li *et. al.* [55] show that super-pixels (segmented regions in an image) give important low-level cues for grouping. They propose an algorithm for aggregating results from multi-layer super-pixels obtained from several algorithms and arrive at a final result in a principled manner.

Mean Shift based segmentation [56] is a non-parametric clustering technique which segments an incoming image based on which mode (a frequently occurring feature value) it belongs to. The red dots in figure 2.5 shows the modes. On convergence, each pixel in the image gets clustered with one of the modes, which results in a segmented image.
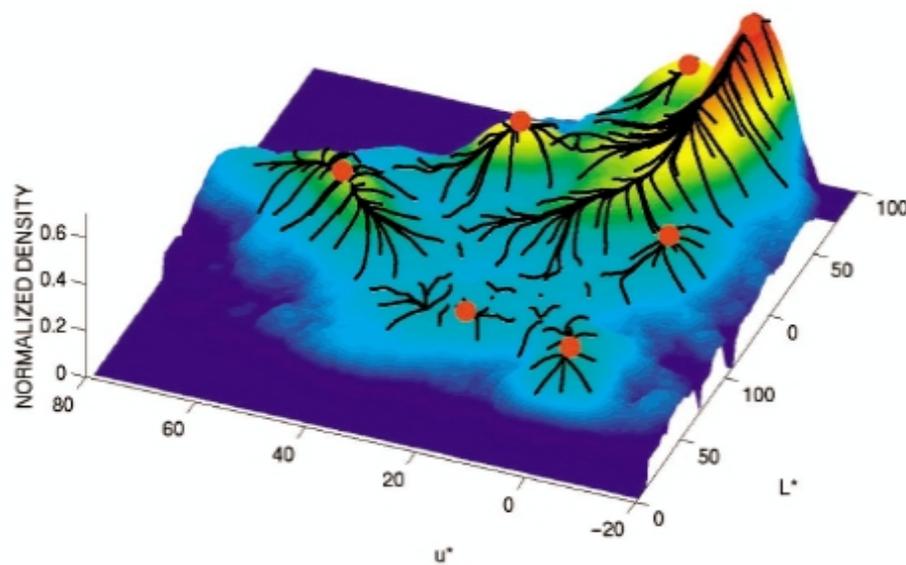
Figure 2.5: Illustration of Mean Shift Segmentation

2.2.3. ITERATIVE CLOSEST POINT

Iterative closest point (ICP) was first introduced in [13] as a way of performing registration between $3D$ shapes. Many variants of ICP has been proposed over the years. The variations come from how the points are sampled across the shapes, how they are matched and how the different matches are weighted to arrive at a fitness value and how the optimization is done for converging at an optimum transformation. A complete overview of different variants of ICP and comparison of their performances is given in [57]. An illustration of the point clouds, before and after alignment using ICP is given in figure 2.6.

In the original paper Besl and McKay used all the points in the shapes for alignment. When the number of points in the shape become very high, to reduce the computational complexity, uniform sampling [58] or random sampling of points in each iteration [59] may be used.
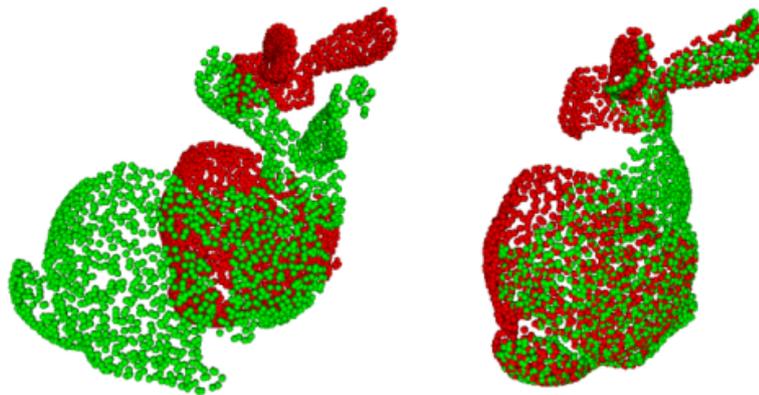
Figure 2.6: Illustration of establishment of point correspondence [1]

Matching of points from the two sampled point-clouds can be done based on nearest neighbour estimation or based on the point of intersection on the second point-cloud of a normal which originates from one point cloud (at a particular point) [60]. Nearest neighbour estimation is usually done by creating a k-d tree of the points in the point clouds [61]. The k-d tree based estimation of nearest neighbour is seen to be computationally more efficient than a brute-force nearest neighbour approach.

Weighing different pairs of matched points is an important step in an ICP algorithm. Presence of outliers can bias the result unfairly towards the outliers. Also, certain types of matches are more reliable than the others. In [62], the matches were ranked based on the distance between the points and the worst 10% of the matches were termed as outliers and rejected. In [59], any match worse than a certain predefined threshold (of distance) was termed as an outlier. Godin *et. al.* [63] weighs the matches based on the alignment of the normal direction between the points matched.

Once the fitness value for a particular alignment between the two point clouds is estimated, different optimization algorithms as covered in section 2.1 can be used for obtaining the optimally registered point-clouds.

2.3. Non-Rigid Registration

Non-Rigid registration or deformable registration can broadly be classified by the way in which regularization is applied for the deformable transformation. Implicit regularization is usually achieved by restricting the transformation space either by parameterization or by considering special spaces such as sobolov space. Explicit regularization can be achieved either by using regularizes such as a penalty term varying with the amount of deformation or by explicitly smoothening the deformation field. The deformation field of a typical non-rigid registration algorithm can be seen in figure 2.7.

fixed   moving   deformed moving
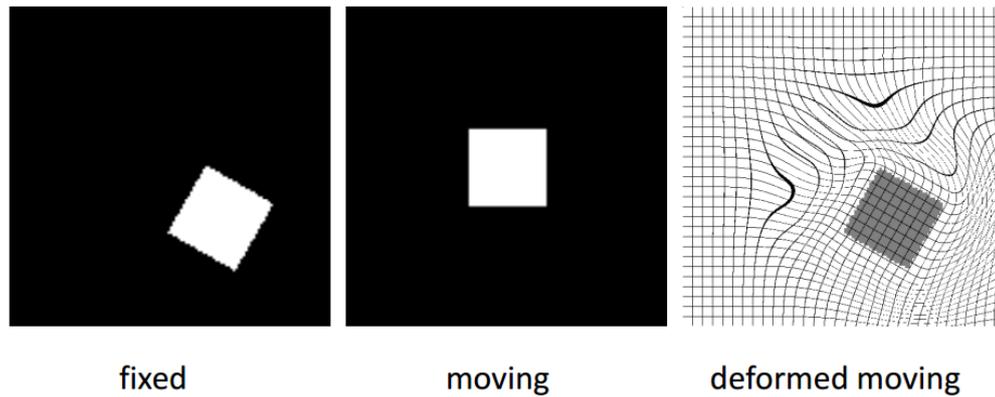
Figure 2.7: Deformation mesh of a non-rigid registration

2.3.1. IMPLICIT REGULARIZATION

Parameterization of transformation can either be done by thin plate splines [64] [65], B-splines [66] [67], or radial basis functions [68]. Thin plate splines bring about the deformation in the entire image as they have a global support region where as in case of B-splines and radial basis functions, the deformation is local and the effect can be seen in a small neighbourhood around the spline. Hence while using B-spline or radial basis function based kernel for parameterization, a swarm of similar kernels are considered throughout the image so that local deformations throughout the entire image can be considered as a possibility while registering the image. Each kernel is controlled by a control point. Since the transformations are defined only in the area local to each control point, in the area in-between two control points transformations are propagated. This propagation can either be linear interpolation [69], Gaussian kernels [70] or any other symmetrical convex kernels.

Implicit regularization of the deformable registration can also be brought about by defining the transformation in a regularized space. Thus all the transformations are by default regularized. This was first introduced in [71] and later extended in [72] and [73].

2.3.2. EXPLICIT REGULARIZATION

Regularization can also be obtained by explicitly filtering the deformation field. This is typically used in algorithms where a stationary vector field is obtained as a deformation. Demon registration [74] is a nonparametric deformable registration modeled as a diffusion process. The forces responsible for motion field estimation are inspired from optical flow equations.

Demon algorithm uses image gradient as the image based force. Equation 2.5 gives the image force ($\vec{u}$). Intuitively, demons algorithm pushes the pixels towards the gradient in case the pixel value in the fixed image (I_F) is lower than the pixel value of the corresponding pixel in the moving image(I_M) and away from the gradient if it is higher.

$$\vec{u} = \frac{(I_M - I_F)\nabla I_F}{|\nabla I_F|^2 + (I_M - I_F)^2} \tag{2.5}$$

The obtained displacement field is Gaussian smoothed which serves as a regularization function. Even though this gives a smooth displacement field, deformable registration is not just about intensity differences. Ideally, the transformation from the moving image to the fixed image should be invertible and without foldings. This is not guaranteed by the above approach, where the deformations are additive. Vercauteren *et.al.* [75] introduced a diffeomorphic demon registration algorithm which fixes the above problems. Vercauteren *et.al.* [76] has also proposed demon-based algorithm in a spherical geometry which results in implicit smoothness (because of the geometry) as well as explicit smoothness (because of the regularization term).

$$E(\vec{u}) = ||I_F - I_M \circ \vec{s} \circ exp(\vec{u})||^2 + ||\vec{u}||^2 \tag{2.6}$$

The difference between the results of an additive demon algorithm and diffeomorphic demon algorithm can be seen in figure 2.8.
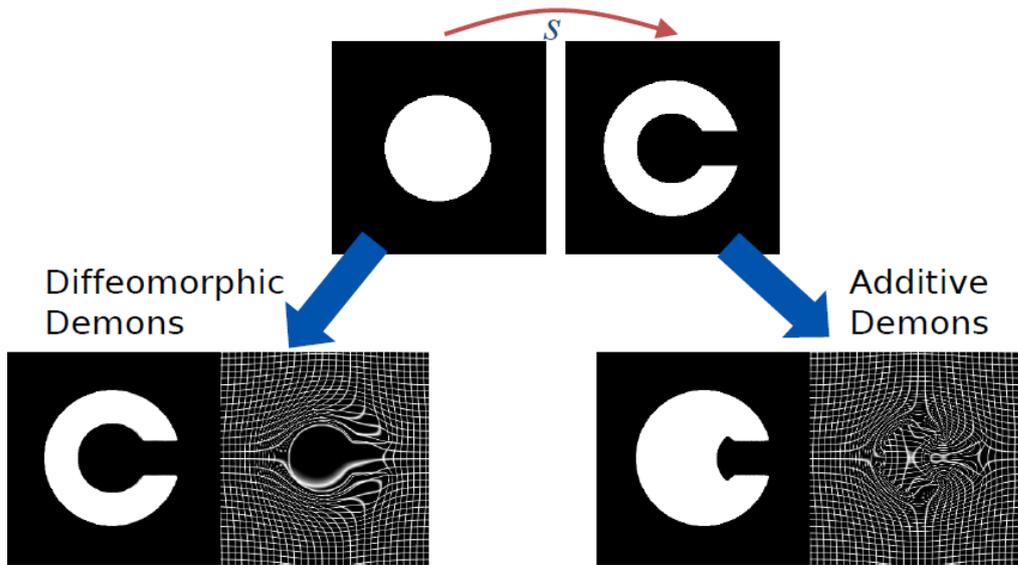
Figure 2.8: Difference in the deformation field of additive and diffeomorphic demon algorithm

Instead of having a moving image and a fixed image like most registration algorithms, a symmetric optical flow based approach has been proposed in [77], where both the images are deformed symmetrically towards each other.

At the core of every Demon-based registration algorithm is a step which finds a vector field which mimics the deformation of the image under consideration to obtain the target image. This can be seen as finding a dense correspondence between the two images. However, there are a few algorithms which figures this out

outside of the demon registration framework. Hermosillo *et.al.* [78] devise a local statistical criteria based method for coming up with dense matching of two images. Adding a regularization term to constrain local deformations is also not specific to demon based algorithms. [79] provides an alternative approach with similar goal.

3

PROPOSED FRAMEWORK

This chapter is dedicated to providing the details of the proposed framework. With an incoming *ex-vivo* FF-OCT image and a histology image captured using a whole slide scanner, it is of foremost importance to establish the correspondence between the FF-OCT image and one of the several tissues in the histology image. Multiple tissues in a histology image digitized using a whole slide scanner can be seen in figure 3.1. FF-OCT image corresponding to one of these tissues is in figure 3.2. Manually selecting the corresponding tissue in the histology image is not convenient for the user. A deformable registration with each one of these tissues is computationally too expensive and can be avoided. Hence a multi-stage registration process has been proposed here.

To establish the correspondence between the FF-OCT image and histology image, we use a feature based registration algorithm where significant edge points are taken as features. The fitness value of the registration is used for establishing the correspondence. Once the correspondence is established, an area based affine registration is done to refine the result further. This further serves as an initializing step for the deformable registration using a novel deep diffeomorphic demon registration algorithm. The following sections provide the specifics for various aspects of the proposed multi-stage registration framework.

3.1. POINT CLOUD BASED AFFINE REGISTRATION

A feature based registration algorithm with significant edge points used as features, is used for the purpose of establishing correspondence. The choice of feature based registration stems from the fact that feature based
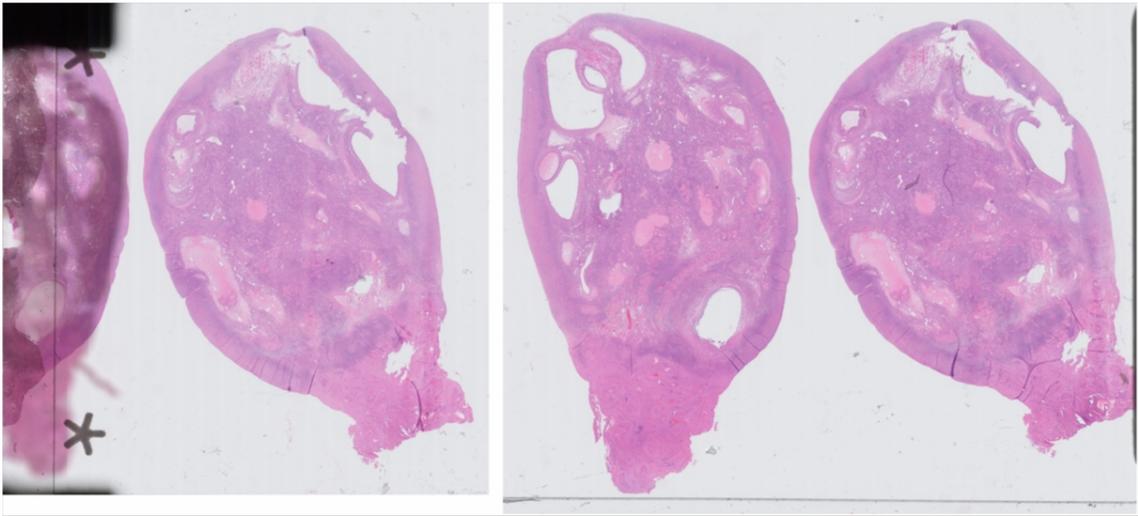
Figure 3.1: Histology Image with Multiple Tissues

Figure 3.2: FF-OCT image

algorithms usually perform better in large-scale misalignment correction. Area based registration algorithms have a tendency to get stuck in a local optima unless the initial position before registration is indeed close to the final solution. Figure 3.3 gives a subset of the figure 1.5 and concerns only with the point cloud based affine registration. As seen in this figure, there are three main functional blocks in the point cloud based affine registration module of the algorithm. Those are:

1. A Robust Edge Detection Algorithm to detect the feature points in the images from both the modalities (section 3.1.1).

2. Color Image Segmentation to Separate the different tissues in histology image (section 3.1.2).

3. Point Cloud Registration (section 3.1.3).

The algorithmic details of these functional blocks are covered in the following subsections.
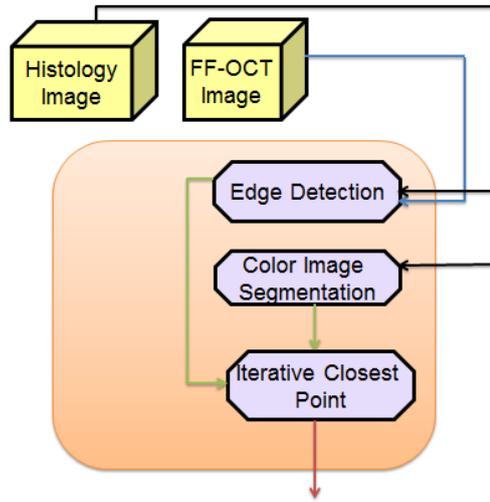
Figure 3.3: Overview of the point cloud based registration algorithm

3.1.1. STRUCTURED FOREST BASED EDGE DETECTION

Robust edge detection is extremely important for the success of this registration. Hence we use Structured Forest Based edge Detection (SFED) [44], a state-of-the-art supervised edge detection algorithm for this purpose. The current sub-section explains the SFED algorithm in detail.

RANDOM FOREST FOR STRUCTURED OUTPUT:

Random forest is an ensemble classifier that consists of T decision trees $(f_1, f_2, ..., f_T)$. A decision tree $f_t x$ classifies a sample x by recursively branching left or right (until it reaches a leaf node) based on whether the sample value is less than or greater than a threshold θ_j associated with the j-th node. Even though the incoming sample x could have multiple dimensions, the threshold at each node usually corresponds to a single feature dimension chosen at random for each node at the time of training. The output of the tree is the prediction stored at the leaf node reached by incoming x. Figure 3.4 illustrates this process diagrammatically.

Given an input sample, each tree in the decision forest comes up with its own prediction and an ensemble model combines these decisions into a single output. A commonly used ensemble model is majority voting. For a detailed compilation of different ensemble models, one may refer to Criminisi *et. al.*'s work [2]. The aforementioned process describes testing of an already trained forest. The remainder of the sub-section is on training the decision forest.

Each tree is trained independently to select a threshold θ_j at each node for a randomly selected feature out of a multi-dimensional feature vector. The selection of this threshold is based on information gain criterion. Given a feature vectors in a multi-class classification problem, the selection of threshold along a particular
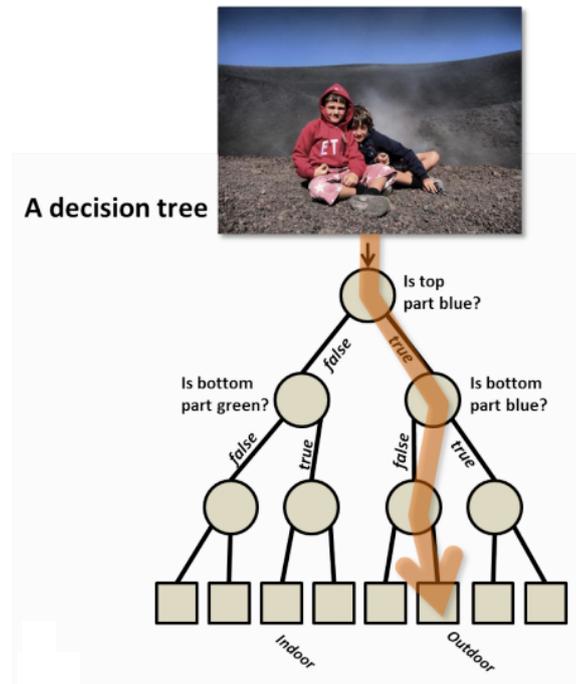
Figure 3.4: Illustration of Decision Tree [2]

dimension (which is randomly selected for a node) is done such that the resultant information gain is maximized. Figure 3.5 illustrates the net information gain as a result of choosing an optimal threshold (c) and compares this with the gain as a result of choosing a sub-optimal threshold (b).
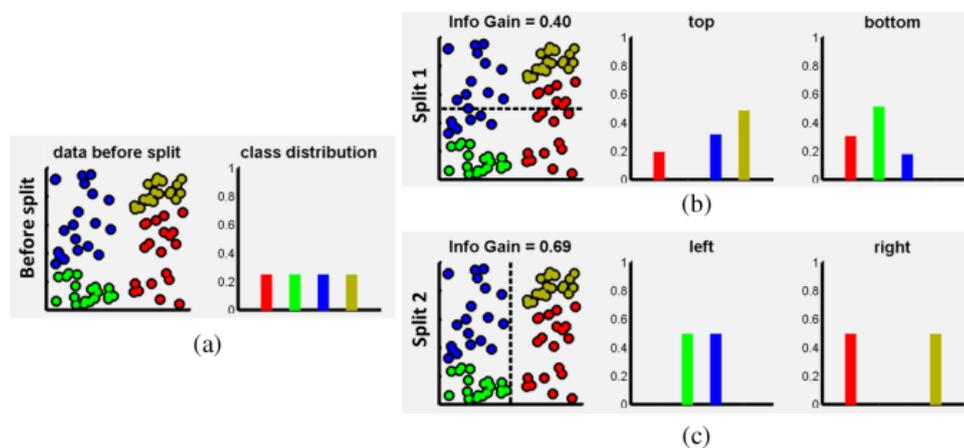
Figure 3.5: Information gain due to selection of a threshold [2]

TRAINING A RANDOM FOREST FOR STRUCTURED OUTPUT:

The outputs of the multi-class classification problem considered previously is a scalar. The output is considered a structured output when complex objects such as sentences (in natural language processing) or scalars along with spatial information between them (in object localization) is the desired result. For edge detection, this structured output used is an image (with edges). The two main difficulties in training a random

forest with a structured output is defining information gain and the second is prohibitively high complexity of candidate splits at each node.

Dollar and Zitnick [44] propose a two step procedure for computing an approximate information gain measure. At the time of training, the structured output is considered to be a 16×16 ground-truth segmentation patch. Each pixel pair is evaluated to check if they are in the same segment or different ones in the ground-truth. However evaluating $\binom{16\times16}{2}$ pairs is computationally expensive. Hence, 256 randomly selected pixel pairs are evaluated to reduce the dimensionality of the problem. After evaluation, 256 long binary vector for all the training samples are aggregated and principal component analysis [80] is used to further reduce the dimensionality to at most 5. Based on the 5 dimensional binary feature space, structured output is clustered to k different labels. The high complexity of candidate splits can be effectively managed by restricting the value of k. For $k = 2$, the splits become binary and the approximate information gain gets reduced to its scalar output counterpart. Based on k different labels, information gain at a node (I_j) can be computed using Shannon entropy. Equation 3.1 shows this relation.

$$I_j = H(S_j) - \sum_{k\in\{L,R\}} \frac{|S_j^k|}{|S_j|} H(S_j^k)$$

(3.1)

Where, L and R are the left and right split in a binary split at a node, and $H(S)$ denotes the Shannon entropy given in equation 3.2, where p_y is the fraction of elements in S that belongs to label y. Intuitively, this gets maximized when all the elements belong to the same label.

$$H(S) = -\sum_{y} p_y log(p_y)$$

(3.2)

FEATURE EXTRACTION FOR EDGE DETECTION:

Feature extraction is an important part of any supervised classification problem. Feature extraction for edge detection is done in each 32×32 image block to provide a sub-sampled structured output of 16×16 edge map. For edge detection, the following features are used:

- Three color channels of the image in CIE-LUV color space.

- Normalized gradient magnitude at 2 scales, the original and half resolution

- Four gradient orientation channel for each gradient magnitude channel.

- Pairwise difference feature for each of the aforementioned 13 feature dimensions. Since computing pairwise difference features in 32×32 image patch leads to $\binom{32\times32}{2}$ features, to reduce the number of features, each image patch (including gradient and orientation channels) is sub-sampled to 5×5 images. Thus, for each feature dimension, a pairwise feature vector of length $\binom{5\times5}{2}$ is computed.

As it can be inferred from the above explanation, SFED edge detector is defined for color images. This is not a problem for histology image. FF-OCT image, on the other hand, is a gray scale image. To apply the SFED edge detector, FF-OCT is converted to an image with three color planes with each plane being a replica of each other.

3.1.2. MEAN SHIFT CLUSTERING FOR IMAGE SEGMENTATION

Histological images typically contain multiple tissues of the subject. Hence, there is a need to separate the point clouds (significant edge points) generated from histological images based on the tissue to which it belongs. As the tissues are non-overlapping, it is safe to assume that there is some amount of background separating the tissues. Thus, by segmenting the image into foreground (tissues) and background, one can take point clouds in each connected component of the foreground separately. For this purpose, color based segmentation technique of mean shift clustering [56] was used.

Mean shift clustering is a non-parametric clustering algorithm introduced in [81]. Its application to segmenting images and a theoretical background for mode detection in an image using mean-shift clustering was given by Comaniciu and Meer in [56]. Given a feature space, mean shift clustering computes the mean of features in a pre-defined window. The newly computed mean is taken as the center of the search window for the next iteration. This procedure goes on until convergence is achieved. Mean shift clustering in a $2D$ feature space is diagrammatically illustrated in figure 3.6.

A typical window, being similar to a step function gives unfair advantage to something that has barely made into the window while completely ignoring the data points that are outside. This is not completely fair and can sometimes lead to drastic change in the computed mean. Hence a radially symmetric kernel is often used to estimate the probabilistic densities. For a d-dimensional feature space with n samples, probabilistic density $P(x)$ is given in equation 3.3.

$$P(x) = \frac{1}{nh^d} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right) \tag{3.3}$$

Where K is the radially symmetric kernel with h as the kernel parameter. Now, going back to the original problem of computing mean around a point x_i, it can be done with the use of a kernel function as given in equation 3.4. Here, K' denotes the derivative of the kernel. The most commonly used kernel function for mean computation is a Gaussian kernel.

$$m(x_i) = \frac{\sum_{j=1}^{n} K'\left(\frac{x_i - x_j}{h}\right) x_j}{\sum_{j=1}^{n} K'\left(\frac{x_i - x_j}{h}\right)} \tag{3.4}$$

As mentioned before, converging on the cluster centers is an iterative procedure. To begin with, the mean
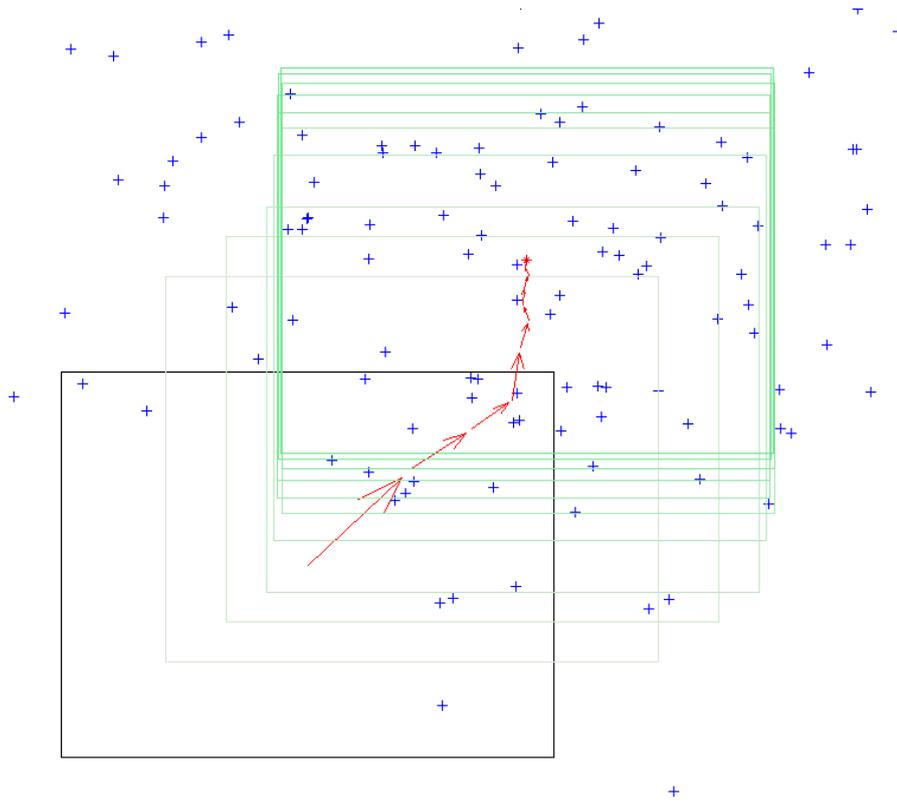
Figure 3.6: Illustration of convergence in mean-shift clustering.

at each point x_i is assumed to be x_i. After each iteration, the kernel center is shifted to the updated value of $m(x_i)$, much like the shift of the center of the rectangle in figure 3.6.

For color image segmentation, the crucial task is to select the feature space which amply describes the affinity of a pixel with respect to other pixels. This affinity can then be used for making a decision as to whether the pixels belong to the same region or different ones. There are two parts to this affinity measure, the range and the spatial affinities.

Range affinity is the perceptual distance between the two pixel intensities. If it were a gray scale image, the distance would be equal to the euclidean distance between their grey scale intensities. For a color image, the euclidean distance between their (R,G,B) values is not a measure of perceptual color difference. Hence the color space of the input image is converted to CIELUV color space before computing the euclidean distance. Spatial affinity is the spatial proximity of the two pixel locations.

For computing the range affinity in $2D$ color images such as histological images, 3 color planes from CIELUV color space are used. Spatial affinity is computed using the cartesian co-ordinates of the pixel locations in the $2D$ image. A $5D$ feature vector for each pixel is created and mean shift clustering is performed in this space. For the purpose of mean computation, a $5D$ Guassian kernel is chosen. As the first three range dimensions are fundamentally different from the last two spatial dimensions, the kernel bandwidth parameter h is also different for the two. The two parameters which control the kernel bandwidth are spatial bandwidth

parameter and the range bandwidth parameter. This is quite similar to the the homogeneity scale and the spatial scale parameters as used in the computation of attraction force field [82]. The reader is advised to refer to the original paper on attraction force field [82] for an in-depth analysis of the interdependence of spatial and range bandwidth parameter values.

Mean shift based image segmentation often leads to over-segmentation. As our aim is to separate the foreground from the background, clusters are merged based on their range and spatial proximity to each other, until there are only two cluster centres.

3.1.3. Iterative Closest Point

In the histology image, only the edge points within the foreground regions are considered for further steps in registration. All the foreground regions (connected components) are taken separately and are used to separate the point-clouds of different tissues in the same histology image.

After separating the point-clouds from each other, each one of them is processed separately with FF-OCT point-cloud. To minimize the distance between the two point clouds, we use Iterative Closest Point (ICP) algorithm. Although ICP is capable of finding an optimum affine transformation to align the point-clouds, it is quite possible that it could get stuck in a local optima. A robust initialization can go a long way in avoiding such mishaps. Hence we use a two pronged pre-processing step for coming up with a good initialization.

Pre-processing of the point-clouds consists of Orientation alignment and scale & translation normalization. Algorithm for orientation alignment is given in algorithm 1.

Input: Point Cloud: $\left(X^{(1)}, X^{(2)}, ..., X^{(K)}\right)$
Output: Point Cloud: $\left(Y^{(1)}, Y^{(2)}, ..., Y^{(K)}\right)$
Parameters: θ_0
$\theta \leftarrow \text{PCA}\left(X^{(1)}, X^{(2)}, ..., X^{(K)}\right)$;
$Y \leftarrow \text{RotateCloud}\left(X, \theta_0 - \theta\right)$;

Algorithm 1: Orientation Alignment Using Principal Component Analysis

In the above pseudo-code, the first step is finding out the orientation (θ) of major axis of the point cloud using principal component analysis (PCA). This is done by finding the mean of co-ordinate positions of the points in the cloud and subtracting the mean from each of these positions. This is followed by finding the covariance matrix of the zero-mean point cloud. Eigen values and eigen vectors are found out for the co-variance matrix. The vector that corresponds to the most dominant eigen value is the major eigen vector (or major axis). The angle of orientation of the major eigen vector with respect to one of the pre-defined co-ordinate axis (e.g.: x-axis), is the final output θ. Pictorial representation of major and minor eigen vector of a point-cloud can be seen in figure 3.7.

Once this is done, the point cloud is rotated such that the major axis has a pre-decided orientation(θ_0). The choice of this angle is arbitrary. As long as the point clouds are all rotated to a pre-decided orientation,
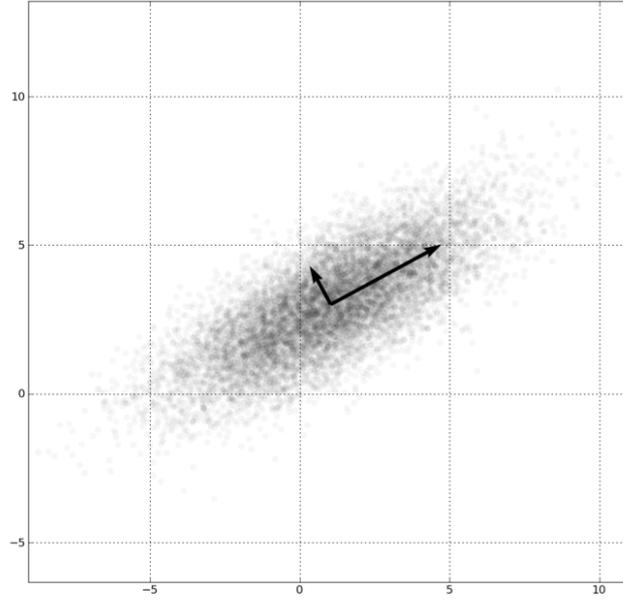
Figure 3.7: Illustration of Principal Component Analysis.

the exact choice of this orientation is immaterial.

After orientation alignment of the point clouds based on their major eigen vector is completed, each point-cloud is also scaled and translated such that the coordinates have zero mean and unit variance. The equation for scale and translation normalization of a D dimensional point cloud with K points in the point cloud, is given in equation 3.5.

$$P(d)^{(k)} = \frac{Y(d)^{(k)} - \mu(d)}{\sigma(d)} \tag{3.5}$$

Where, $d \in [1, D]$ and $k \in [1, K]$. $\mu(d)$ is the mean location of the points in the dth dimension and $\sigma(d)$ is its standard deviation. Once the point clouds are normalized, the algorithm matches each point to its closest point and figure out an affine transformation such that this distance decreases in a mean-squared sense. The pseudo-code for the ICP algorithm used for this purpose, is given in algorithm 2.

As it can be seen, this is an iterative algorithm. As discussed in section 2.1, this needs an optimizer. Levenberg-Marquardt [28] algorithm was used for this purpose. As discussed before in section 2.2.3, it is important to separate the outliers while computing the total error. $ErrorThreshold$ is hence needed as one of the user given parameters for detecting the outliers. $MaxIterations$ and δ provide two different exit condition for the iterative algorithm. The former is needed when the algorithm fails to converge properly, in order to make sure the optimization doesn't continue indefinitely. The latter is needed as a threshold to decide if the algorithm has converged.

Input: Normalized Point Cloud: $\left(P_1^{(1)}, P_1^{(2)}, ..., P_1^{(n)}\right)$

Input: Normalized Point Cloud: $\left(P_2^{(1)}, P_2^{(2)}, ..., P_2^{(m)}\right)$

Input: Gradient Orientation for points in cloud P_1: $\left(G_1^{(1)}, G_1^{(2)}, ..., G_1^{(n)}\right)$

Input: Gradient Orientation for points in cloud P_2: $\left(G_2^{(1)}, G_2^{(2)}, ..., G_2^{(m)}\right)$

Output: Normalized Point Cloud P_2 aligned with P_1

Parameters: $MaxIterations, ErrorThreshold, WeightLUT, \delta$

Initialization: counter $\leftarrow$ 0;

Initialization: TransformationMatrix $T_{(0)} = I$; % Identity Matrix

while ($counter < MaxIterations$) & $\left(||T_{(counter)} - T_{(counter-1)}|| \geq \delta\right)$ **do**

    counter $\leftarrow$ counter+1;

    $P_2 \leftarrow$ MovePointCloud($P_2, T_{(counter)}$);

    TotalDistanceError $\leftarrow$ 0;

    for $i = 1$ **to** n **do**

        $Error^{(i)} \leftarrow$ FindClosestPoint ($P_1^{(i)}, P_2, WeightLUT$);

        if $Error^{(i)} < ErrorThreshold$ **then**

            TotalError $\leftarrow$ TotalError + $Error^{(i)}$;

        end

    end

    $T_{(counter)} \leftarrow$ Optimizer ($T_{(counter-1)}$, TotalError, P_1, P_2);

end

Algorithm 2: Edge Orientation Alignment Weighted Iterative Closest Point

One of the minor refinement in the well researched ICP algorithm that has been added in this work is the way in which error is computed. It is not just based on distance between the corresponding points, but also based on alignment of the edge orientations between the points that is selected as the corresponding points. The relation is given in equation 3.6.

$$Error = D \times 0.5(1 + sin^2(\Delta\theta)) \tag{3.6}$$

Where, D is the physical distance between the two corresponding points. $\Delta\theta$ is the difference in edge orientation between those two points. Since we are considering images from multiple modalities, contrast reversal could cause aligned edges to either have $\Delta\theta = 0°$ or $180°$. Hence in cases where $\Delta\theta > 90°$, it is updated as $\Delta\theta = 180° - \Delta\theta$ before the computation of error metric. Even though the addition of edge-orientation based metric for error calculation was based on intuition, the exact choice of the function defined in equation 3.6 is empirical.

As the computation of $0.5(1 + sin^2(\Delta\theta)$ for different $\Delta\theta$ can be avoided at the time of optimization, a look up table is created for finely quantized values of $\Delta\theta$. This look up table serves as one more user defined parameter($WeightLUT$) in the algorithm.

Since the point clouds were created using SFED algorithm (section 3.1.1), gradient orientation measure is not readily available at each edge point. To circumvent this problem, gradient orientation is computed based on Sobel edge detector. Unsupervised edge detector such as Sobel operator, while not being able to detect

a visual edge point robustly, seems to quite accurately represent the edge orientation at an already selected edge-point.

Point cloud registration of an incoming FF-OCT image is performed with all the tissues in a histology slide. The one which gives the least error of registration is taken to be the best-fit tissue and the correspondence between FF-OCT image and the histology tissue is established.

3.2. AREA BASED NON-RIGID REGISTRATION

Area based registration techniques need good initialization to converge to a favourable optima and not get stuck at a local optima somewhere else. Once the match between FF-OCT and one of the several histology tissue samples of the subject is established, it provides a good initialization for the next step in multi-stage registration. Figure 3.8 (a subset of the figure 1.5) gives an overview of different functional blocks involved in the second stage in this registration framework culminating in a robust non-rigid registration solution.
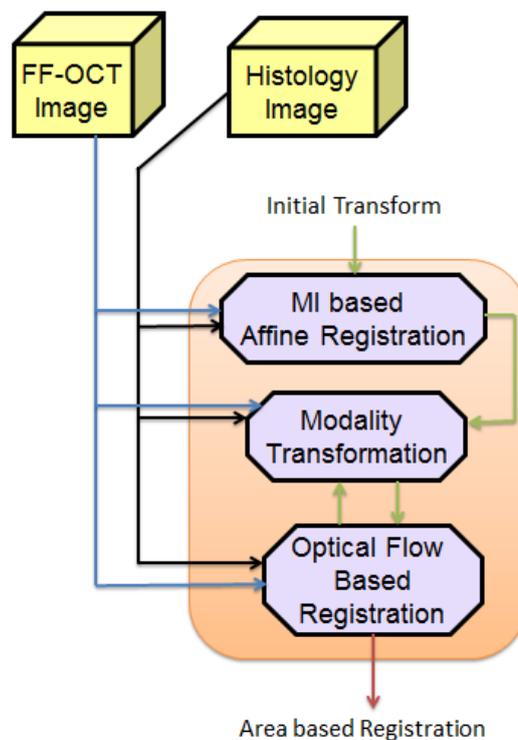
Figure 3.8: Overview of the approach used for Non-rigid Registration

3.2.1. MULTI-SCALE MUTUAL INFORMATION BASED AFFINE REGISTRATION

As discussed in section 2.1, mutual information is the most obvious choice for inter-modality affine registration. For the sake of completeness, the metric used for registration is reiterated in equation 3.7. For the computation of MI, we only use the gray level intensity values of the histology image and ignore the color

information. Optimizer used for this registration is Adaptive Stochastic Gradient Descent algorithm [30].

$$MI(I_M, I_F) = H(I_M) + H(I_F) - H(I_M, I_F) \qquad (3.7)$$

Even though the result of affine registration using edge-points is used for initializing this algorithm, edge points account for a very negligible percentage of the overall image area. Hence, it would be wise to expect that the overall optima obtained after MI based afine registration would be considerably different from point-cloud based affine registration. To help the algorithm cope with large transformations (and not get stuck in the wrong optima), we choose a multi-scale registration approach as illustrated in figure 3.9. To be precise, we perform registration in 3 different scales $\frac{1}{4}, \frac{1}{2}$ and 1 respectively.
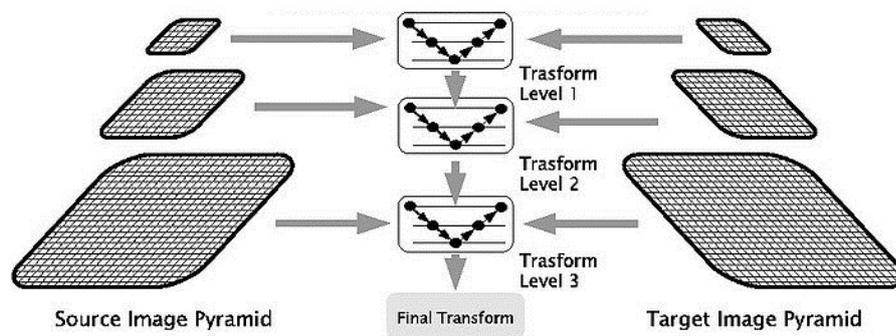
Figure 3.9: Multi-scale strategy for MI based Affine Registration

The results of registration from each previous scale is used for initializing the transformation for the next step. The result of the last stage in this multi-scale registration is later used for initializing the non-rigid registration. This sort of multi-scale approach is known to help in speeding up the convergence as well as in improving the accuracy of converged result.

3.2.2. INTRODUCTION TO OPTICAL FLOW IN REGISTRATION

Demon registration as introduced by Thirion *et.al.* [74] views the images to be registered as the same physical image moving across time. In this paradigm, the problem of finding a transformation for the moving image to best fit the fixed image is equivalent to that of finding a dense correspondence between the pixels of moving image and fixed image. In this sense, demon registration algorithm can be thought of an offshoot of optical flow algorithm.

Demon algorithm uses image gradient as the image based force. Equation 3.8 gives the image force ($\vec{u}$). Intuitively, demons algorithm pushes the pixels towards the gradient in case the pixel value in the fixed image (I_F) is lower than the pixel value of the corresponding pixel in the moving image(I_M) and away from the gradient if it is higher.

$$\vec{u} = \frac{(I_M - I_F)\nabla I_F}{|\nabla I_F|^2 + (I_M - I_F)^2} \quad (3.8)$$

As it can be seen from the above equation, it thrives mainly on edge based force of the fixed image. The image forces push the initial displacement vector $\vec{s}$ such that $I_M \circ (\vec{s} + \vec{u})$ becomes more like I_F. The newly obtained field $(\vec{s} + \vec{u})$ is Gaussian smoothed which serves as a regularization function. The smoothed displacement vector is taken as $\vec{s}$ for the next iteration. The deformation field is computed and regularized iteratively until convergence to register the two images.

Wang *et.al.* [83] proposed a modification in which the edge based force of both fixed and moving image play vital roles in image force. The modified equation is given in equation 3.9

$$\vec{u} = \frac{(I_M - I_F)\nabla I_F}{|\nabla I_F|^2 + \alpha^2 (I_M - I_F)^2} + \frac{(I_M - I_F)\nabla I_M}{|\nabla I_M|^2 + \alpha^2 (I_M - I_F)^2} \quad (3.9)$$

where α is a coefficient to decide the effect of edge based force. Although the above formulations converges to a solution, there exists no theoretical foundation which guarantees a convergence. Hence Cachier *et.al.* [84] modified the demons algorithm to that of a global minimization problem. The error metric for minimization is given in equation 3.10.

$$E(\vec{u}) = ||I_F - I_M \circ (\vec{s} + \vec{u})||^2 + \frac{\sigma_i^2}{\sigma_x^2} ||\vec{u}||^2 \quad (3.10)$$

The added term $||\vec{u}||^2$ acts as a implicit regularizer in the above equation. σ_x accounts for spatial uncertainty in the correspondence and σ_i accounts for noise in the intensity level of the image. The term $\frac{\sigma_i^2}{\sigma_x^2}$ weighs the regularizing term with the image based similarity term in the error function. The net effect of this term is quite similar to that of Lagrange's multiplier in a minimization problem. Apart from this, an explicit regularization of the displacement field obtained after each iteration using Gaussian kernel is also performed.

All the variants of Demon registration algorithm discussed so far work on the assumption that the two images that needs to be registered are obtained from the same imaging modality. For images from multiple modality, optical flow based registration algorithms are not commonly used. However, if the images from multiple modalities can somehow be transformed to pseudo-similar modality images, optical flow based registration algorithms including demon and its variants can be applied to these transformed images.

3.2.3. EDGE PRESERVING MUTUAL TRANSFORMATION

Modality Transformation is a way of transforming an image in one modality to an image in another modality based on the joint-occurrence of gray-level intensities. This transformation assumes that the input images have the same number of color-channels. In this case, the histology image is a color image where as the FF-

OCT image is a gray scale image. For computation of joint occurence of gray-level intensities as well as for further deformable registration process, we only consider gray-level intensities of histology image and ignore the color information. Mutual transform as introduced by Kroon *et.al.* [14] constructs a joint histogram using intensity values of images from different modalities. Let $N_1 \times N_2$ be the size of the joint histogram. The transformed image after modality transformation (I_{1T}) at a pixel (x) is given by:

$$I_{1T}(x) = \arg\max_{j}(H(I_1(x), j)) \qquad \forall j \in [1, N_2] \tag{3.11}$$

However, a gray-level intensity in one image modality often correspond to multiple gray-level intensities. Therefore, if I_1 and I_2 are from 2 different modalities, intensity level L_1 in I_1 may not always get uniquely mapped to L_2 in I_2. This was handled by Kroon *et.al.* using local modality transformation where the transformation was local to each pixel and was calculated using its corresponding neighbourhood. The drawback of this approach is that the edges are not properly defined in the transformed domain.

We argue that the joint histogram should not be weighed equally for each joint occurrence of intensities (L_1, L_2) in I_1 and I_2 respectively, but should be based on the confidence of each occurrence. For a perfectly registered pixel, we know that the gradient orientation in both the images would be identical [85]. This is used to measure the confidence (C) if a gray-level L_1 in I_1 corresponds to a gray-level L_2 in I_2. Thus the histogram is computed as:

$$H(I_1(x), I_2(x)) = H(I_1(x), I_2(x)) + C(x) \tag{3.12}$$

Since we are considering images from multiple modalities, contrast reversal could cause aligned edges to either have $\Delta\theta = 0°$ or $180°$. Hence the confidence measure is defined as follows:

$$C_{edge}(x) = \begin{cases} \cos^2(\Delta\theta), & for\,\Delta\theta \le 90° \\ \cos^2(180 - \Delta\theta), & for\,\Delta\theta > 90° \end{cases} \tag{3.13}$$

Confidence is computed for only those points gradient magnitude is significant and for all the other points, confidence is induced based on proximity to an edge point. Equation 3.14 brings out this relationship.

$$C_{non-edge}(x) = \{0.1 + 0.9 \times \exp\left(\frac{-d^2}{d_0^2}\right)\} \times C'_{edge} \tag{3.14}$$

where, d is the euclidean distance from the nearest edge point and d_0 is a parameter which decides the range in which an edge point has an effect on its neighbourhood and C'_{edge} is the confidence of the nearest edge point. The value of d_0 is selected based on the scale which we are operating in.

Accordingly for a perfectly aligned edge, the joint histogram is incremented by 1, whereas if the edge directions are completely misaligned ($\Delta\theta = 90°$), the joint histogram is not incremented at all. Thus, the occurrence of L_1 in I_1 and L_2 in I_2 is not given any weight at all in such a scenario. Once the histogram is constructed, equation 3.11 is used for modality transformation. This transformation preserves aligned edges completely.

3.2.4. INTRODUCTION TO DEEP FLOW

The Demon registration algorithm and its variants are known to work well only for small deformations. Large displacements are completely missed by both Demon registration algorithms and by the traditional optical flow algorithm on which it is based. With advent of optical flow algorithms which cope with large displacements, it is time we revisited the problem formulation of the Demon registration algorithm. But before we do that, this section is dedicated to the introduction of algorithm called Deep Flow [86] which computes accurate optical flow displacement vectors even in case of large displacements.

DEEP FLOW

Traditionally, variational approach has been the most sought out one for optical flow algorithms [87]. Brox and Malik [88] devised the first optical flow algorithm which combines descriptor based matching with the variational approach. Weinzaepfel *et. al.* [89] modified Brox and Malik's approach with a better descriptor based matching technique but the core idea remains the same. This new algorithm was named as Deep Flow, because of the connection the matching technique has with deep convolutional networks.

The energy function optimized in Deep Flow (equation 3.15) is a weighted summation of data term (E_D), smoothness term (E_S) and the matching term (E_M). Ω is the space in which the images exist and $\vec{w}$ is the flow vector.

$$E(\vec{w}) = \int_\Omega \left(E_D + \alpha E_S + \beta E_M \right) d\Omega \tag{3.15}$$

The data term is based on brightness constancy (E_B) and gradient constancy(E_G) as defined in [90]. Brightness constancy criteria as defined in [87] is given by $\left((\nabla_3^T I) \vec{w} = 0 \right)$, where ∇_3^T is the spatio-temporal gradient $(dx, dy, dt)^T$. If $B = (\nabla_3^T I) \vec{w}$, $B^T B$ gives the penalty term used for minimization. This is known to add higher weight in locations with high image gradient. Hence, the penalty term is divided by $\epsilon + ||\nabla_2 I||^2$ to compensate for the high weight in the numerator near gradient locations. ϵ is added to avoid divide by 0. This is summarized in equation 3.16.

$$E_B = \frac{B^T B}{\epsilon + ||\nabla_2 I||^2} \tag{3.16}$$

The gradient constancy term was defined in [90] for color images. Since we are dealing with gray-scale images, we only are only interested in the simplified version of the term. Equation 3.17 and 3.18 gives the gradient constancy part of the deep flow energy metric.

$$J_{xy} = \frac{\nabla_3^T I_x}{\epsilon + ||\nabla_2 I_x||^2} + \frac{\nabla_3^T I_y}{\epsilon + ||\nabla_2 I_y||^2} \tag{3.17}$$

$$E_G = \vec{w}^T J_{xy} \vec{w} \tag{3.18}$$

Combining the brightness constancy term and the gradient constancy term, the data term E_D is given by:

$$E_D(\vec{w}) = \delta E_B(\vec{w}) + \gamma E_G(\vec{w}) \tag{3.19}$$

The smoothness term acts as a regularizing term in the optical flow optimization problem. Equation 3.20 gives the relation between the smoothness constraint and the flow components in the two co-ordinate directions.

$$E_S(\vec{w}) = ||\nabla u||^2 + ||\nabla v||^2 \tag{3.20}$$

The matching term is based on the score obtained from Deep Matching. This part is covered in the next sub-section.

DEEP MATCHING

Rigid matching between the template and any subsequent images does not account for any relative motion between different objects in the template. This problem is addressed in the deep matching algorithm. Deep Matching is a robust matching technique capable of matching an object even after it has gone through major transformations, including deformation. It is based on a descriptor constructed using histogram of orientaed gradients [17].
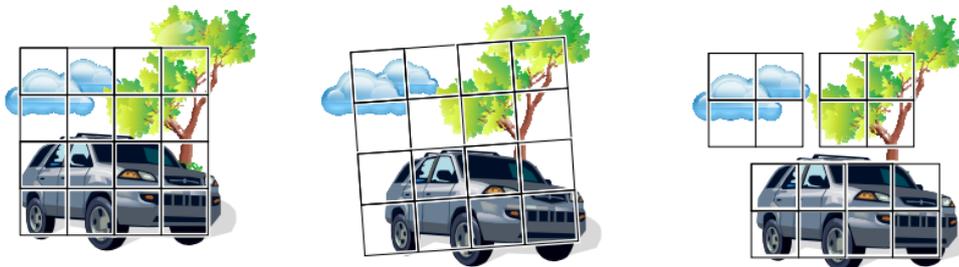

Figure 3.10: Illustration of moving quadrant based matching

Consider an image such as the one shown in 3.10. The template (the object to be matched, usually defined

by a bounding rectangle) is divided into 4×4 cells. In each cell, a weighted histogram of gradients is created for 8 different orientations. The weighing of these histograms is based on gradient magnitude at each location. This gives rise to a descriptor $R \in \mathbb{R}^{4 \times 4 \times 8}$. This can be seen as concatenation of descriptors from 4 quadrants of the template, each with size 2×2 cells. Thus, $R = [R_0, R_1, R_2, R_3]$.

Each quadrant R_i can be thought to have a center at p_i. Matching an object while also accounting for relative motion between objects is effectively allowing for relative motions between various p_i. Similarity between a template R and R' is given by equation 3.21. This quad-tree like sub-division of the templates (also known as patches), is done hierarchically until a cell size reaches 4×4 pixels.

$$sim(R, R') = \frac{1}{4} \sum_{i=0}^{3} \max_{p_i} sim(R_i, R'_i) \tag{3.21}$$

To determine matching regions, we need to compute matching scores for all probable locations. Since the matching of the overall template is defined recursively, it is important to create these matching scores for the smallest patches first and then work our way up. Figure 3.11 diagrammatically illustrates the correlation map creation process. As the correlation map creation is effectively bottom-up, aggregation method is needed to use the correlation maps created for smallest patches and combine them to find correlation maps for level-2, level-3 and so on. Figure 3.12 illustrates this aggregation process. The reader is advised to refer to [86] for a detailed explanation of the correlation map creation and aggregation algorithms.
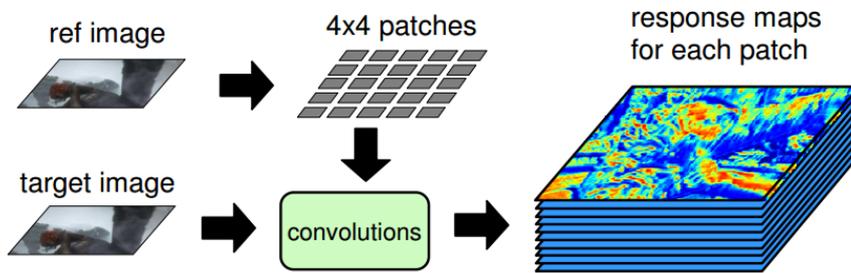
Figure 3.11: Correlation Map in Deep Matching
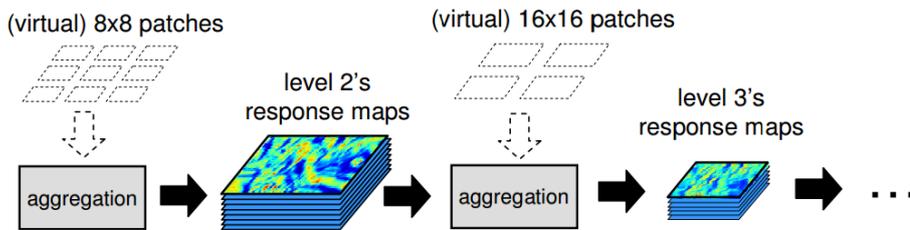
Figure 3.12: Aggregation to create Correlation Map in Bottom-Up fashion

3.2.5. Using Deep Flow and edge preserving Mutual Transform for Registration

When we do a mutual transform using I_M and I_F, we create two sets of pseudo-similar modality images. I_{MT} & I_F belong to the same modality and so does I_M & I_{FT}. This gives rise to two spaces where the deep flow algorithm can be applied. Let the space with I_{MT} & I_F be termed as $\Omega^{(MTF)}$ and the space with I_M & I_{FT} be termed as $\Omega^{(FTM)}$. The problem of finding dense correspondence between two different modality images can be written as equation 3.22. This objective function is fed to the optimizer for obtaining a deformation vector field which registers the images I_M and I_F.

$$E(\vec{w}) = \int_{\Omega^{(MTF)}} \left(E_D^{(MTF)} + \alpha E_S + \beta E_M^{(MTF)} \right) d\Omega^{(MTF)} + \int_{\Omega^{(FTM)}} \left(E_D^{(FTM)} + \alpha E_S + \beta E_M^{(FTM)} \right) d\Omega^{(FTM)} \tag{3.22}$$

<div align="right">

4

</div>

EXPERIMENTS AND RESULTS

4.1. MATERIALS AND METHODS FOR IMAGING

The subjects chosen for this *ex-vivo* registration of Histology and FF-OCT images were the ones who were clinically screened for breast cancer and ovarian cancer. 15 histology and FF-OCT pairs for both breast cancer and ovarian cancer were taken for evaluating the several registration algorithms. The ages of the subject were between 19 and 74. Histogram of ages of all 30 subjects can be seen in figure 4.1.

Figure 4.1: Histogram of Ages of Patients

After the breast / ovarian tissue was taken out during surgery, the tissue was transported to the pathology department. Here it was examined with the aim of diagnosis, prognosis and treatment of the patient. The
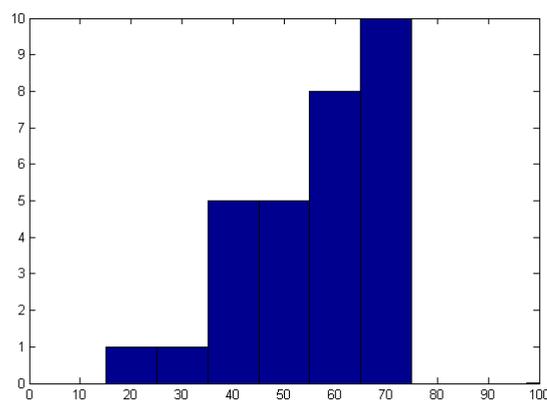
tissue was inked for orientation purposes, and shortly frozen to facilitate slicing of the tissue. After slicing the specimen in $\pm 5mm$ thick slices, and macroscopic analysis by the pathologist, the tissue was placed in formalin overnight. The next day, the slices were again macroscopically analyzed, and small tissue pieces, around $2 \times 3cm$, were selected for further microscopic examination. Several pathological areas, non-pathological areas and suspicious resection margin areas were selected. These pieces were degreased (in alcohol), embedded into candle-wax, sliced very thin ($5\mu m$). The slices were further stained using hematoxylin and eosin stain. These were then digitized using Philips' Digital pathology slide scanner called - IntelliSite Ultra Fast Scanner.

FF-OCT images were collected from either fresh or formalin-fixed tissues. For the purpose of this study, histology Images and FF-OCT images were created using two different slices from the same biopsy sample. This makes establishing the correspondence between FF-OCT and histology possible. The FF-OCT images were created with the LLTech's Light-CT scanner.

The magnification and resolution (in mm) of histology images varied significantly across different images in the dataset. In some of the images, the physical resolution was even unknown. However, the resolution of FF-OCT images were consistent across different images in the dataset. FF-OCT was a square pixel of size $1.4\mu m \times 1.4\mu m$.

Several anatomically similar structures from histology and FF-OCT of the same subject are marked manually to correspond to each other. On an average, 10 manually marked point pairs spanning almost the entire tissue under consideration is marked per image-pair. This serves as a ground truth of comparison of different registration algorithms and for establishing accuracy of the proposed algorithm. The accuracies reported in this work are average distances between these 300 point-pairs (30 image-pairs with 10 corresponding points each) after the registration is complete. Ideally, the corresponding points should fall over each other making the ideal distance as 0.

All the algorithms used in this work were either implemented in MATLAB or called from MATLAB either through a mex interface (for Mutual Transform) or through system call to a pre-compiled executable (for Elastix [91]). The original resolution of the images in pixels were extremely huge with the size exceeding 50000×30000 in several images. Since it is mandatory for the images to be loaded fully in memory before being able to process it in MATLAB, the RAM requirement for processing is also very high (greater than 24 GB RAM). To circumvent this problem and work with regular workstations (6 GB RAM), all the images used for registration were re-sized to $\frac{1}{4}$ of the original size. Hence the physical dimensions of the FF-OCT pixels used for registration were $5.6\mu m \times 5.6\mu m$.

4.2. POINT CLOUD BASED AFFINE REGISTRATION

This section gives detailed results of the first stage in the registration framework where the correspondence between tissues are established and a feature based registration between the two modality images are established using edge-points as features.

4.2.1. STRUCTURAL FOREST BASED EDGE DETECTION

As Structured Forest based Edge Detection (SFED) is a supervised edge detector, it is mandatory to train the classifier using a database which has similar feature statistics as that of the database that is intended as the test set. This ideally means training the detector with histology and FF-OCT images independently. However, creating ground truth for visual edges can be quite tricky and labour intensive. Thus as an alternative, BSDS500 database which contains a variety of natural color images along with their visually prominent edges, was used training the detector [38]. As the training and test data-set are quite different, we compare feature statistics for different features extracted during the classification procedure.

The features used for this classifier mainly consists of normalized gradient values (in multiple scales and orientations), LUV color values and pairwise difference of these features in a small neighbourhood 3.1.1. Figure 4.2 compares the histogram of normalized gradient values in 32×32 image patches selected in training dataset as well as in Histology and FF-OCT images. Since a histogram is computed for each 32×32, a box-plot is used to represent the variation in the feature statistic across different samples. Pairwise difference in 32×32 resized to 5×5 image patch consists of $\binom{5 \times 5}{2}$ values for each patch. Histogram of these values are computed and compared in figure 4.3. From figures 4.2 and 4.3, it can be inferred that although there are minor variations in the feature set, the overall trend in the feature set remains the same.

Comparison of LUV features in BSDS500 dataset and histology image is carried out in figure 4.4. As the training dataset is much more vast as compared to the testing dataset, the color features correspond to only a subset of the training dataset. Since the classifier has been trained using a dataset with lot more variations, it is quite safe to use the trained classifier for histology images. However, using SFED detector for FF-OCT image is quite tricky. FF-OCT being a gray-scale image, it consists of only L values in the LUV feature space. However it was observed (qualitatively) that SFED detector provides almost identical results for a color image and its gray scale counter-part. Hence we use a SFED detector which was trained with BSDS500 dataset.

Figures 4.5 and 4.6 shows the result of using SFED Edge Detector on a histology image and FF-OCT image respectively.
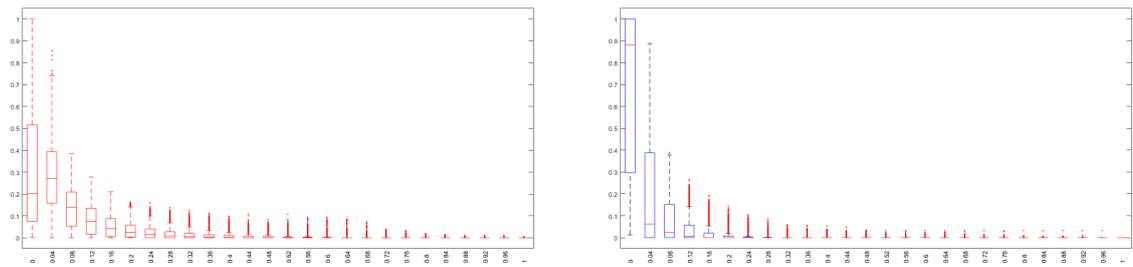
(a) Training Dataset

(b) Testing Dataset - Histology Images
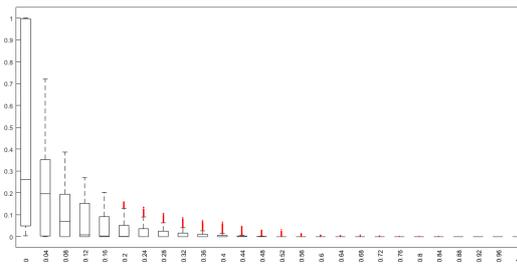
(c) Testing Dataset - FF-OCT Images

Figure 4.2: Comparison of Normalized Gradient Values between training and testing for SFED

4.2.2. Color Image Segmentation

Mean Shift Clustering has an advantage over k-means clustering and its variants in that it works well even in cases where the feature space of different clusters form non-elliptical shapes. To justify the need to use a computationally expensive clustering technique of mean shift clustering, we plotted LUV feature values for foreground and background in histology images. Although the actual feature space used for clustering is a $5D$ feature space with 2 spatial dimensions as well, a $3D$ subset of the feature space can give a pretty good idea of the shape of the cluster and it can be easily visualized as well (as compared to its $5D$ counterpart). This feature space subset can be seen in figure 4.7 and it is quite evidently non-elliptical.

Mean-shift clustering for high dimensional feature vector is known to give a lot of fragments. The clustered output of the image seen in figure 4.5 is shown in figure 4.8 (a). This is clearly over-segmented as we can see multiple segmented regions within the intended tissue region. Hence there is a need for merging the clusters based on spatial and color based proximity. The results of segmentation after cluster merging can be seen in figure 4.8 (b).

4.2.3. Iterative Closest Point for Affine Registration

As mentioned in section 3.1.3, orientation correction is needed to get rid of large-scale rotation mis-alignments. This is done based on PCA of significant edge points as detected by SFED detector to find the major eigen vector and rotating the point-cloud such that major eigen vector has a pre-determined orientation. This pre-

(a) Training Dataset

(b) Testing Dataset - Histology Images
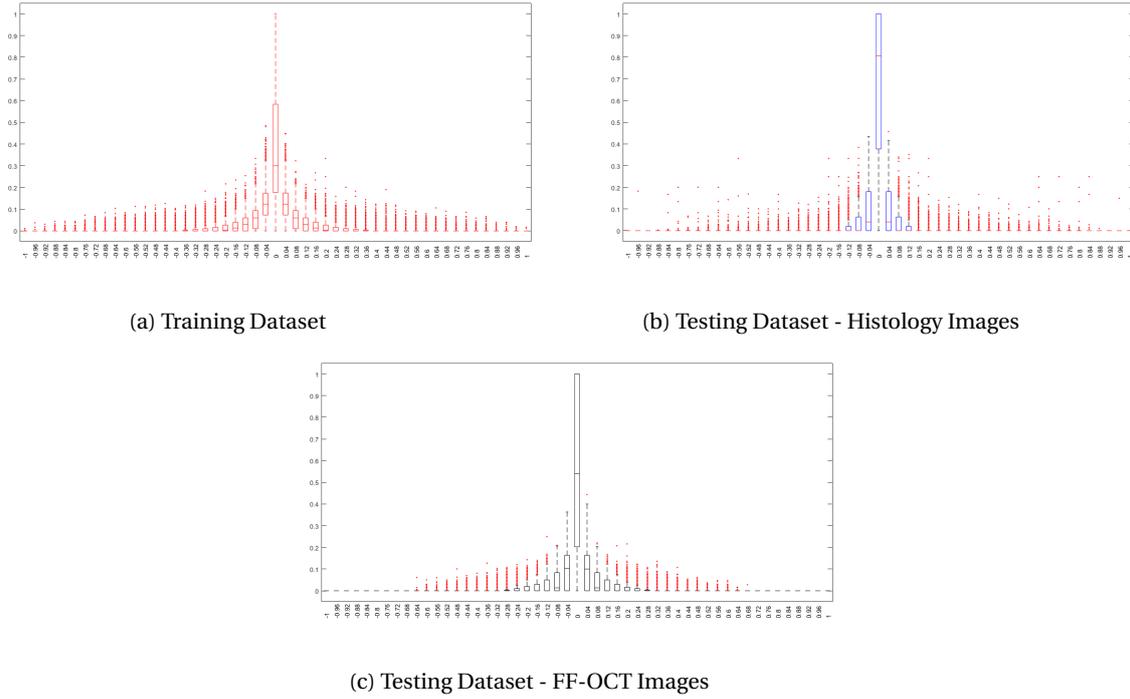

(c) Testing Dataset - FF-OCT Images

Figure 4.3: Comparison of Pairwise Difference of Normalized Gradient Feature Values between training and testing for SFED

determined orientation is taken to be along the x-axis. Figure 4.9 (a) shows a FF-OCT image with significant edge points overlaid on it, and (b) shows the overlaid point cloud after correction. In this illustration, the histology image (c) is already horizontal and correction mechanism for FF-OCT aligns the point cloud with that of histology image.

Point cloud based registration aligns only the significant edge points and not the area within it. Hence the alignment can only be seen as a crude initialization for more accurate registration algorithms. The edge orientation based weightage in the computation of objective function used for ICP algorithm was suggested in section 3.1.3. The result of this minor modification on 15 pairs of FF-OCT and histology images for both ovarian cancer data and breast cancer data can be seen in table 4.1. As it can be seen from the table, it improves the alignment between the point-clouds ever-so-slightly. This improvement is the result of penalizing proximities between edge-points which are not properly aligned for the sake of objective measure computation and rewarding the ones that does align well. In the bigger picture of multi-stage registration, this improvement may not be considered significant enough to affect the end result, but this finding can be quite useful for applications in other domains.

Table 4.1: Point Cloud Registration: Error with and without edge orientation based weight

Error (in pixels)	Standard ICP		Edge Orientation Weighted ICP	
	Breast Cancer Data	Ovarian Cancer Data	Breast Cancer Data	Ovarian Cancer Data
Mean	9.4861	8.6055	9.0337	8.3559
Std. Dev.	2.5551	2.3799	2.5208	2.3261
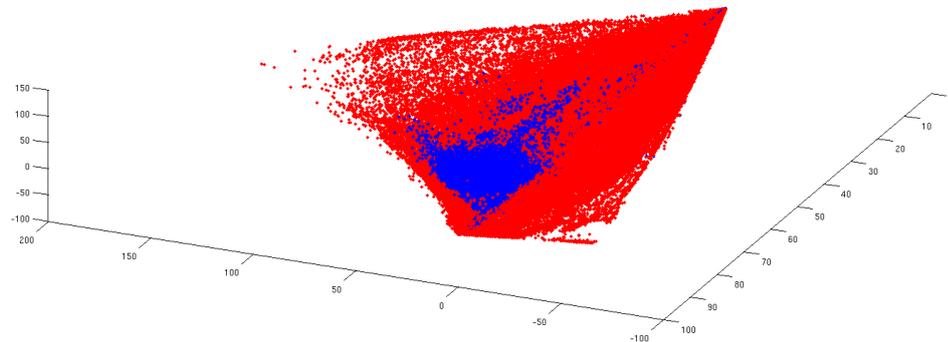
Figure 4.4: LUV features in BSD 500 Training dataset and Histological Images in testing dataset
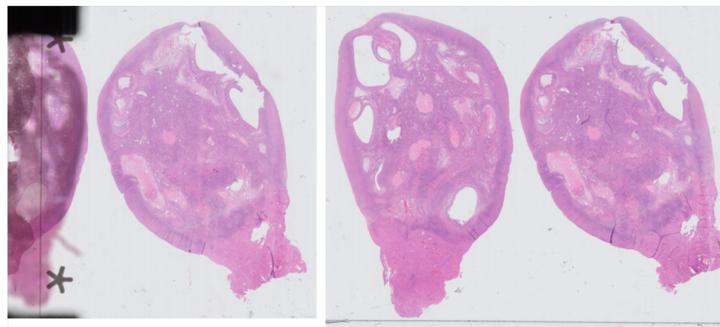
(a) Histology Image

Figure 4.5: Illustration of SFED Edge Detection in Histology Images

In cases where there are multiple tissues in the same histology image, the objective function used for ICP is also used as a deciding metric for establishing correspondence between FF-OCT and one of the several tissues in histology image. Visual results of point-cloud alignment can be seen in figures 4.10 and 4.11.

It can also be seen from figure 4.11 that the FF-OCT point cloud establishes correspondence with the appropriate tissue in the histology image. However this cannot always be taken for granted. Establishing a correspondence between point-clouds becomes even more difficult if only a fraction of the tissue is present in the field of view. However, this is a realistic use case scenario as sometimes only part of the tissue is imaged using FF-OCT. This is similar to occlusion problem in many computer vision applications. Table 4.2 shows the variation of objective function with respect to percentage of tissue captured in FF-OCT in a particular case of registration. The table shows that it is quite possible for some other tissue in the same histology image to be mistakenly selected as the corresponding one, specially when only a part of the FF-OCT image is visible. Hence we intend this selection of corresponding tissue as an automatic initialization steps which needs to confirmed manually with a click of a button in the implementation intended for clinical use. While

(a) FF-OCT Image   (b) SFED Edge Detection of FF-OCT Image

Figure 4.6: Illustration of SFED Edge Detection in FF-OCT Image
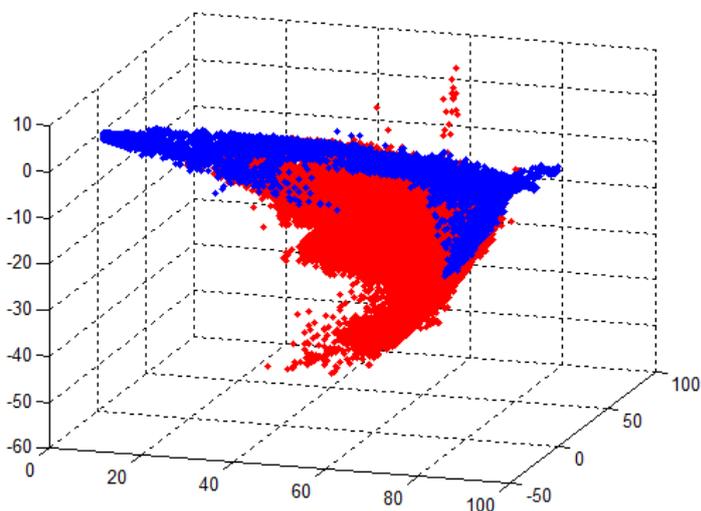
Figure 4.7: LUV feature space: Non-elliptical cluster shapes of foreground and background

the algorithm would select the right tissue in majority of cases, in a few cases where it fails, a simple user click can prevent error propagation in the subsequent steps of registration.

To check the robustness of the point-cloud based registration with respect to 'occlusion', we varied the percentage of FF-OCT tissue in the field of view and studied the translation error of the registered point-cloud. Errors in individual cases varied depending on the how unique the shape of the 'occluded' part was. Figure 4.12 shows the mean trend in error. The errors are measured by treating the case with the complete tissue visibility as ground truth. While this is not completely accurate, for the sake of studying the variation with respect to tissue visibility in the field of view this assumption can be made. It can be seen from the plot that the mean error is well within 10 pixels even when 50% of the tissue is outside the field of view.
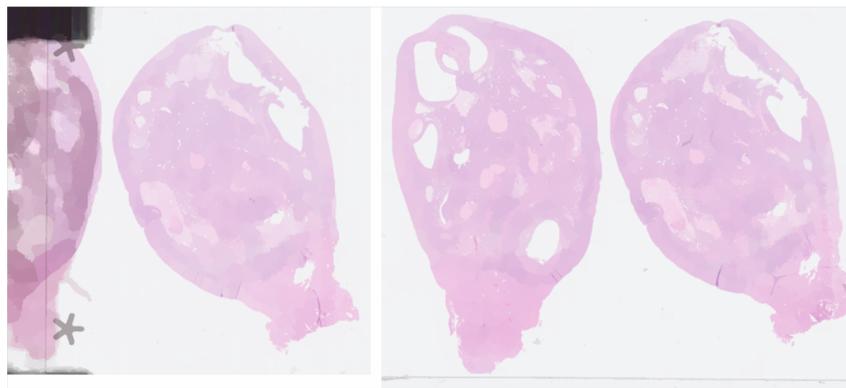
(a) Mean Shift Clustering of Histology Image

(b) Segmented Foreground after Cluster Merging

Figure 4.8: Segmentation of Tissues in Histology Image based on Color Information

Table 4.2: Variation of Final Objective Function while registering FF-OCT point cloud with different histology tissues of the same subject

Percentage of Tissue Captured in FF-OCT	Objective Function with Tissue 1	Objective Function with Tissue 2	Objective Function with Tissue 3	Objective Function with Tissue 4
100	776.4923	370.3727	222.1450	420.2374
80	641.1938	236.2982	177.7160	298.1899
60	565.8954	152.2236	133.2870	184.1424
40	360.5969	78.1491	88.8580	108.0950

4.3. Area Based Non-rigid Registration

Once the correspondence is established and a crude point-cloud based affine registration is performed, area based registration algorithms can be initialized. The advantage of having a crude algorithm initialize a more accurate and computationally intensive algorithm are two-fold. Firstly, it reduces the time taken for the computationally intensive algorithm to converge by initializing it quite close to the final result. Secondly, it also reduces the chances of the algorithm with more degrees of freedom from getting stuck in a local optima. With this in mind, ICP algorithm is used to initialize MI based multi-scale rigid registration algorithm which in turn initializes non-rigid registration algorithm.

Convergence of MI in Multiscale Registration is illustrated in figure 4.13. It can also be noted that the argument about faster convergence with a good initialization can actually be observed in these plots. In (a) -

(a) FF-OCT image with significant edge points overlaid on it

(b) Orientation Normalization using PCA of edge points and Rotation

(c) Corresponding Histology Image with Horizontal Orientation
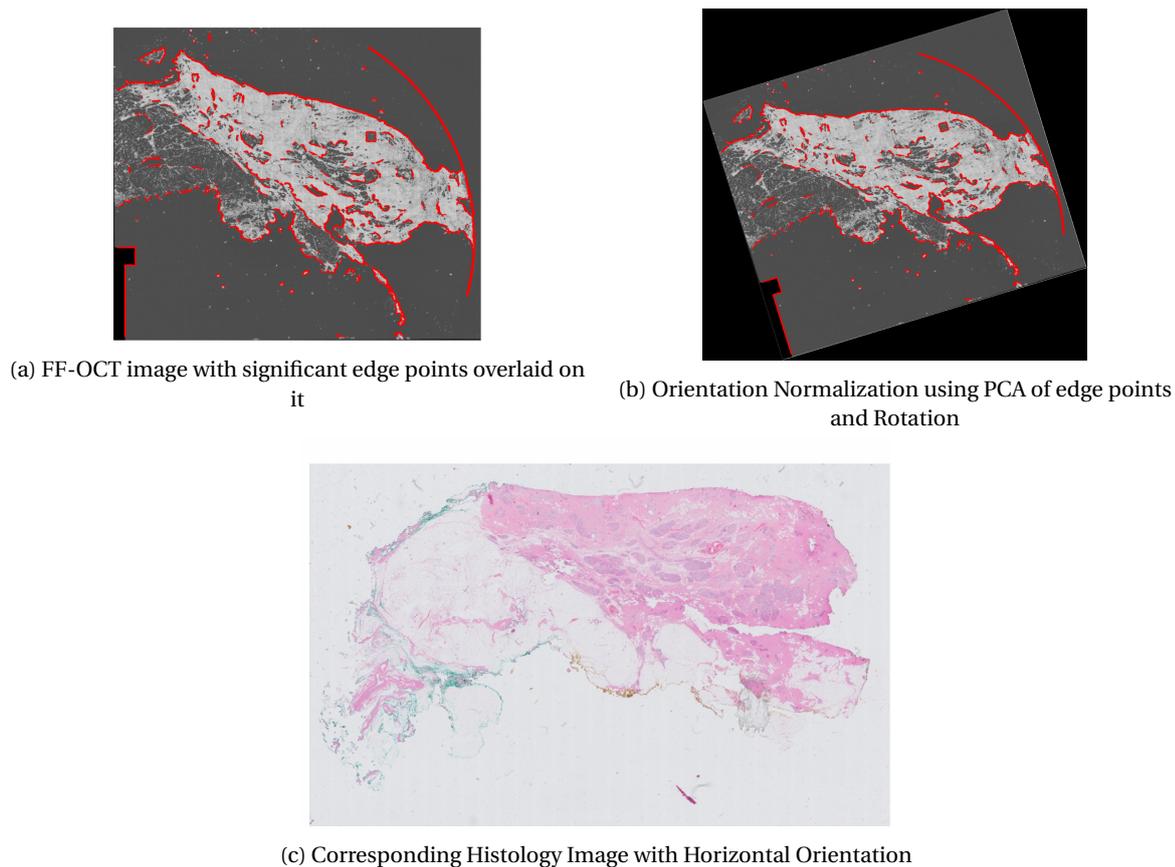
Figure 4.9: Orientation Correction

the lowest scale - where the initialization is from ICP algorithm, it takes quite a lot of iterations to converge to a solution. However, in (c) - the highest scale - where the initialization is from MI based affine registration in an intermediate scale, the convergence is very rapid.

The results of edge preserving mutual transform introduced in section 3.2.3 can be seen in figures 4.14 and 4.15.

The results of Deep Flow based registration technique as discussed in section 3.2.5 is shown in figure 4.16. Figures (a) and (b) shows the Histology and FF-OCT images used for registration where as (c) depicts the Flow Vector using a pseudo coloring that is generally used in optical-flow based literature. The code for pseudo-coloring the vector flow can be obtained from the webpage with Middlebury optical flow dataset[1].

The maximum magnitude of the flow vector is 19.92 pixels which means that between FF-OCT and histology image, there is a movement of as much as approximately 20 pixels in some places. The magnitude of flow vectors at different places is depicted as a gray scale image in figure 4.17 (a). A small region in the image is taken in (b) where a surface plot of the magnitude in the chosen region is shown. While the maximum displacement is seen in lower left region of the image where the tissues are absent and it is part of the background, the selected region in (b) shows as much as up to 12 pixel movement between the two images. This

---

[1]http://vision.middlebury.edu/flow/data/
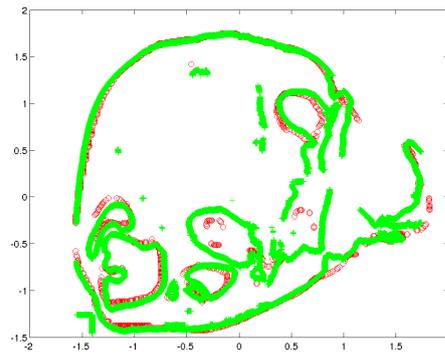
Figure 4.10: Overlay of Point Clouds after Weighted Iterative Closest Point
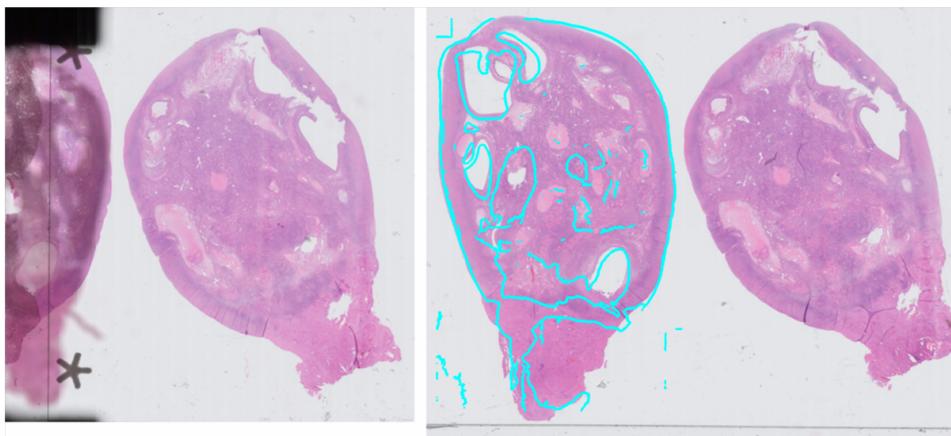
Figure 4.11: Overlay of FF-OCT Point Cloud on Histology Tissue after Weighted Iterative Closest Point

results in large deformation which is often not very well captured in demon based algorithms.

Table 4.3 shows the comparative results where FF-OCT and Histology images were registered using several non-rigid registration algorithms. The reported error is the mean error in pixels based on 15 registration results for both ovarian cancer and breast cancer data. As expected (due to the aforementioned reasons), Deep Flow based techniques quite easily outperform Demon registration based algorithm. Also, the edge preserving mutual transform, which was proposed to increase the gradient based image force in the objective function used for minimization, is seen to have a positive effect on the final result. It can also be observed that Vector flow based algorithms operating on the transformed domain consistently outperform MI and NMI based registration algorithms.

Finally, the result of the proposed framework on an incoming FF-OCT and Histology image is shown in a checker board pattern in figures 4.18 and 4.19. It can be easily observed that similar regions of these images are very well aligned as in the checker board pattern there is hardly any observable spatial discontinuity.
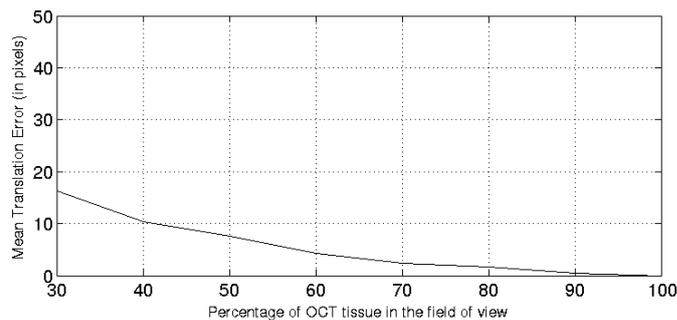
Figure 4.12: Variation of mean translation error

(a) Lowest Scale   (b) Intermediate Scale   (c) Original Scale
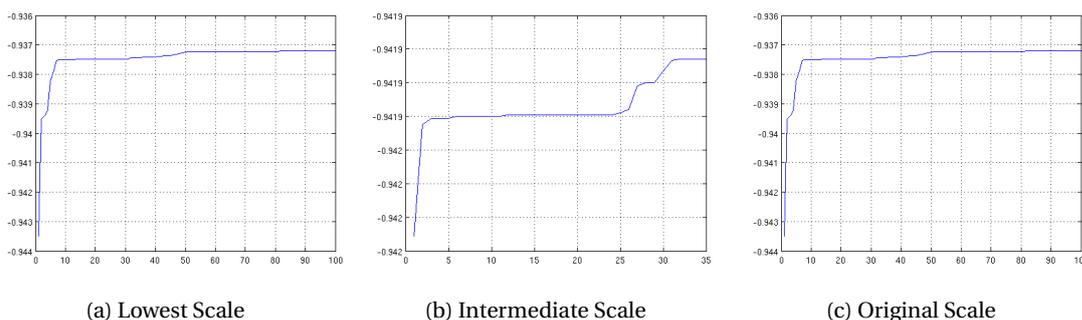
Figure 4.13: Convergence of MI in Multiscale Registration

4.4. IMPLEMENTATION DETAILS

All the algorithms used in this work were either implemented in MATLAB or called from MATLAB either through a mex interface (for Mutual Transform) or through system call to a pre-compiled executable (for Elastix [91]). Several open source codes were used for different parts of the work. P. Dollár's MATLAB toolbox for Structured Forest based Edge Detection[2] was used for edge detection. Although Mean Shift Clustering was implemented in MATLAB, it took over an hour for clustering the histology images in $5D$ feature space. Edge Detection and Image SegmentatiON (EDISON)[3] system's C++ implementation of mean shift clustering was used for speeding up the registration process. Edge Orientation Alignment based Weighted Iterative Closest Point was implemented wholly in MATLAB and the optimizer used for this purpose was MATLAB's inbuilt Levenberg Marquardt algorithm.

Multi-scale affine registration based on MI was implemented using Elastix. The parameter file for this registration is given in appendix A. The inspiration for using modality transformation based on Mutual Transform was obtained based on trying out a open source MATLAB code[4]. The suggested changes in modality transformation were implemented in C++ and a mex wrapper for mutual transform was made to speed up the computation. Open Source C++ implementation of Deep Flow[5] was used for non-rigid registration sug-

---

[2]http://research.microsoft.com/en-us/downloads/389109f6-b4e8-404c-84bf-239f7cbf4e3d/
[3]http://coewww.rutgers.edu/riul/research/code/EDISON/index.html
[4]http://nl.mathworks.com/matlabcentral/fileexchange/21451-multimodality-non-rigid-demon-algorithm-image-registration
[5]http://lear.inrialpes.fr/src/deepflow/
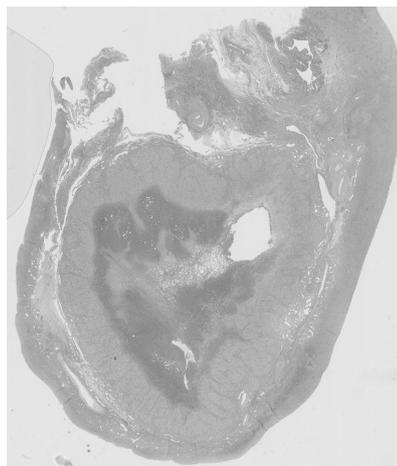
(a) Histology Image   (b) Psuedo Histology Image Created from FF-OCT Image

Figure 4.14: Edge Preserving Mutual Transform: Histology Image Pair
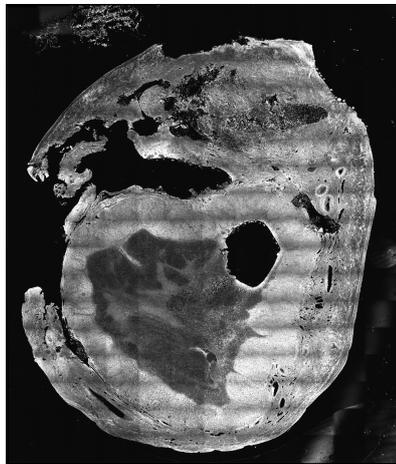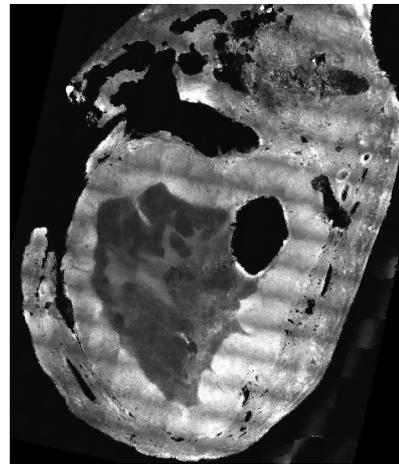
(a) FF-OCT Image   (b) Psuedo FF-OCT Image Created from Histology Image

Figure 4.15: Edge Preserving Mutual Transform: FF-OCT Image Pair

gested in section 3.2.5. However the objective function used to optimize in the implementation was tweaked to suit the purpose of inter-modaility registration based on mutual transform as given in equation 3.22.

Elastix was also used for generating comparative results in table 4.3 for MI and NMI based non-rigid registration with B-spline kernel based deformation field. The parameter file used for Elastix calls for these registrations are given in appendix B (MI) and C (NMI).

(a) Histology Image　　　(b) FF-OCT Image　　　(c) Pseudo Coloring of Flow Vector
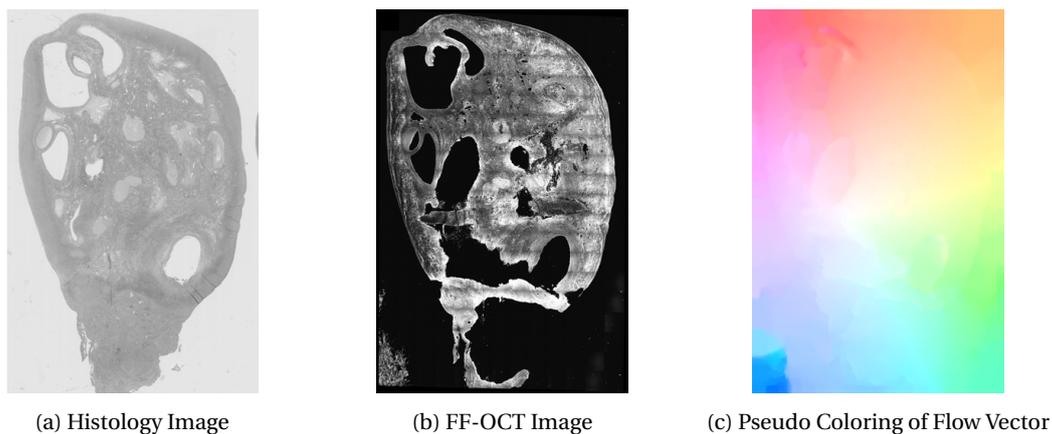
Figure 4.16: Stationary Vector Flow between Histology and FF-OCT image

(a) Magnitude of Flow Vectors [0 - 19.92] mapped to gray scale values of [0 - 255]

(b) Surface Plot of a Smaller Region
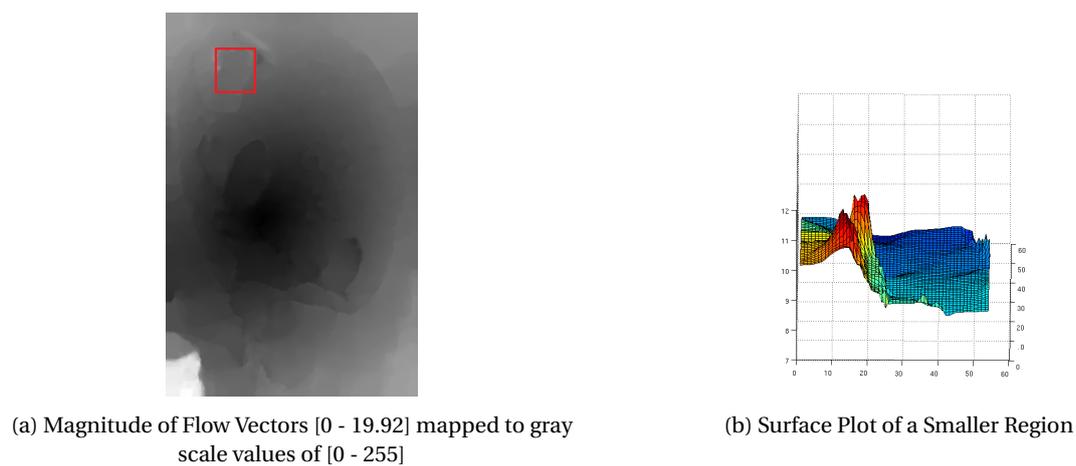
Figure 4.17: Magnitude of Flow Vector

Table 4.3: Non-Rigid Registration: Error (in pixels) using different registration techniques

Method	Ovarian Cancer Data	Breast Cancer Data
Deep Flow + Edge Preserving Mutual Transformation	3.4277	4.6215
Deep Flow + Mutual Transformation	3.7922	4.8507
Demon + Edge Preserving Mutual Transformation	4.8838	5.7007
Normalized Mutual Information	5.2402	6.0479
Mutual Information	5.4193	6.2345
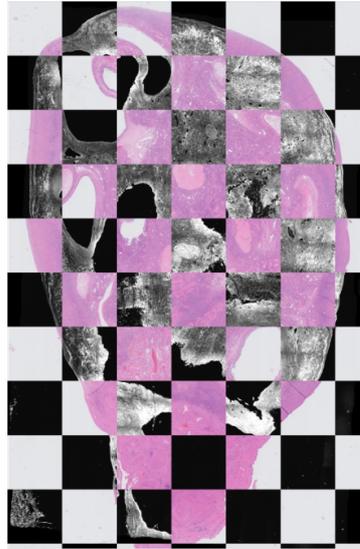
Figure 4.18: Checker Board Pattern of the Non-Rigid Registration Result in one of the images in Ovarian Cancer Data
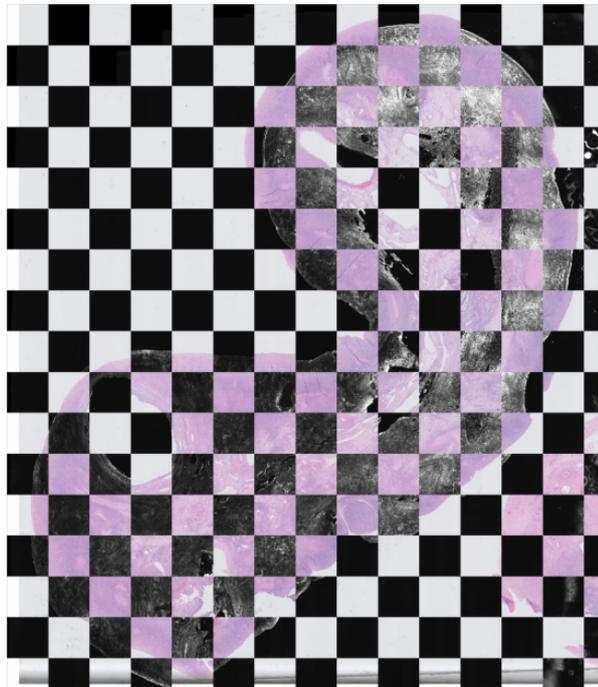
Figure 4.19: Checker Board Pattern of the Non-Rigid Registration Result in one of the images in Breast Cancer Data

5

CONCLUSIONS AND FUTURE WORK

5.1. DISCUSSION

A multi-stage framework for registration between histology and *ex-vivo* FF-OCT has been presented here. The biopsy section used for creating FF-OCT and histology images, can be seen as made of several thin slices. The scene that can be observed in each slice changes continuously. Any two slices, however minute the separation between them, are different from each other in some sense. This results in deformation between the two corresponding structures. Thus, the deformation we observe here is the spatial change in the scene under observation. Optical flow algorithms deal with motion in consecutive frames of a video. The deformation we see in such scenarios is because of temporal change in the scene under observation. The realization of this similarity between the registration problem in hand and optical flow algorithms was instrumental in arriving at a robust framework for registration.

Optical flow based algorithms are generally not used for inter-modality registration problems. This is becuse the brightness constancy assumption (in optical flow algorithms) between the two images in different modalities does not hold true. This problem was solved by the introduction of modality transformation in [14]. The problem with this mutual transform is that it does not preserve the edges well and hence it affects the gradient based image force used in the objective function of many registration algorithms such as Demon algorithm. The proposed modification to the modality transformation addresses this and this results in better accuracy of registration.

Demon registration algorithm however does not quite work that well for large deformations. Since the slices of tissues used for histology and FF-OCT can be physically well separated (several μm in the axial direction), large deformations should be anticipated and an appropriate algorithm should be used. Deep flow based optical flow algorithm is known to address this problem. The algorithm copes with large deformation by introducing a matching based term in the objective function used for minimization. Matching based term is nothing but a coarse deformation field obtained based on texture feature-based dense matching. The regularizing term in objective function increases the smoothness of the deformation field where as the matching based term increases the similarity of the final deformation field to that of matching based deformation field (which can be non-smooth). This addition results in much better estimation of deformation field even when the deformations are large. This can be conclusively seen in the registration accuracies.

Presence of significant edges in the images is an important information we use across different algorithms in the multi-stage registration framework. The absence of such edges causes the accuracy to reduce significantly. This is one of the reasons why the results in ovarian cancer data is consistently better than the ones in breast cancer data. Presence of the tissue edges in FF-OCT is almost a mandatory feature without which the first stage (ICP based affine registration) would completely fail. Even while experimenting with varying amounts of tissue captured in FF-OCT, at least 50% of the tissue edges were visible. Instead of taking a part of the tissue from one of sides if only the center portion of the tissue is captured in FF-OCT, the algorithm fails to align properly in the first stage of the algorithm and subsequently the entire registration fails to achieve its final goal.

The reason for Mutual Transformation followed by Vector Flow based algorithms to outperform MI and NMI based non-rigid registration algorithm is slightly puzzling. Unlike other inter-modality registration applications such as CT-MRI registration, where a one-to-one correspondence between each point in both the modalities exists (as it is from the same patient), FF-OCT-Histology registration is between two different slices prepared from the same biopsy sample which are spatially in slightly different places in the sample. Corresponding structures may not exist for some regions. Mutual Transformation of a image with a blank image transforms the blank image to be identical to the input image. Hence in regions where the corresponding structures are absent and we instead find blank regions of the background, mutual transformation duplicates the image under consideration and the net deformation in the region is nil. In other regions where the corresponding structures exists, it finds the deformation well. This is not the case in MI based registration algorithms. Even in places where the corresponding structures do not exist, MI based algorithm computes the metric and uses it in finding an overall deformation field. This might be one of the reasons which affects the overall performance of the algorithm.

5.2. Conclusion and Future Work

The automated framework can significantly reduce the pathologist's involvement and effort in finding correspondence. Point-cloud based affine registration is invariant to tissue orientation and is also quite robust with respect to the amount of tissue present in the field of view. Modification to modality transformation that has been proposed here. Instead of giving equal weights to each joint-occurrence of gray levels, a confidence measure based on edge alignment has been introduced here. The modified transformation has more prominent edges in places where the edges in both modalities align well and this has been shown to lead to a faster convergence of demon registration algorithm. This is seen to have a positive effect on gradient based image force used in vector flow based registration techniques such as Demon Registration and Deep Flow algorithm. This is also the first reported use of Deep Flow based optical flow technique in inter-modality registration.

The fields of point-cloud based registration, color image segmentation, non-rigid registration using vector flow based techniques are all evolving continuously. The proposed framework of multi-stage registration between FF-OCT and histology image recognizes these ever-evolving fields and is hence divided into several distinct modules each of which can be replaced by an improved algorithm as and when they are proposed.

Extension of the work presented here can be done in many directions. From algorithm point of view, modality transformation for inter-modality registration is not as widely used as MI based techniques. The use of edge-preserving transformation in other inter-modality registration techniques can be investigated. Deep Flow and other such optical flow algorithms designed to compute large-deformations can be investigated more rigorously for several registration applications.

From an application point of view, *in-situ* FF-OCT is still a distant dream. This is because the current machine is quite bulky and the imaging technique cannot deal with motion of the subject yet. Research in these domains can make intra-operative pathology a very realistic option in the future. Increasing the dataset to several hundreds.

Also, from an implementation point of view, the registration algorithms used here require several GBs of RAM even while working with sub-sampled images of $\frac{1}{4}$ resolution. This can be significantly improved upon by using streaming feature in ITK library wherein it would not be required for the entire image to be loaded in memory for the algorithms to commence. Different parts of the image can be loaded as and when they are required for computation. This would help in reducing the RAM requirement for processing the images and can help in working with original resolution images. This in-turn might have a direct implication on the accuracy of the presented methods (it might improve). GPU implementation of the presented algorithms can be done to speed-up the registration process before wide-spread use of such registration algorithms.

BIBLIOGRAPHY

[1] M. Sofka, G. Yang, and C. V. Stewart, *Simultaneous covariance driven correspondence (CDC) and transformation estimation in the expectation maximization,* in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (Minneapolis, MN, USA, 2007).

[2] A. Criminisi, J. Shotton, and E. Konukoglu, *Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning,* Foundations and Trends in Computer Graphics and Vision: Vol. 7: No 2-3, pp 81-227 (2012).

[3] C. CULLING, *Chapter 2 - methods of examination of tissues and cells,* in *Handbook of Histopathological and Histochemical Techniques (Third Edition),* edited by C. CULLING (Butterworth-Heinemann, 1974) third edition ed., pp. 19 – 25.

[4] C. CULLING, *Chapter 6 - section cutting,* in *Handbook of Histopathological and Histochemical Techniques (Third Edition),* edited by C. CULLING (Butterworth-Heinemann, 1974) third edition ed., pp. 111 – 147.

[5] C. CULLING, *Chapter 11 - routine stains,* in *Handbook of Histopathological and Histochemical Techniques (Third Edition),* edited by C. CULLING (Butterworth-Heinemann, 1974) third edition ed., pp. 211 – 220.

[6] W. Jung and S. Boppart, *Optical coherence tomography for rapid tissue screening and directed histological sectioning,* Analytical Cellular Pathology **35**, 129 (2012).

[7] A. Douplik, *Intraoperative optical coherence tomographic monitoring,* in *Handbook of Biophotonics* (Wiley-VCH Verlag GmbH and Co., 2013) pp. 377–400.

[8] K. Wiesauer, M. Pircher, E. Götzinger, S. Bauer, R. Engelke, G. Ahrens, G. Grützner, C. K. Hitzenberger, and D. Stifter, *En-face scanning optical coherence tomography with ultra-high resolution for material investigation,* Opt. Express **13**, 1015 (2005).

[9] A. Dubois, K. Grieve, G. Moneron, R. Lecaque, L. Vabre, and C. Boccara, *Ultrahigh-resolution full-field optical coherence tomography,* Appl. Opt. **43**, 2874 (2004).

[10] L. Vabre, A. Dubois, and A. C. Boccara, *Thermal-light full-field optical coherence tomography,* Opt. Lett. **27**, 530 (2002).

[11] O. Assayag, M. Antoine, B. Sigal-Zafrani, M. Riben, F. Harms, A. Burcheri, K. Grieve, E. Dalimier, B. Le Conte de Poly, and C. Boccara, *Large field, high resolution full-field optical coherence tomography: A pre-clinical study of human breast tissue and cancer assessment,* Technology in Cancer Research & Treatment **13**, 455 (2014).

[12] A. Latrive and A. C. Boccara, *In vivo and in situ cellular imaging full-field optical coherence tomography with a rigid endoscopic probe,* Biomed. Opt. Express **2**, 2897 (2011).

[13] P. Besl and N. D. McKay, *A method for registration of 3-d shapes,* Pattern Analysis and Machine Intelligence, IEEE Transactions on **14**, 239 (1992).

[14] D.-J. Kroon and C. H. Slump, *Mri modalitiy transformation in demon registration,* in *Biomedical Imaging: From Nano to Macro, 2009. ISBI '09. IEEE International Symposium on* (2009) pp. 963–966.

[15] B. Zitova and J. Flusser, *Image registration methods: a survey,* Image and Vision Computing **21**, 977 (2003).

[16] B. Sturm, K. Powell, A. Stillman, and R. White, *Registration of 3d ct angiography and cardiac mr images in coronary artery disease patients,* The International Journal of Cardiovascular Imaging **19**, 281 (2003).

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

[18] G. jian Wen, J.-j. Lv, and W. xian Yu, *A high-performance feature-matching method for image registration by combining spatial and similarity information,* Geoscience and Remote Sensing, IEEE Transactions on **46**, 1266 (2008).

[19] J. Pluim, J. Maintz, and M. Viergever, *Mutual-information-based registration of medical images: a survey,* Medical Imaging, IEEE Transactions on **22**, 986 (2003).

[20] P. Viola and W. Wells, *Alignment by maximization of mutual information,* in *Computer Vision, 1995. Proceedings., Fifth International Conference on* (1995) pp. 16–23.

[21] C. Shannon, *A mathematical theory of communication,* Bell System Technical Journal **27**, 379 (1948).

[22] C. E. Rodriguez-Carranza and M. H. Loew, *Weighted and deterministic entropy measure for image registration using mutual information,* in *Proc. SPIE Medical Imaging,* Vol. 3338 (1998) pp. 155–166.

[23] C. Studholme, D. Hill, and D. Hawkes, *An overlap invariant entropy measure of 3d medical image alignment,* Pattern Recognition **32**, 71 (1999).

[24] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, *Multimodality image registration by maximization of mutual information,* Medical Imaging, IEEE Transactions on **16**, 187 (1997).

[25] J. Meyer, *Histogram transformation for inter-modality image registration,* in *Bioinformatics and Bioengineering, 2007. BIBE 2007. Proceedings of the 7th IEEE International Conference on* (2007) pp. 1118–1123.

[26] J. Nocedal and S. J. Wright, *Numerical Optimization,* 2nd ed. (Springer, New York, 2006).

[27] J. J. E. Dennis and J. J. Moré, *Quasi-newton methods, motivation and theory,* SIAM Review **19**, 46 (1977).

[28] J. Moré, *The levenberg-marquardt algorithm: Implementation and theory,* in *Numerical Analysis,* Lecture Notes in Mathematics, Vol. 630, edited by G. Watson (Springer Berlin Heidelberg, 1978) pp. 105–116.

[29] H. J. Kushner and G. Yin, *Stochastic approximation and recursive algorithms and applications,* Applications of mathematics (Springer, New York, 2003).

[30] S. Klein, J. Pluim, M. Staring, and M. Viergever, *Adaptive stochastic gradient descent optimisation for image registration,* International Journal of Computer Vision **81**, 227 (2009).

[31] N. Hansen and A. Ostermeier, *Completely derandomized self-adaptation in evolution strategies,* Evol. Comput. **9**, 159 (2001).

[32] S. Klein, M. Staring, and J. Pluim, *Evaluation of optimization methods for nonrigid medical image registration using mutual information and b-splines,* Image Processing, IEEE Transactions on **16**, 2879 (2007).

[33] L. Roberts, *Machine perception of three-dimensional solids,* Ph.D. thesis, Massachusetts, United States of America (1963).

[34] R. O. Duda and P. E. Hart, *Pattern Classification and scene analysis* (John Wiley & Sons, Inc., 1973).

[35] D. Ziou and A. Koukam, *The selection of edge detectors using local image structure,* in *Tools with Artificial Intelligence, 1995. Proceedings., Seventh International Conference on* (1995) pp. 366–370.

[36] D. Ziou and S. Tabbone, *Edge detection techniques - an overview,* INTERNATIONAL JOURNAL OF PATTERN RECOGNITION AND IMAGE ANALYSIS **8**, 537 (1998).

[37] D. Martin, C. Fowlkes, and J. Malik, *Learning to detect natural image boundaries using local brightness, color, and texture cues,* Pattern Analysis and Machine Intelligence, IEEE Transactions on **26**, 530 (2004).

[38] D. Martin, C. Fowlkes, D. Tal, and J. Malik, *A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,* in *Proc. 8th Int'l Conf. Computer Vision*, Vol. 2 (2001) pp. 416–423.

[39] P. Dollar, Z. Tu, and S. Belongie, *Supervised learning of edges and object boundaries,* in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, Vol. 2 (2006) pp. 1964–1971.

[40] S. Zheng, Z. Tu, and A. Yuille, *Detecting object boundaries using low-, mid-, and high-level information,* in *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on* (2007) pp. 1–8.

[41] J. Mairal, M. Leordeanu, F. Bach, M. Hebert, and J. Ponce, *Discriminative sparse image models for class-specific edge detection and image interpretation,* in *Computer Vision – ECCV 2008*, Lecture Notes in Computer Science, Vol. 5304, edited by D. Forsyth, P. Torr, and A. Zisserman (Springer Berlin Heidelberg, 2008) pp. 43–56.

[42] R. Xiaofeng and L. Bo, *Discriminatively trained sparse code gradients for contour detection,* in *Advances in Neural Information Processing Systems 25*, edited by F. Pereira, C. Burges, L. Bottou, and K. Weinberger (Curran Associates, Inc., 2012) pp. 584–592.

[43] J. Lim, C. L. Zitnick, and P. Dollár, *Sketch tokens: A learned mid-level representation for contour and object detection,* in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on* (2013) pp. 3158–3165.

[44] P. Dollár and C. L. Zitnick, *Structured forests for fast edge detection,* in *Proceedings of the 2013 IEEE International Conference on Computer Vision*, ICCV '13 (IEEE Computer Society, Washington, DC, USA, 2013) pp. 1841–1848.

[45] S. Hallman and C. Fowlkes, *Oriented edge forests for boundary detection,* in *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on* (2015) pp. 1732–1740.

[46] J. J. Kivinen, C. K. I. Williams, and N. Heess, *Visual boundary prediction: A deep neural prediction network and quality dissection,* in *Proceedings of the Seventeenth International Conference on Artificial Intelligence and Statistics, AISTATS 2014, Reykjavik, Iceland, April 22-25, 2014* (2014) pp. 512–521.

[47] G. Bertasius, J. Shi, and L. Torresani, *Deepedge: A multi-scale bifurcated deep network for top-down contour detection,* in *Computer Vision and Pattern Recognition (CVPR), 2015 IEEE Conference on* (2015) pp. 4380–4389.

[48] N. Otsu, *A threshold selection method from gray-level histograms,* Systems, Man and Cybernetics, IEEE Transactions on **9**, 62 (1979).

[49] A. K. Jain and R. C. Dubes, *Algorithms for Clustering Data* (Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1988).

[50] R. Adams and L. Bischof, *Seeded region growing,* Pattern Analysis and Machine Intelligence, IEEE Transactions on **16**, 641 (1994).

[51] X. Wu, *Adaptive split-and-merge segmentation based on piecewise least-square approximation,* Pattern Analysis and Machine Intelligence, IEEE Transactions on **15**, 808 (1993).

[52] Z. Kato and T.-C. Pong, *A markov random field image segmentation model for color textured images,* Image and Vision Computing **24**, 1103 (2006).

[53] J. Shi and J. Malik, *Normalized cuts and image segmentation,* IEEE Trans. Pattern Anal. Mach. Intell. **22**, 888 (2000).

[54] H. Cheng and J. Li, *Fuzzy homogeneity and scale-space approach to color image segmentation,* Pattern Recognition **36**, 1545 (2003).

[55] Z. Li, X.-M. Wu, and S.-F. Chang, *Segmentation using superpixels: A bipartite graph partitioning approach,* in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on* (2012) pp. 789–796.

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

[57] S. Rusinkiewicz and M. Levoy, *Efficient variants of the ICP algorithm,* in *Third International Conference on 3D Digital Imaging and Modeling (3DIM)* (2001).

[58] G. Turk and M. Levoy, *Zippered polygon meshes from range images,* in *Proceedings of the 21st Annual Conference on Computer Graphics and Interactive Techniques,* SIGGRAPH '94 (ACM, New York, NY, USA, 1994) pp. 311–318.

[59] T. Masuda, K. Sakaue, and N. Yokoya, *Registration and integration of multiple range images for 3-d model construction,* in *Pattern Recognition, 1996., Proceedings of the 13th International Conference on,* Vol. 1 (1996) pp. 879–883 vol.1.

[60] Y. Chen and G. Medioni, *Object modeling by registration of multiple range images,* in *Robotics and Automation, 1991. Proceedings., 1991 IEEE International Conference on* (1991) pp. 2724–2729 vol.3.

[61] D. A. Simon, *Fast and Accurate Shape-based Registration,* Ph.D. thesis, Pittsburgh, PA, USA (1996).

[62] K. Pulli and L. G. Shapiro, *Surface reconstruction and display from range and color data,* Graphical Models **62**, 165 (2000).

[63] G. Godin, M. Rioux, and R. Baribeau, *Three-dimensional registration using range and intensity information,* in *Proc. SPIE 2350, Videometrics III*, Vol. 2350 (1994) pp. 279–290.

[64] C. R. Meyer, J. L. Boes, B. Kim, P. H. Bland, K. R. Zasadny, P. V. Kison, K. Koral, K. A. Frey, and R. L. Wahl, *Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations.* Medical Image Analysis **1**, 195 (1997).

[65] M. A. Horsfield, *Demonstration of accuracy and clinical versatility of mutual information for automatic multimodality image fusion using affine and thin-plate spline warped geometric deformations.* Magnetic Resonance Imaging **17**, 1335 (1999).

[66] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, and D. Hawkes, *Nonrigid registration using free-form deformations: application to breast mr images,* Medical Imaging, IEEE Transactions on **18**, 712 (1999).

[67] C. Studholme, R. Constable, and J. Duncan, *Accurate alignment of functional epi data to anatomical mri using a physics-based distortion model,* Medical Imaging, IEEE Transactions on **19**, 1115 (2000).

[68] M. Fornefett, K. Rohr, and H. Stiehl, *Elastic registration of medical images using radial basis functions with compact support,* in *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, Vol. 1 (1999) p. 407 Vol. 1.

[69] M. S. Baker, H. Busse, and M. Vogt, *Automatic registration and segmentation algorithm for multiple electrophoresis images,* in *Proc. SPIE Medical Imaging*, Vol. 3979 (2000) pp. 426–436.

[70] T. Gaens, F. Maes, D. Vandermeulen, and P. Suetens, *Non-rigid multimodal image registration using mutual information,* in *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98*, Lecture Notes in Computer Science, Vol. 1496, edited by W. M. Wells, A. Colchester, and S. Delp (Springer Berlin Heidelberg, 1998) pp. 1099–1106.

[71] A. Trouvé, *Diffeomorphisms groups and pattern matching in image analysis,* International Journal of Comput. Vision **28**, 213 (1998).

[72] C. Chefd'hotel, G. Hermosillo, and O. Faugeras, *Flows of diffeomorphisms for multimodal image registration,* in *Biomedical Imaging, 2002. Proceedings. 2002 IEEE International Symposium on* (2002) pp. 753–756.

[73] D. Zikic, M. Baust, A. Kamen, and N. Navab, *Generalization of deformable registration in riemannian sobolev spaces,* in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010*, Lecture Notes in Computer Science, Vol. 6362, edited by T. Jiang, N. Navab, J. Pluim, and M. Viergever (Springer Berlin Heidelberg, 2010) pp. 586–593.

[74] J.-P. Thirion, *Image matching as a diffusion process: an analogy with maxwell's demons,* Medical Image Analysis **2**, 243 (1998).

[75] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, *Non-parametric diffeomorphic image registration with the demons algorithm,* in *Proceedings of the 10th International Conference on Medical Image Computing and Computer-assisted Intervention*, MICCAI'07 (Springer-Verlag, Berlin, Heidelberg, 2007) pp. 319–326.

[76] T. Vercauteren, X. Pennec, A. Perchant, and N. Ayache, *Symmetric log-domain diffeomorphic registration: A demons-based approach,* in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008*, Lecture Notes in Computer Science, Vol. 5241, edited by D. Metaxas, L. Axel, G. Fichtinger, and G. Székely (Springer Berlin Heidelberg, 2008) pp. 754–761.

[77] D. Yang, H. Li, D. A. Low, J. O. Deasy, and I. E. Naqa, *A fast inverse consistent deformable image registration method based on symmetric optical flow computation,* Physics in Medicine and Biology **53**, 6143 (2008).

[78] G. Hermosillo and O. Faugeras, *Dense image matching with global and local statistical criteria: a variational approach,* in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, Vol. 1 (2001) pp. I–73–I–78 vol.1.

[79] P. Hellier and C. Barillot, *Multimodal non-rigid warping for correction of distortions in functional mri,* in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2000*, Lecture Notes in Computer Science, Vol. 1935, edited by S. Delp, A. DiGoia, and B. Jaramaz (Springer Berlin Heidelberg, 2000) pp. 512–520.

[80] I. Jolliffe, *Principal Component Analysis* (Springer Verlag, 1986).

[81] K. Fukunaga and L. Hostetler, *The estimation of the gradient of a density function, with applications in pattern recognition,* Information Theory, IEEE Transactions on **21**, 32 (1975).

[82] M. Tabb and N. Ahuja, *Multiscale image segmentation by integrated edge and region detection,* Image Processing, IEEE Transactions on **6**, 642 (1997).

[83] H. Wang, L. Dong, J. O'Daniel, R. Mohan, A. S. Garden, K. K. Ang, D. A. Kuban, M. Bonnen, J. Y. Chang, and R. Cheung, *Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy,* Physics in Medicine and Biology **50**, 2887 (2005).

[84] P. Cachier, X. Pennec, and N. Ayache, *Fast Non Rigid Matching by Gradient Descent: Study and Improvements of the 'Demons' Algorithm*, Tech. Rep. RR-3706 (INRIA, 1999).

[85] Y. S. Kim, J. H. Lee, and J. B. Ra, *Multi-sensor image registration based on intensity and edge orientation information*, Pattern Recognition **41**, 3356 (2008).

[86] J. Revaud, P. Weinzaepfel, Z. Harchaoui, and C. Schmid, *Deepmatching: Hierarchical deformable dense matching*, Submitted to International Journal of Computer Vision (IJCV) (2015).

[87] B. K. Horn and B. G. Schunck, *Determining Optical Flow*, Tech. Rep. (Cambridge, MA, USA, 1980).

[88] T. Brox and J. Malik, *Large displacement optical flow: Descriptor matching in variational motion estimation*, IEEE Trans. Pattern Anal. Mach. Intell. **33**, 500 (2011).

[89] P. Weinzaepfel, J. Revaud, Z. Harchaoui, and C. Schmid, *Deepflow: Large displacement optical flow with deep matching*, in *Computer Vision (ICCV), 2013 IEEE International Conference on* (2013) pp. 1385–1392.

[90] A. Bruhn, J. Weickert, C. Feddern, T. Kohlberger, and C. Schnorr, *Variational optical flow computation in real time*, Image Processing, IEEE Transactions on **14**, 608 (2005).

[91] S. Klein, M. Staring, K. Murphy, M. Viergever, and J. Pluim, *elastix: A toolbox for intensity-based medical image registration*, Medical Imaging, IEEE Transactions on **29**, 196 (2010).

A

ELASTIX: AFFINE REGISTRATION (MI)

//Description: affine
(FixedInternalImagePixelType "float")
(FixedImageDimension 2)
(MovingInternalImagePixelType "float")
(MovingImageDimension 2)
//Components
(Registration "MultiResolutionRegistration")
(FixedImagePyramid "FixedSmoothingImagePyramid")
(MovingImagePyramid "MovingSmoothingImagePyramid")
(Interpolator "BSplineInterpolator")
(Metric "AdvancedMattesMutualInformation")
(Optimizer "AdaptiveStochasticGradientDescent")
(ResampleInterpolator "FinalBSplineInterpolator")
(Resampler "DefaultResampler")
(Transform "AffineTransform")
(NumberOfResolutions 3)
(ImagePyramidSchedule 4 4 2 2 1 1)
(HowToCombineTransforms "Compose")
(AutomaticTransformInitialization "true")
(AutomaticScalesEstimation "true")
(WriteTransformParametersEachIteration "false")
(WriteResultImage "true")
(CompressResultImage "true")
(WriteResultImageAfterEachResolution "true")

(ShowExactMetricValue "true")

//Maximum number of iterations in each resolution level:

(MaximumNumberOfIterations 40)

//Number of grey level bins in each resolution level:

(NumberOfHistogramBins 32)

(NumberOfFixedHistogramBins 32)

(NumberOfMovingHistogramBins 32)

(FixedLimitRangeRatio 0.01)

(MovingLimitRangeRatio 0.01)

(FixedKernelBSplineOrder 3)

(MovingKernelBSplineOrder 3)

(ImageSampler "Grid")

(SampleGridSpacing 1 1 2)

(UseRandomSampleRegion "false")

(NumberOfSpatialSamples 2000)

(NewSamplesEveryIteration "true")

(CheckNumberOfSamples "true")

//Order of B-Spline interpolation used in each resolution level:

(BSplineInterpolationOrder 1)

//Order of B-Spline interpolation used for applying the final deformation:

(FinalBSplineInterpolationOrder 3)

//Default pixel value for pixels that come from outside the picture:

(DefaultPixelValue 0)

//SP: Param-A in each resolution level. $a_k = a/(A+k+1)^\alpha$

(SP_A 20.0)

//Result Image Format

(ResultImagePixelType "unsigned char")

(ResultImageFormat "bmp")

B

ELASTIX: NON-RIGID REGISTRATION (MI)

//Description: BSpline NonRigid Transform with MI

(FixedInternalImagePixelType "float")

(FixedImageDimension 2)

(MovingInternalImagePixelType "float")

(MovingImageDimension 2)

//Components

(Registration "MultiResolutionRegistration")

(FixedImagePyramid "FixedSmoothingImagePyramid")

(MovingImagePyramid "MovingSmoothingImagePyramid")

(Interpolator "BSplineInterpolator")

(Metric "AdvancedMattesMutualInformation")

(Optimizer "AdaptiveStochasticGradientDescent")

(ResampleInterpolator "FinalBSplineInterpolator")

(Resampler "DefaultResampler")

(Transform "BSplineTransform")

(ErodeMask "false")

(NumberOfResolutions 3)

(ImagePyramidSchedule 4 4 2 2 1 1)

(HowToCombineTransforms "Compose")

(WriteTransformParametersEachIteration "false")

(WriteResultImage "false")

(CompressResultImage "true")

(WriteResultImageAfterEachResolution "false")

(ShowExactMetricValue "false")

// Option supported in elastix 4.1:

(UseFastAndLowMemoryVersion "true")

//Maximum number of iterations in each resolution level:

(MaximumNumberOfIterations 100)

//Number of grey level bins in each resolution level:

(NumberOfHistogramBins 32)

(NumberOfFixedHistogramBins 32)

(NumberOfMovingHistogramBins 32)

(FixedLimitRangeRatio 0.01)

(MovingLimitRangeRatio 0.01)

(FixedKernelBSplineOrder 3)

(MovingKernelBSplineOrder 3)

(UseDirectionCosines "true")

//Number of spatial samples used to compute the mutual information in each resolution level:

(ImageSampler "Grid")

(SampleGridSpacing 2 3 5)

(UseRandomSampleRegion "false")

(NumberOfSpatialSamples 2000)

(CheckNumberOfSamples "true")

//Order of B-Spline interpolation used in each resolution level:

(BSplineInterpolationOrder 1)

//Order of B-Spline interpolation used for applying the final deformation:

(FinalBSplineInterpolationOrder 3)

//Default pixel value for pixels that come from outside the picture:

(DefaultPixelValue 0)

(SP_a 10000.0)

(SP_A 100.0)

(SP_{alpha} 0.6)

//Result Image Format

(ResultImagePixelType "unsigned char")

(ResultImageFormat "bmp")

ELASTIX:NON-RIGID REGISTRATION(NMI)

//Description: BSpline NonRigid Transform with NMI

(FixedInternalImagePixelType "float")

(FixedImageDimension 2)

(MovingInternalImagePixelType "float")

(MovingImageDimension 2)

//Components

(Registration "MultiResolutionRegistration")

(FixedImagePyramid "FixedSmoothingImagePyramid")

(MovingImagePyramid "MovingSmoothingImagePyramid")

(Interpolator "BSplineInterpolator")

(Metric "NormalizedMutualInformation")

(Optimizer "AdaptiveStochasticGradientDescent")

(ResampleInterpolator "FinalBSplineInterpolator")

(Resampler "DefaultResampler")

(Transform "BSplineTransform")

(ErodeMask "false")

(NumberOfResolutions 3)

(ImagePyramidSchedule 4 4 2 2 1 1)

(HowToCombineTransforms "Compose")

(WriteTransformParametersEachIteration "false")

(WriteResultImage "false")

(CompressResultImage "true")

(WriteResultImageAfterEachResolution "false")

(ShowExactMetricValue "false")

// Option supported in elastix 4.1:

(UseFastAndLowMemoryVersion "true")

//Maximum number of iterations in each resolution level:

(MaximumNumberOfIterations 100)

//Number of grey level bins in each resolution level:

(NumberOfHistogramBins 32)

(NumberOfFixedHistogramBins 32)

(NumberOfMovingHistogramBins 32)

(FixedLimitRangeRatio 0.01)

(MovingLimitRangeRatio 0.01)

(FixedKernelBSplineOrder 3)

(MovingKernelBSplineOrder 3)

(UseDirectionCosines "true")

//Number of spatial samples used to compute the mutual information in each resolution level:

(ImageSampler "Grid")

(SampleGridSpacing 2 3 5)

(UseRandomSampleRegion "false")

(NumberOfSpatialSamples 2000)

(CheckNumberOfSamples "true")

//Order of B-Spline interpolation used in each resolution level:

(BSplineInterpolationOrder 1)

//Order of B-Spline interpolation used for applying the final deformation:

(FinalBSplineInterpolationOrder 3)

//Default pixel value for pixels that come from outside the picture:

(DefaultPixelValue 0)

(SP_a 10000.0)

(SP_A 100.0)

(SP_{alpha} 0.6)

//Result Image Format

(ResultImagePixelType "unsigned char")

(ResultImageFormat "bmp")

DEMON REGISTRATION OF OCT AND HISTOLOGY IMAGES THROUGH EDGE ORIENTATION-WEIGHTED MODALITY TRANSFORMATION

Smruti Rekha[⋆†]   *Jouke Dijkstra*[†]

⋆ Department of Computer Science, TU Delft
[†]Department of Radiology, Division of Image Processing, Leiden University Medical Center

ABSTRACT

Registration between histology and *ex-vivo* OCT can help in many clinical applications including identifying surgical margin in a tumour tissue during intra-operative pathological diagnosis, quantifying features in OCT for diagnosis and reduction of tissue processing time. In this paper, we present a framework for non-rigid registration between Histology and *ex-vivo* OCT images. The proposed framework consists of a two-stage registration process. The first step consists of large-scale misalignment correction while also establishing the match between OCT and one of the several histology tissue samples based on iterative closest point of prominent edge points. The second step consists of Demon registration based on modality transformation. A modification to modality transformation [1] which preserves the sharpness of the edges in the resultant transformation is also suggested here. This is seen to help in faster convergence of demon registration.

Index Terms— Inter-modality Registration, Iterative Closest Point, Demon Registration

1. INTRODUCTION

Histological examination of a tissue plays a central role and is often considered as a gold standard for diagnosis of various diseases. However, the process of fixing, sectioning and staining a tissue is long and time consuming. Also, there is chance of rendering the tissue unusable with a human error at any of the numerous steps in the specimen preparation process. There has been a growing interest in the use of Optical Coherence Tomography (OCT) in intra-operative pathology as the procedure is quite fast and needs very little preparation time [2] [3]. Establishing a dense correspondence between histology and OCT can help in identifying surgical margin in a tissue using OCT. Identifying descriptors for diagnosis using OCT is also very much at its nascent stage and often involves observing corresponding sections from histology and OCT. Registration between histology and OCT for establishing this correspondence can help speed up the aforementioned clinical procedures. Extensive use of OCT in pathological diagnosis would require training of personnel to extract similar

visual information from OCT images as they would from histological slides. Having an automated framework for registration can also help in this transition.

Registration between images of two different modalities is a well researched topic. A complete overview of area-based and feature-based registration techniques is given in Zitova *et.al.* [4]. Features used in feature-based registration could be one of edges, contours, corners or salient points. The choice of features is essentially application specific. Metrics for area-based registration could either be based on Correlation or Mutual Information(MI). However MI based metric is often preferred in inter-modality registration. Pluim *et.al.* [5] provides a complete review of MI based registration in medical image analysis.

This paper provides a framework for the problem of inter-modality non-rigid registration between Histology and *ex-vivo* OCT images. Since minor deformations are unavoidable at the time of tissue preparation we assume that the corresponding sections could be deformed and hence there is a need for non-rigid registration. A slide prepared for histological examination could contain multiple tissue samples from the same subject. We consider establishing a match between an incoming *ex-vivo* OCT image and one of the several tissue samples present of the same subject as an integral part of the problem as it is in the spirit of speeding up of clinical procedure as discussed before.

The remainder of the paper is organized as follows: The proposed framework for registration of histology and *ex-vivo* OCT images is detailed in section 2. Results are provided in section 3, followed by concluding remarks in section 4.

2. THE PROPOSED APPROACH

As a starting point we present a bird's eye view of the approach followed by delving into the details of each aspect of it. The proposed framework consists of a two-stage registration process. The first step consists of large-scale misalignment correction while also establishing the match between OCT and one of the several histology tissue samples based on iterative closest point [6] of prominent edge points. The second step consists of Demon registration based on edge orientation and gray-level based modality transformation. This

is motivated by [1] but is different from it in the sense that it uses edge-orientation information as a means for establishing confidence of joint occurrence of gray-levels used in modality transformation. This modification is seen to preserve the edges in the resultant transformation.

2.1. Point Cloud Based Rigid Registration

The process of establishing a match between an incoming *ex-vivo* OCT image and one of the histology tissues of the subject is posed as a feature-based registration problem where the features are significant edges in the images. The detected edge points are then registered with each other based on iterative closest point.

Point Cloud Generation: Edge detection is a fundamental low-level task in computer vision and have been widely used in various applications. Learning techniques for edge detection have recently been explored as visually salient edges do not always correspond to color gradient as observed by Martin *et.al.* [7]. We use structured forest based edge detection (SFED) in our work [8]. Classification of each pixel is based on individual predictions of decision trees for the pixel in question as well as the ones in its neighbourhood. Features used for this classification include 3 color channels in CIE-LUV color space, gradient magnitude in original scale as well as half its scale and orientation of the aforementioned gradients.

Point Cloud Pre-processing: The edge points in point clouds could correspond to multiple tissues of the subject. These point clouds are separated from each other by segmenting the foreground(tissue) from the background. For this purpose, color based segmentation technique of mean shift clustering [9] was used. Once the cluster centres are found, clusters are merged based on their proximity to each other until there are only two cluster centres. After separating the point-clouds from each other, each one of them is processed separately with OCT point-cloud. Major orientation of each of these point clouds are found out using principal component analysis and are aligned before registration. Each point-cloud is also scaled and translated such that the coordinates have zero mean and unit variance. This normalization procedure corrects for large-scale translation, scaling and orientation differences.

Iterative Closest Point (ICP): To minimize the distance between the two normalized point clouds, we use ICP. The algorithm matches each point to its closest point and figure out an affine transformation such that this distance decreases in a mean-squared sense. Mean-squared error cost computation excludes outliers so that the error function is not overly affected by a few far away points. This is done iteratively until convergence is reached. The histology point cloud which best matches the OCT point cloud is selected as the corresponding tissue.

2.2. Demon Registration and Modality Transformation

Demon Registration as introduced by Thirion *et.al.* [10] and the improved version by Cachier *et.al.* [11] thrives on edge based force for finding the deformation. Displacement field (u) is defined by:

$$u = \frac{(I_2 - I_1)\nabla I_1}{|\nabla I_1|^2 + \alpha^2(I_2 - I_1)^2} + \frac{(I_2 - I_1)\nabla I_2}{|\nabla I_2|^2 + \alpha^2(I_2 - I_1)^2} \quad (1)$$

where I_1 and I_2 are the two images to be registered and α is a coefficient to decide the effect of edge based force. The obtained field is Gaussian smoothed which serves as a regularization function. The deformation field is computed and regularized iteratively until convergence to register the two images. This works on the assumption that the two images that needs to be registered are obtained from the same imaging modality.

Mutual transform as introduced by Kroon *et.al.* [1] constructs a joint histogram using intensity values of images from different modalities. Let $N_1 \times N_2$ be the size of the joint histogram. The transformed image after modality transformation (I_{1T}) at a pixel (x) is given by:

$$I_{1T}(x) = \arg\max_j(H(I_1(x), j)) \qquad \forall j \in [1, N_2] \quad (2)$$

Since I_1 and I_2 are from 2 different modalities, intensity level L_1 in I_1 may not always get uniquely mapped to L_2 in I_2. This was handled by Kroon *et.al.* using local modality transformation where the transformation was local to each pixel and was calculated using its corresponding neighbourhood. The drawback of this approach is that the edges are not properly defined in the transformed domain.

We argue that the joint histogram should not be weighed equally for each joint occurrence of intensities (L_1, L_2) in I_1 and I_2 respectively, but should be based on the confidence of each occurrence. For a perfectly registered edge, we know that the edge orientation in both the images would be identical [12]. This is used to measure the confidence (C) if a gray-level L_1 in I_1 corresponds to a gray-level L_2 in I_2. Thus the histogram is computed as:

$$H(I_1(x), I_2(x)) = H(I_1(x), I_2(x)) + C(x) \quad (3)$$

Since we are considering images from multiple modalities, contrast reversal could cause aligned edges to either have $\Delta\theta = 0°$ or $180°$. Hence the confidence measure is defined as follows:

$$C_{edge}(x) = \begin{cases} \cos^2(\Delta\theta), & for \Delta\theta \leq 90° \\ \cos^2(180 - \Delta\theta), & for \Delta\theta > 90° \end{cases} \quad (4)$$

Confidence is computed for only those points which were selected as edges as defined in section 2.1 and for all the non-edge points, confidence is induced based on proximity to an edge point. Equation 5 brings out this relationship.

$$C_{non-edge}(x) = \{0.1 + 0.9 \times \exp\left(\frac{-d^2}{d_0^2}\right)\} \times C'_{edge} \quad (5)$$

where, d is the euclidean distance from the nearest edge point and d_0 is a parameter which decides the range in which an edge point has an effect on its neighbourhood and C'_{edge} is the confidence of the nearest edge point. The value of d_0 is selected based on the scale which we are operating in.

Accordingly for a perfectly aligned edge, the joint histogram is incremented by 1, whereas if the edge directions are completely misaligned ($\Delta\theta = 90°$), the joint histogram is not incremented at all. Thus, the occurrence of L_1 in I_1 and L_2 in I_2 is not given any weight at all in such a scenario. Once the histogram is constructed, equation 2 is used for modality transformation. This transformation preserves aligned edges completely.

3. RESULTS

Figure 1 shows an illustration of SFED based edge detection used in this work. To test the performance of our algorithms, 30 Histology and *ex-vivo* Full-Field OCT [13] pairs involving ovarian and breast cancer were collected from Leiden University Medical Center.

Establishing a correspondence between point-clouds becomes difficult if only a fraction of the tissue is present in the field of view. However, this is a realistic use case scenario as sometimes only part of the tissue is imaged using OCT. This is similar to occlusion problem in many computer vision applications. To check the robustness of the point-cloud based registration with respect to 'occlusion', we varied the percentage of OCT tissue in the field of view and studied the translation error of the registered point-cloud. Errors in individual cases varied depending on the how unique the shape of the 'occluded' part was. Figure 2 shows the mean trend in error. The errors are measured by treating the case with the complete tissue visibility as ground truth. While this is not completely accurate, for the sake of studying the variation with respect to tissue visibility in the field of view this assumption can be made. It can be seen from the plot that the mean error is well within 10 pixels even when 50% of the tissue is outside the field of view. Result of the point-cloud based registration as discussed in section 2.1 is shown in figure 3.

Modified modality transformation has a stronger effect on Demon Registration motion field estimation because of the presence of more prominent edges. This makes the algorithm converge faster than when original modality transformation
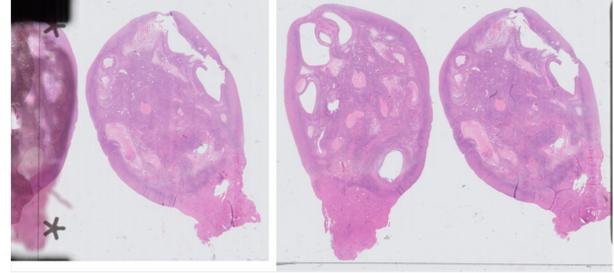
(a) Digitized histology slide with multiple tissues from the same subject
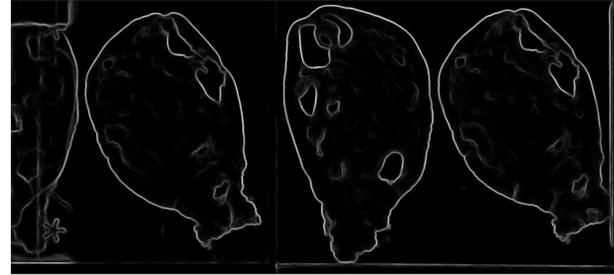
(b) SFED based edge detection of Histology image

(c) *ex-vivo* OCT Image of a tissue (d) SFED based edge detection of OCT image

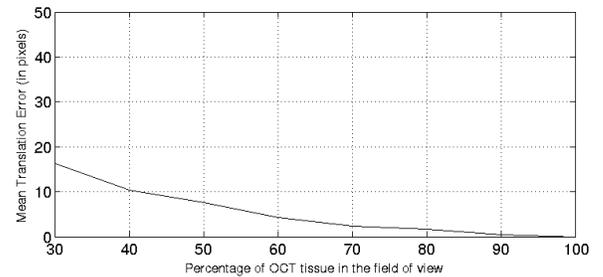**Fig. 1**: Illustration of structured forest based edge detection

Fig. 2: Variation of mean translation error

was used. Figure 4 shows the variation of negated mutual information evaluated after each iteration of demon registration algorithm.
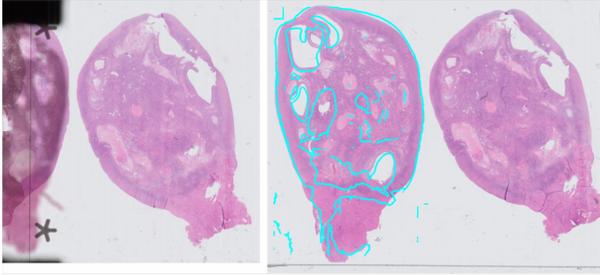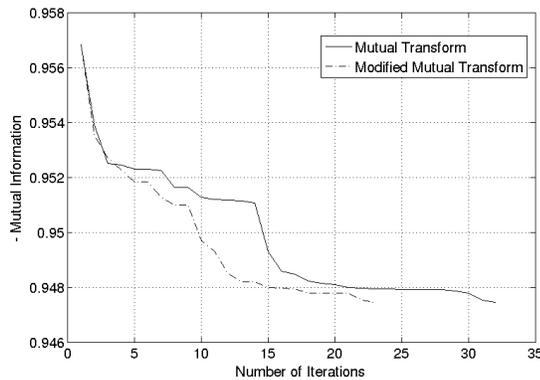
Fig. 3: Rigid Registration

Fig. 4: Convergence of Demon Registration

4. CONCLUSION

A multi-stage framework for registration between histology and *ex-vivo* OCT has been presented here. The automated framework can significantly reduce the pathologist's involvement and effort in finding correspondence. Point-cloud based affine registration is invariant to tissue orientation and is also quite robust with respect to the amount of tissue present in the field of view. Modification to modality transformation has also been proposed here. Instead of giving equal weights to each joint-occurrence of gray levels, a confidence measure based on edge alignment has been introduced here. The modified transformation has more prominent edges in places where the edges in both modalities align well and this has been shown to lead to a faster convergence of demon registration algorithm.

5. REFERENCES

[1] D.-J. Kroon and Cornelis H. Slump, "Mri modalitiy transformation in demon registration," in *Biomedical Imaging: From Nano to Macro, 2009. ISBI '09. IEEE International Symposium on*, June 2009, pp. 963–966.

[2] W. Jung and S.A. Boppart, "Optical coherence tomography for rapid tissue screening and directed histological sectioning," *Analytical Cellular Pathology*, vol. 35, pp. 129–143, 2012.

[3] A. Douplik, "Intraoperative optical coherence tomographic monitoring," in *Handbook of Biophotonics*, pp. 377–400. Wiley-VCH Verlag GmbH and Co., 2013.

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

[5] J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever, "Mutual-information-based registration of medical images: a survey," *Medical Imaging, IEEE Transactions on*, vol. 22, no. 8, pp. 986–1004, Aug 2003.

[6] P.J. Besl and Neil D. McKay, "A method for registration of 3-d shapes," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 14, no. 2, pp. 239–256, Feb 1992.

[7] D.R. Martin, C.C. Fowlkes, and J. Malik, "Learning to detect natural image boundaries using local brightness, color, and texture cues," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 26, no. 5, pp. 530–549, May 2004.

[8] Piotr Dollár and C. Lawrence Zitnick, "Structured forests for fast edge detection," in *Proceedings of the 2013 IEEE International Conference on Computer Vision*, Washington, DC, USA, 2013, ICCV '13, pp. 1841–1848, IEEE Computer Society.

[9] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 5, pp. 603–619, May 2002.

[10] J.-P. Thirion, "Image matching as a diffusion process: an analogy with maxwell's demons," *Medical Image Analysis*, vol. 2, no. 3, pp. 243 – 260, 1998.

[11] Pascal Cachier, Xavier Pennec, and Nicholas Ayache, "Fast Non Rigid Matching by Gradient Descent: Study and Improvements of the 'Demons' Algorithm," Tech. Rep. RR-3706, INRIA, June 1999.

[12] Yong Sun Kim, Jae Hak Lee, and Jong Beom Ra, "Multi-sensor image registration based on intensity and edge orientation information," *Pattern Recognition*, vol. 41, no. 11, pp. 3356 – 3365, 2008.

[13] A. Dubois, K. Grieve, G. Moneron, R. Lecaque, L. Vabre, and C. Boccara, "Ultrahigh-resolution full-field optical coherence tomography," *Appl. Opt.*, vol. 43, no. 14, pp. 2874–2883, May 2004.