

Template-based 3D-2D Rigid Registration of Vascular Structures in  
Frequency Domain from a Single View

by  
Timur Aksoy

Submitted to the Graduate School of Engineering and Natural Sciences  
in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy  
in Computer Science and Engineering

Sabancı University

February 2015

Template-based 3D-2D Rigid Registration of Vascular Structures in Frequency  
Domain from a Single View

APPROVED BY

Assoc. Prof. Dr. Gözde ÜNAL .....  
(Thesis Supervisor)

Prof. Dr. Mustafa Ünel .....

Assoc. Prof. Dr. Selim Balcısoy .....

Assoc. Prof. Dr. Hüsnü Yenigün .....

Prof. Dr. Franjo Pernus .....

DATE OF APPROVAL: .....



©Timur Aksoy 2015

All Rights Reserved



## Acknowledgements

I would like to thank my advisor, Dr. Gözde Ünal for her guidance, sharing of knowledge, generosity with resources and dedicated support for my PhD study and research in our group.

I would like to thank TUBITAK for its support during my PhD studies through the following two projects:

(1) TUBITAK-BMBF Germany Intense Cooperation Grant, Project No: 108E162, titled “Assessment of Fluid Tissue Interaction Using Multi-Modal Image Fusion for Characterization and Progression of Coronary Atherosclerosis”;

(2) TUBITAK-ARRS Slovenia Collaboration Project No: 113E640, titled “Advanced image-guided endovascular interventions by segmentation and 3D/2D registration of cerebral angiograms”.

I would like to thank my family for their extensive support to my PhD study.

I would like to thank my colleagues Andaç Hamamcı, Süheyla Çetin, Devran Uğurlu, Alp Güler for their comments and support during my studies.

I would like to thank our collaborators from Technical University of Munich, Dr. Nassir Navab and Dr. Stefanie Demirci for their contributions to the study in Chapter 3.

I would like to thank our collaborators from University of Ljubljana, Dr. Franjo Pernus and Dr. Ziga Spiclin for their contributions to the study in Chapter 4.

I would like to thank my jury members Dr. Mustafa Ünel, Dr. Selim Balcısoy, Dr. Hüsnü Yenigün, Dr. Franjo Pernus for their helpful feedback on my thesis.

I also would like to thank Dr. Bahattin Koç and his student Can Küçükgül for helping with the 3D printing of the phantom vessel used in our first study at the Sabancı University Nanotechnology Research and Application Center (SUNUM).

Template-based 3D-2D Rigid Registration of Vascular Structures in Frequency  
Domain from a Single View

Timur Aksoy

CS, PhD Thesis, 2015

Thesis Supervisor: Gözde ÜNAL

**Keywords: Medical 3D-2D Registration, Rigid Registration, Image Guided Interventions, Angiography, Coronary Vessels, Cerebral Vessels**

## **Abstract**

Image guided interventions in angiography are performed with a real-time X-ray sequences acquired by a C-arm device which provides the surgeon two dimensional visualization needed to guide the surgical instruments. This visualization may be augmented by registering a three dimensional preoperative volume with the interventional images to provide additional information such as depth, removal of occlusions and alternative views of vessel paths. This thesis presents two novel methods for rigid registration of vascular structures in the preoperative volume to the interventional X-ray image for enhancing visualization in Image Guided Interventions. In the first part of this thesis, estimation of rotation and translation are decoupled. Rotation is estimated by comparing rotated projections of the segmented vessels of the volume with segmented X-ray vessels in frequency domain. Translation is then estimated by minimizing the distances and maximizing the overlap ratio between segmented vessels. The registration results are reported in mean Projection Distances. The second part of the thesis adds separation of out-of-plane translation estimation to the first part and replaces segmentation by gradients. Rotation and out-of-plane translation are estimated by comparing rotational projected templates of volume with depth templates formed by scaling the X-ray image in the Fourier Magnitude Domain. The in-plane translation is then estimated by a Fourier Phase

correlation. The registration results are evaluated by a Gold Standard dataset on cerebral arteries. This method is robust against occlusions and noises due to its usage of gradients and frequency domain similarity, has high capture range and fast, fixed computation times for every step due to template based framework.

Damarların Frekans Uzayında Şablona Dayalı Tek Görüntüden 3B-2B Katı  
Çakıştırılması

Timur Aksoy

CS, Doktora Tezi, 2015

Tez Danışmanı: Gözde ÜNAL

**Anahtar Kelimeler:** Tıbbi 3B-2B Çakıştırma, Katı Çakıştırma, Görüntü kılavuzlu Ameliyatlar, Anjiyografi, Koroner Damarlar, Beyin Damarları

## Özet

Anjiyografideki görüntü kılavuzlu ameliyatlarda C-arm cihazından alınan gerçek zamanlı X-ışını görüntüleri doktorun ameliyat cihazlarını yönlendirmesini sağlar. Bu iki boyutlu görüntüler üç boyutlu ameliyat öncesi alınan görüntülerle çakıştırılarak derinlik, örtünmelerin kaldırılması ve daha detaylı görünüm gibi ek bilgiler sağlar. Bu tezde ameliyat öncesi görüntülerdeki damarlarla ameliyat sırasında alınan X-ışını görüntülerindeki damarların katı çakıştırılması için iki yeni yöntem sunulmuştur. İlk yöntemde dönme ve ötelenme bulunması ayrılır. Dönme üç boyutlu görüntüden alınan bölütlenmiş damarların döndürülmüş projeksiyonlarının bölütlenmiş X-ışını damarlarıyla frekans uzayında karşılaştırılmasıyla bulunur. Ötelenme ise bölütlenmiş damarlar arasındaki uzaklıkların en aza indirilmesi ve örtüşmenin en yükseğe çıkarılması ile bulunur. İkinci yöntem ilk yönteme ayrı derinlik hesaplaması getirir ve bölütlenmiş görüntüler yerine gradyanları kullanır. Dönme ve derinlik hacmin dönme şablonlarıyla derinlik şablonlarının Fourier Büyüklük uzayında karşılaştırılmalarıyla bulunur. Düzlem üzerindeki ötelenme ise Fourier Faz korelasyonu ile kestirilir. Çakıştırma sonuçları beyin damarlarının Altın Standardı veri kümesiyle ölçülür. Bu yöntem örtünme ve gürültülere gradyanlar ve frekans uzayındaki benzerlik sayesinde dayanıklıdır. Şablonlar sayesinde büyük yakalama menzili ve her adımı için sabit, hızlı hesaplama zamanlarına sahiptir.

# Table of Contents

<b>Acknowledgements</b>	<b>vi</b>
<b>Abstract</b>	<b>vii</b>
<b>Özet</b>	<b>ix</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Problem Statement . . . . .	2
1.3 Overview of Approach . . . . .	2
1.4 Contributions . . . . .	4
1.5 Thesis Organization . . . . .	5
<b>2 Background</b>	<b>6</b>
2.1 Overview of Image Registration . . . . .	6
2.1.1 Spatial Transform Types . . . . .	8
2.2 Medical Image Registration Types . . . . .	9
2.2.1 Intrinsic Registration Review . . . . .	9
2.2.2 Medical 3D-2D Registration Review . . . . .	11
2.3 Registration in Frequency Domain . . . . .	14
2.4 Clinical Setup in Angiography . . . . .	15
2.5 Volume Rendering . . . . .	17
2.6 Error Metrics . . . . .	18
<b>3 Template-based CTA to X-ray Angio Rigid Registration of Coronary Arteries in Frequency Domain with Automatic X-Ray Segmentation</b>	<b>20</b>
3.1 Introduction . . . . .	20
3.2 Methodology . . . . .	22
3.2.1 Segmentation . . . . .	22
3.2.2 3D-2D Registration of Coronary Arteries . . . . .	24
3.2.3 Rotation Recovery . . . . .	25
3.2.4 Similarity Measures for Rotation Recovery . . . . .	29
3.2.5 Optimization for Translation . . . . .	29
3.3 Experiments and Results . . . . .	30
3.3.1 Rotation Recovery . . . . .	30
3.3.2 Validation on Phantom Dataset . . . . .	30
3.3.3 Validation on Synthetic Projections . . . . .	32

3.3.4	Validation on Deformed Phantom Dataset . . . . .	33
3.3.5	Validation on Clinical Dataset . . . . .	38
3.3.6	Effect of $\lambda$ on Registration Results . . . . .	39
3.3.7	Computation Time . . . . .	39
3.4	Discussion and Conclusions . . . . .	43
<b>4</b>	<b>Monoplane 3D-2D Rigid Registration based on Stratified Parameter Optimization Method</b>	<b>45</b>
4.1	INTRODUCTION . . . . .	45
4.1.1	Contributions . . . . .	48
4.2	METHODOLOGY . . . . .	49
4.2.1	Geometry of 3D-2D Registration . . . . .	49
4.2.2	3D-2D Registration Method . . . . .	50
4.2.3	Gradient Images . . . . .	51
4.2.4	Invariance to In-plane Translation . . . . .	52
4.2.5	Estimation of Rigid-body Rotations and Out-of-plane Translation . . . . .	53
4.2.6	Estimation of In-plane Translation . . . . .	55
4.3	Experiments and Results . . . . .	56
4.3.1	Experiment Description . . . . .	56
4.3.2	Registration Evaluation . . . . .	57
4.3.3	Registration Results . . . . .	58
4.4	DISCUSSION . . . . .	59
<b>5</b>	<b>Conclusions and Outlook</b>	<b>64</b>
5.1	Summary and Conclusions . . . . .	64
5.2	Outlook . . . . .	65

## List of Figures

1.1	Illustration of First Contribution . . . . .	3
1.2	Illustration of Second Contribution . . . . .	4
2.1	CT Scanner [1] . . . . .	16
2.2	C-arm Device [2] . . . . .	17
2.3	DRR Projection . . . . .	18
3.1	3D-2D Registration Framework of Chapter 3 . . . . .	22
3.2	X-ray Segmentation Pipeline . . . . .	23
3.3	X-ray Segmentation Steps (left to right): Original X-ray, Weickert and Frangi Filtered, Histogram Thresholded and Median Filtered, Skeleton Map, Skeleton Map Catheter Removed, Segmented X-ray . .	24
3.4	C-arm Setup [3] . . . . .	26
3.5	Phantom Vessel . . . . .	32
3.6	Phantom Image Registration Results: X-ray Images and Difference Images . . . . .	33
3.7	Phantom DRR Image Registration Results: DRR Images and Differ- ence Images after Registration . . . . .	36
3.8	Deformed Phantom Image Registration Results: Deformed CTA and Difference Images . . . . .	37
3.9	Patient X-ray Registration Results: X-ray Image and Difference Image	40
3.10	Box Whisker Plots for Phantom X-ray . . . . .	41
3.11	Box Whisker Plots for Patient X-ray . . . . .	42
4.1	Illustration of a C-arm imaging system with indicated geometric pa- rameters in (4.4) and axes $x_w, y_w, z_w$ of the world coordinate system, axes $x_v, y_v, z_v$ of the 3D image and the axes $u, v, w$ of the 2D detector.	50

4.2	Flow chart of the 3D-2D registration method. . . . .	51
4.3	Results for the proposed, iterative and combined registration on each of 10 clinical image datasets of cerebral angiograms, reported for 3D- and 2D-DSA image pairs involving AP ( <i>top</i> ) and LAT ( <i>bottom</i> ) projection views. Registration accuracy as mRPD is shown on the <i>left</i> and success rate on the <i>right</i> . . . . .	60
4.4	Cumulative success rate (cSR) as fraction of registration trials, which had mRPD < 2 mm, with respect to initial mTRE. The cSR is computed over 10 clinical image datasets that involved registrations of 3D-DSA to 2D-DSA in either AP ( <i>left</i> ) or LAT ( <i>right</i> ) projection views. . . . .	61

## List of Tables

3.1	Numerical Results of Phantom Registration . . . . .	34
3.2	Landmark Errors of Phantom Registration . . . . .	34
3.3	Numerical Results of Phantom DRR Registration . . . . .	35
3.4	Landmark Errors of Phantom DRR Registration . . . . .	37
3.5	Numerical Results of Deformed Phantom Registration . . . . .	37
3.6	Numerical Results for Patient Registration . . . . .	38
4.1	The overall results for registrations on 10 patient datasets of cerebral angiograms [4] between 3D- and 2D-DSA in either AP or LAT projection views. Registration accuracy is reported as MEAN $\pm$ STD of mRPD of successful registrations (mRPD<2 mm), and success rates (SR) computed for registrations over 1000 registrations of 3D and 2D image pairs. Mann-Whitney-Wilcoxon non parametric test was used to evaluate the difference mRPD between of successful registration trials. . . . .	62
4.2	Execution times (MEAN $\pm$ STD) of the proposed method in each of three resolution stages for the generation of rotation and depth templates and registration execution time computed over all 10 datasets for registration to AP and LAT projection views. . . . .	63

# Chapter 1

## Introduction

### 1.1 Motivation

Medical 3D-2D registration is generally defined as bringing voxels in a volume and pixels in a 2D image into correspondence. 3D-2D registration is mainly related to angiography procedures, where blood vessels, particularly arteries and veins are visualized by medical imaging techniques such as X-ray, Computed Tomography (CT) or Magnetic Resonance Imaging (MRI). Typically, a radio-opaque contrast agent is injected into the blood vessel and imaging is performed through an X-ray based technique such as fluoroscopy. The aim of 3D-2D rigid registration for angiographic surgical planning is to enable the physician to apply 3-dimensional depth information and other occluded structures embedded in pre-interventional volumes such as CT and MRI to 2-dimensional intra-interventional images (fluoroscopy) and thereby improve visualization and navigation during the intervention. Registration of the pre-interventional volume and interventional images provides additional valuable information to the surgeon in Image Guided Interventions (IGI) such as precise localization and clearer visualizations of vessel paths during navigation which are critical in assessment of structure and location of pathological conditions such as aneurysms, vessel narrowing, clotting and atherosclerosis. The registration should take place in ideally real-time or near real-time to enable physicians to plan and act accordingly. The benefits of IGIs are shorter procedure times and prevention of tissue damage that may result from misplaced catheters, wires and stents as well as reduced exposure to X-ray radiation of the patient due to potential reduction in the

intervention time.

## 1.2 Problem Statement

Registration is defined as bringing two images into a common coordinate system. The aim of 3D-2D registration is to find spatial transformation of a structure in three dimensional space to align it with the projection on two dimensional image plane. In medical context the volume usually corresponds to a preoperative scan such as CT, MR, PET (Positron Emission Spectroscopy) or SPECT (Single Photon Emission Computed Tomography) and the 2D image can be X-ray or Ultrasound acquired during an intervention. The purpose of 3D-2D Registration in this thesis of either 3D CTA and 2D X-ray images or 3D Rotational Digitally Subtracted Angiography (RDSA) and 2D X-ray images is estimating the rotation and translation of pre-interventional volume (CTA or RDSA) through alignment of its projection with a single X-ray acquired during an intervention. The projection type is perspective for the purpose of simulating the X-ray machine imaging geometry. In the framework of this thesis, all other camera parameters such as position of the detector are assumed to be correct and loaded from the DICOM headers of the given data.

## 1.3 Overview of Approach

Estimation of rotation and translation parameters of the CTA volume are decoupled in the first part of this thesis. In this part, the anatomy of interest is the coronary artery and the medical image modalities are the patient's CTA image volume in 3D, and the X-ray images in 2D. Rotation is recovered by matching rotated CTA Digitally Reconstructed Radiography (DRR) templates (described in Chapter 2) to the segmented X-ray image in frequency domain (See Figure 1.1). In the second step, 3D translation is recovered in spatial domain by minimizing the distance and maximizing the overlap ratio between the 3D vessel model and the 2D vessels. Only the translation component was optimized in the image plane, by minimizing a distance-based cost functional. The results are reported as mean Projection Distance Error.

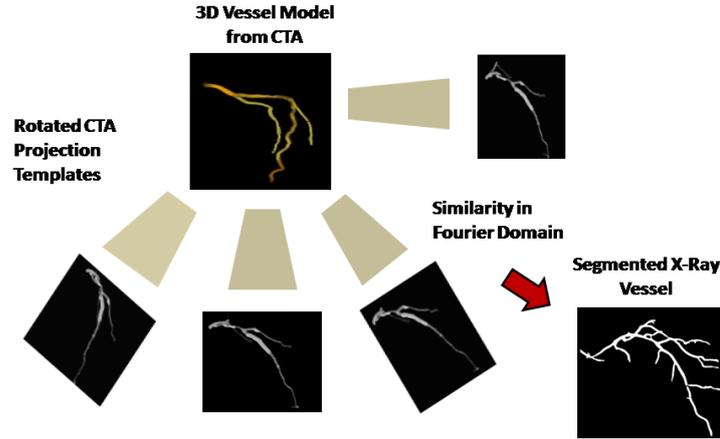


Figure 1.1: Illustration of First Contribution

The second contribution of the thesis presents stratified estimation of rotation, depth and in-plane translation parameters. The medical imaging modalities are the Rotational Digitally Subtracted Angiography (RDSA) in 3D, and the Digitally Subtracted Angiography (DSA) images in 2D of cerebral vessels. This part of the thesis makes use of gradients of the DSA images as inputs instead of segmentations. Rotational DRRs of the RDSA image are compared to depth templates formed by scaled X-ray images in Fourier Magnitude domain which is invariant to in-plane translation. Rotation and depth are recovered from the rotation and scale of the highest correlated DRR template and X-ray, respectively (See Figure 1.2). The in-plane translation is found by Fourier Phase Correlation. The 3D translation is computed by projection matrix equations from in-plane and out-of-plane translations. The discrete estimates of scale and rotation are interpolated to continuous values by linear regression and linear interpolation, respectively. This procedure is repeated twice by changing the resolution of the registration parameters. The results are compared to the Gold Standard Registration and errors are reported in three dimensional distance metrics.

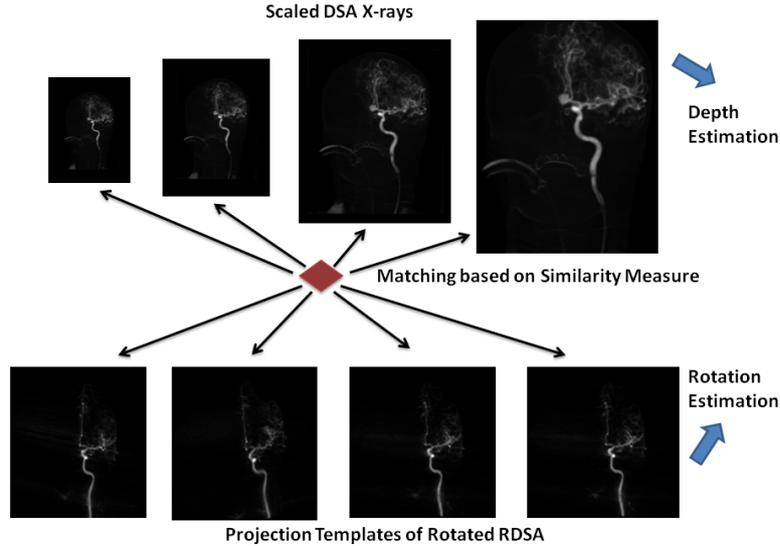


Figure 1.2: Illustration of Second Contribution

## 1.4 Contributions

Rotation and depth estimation is difficult in 3D space from a single view with traditional optimizers due to abundance of local minima in 3D-2D registration. In the first contribution of the thesis, a novel library and segmentation-based 3D-2D registration scheme that decouples rotation and translation estimations is presented. Unlike other segmentation-based approaches, the method requires only a rough translation initialization since the capture range is increased by rotational CTA DRR templates. Iterative optimization is applied only in the translation search. The main contributions of this part are the separation of rotation and translation estimation, use of rotational templates for estimation of rotation and almost fully automatic X-ray segmentation method.

The critical issue of monoplane 3D-2D registration is the correct estimation of the depth component of translation, which was not addressed separately in any of the previous methods. The second contribution of the thesis adds an improved depth estimation to the previous work. The estimation of out-of-plane translation is separated from in-plane translation by creating a discrete scale space of X-rays for the former and using Fourier phase correlation for the latter. Furthermore a possible source of error, the segmentation is replaced by gradients of DSAs. The main con-

tributions of this method are separate depth, rotation and translation estimations, fast and fixed computation times and comparison to Gold Standard Registration.

## **1.5 Thesis Organization**

The organization of the thesis is as follows. In chapter 2 background information on medical registration, clinical setup and related literature review are provided. Chapter 3 presents a novel segmentation based 3D-2D registration method of coronary arteries using rotational templates. Chapter 4 presents a novel segmentation free stratified 3D-2D registration method of vessels by separating estimation of rotation, translation and depth in frequency domain. Chapter 5 contains conclusions and possible future directions in the field.

## Chapter 2

# Background

### 2.1 Overview of Image Registration

Image registration is the process of finding the mapping that brings similar scenes in different images to the same coordinate space. Three main sources of differences are such that the two images are: [5]:

1. recorded from the same scene at different times;
2. recorded by different acquisition devices (intermodal);
3. undergone a spatial transformation such as a rigid transformation or a deformation.

The above three scenarios may occur alone or in combination. Examples for registering images recorded at different times are establishing correspondences between video frames of a moving object or between two snapshots of a scene that has undergone a change from its initial state. In medical imaging, changes of tissues in a pathological region of a patient can be traced by registering images recorded at different times. If these images are from the same modality, they would be considered as a time series.

Intermodal registration has been most commonly performed among images from Computer Tomography (CT), Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), X-ray and Ultrasound devices [6] in literature. The goal is to combine and fuse data from devices that capture the same region of interest with a different sensor hence different response characteristics. Fusion of data from

different devices can provide additional information such as higher precision, depth gain and removal of occlusions by other organs. Because different modalities produce different responses to anatomical structures in the same field of view, finding correspondences between them may require complex operations.

Intermodal registration can aid in clinical imaging of treatment planning through combining functional and structural imaging, for instance, in assessment and treatment of brain tumors by registration of MRI and PET modalities. The active regions of the tumor which may not be enhanced in MRI are highlighted in PET. Intermodal registration can also aid in planning radiation treatment by combining data from previously acquired high resolution non-planning study such as MR, PET, SPECT to planning study such as Computer Tomography Scan. Image Guided Surgery (IGS) combines 3D preoperative volumes with 2D fluoroscopic images during minimal interventional procedures. It provides benefits of planning optimal trajectory of catheter for navigation, reduced risk of tissue damage, shorter procedure times and more accurate tissue resection, ablation or biopsy [7].

The general form of any registration problem entails an optimization function, which is a function of the unknown geometric transform  $\mathcal{T}$ . It penalizes the cost of a distance between the two images when a second image is mapped by the given transformation onto the first one. This term is often called a "data fidelity term". In addition to the data term, typically a regularizer term is added to the cost function in order to constrain the space of solutions, e.g. through a smoothness constraint. Mathematically, the problem of registration is expressed as:

$$\arg \min_{\hat{\mathcal{T}}} C(\mathcal{T}) = S(I_1(\mathbf{x}), \mathbf{I}_2(\mathcal{T}(\mathbf{x}))) + \alpha \mathbf{Regularizer}(\mathcal{T}) \quad (2.1)$$

where  $S(\cdot)$  is the similarity measure,  $\hat{\mathcal{T}}$  is the unknown geometric transform between the two images  $I_1$  and  $I_2$ , which are defined as:  $I : \mathbf{x} \in \mathbb{R}^3 \rightarrow \mathbb{R}$ , as medical imaging modalities in CT, X-ray, structural MRI, PET and ultrasound yield scalar intensity measurements in 3-D (or 2-D) space.

Spatial transform registrations can be categorized as 3D-3D, 2D-2D and 3D-2D with respect to dimension. The most complex type of registration in this scenario is 3D-2D registration which has been studied extensively in the recent literature. 3D-2D registration can be defined as finding the spatial transformation of a structure

in three dimensional space to align it with a projection on two dimensional image plane. There are two main approaches for registering a 3D structure to 2D image(s). One method is by projecting the structure onto image plane and computing the similarity. This can be expressed as:

$$\arg \min_{\mathcal{T}} C(\mathcal{T}) = S(I_1(\mathbf{x}), \mathbf{P}_F(\mathbf{V}(\mathcal{T}(\mathbf{x}))) + \alpha \mathbf{Regularizer}(\mathcal{T}) \quad (2.2)$$

where  $V$  is the 3D structure and  $P_F(.)$  is the projection operator. This thesis uses this approach. Another method is by back projecting the 2D image(s) into 3D space and maximizing the similarity in that space. This can be expressed as:

$$\arg \min_{\mathcal{T}} C(\mathcal{T}) = S(P_B(I_1(\mathbf{x})), \mathbf{V}(\mathcal{T}(\mathbf{x}))) + \alpha \mathbf{Regularizer}(\mathcal{T}) \quad (2.3)$$

where  $P_B(.)$  is the back projection operator. The regularizer terms in Equations (2.2) and (2.3) are typically not required for the specific case of rigid registration, which is the problem of interest in this thesis.

Spatial transform registrations are also classified by their spatial transformation type which is explained in the next section.

### 2.1.1 Spatial Transform Types

Main spatial transforms studied in medical imaging are listed from the most restricted to general as:

1. Rigid Transform;
2. Similarity Transform;
3. Affine Transform;
4. Deformable or Elastic Transform.

Rigid transform occurs by translation and rotation of an object. Similarity transform additionally includes a scale difference between objects. Affine transform adds a shearing effect to the similarity transform where scaling is anisotropic along the cardinal axes [7]. Affine transform is the most general form of linear transformations. Deformable or elastic transforms are nonlinear transforms where constraints are usually needed to improve estimation results and speed. For instance for a

smooth and realistic mapping, length and smoothness preservation terms may be added to the cost function. The function may also be written as combination of polynomials with compact support such as B-Splines, which are inherently smooth. Thin Plate Splines also have inherent smoothness and length constraints. For a recent review on deformable registration techniques, see [8].

## 2.2 Medical Image Registration Types

Medical image registration is divided generally into two main types as intrinsic and extrinsic [5].

Extrinsic methods mark a certain anatomical part of the patient with foreign objects. Correspondences between images are derived from the locations of the markers. This way, calibration and transformation parameters are computed without the need for complex algorithms. Drawback of this type of methods is its invasive nature that includes for instance screw-mounted markers, which may not be always permissible. There are markers that are less invasive such as glues attached to skin but their accuracy drops compared to invasive procedures [6].

Intrinsic registration methods require no introduction of foreign objects and work only by the content of images. These methods are classified into further categories as described in the next subsection.

### 2.2.1 Intrinsic Registration Review

Two main approaches to intrinsic registration are: (i) feature-based and segmentation-based; (ii) voxel-property or intensity-based.

(i) In feature based registration, points or regions with high visibility are extracted and then matched. These features are generally very descriptive and distinctive elements. Features may be geometric or structural. Features can be determined either fully-automated or semi-automatic [9]. Full-automatic methods search and match features without any intervention meanwhile semi-automatic methods may require the user to initialize a few points or approximately mark ROIs. The coordinates, orientation and curvature of the features may be used as the data term in the estimation of the spatial transformation between images. Detection and cor-

response of features are implemented by automatic algorithms such as Scale Invariant Feature Transform, SURF, Laplacian of Gaussian and Hough transform [10, 11]. Line features are composed of edges, ridges and contours. Canny Edge Detector and Laplacian of Gaussian are some of the well known edge finders [12]. Point features may be intersections, center of mass, corners, curvature, extremum responses to signal transforms, inflection points, and so on [9]. Corner detectors are preferred for their invariance to spatial transformation such as rotation, translation and scaling. One of the popular signal transforms is wavelets due to their ability to represent both frequency and spatial information. By choosing the scaling or wavelet functions appropriately, specific details in images can be extracted [13].

Segmentation based registration separates the Region of Interest (ROI) in medical images and then uses them as the sole inputs to the process [6]. While some methods use the segmented regions as a whole, others try to fit geometric elements such as surfaces or contours. Then these elements are matched by an optimization process. One of the most popular tools in this area is Chamfer Distance Functions. By measuring the distances between geometric entities or shapes, they may be overlaid. Gradients of shape contours may also be matched by searching for their best alignment. Moments of segmented structures describe position, orientation and some geometric qualities of objects [14]. For instance second central moments return the principle axes which may indicate the orientation of the shape. Many moment types have been derived which are invariant to noise, distortions and some spatial transforms. Generally moments up to 3rd degree are used. Hu moments derived from classical centralized moments are invariant against translation, rotation, scaling and reflection [15]. Orthogonal moments are preferred for their uncorrelated responses. Affine invariant moments are also found in literature in the field of object recognition. In registration framework, moments must be invariant to certain spatial transforms to determine registration parameters that it is variant to [16]. A drawback of these methods is their dependence on the segmentation success.

(ii) Voxel property based registration uses voxel or pixel intensities directly after applying enhancement operations on images [6]. The methods in this category either use the entire image or the segmented shape intensities. Methods using entire image content were until recently not very popular in 3D due to high computational

load and complexity. After the processor speeds, parallel architectures and memory sizes grew, they became more feasible. Some of the similarity measures are Sum of Squared Differences, Normalized Cross Correlation, Gradient Correlation, Difference of Gradients and Mutual Information Quantity [17]. Moments of voxel intensities may also provide information on the position of shapes. Moments of objects with dense intensity patterns may indicate center of mass, orientation and other clues about its spatial state [15]. Moments up to degree 3 are again commonly used.

### 2.2.2 Medical 3D-2D Registration Review

Markelj et al.[18] categorize 3D-2D medical image registration as intrinsic and extrinsic similarly to the general registration categories in § 2.2. Intrinsic methods of 3D-2D registration are further divided as feature, intensity and gradient based.

The following work have utilized extrinsic fiducials for 3D-2D registration and tracking. Varnavas et al.[19] insert Virtual Fiducial Markers (VFM) to a reference point in preoperative 3D vertebrae data using the markers physically present in 2D intraoperative data. The 3D coordinates of VFM are reconstructed by triangulation of 2D coordinates of fluoroscopy images. Using VFMs, they identify the vertebra of interest, obtain an initial pose estimation and verify the registration results. Initial pose estimation includes estimates of rotation and out-of-translation parameters. Otake et al. [20] estimate relative pose between X-ray images and the 3D anatomical structure using an in-image fiducial and then register CTA with multiple X-rays by intensity-based mutual information and gradient information similarity measures. DRR generation and similarity measure computation algorithms were implemented in GPU, and high success rates were reported in terms of measured mean Target Registration Errors (mTREs).

In the intrinsic feature-based category of 3D-2D registration, the following papers utilized either centerlines, the binary mask of the vessels, or a certain point set extracted from the vessels in their formulations. Metz et al. [21] align coronary centerlines of CTA and X-ray images using the distance transform on the projected model and a fuzzy segmentation of X-ray. Their method also accounts for heart beat and respiratory motion. Rivest-Henault et al. [22] minimize the distances between centerline points of the projected CTA and X-ray images using different optimization

algorithms for translation, rigid, and affine transformations. Their method is further able to perform multi-frame and non-rigid alignment. Turgeon et al. [23] register binarized DRRs of segmented 3D coronary models with segmented X-ray angiography images using mutual information on both single and dual-plane angiography where two images are acquired simultaneously from two different viewpoints. Ruijters et al. [24] compare the distance transform of segmented CTA projection and vesselness of the X-ray image with Powell optimizer. Groher et al.'s [25] graph-based approach specifically developed for liver vasculature, registers a 3D segmented vascular model to the enhanced intra-operative image and simultaneously derives a segmentation of the 2D image. The approach has been advanced to perform non-rigid alignment [26]. Baka et al. [27] train a population CTA coronary model by measuring landmark coordinates on cardiac surfaces and estimating cardiac motion. They register the 3D CTA model to X-ray sequence based on distances and orientation differences between extrapolated 3D vessel points and extracted 2D centerlines and minimized the cost for all 3D and 2D frames. Temporal alignment between CTA and X-ray is modeled by a piecewise linear function and respiratory motion was constructed by quadratic interpolation of poses in first, center and last frames of the sequence. Continuing with more recent feature-based 3D-2D registration approaches, Metz et al [28] have proposed a 3D+t/2D+t (t: time) registration method that uses a patient specific dynamic coronary model derived from the CTA scan by a centerline extraction and motion estimation. The model is aligned by time varying rigid transformations to the X-ray sequence which takes breathing motion (which is also rigid) and temporal relation between CTA and X-ray time points into account. The cost function at any time point is measured as average centerline distance between the CTA and the X-ray centerlines. Baka et al. [29] construct Gaussian Mixture Models (GMM) from moving and scene point-sets of vessel centerlines and use a similarity metric that minimizes the difference between Gaussian mixture probability distributions of both point sets for a 3D-2D registration. Jacobian matrix of the cost function have been analytically computed. Later, orientations are added to the point sets to create 4D GMM distributions. Fully automatic feature-based methods are highly dependent in accurate detection and correspondence of features thus may not satisfy reliability criteria for IGI applications where conditions may vary.

Intensity-based 3D-2D registration uses solely information of pixels and voxels. Digitally Reconstructed Radiographs (DRR) or Maximum Intensity Projections (MIP) are commonly used to project the pre-operative 3D volume into 2D planar space and compare this artificial 2D image to the intra-operative X-ray image(s) by similarity measures such as mutual information [30], gradient difference [17], pattern intensity [31], gradient correlation [32] and sum of squared differences [6]. Dong et al. [33] compared coefficients of orthogonal Zernike moment decompositions instead. Intensity-based methods generally require close initialization due to presence of local minima and suffer from high computational complexity despite speed improvements in new approaches.

In gradient-based methods [34, 35, 36] a small subset of high-magnitude gradients (the edges of structures of interest) is corresponded between 3D and 2D images through projection [34], backprojection [35] or reconstruction [36]. The advantage is that corresponding two subsets of gradients, one of 3D and the other of 2D image, is more efficient than generating the projection images like DRRs or MIPs, while a segmentation or feature extraction that may otherwise be modality- or anatomy-dependent is generally not required. Hybrid feature- and gradient-based method also emerged, for example, to register 3D and 2D cerebral angiograms Mitrović et al. [4] performed a model-to-image 3D-2D registration by matching geometric primitives of the 3D vessels like centerlines, radii and orientations of the vessels to high-magnitude intensity gradients of biplane X-rays. The aforementioned methods, however, were mainly used in registration of 3D to biplane or even multiplane X-ray images.

Reconstruction-based approaches [37, 38] use several X-ray images acquired from different view points, to generate a 3D model that is then registered to the preoperative volume via 3D-3D registration algorithms. However, in particular for cardiac applications, the 2D X-ray images need to be acquired at the same time in order not to create reconstruction artifacts induced by heart motion. Further, as the number of fluoroscopic images decrease, so does the quality of the reconstructed model. Serradel et al. [39] use a generative model for CTA from synthetic samples and simultaneously reconstruct the 3D structure of a non-rigid coronary tree by estimating point correspondences between an input X-ray image and a reference

3D shape. Features are nodes generated in 3D and points of interest are extracted by the vesselness filter in X-ray [40]. The cost function minimizes the reprojection error by alternating matches of corresponding features and perturbing non-rigid parameters stored as a Principal Component Analysis (PCA) model [41]. Nodes are matched by an optimal assignment problem, which uses position, tangent orientation as features. A 3D generative deformation model in the form of a set of nodes are created from the CTA or the biplanar angiography images in order to model possible deformations by Kremer [42]. The method of alignment is similar to that of Serradel et al.'s [39] except that the curvature information is also used when matching nodes from 3D to 2D.

## 2.3 Registration in Frequency Domain

Because optimization algorithms frequently encounter local minima in spatial domain, some authors have mapped images to log-polar Fourier domains to estimate linear registration parameters for generic 2D image registration.

McGuire [43] transforms the images into the Fourier magnitude domain which is invariant to translation in order to recover the similarity transform. Scaling and rotation are reflected in Fourier domain as inverse scale and negative rotation, respectively. Fourier magnitude is mapped to the log-polar space where both parameters are reduced to 2D translations and decoupled as separate variables. Scale axis is integrated to remove the scale difference to obtain a 1D signal that stores the rotation as translation. Scale factor is recovered integrating the rotation axis. The result is a 1D signal where the scale difference is reduced to a translation and the normalized cross correlation is used to find the amount of shift. However, since several peaks appear in the correlation function of the 1D scale signature signal, scale recovery was considered to be not very reliable.

Wolberg and Zokai [44] map the image into the log-polar space without Fourier transform to recover a 2D affine transform. They attempt to minimize the differences in log-polar space with Levenberg-Marquart optimization [45].

Tzimiropoulos et al [46] apply the log-polar Fourier domain registration for the similarity transform using intensity gradients instead of the image function. Nor-

malized gradient correlation is used rather than the standard correlation for translational displacement in spatial domain. Gradient images are mapped to log-polar Fourier magnitude domain where normalized gradient correlations are used to recover scale and rotation stored as 2D translations.

Sarvaiya et al. [47] map images to Fourier log-polar domain for the similarity transform registration. In addition, Fourier phase correlation is used to find the spatial translation. They report that these methods are robust against noise and partial occlusions.

## 2.4 Clinical Setup in Angiography

Angiography is a medical imaging technique used to visualize blood vessels by fluoroscopy. It is performed by inserting a needle into a blood vessel. A catheter is then selectively advanced inside it to the area of the body that needs to be imaged. Contrast material is injected into the blood vessel to highlight the vascular structures and check for medical conditions. If a condition such as clotting, occlusions and dilation is detected that requires further treatment, a minimally invasive endovascular intervention is usually performed at the same time. A surgical instrument is inserted inside the catheter to treat the condition [48]. Atherosclerosis is a vascular disease where arteries are blocked due to buildup of plaques on the vessel walls. The minimally invasive intervention for this disease would be carried out by inserting a stent through the catheter to the narrow section of the artery for allowing the blood to flow. Before stenting, a balloon angioplasty is performed where a collapsed balloon is inserted into the vessel, is advanced to the stenotic region and is inflated at the site to widen the narrowed artery [48].

Preoperative images for angiography in this work is CT or RDSA. CT technology uses computer processing of X-rays to produce tomographic images of a scanned object. CT device consists of an X-ray source and a detector mounted on a gantry [49] (See Figure 2.1). The gantry rotates around the object acquiring several X-ray images by fan beam projection. Images in each rotation form a cross section of the patient anatomy. Spiral CT machines rotate around the object several times by sliding the table on which the patient is lying slowly. Multi-slice CTs have

multiple rows of detectors instead of single one allowing acquisition of multiple cross sections in each rotation by cone beam projection. The raw data obtained from the X-ray projections is a sinogram which is reconstructed to an interpretable volume by inverse Radon Transform. The value of each voxel in the CT image volume corresponds to the attenuation of the tissue on the Hounsfield scale [49]. In angiography applications, a contrast material is administered in order to highlight vascular structures.



Figure 2.1: CT Scanner [1]

The intra-operative imaging device used is called the C-arm (See Figure 2.2). A C-arm device has a c-shaped body consisting of aligned X-ray source and detector attached to both ends. The patient undergoing surgery is placed on a moving table between the source and the detector. There are two possible rotation axes of the gantry, which are cranial-caudal and from left-right. High rotational capability of C-arms enables imaging of the subject from different viewpoints and its use in minimally invasive surgeries in cardiac and neurovascular applications [50].

C-arms emit cone-beam of X-rays that travel through the subject, which are captured and digitized by the detector. When an x-ray beam passes through the body, some of the radiation is absorbed in a process known as attenuation. Anatomy that is denser has a higher rate of attenuation than anatomy that is less dense. The remnant energy of the beam is captured and quantized by the detector which then determines the intensity value of the X-ray image [49]. C-arms are widely used in

hospitals for fluoroscopy. Fluoroscopic imaging creates a sequence of images during interventional procedures. Radio-opaque contrast agents are injected in angiography applications to highlight the vascular systems during angiography interventions.



Figure 2.2: C-arm Device [2]

C-arms can either be stationary or mobile. Stationary machines have the higher image quality and need to be calibrated less frequently while mobile C-arms have the flexibility of movement. Biplane C-arms are equipped with two X-ray detector systems whereas monoplanes C-arms have one detector. Two image planes in biplanar C-arms have usually 90 degrees angle difference between them. Biplane C-arms are usually employed in neurological surgeries whereas monoplanes C-arms are more common in cardiac and abdominal procedures [3].

The integrated stationary C-arms have the ability to rotate around the patient to acquire series of images for cone beam CT reconstruction. These tomograms are referred to as Rotational Angiograms (RA) [51]. In Digitally Subtracted Angiography (DSA), the non-contrasted X-ray image is subtracted from the contrasted one to visualize the vessels only.

## 2.5 Volume Rendering

Volumes such as CT and Rotational Angiography are visualized by a ray casting operation called Digitally Reconstructed Radiography (DRR) [7]. DRR rendering is formed by a ray traveling a straight line from a virtual source through the volume

to the pixel on the screen as depicted in Figure 2.3.

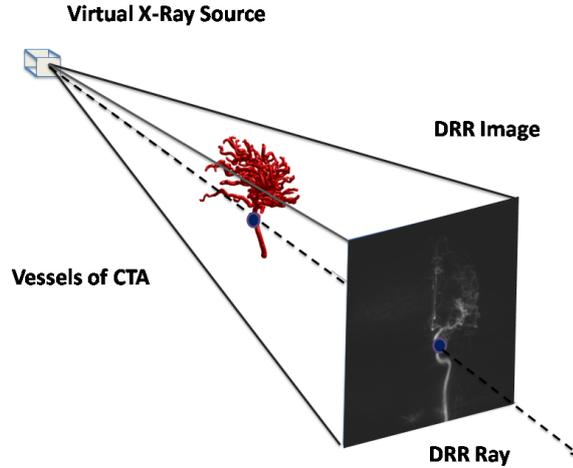


Figure 2.3: DRR Projection

The pixel value is determined by an integrated function of voxels along the ray's path. There are two such functions that determine the value of the intensity of the pixel in this thesis. Composite DRRs process all the voxel intensities which the ray hits along its path. Each voxel is assigned an opacity and a color value based on the tissue type and those values are summed along the ray to determine the final pixel intensity value. Different assignments for opacity and color maps are used for different visualization applications [7]. Minimum Intensity Projection (MinIP) and Maximum Intensity Projection (MIP) put only the value of the minimum and the maximum voxel intensity along the path to the screen, respectively.

## 2.6 Error Metrics

Definitions of error metrics used in this thesis to evaluate the results of registration methods are given below:

**mean Projection Distance (mPD):** The average of distances between projected points and corresponding gold standard points on the image plane.

**mean Reprojection Distance (mRPD):** The average of minimum distances between the 3D target points and lines back projected to source from the target points on the image plane.

**mean Target Registration Error (mTRE):** The average of distances between 3D gold standard point and corresponding target point.

**False Positive:** The percentage of area of vessels in the volume projection not matched by the X-ray vessels.

**False Negative:** The percentage of area of X-ray vessels not matched by vessels in the volume projection.

mPD, False Positive and False Negative metrics are used in Chapter 3 whereas mRPD and mTRE metrics are used in Chapter 4.

## Chapter 3

# Template-based CTA to X-ray Angio Rigid Registration of Coronary Arteries in Frequency Domain with Automatic X-Ray Segmentation

1

### 3.1 Introduction

A key challenge for image guided coronary interventions is accurate and absolutely robust image registration bringing together pre-interventional information extracted from a 3D patient scan and live interventional image information. In this chapter, a novel scheme for 3D to 2D rigid registration of coronary arteries extracted from a pre-operative image scan (3D) and a single segmented intra-operative X-ray Angio frame in frequency and spatial domains for real-time angiography interventions by C-arm fluoroscopy is presented.

The applications of 3D-2D registration involve spine surgery [19], endovascular

---

<sup>1</sup>This chapter is based on the article: T. Aksoy, G. Unal, S. Demirci, N. Navab, and M. Degertekin, "Template-based cta to x-ray angio rigid registration of coronary arteries in frequency domain with automatic x-ray segmentation" *Medical Physics*, vol. 40, no. 10, 2013.

IGIs for treatment of pathologies on cardiac [28] and cerebral vasculatures [4], and on vasculatures in the abdomen (liver [26]), but also further in image-guided radiation therapy. In most of these applications, it is sufficient to use rigid-body alignment of the 3D and 2D images, i.e. on rigid anatomy such as bony structures, on cerebral vasculatures, and even on cardiac vasculatures, if image acquisition is gated to ECG signals. For vasculatures in the abdomen a non-rigid alignment may be required, however, the rigid-body alignment is typically used to initialize the pose of 3D image. Finally, to increase the potential for clinical applications, the 3D-2D registration should perform fast enough so as to enable the surgeon to plan and act accordingly.

Most existing rigid registration approaches require a close initialization due to the abundance of local minima and high complexity of search algorithms. This method eliminates this requirement by transforming the projections into translation-invariant Fourier domain for estimating the 3D pose. For 3D rotation recovery, template Digitally Reconstructed Radiographs (DRR) as candidate poses of 3D vessels of segmented CTA are produced by rotating the camera (x-ray detector) around the DICOM angle values with a specific range as in C-arm setup. We have compared the 3D poses of template DRRs with the segmented X-ray after equalizing the scales in 3 domains, namely Fourier magnitude, Fourier phase and Fourier polar. The best rotation pose candidate was chosen by one of the highest similarity measures returned by the methods in these domains. It has been noted in literature that frequency domain methods are robust against noise and occlusion which was also validated by our results. 3D Translation of the volume was then recovered by distance-map based BFGS optimization well suited to convex structure of our objective function without local minima due to distance maps. A novel automatic X-ray vessel segmentation was also performed in this study. Final results were evaluated in 2D projection space for patient data; and with ground truth values and landmark distances for the images acquired with a solid phantom vessel. Results validate that rotation recovery in frequency domain is robust against differences in segmentations in two modalities. Distance-map translation is successful in aligning coronary trees with highest possible overlap. Numerical and qualitative results show that single view rigid alignment in projection space is successful.

## 3.2 Methodology

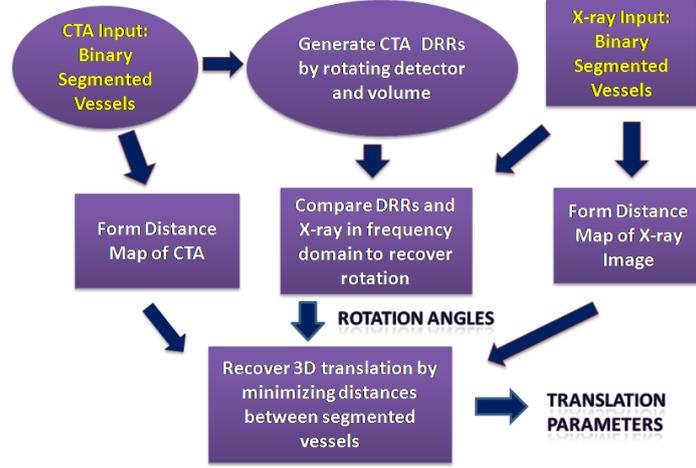


Figure 3.1: 3D-2D Registration Framework of Chapter 3

In this work, 3D-2D registration is performed on segmented vessels. Before describing the details of the 3D-2D registration depicted in the flowchart of Figure 3.1, 3D coronary vessel segmentation algorithm and our almost fully automatic X-ray Angiography segmentation is explained.

### 3.2.1 Segmentation

The most distinct element that determines the position and orientation of vessel projections is their overall shape rather than intensities or features. In our framework, 3D segmentation was achieved by vessel tractography algorithm [52] that successfully segments both left and right coronary artery tree. By employing subsequent connected components, we are able to select the tree of interest where the Image Guided Coronary Intervention (IGCI) is performed. Three or four main branches are extracted by the algorithm; since IGCI is performed only on branches with large diameters, segmentation results are considered sufficient for registration purpose.

The proposed 2D-segmentation of X-ray angiography (Figure 3.2) is almost fully automatic. The process starts up by smoothing Gaussian convolution followed by

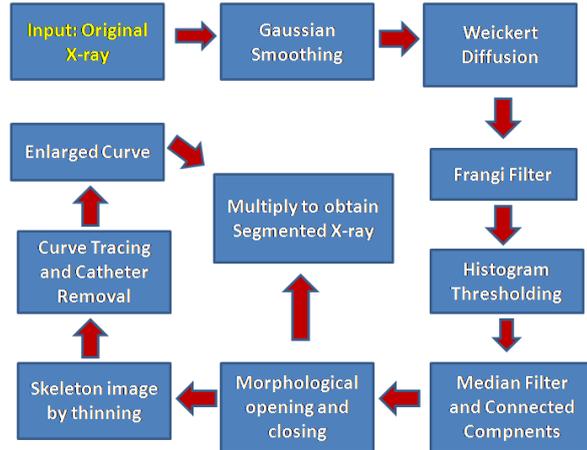


Figure 3.2: X-ray Segmentation Pipeline

Weickert Diffusion [53], which preserves edges and produces segmentation-like results with specific parameters. Frangi filter [40] between scales 2 and 5 was applied on the output. In order to eliminate responses from tubular non-vascular structures in the Frangi filter output, we have applied histogram thresholding that removed 2/5 of foreground pixels, 5x5 median filtering (to erase small artifacts like salt and pepper noise), connected components that remove objects smaller than a ratio of two largest components area, morphological closing and opening with a disk structuring element (to eliminate thin vessels and small artefacts). The output of the last operation still contained some tubular non-vessel structures and the catheter, which also has tubular appearance. The catheter was removed by finding the smoothest segment in the thinned skeleton map of the output image, since it is assumed that real vessels do not follow a smooth path. Initially, the spurs (short outgoing branches) of the skeleton map were removed and the resulting map was traced from every end to every other end of the skeleton curve to divide it into segments. The smoothest segment is found by fitting a conic curve to each segment and measuring the algebraic error. Two smoothest segments with lowest algebraic error were assumed to be catheter or other non-vessel structures to be removed from the map. A limited user interaction is required for deciding whether the smoothest segments are vessel branches because some X-ray frames may not show the catheter. The resulting

skeleton curve was enlarged by 10 pixels in both directions and multiplied by the binary map from the output of the morphological operation step to obtain a segmented vessel image. Figure 3.3 depicts sample images from the described X-ray segmentation process.

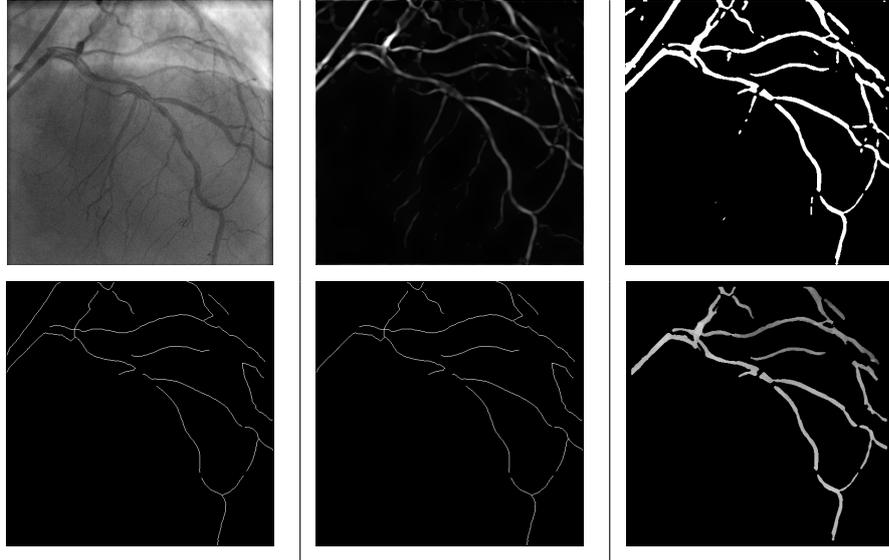


Figure 3.3: X-ray Segmentation Steps (left to right): Original X-ray, Weickert and Frangi Filtered, Histogram Thresholded and Median Filtered, Skeleton Map, Skeleton Map Catheter Removed, Segmented X-ray

### 3.2.2 3D-2D Registration of Coronary Arteries

In general, 3D-2D rigid body registration estimates the optimal projective transformation  $\hat{\mathbf{P}}$  of a 3D volume  $V$  such that its projection perfectly aligns with a 2D image  $I$  in terms of a certain similarity measure  $\mathcal{S}$ :

$$\hat{\mathbf{P}} = \arg \min_{\mathbf{P}} \mathcal{S}(\mathbf{P} \circ V, I) \quad (3.1)$$

where  $\circ$  denotes the application of projection  $\mathbf{P}$  to  $V$ , in particular the multiplication of  $\mathbf{P}$  with every image vector of  $V$ . The projection transformation  $\mathbf{P} = K[R|t]$  consists of the 6-DOF extrinsic parameters  $[R|t]$  for rotation  $(\alpha, \beta, \gamma)$  and translation  $(t_x, t_y, t_z)$  of the 3D volume and the 4-DOF intrinsic imaging parameters

$$K = \begin{pmatrix} f_x & 0 & x_0 \\ 0 & f_y & y_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3.2)$$

of the pinhole projection model [54] with focal length  $f$  in x- and y-dimensions  $f_x = \frac{f}{s_x}, f_y = \frac{f}{s_y}$  ( $s_x, s_y$  is the respective size of a pixel in the target X-ray) and principal point  $(x_0, y_0)$ .

Transferred to the interventional C-arm environment, this theoretical setup has the following practical setting (see also Figure 3.4). The 6-DOF extrinsic parameters are the intensifier primary and secondary rotation angles, the volume rotation along world z-axis around its center, and the volume translation along world coordinate x-, y-, and z-axes. x- and y-axes form the coronal plane of the patient. The intensifier rotation axes are the C-arm detector’s primary (Right Anterior Oblique to Left Anterior Oblique) and secondary (Cranial to Caudal) angles. As model interventional X-ray imaging systems are fully calibrated, prior information on the 4-DOF intrinsic imaging parameters can be directly accessed via respective DICOM header tags (e.g. source-detector distance, source-patient distance, pixel and voxel spacings). In this C-arm setup, the volume origin initially rests at the world origin. Initial x-ray tube and intensifier world coordinates are determined by source-patient distance and source-detector distances respectively.

In the following sections, we describe our novel registration scheme that, in contrast to existing approaches, separates the recovery of rotation parameters  $(\alpha, \beta, \gamma)$  from the translation parameters  $(t_x, t_y, t_z)$ . In order to bring together  $V$  and  $I$  in one world coordinate system as very rough initialization, respective extrinsic parameters are filled with values of DICOM tags for intensifier’s primary and secondary angles and the patient’s orientation. As the intensifier is in motion and patients are displaced during C-arm Angiography interventions, DICOM header values are not entirely reliable and are used only as initial parameters before rotation and translation recovery.

### 3.2.3 Rotation Recovery

The rotation recovery phase aims to determine gantry rotation angles accurately. We estimate the rotation angles from a single X-ray frame where the vessel structure is most visible. Binary DRR projections from CTA are created by rotating the camera and volume within a specific angle range and sampling rate as in the setup explained in section 3.2.2. The goal is to find a candidate DRR, which is most

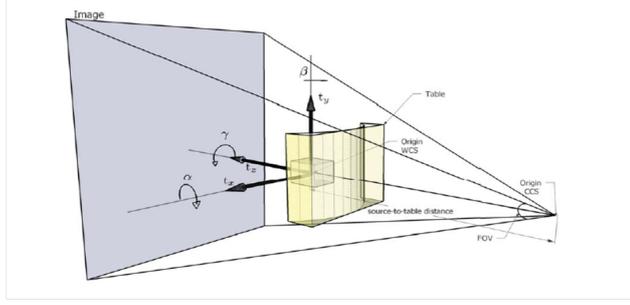


Figure 3.4: C-arm Setup [3]

similar to the real X-ray image with respect to rotation pose. However, since 3D translation difference between the CTA volume and patient is not known yet, the method comparing DRRs and the X-ray image, have to be invariant to translation.

In 2D, the similarity spatial transform between images  $p$  and  $r$  can be expressed as:

$$r(x, y) = p(\delta x + s(x \cos(\theta) - y \sin(\theta)), \delta y + s(x \sin(\theta) + y \cos(\theta))) \quad (3.3)$$

where  $\delta x, \delta y$  are translation distances,  $\theta$  is rotation angle, and  $s$  is the scale difference.

It is well known that *Fourier magnitude or power spectrum* of a digital signal is invariant to translation and large amount of shifts in plane add only noise to the spectrum [43]. *Discrete Fourier transform* (DFT) assumes that finite signals are periodic with their sizes as if the signal repeats itself. So DFT acts as if opposite edges of the image are actually neighbors. This may lead to artifacts in the transform appearing as a plus sign shape. If the vessel in an image is touching an edge as in case of clipping, Fourier transform will respond as if that edge continues at the other end and the vessel structure changes abruptly. Thus, there will be high frequency artifacts due to this unintended abrupt change. To prevent that, a low-pass, blurring, and circular Gaussian filter is applied to remove sudden cutoffs at edges and decrease the values of high frequency components that account for small details and noise. Transferring Equ. 3.3 into Fourier domain, the relation between

$p$  and  $r$  images becomes:

$$F_r(w_x, w_y) = e^{j2\pi(w_x\delta x + w_y\delta y)/s} s^2 F_p((w_x \cos(\theta) + w_y \sin(\theta))/s, (w_x \cos(\theta) - w_y \sin(\theta))/s) \quad (3.4)$$

where  $F_r$  denotes the Fourier transform of the image  $r$ ,  $F_p$  denotes the Fourier transform of the image  $p$ ,  $s$  is the scale difference and  $(\delta x, \delta y)$  denotes translation amount. Coming back to our image registration setup, this method returns the DRR candidate, whose Fourier magnitude is most similar to the real X-ray's, as the closest rotation pose after equalizing the scales of DRRs and X-ray in 2D. Scale differences may result from displacement of arteries along the camera's optical axis or some parts remaining outside of the field of view. In order to equalize the scales, the 0th moments i.e. the areas of binary vessels, then their ratios are computed. The scale difference is removed by inverse scaling of the real X-ray's Fourier magnitude transform as seen in equation 3.4. The reason for scale equalization in Fourier rather than spatial domain is that the enlarged vessels may move out of the frame. After the scale and translation differences are eliminated, the dissimilarity between two spectra can only be due to rotation pose.

A second method for finding potential candidate DRR(s) is analysis of *Fourier phase* relations of DRRs and real X-ray image to conclude whether there are any differences other than pure translation. Images  $p$  and  $r$  having only 2D translation motion  $\delta x, \delta y$  between them, satisfy the following phase,  $\phi$  relation:

$$\phi_r(w_x, w_y) = 2\pi(w_x\delta x + w_y\delta y) + \phi_p(w_x, w_y) \quad (3.5)$$

Since power spectra of these images are the same, their cross correlation in Fourier domain  $R_{rp}$  would yield only the phase difference and can be expressed as:

$$R_{rp}(w_x, w_y) = e^{j2\pi(w_x\delta x + w_y\delta y)} \quad (3.6)$$

As a result, the inverse Fourier transform of  $R_{rp}$  is a delta dirac function translated by  $(\delta x, \delta y)$ .

*Fourier phase correlation* of two shapes differing only by translation would yield a delta dirac function translated by the same amount in the spatial domain [47]. If such output is obtained, then, it will be assumed that the shapes in the projections

differ only by translation. In order to analyze whether this output is a delta dirac function, 1 is subtracted from the largest real element of the output and the *Frobenius norm* is computed. This norm measures the distance of all elements of a matrix from zero. A lower Frobenius norm means that difference from one is closer to a null matrix and hence to the delta dirac function. Frobenius norms of scale equalized Fourier phase correlations are computed for all DRRs. The DRR candidate with the lowest Frobenius norm is selected as the best candidate for rotation pose for this measure. An advantage of this method is that phase correlation in frequency domain is robust to noise and occlusions.

As possible third similarity measure, *Fourier power spectrum* has been mapped into polar space and integrated along the radial axis ( $r$ ). In similarity transform, rotation and scaling in spatial domain result in negative rotation and inverse scaling in Fourier domain respectively. In Fourier polar domain, rotation reflects as translation in angle axis ( $\theta$ ), and scaling reflects as inverse scaling multiplied by a constant amplitude coefficient [44]. This means that rotation and scaling are decoupled into separate axes in this domain. When the radius axis is integrated, a 1D signal invariant to scale and translation differences is obtained. This signal, called rotation signature stores the rotation angle of a shape as a shift. The amplitude coefficients in the rotation signature due to scale difference do not need be normalized since normalized cross correlation between two signals are not affected by constant scaling (variance) and mean values. The highest rotation signature correlation indicates the closest rotation pose. The integration of the polar Fourier transform signal along the radius axis ( $r$ ) is expresses as:

$$\kappa(\theta) = \int M(r\cos(\theta), r\sin(\theta))dr \quad (3.7)$$

where  $M(w_x, w_y) = |F(w_x, w_y)|$  is the Fourier Magnitude,  $r = \sqrt{(w_x)^2 + (w_y)^2}$  and  $\theta = \text{atan}(w_y/w_x)$ .

The Euler angles are recovered from the saved angle of the best candidate DRR returned as described in the next section.

### 3.2.4 Similarity Measures for Rotation Recovery

The three methods we presented in Section 3.2.3, provide ways to obtain various similarity measures between a DRR and the given X-ray image. All binary candidate DRRs were compared with the segmented X-ray image in order to find the closest rotation pose by the following similarity measures:

1. Fourier power spectrum correlation;
2. Integrated Fourier polar signal correlation;
3. Frobenius norm of inverse Fourier phase correlation;

### 3.2.5 Optimization for Translation

Euclidean distance maps of segmented and binarized CTA volume and an X-Ray image are given as the inputs to the optimization program which is developed using QT [55], VTK [56], and ITK [57]. After execution of the first part of the method described in Sections 3.2.3 and 3.2.4, the best DRR candidate's rotation angles and initial translation parameters are entered. Rough translation initialization is necessary because searching algorithms face difficulties in estimating depth in other words finding translation along the optical axis.

Translation-only optimization is less complex, since minima occur only when the shapes are correctly aligned. The objective function value needs to decrease as the overlap measure between the shapes increase. Monotonicity and continuity for gradient-based optimization is provided by the distance maps of binary vessels in 2D and 3D, which are the inputs to our registration algorithm. In the distance maps, the values inside the vessels are 0 and background pixels take the value of the shortest millimeter distance to the vessel structure in single precision floating point format.

Gradient based BFGS optimization algorithm [58] is used for translation registration on x-y-z world coordinates. First order derivatives are computed by central finite difference scheme. During optimization, MinIP (Minimum Intensity Projection) in VTK library [56] which is faster than DRR computation has been used to project the 3D distance map. The energy function to be minimized is the average

shortest distance of all 2D vessel pixels to the projected 3D vessel plus the average shortest distance of all projected 3D vessel pixels to the 2D vessel minus the overlap ratio of both vessels. The overlap ratio measures the percentage of areas of both vessels matched at the end of registration. The energy function is written as:

$$E(t) = \frac{\int \phi(I_1)I_2d\mathbf{x}}{\int \phi(I_1)d\mathbf{x}} + \frac{\int \phi(I_2)I_1d\mathbf{x}}{\int \phi(I_2)d\mathbf{x}} - \lambda \frac{2 \int \phi(I_2)\phi(I_1)d\mathbf{x}}{\int \phi(I_1)d\mathbf{x} + \int \phi(I_2)d\mathbf{x}} \quad (3.8)$$

where  $I_1 = \mathbf{P}_{MIP} \circ D$  expresses the projection of the 3D distance map  $D$  using MIP and incorporating the previously recovered optimal rotation  $\hat{R}$  via  $\mathbf{P}_{MIP} = K[\hat{R}|t]$ .  $I_2$  is the distance map of the real X-ray image,  $\lambda$  a regularization constant set to 2 and  $\phi(\cdot)$  is a one-sided delta dirac function:

$$\phi(y) = \begin{cases} 1 & \text{if } y = 0. \\ 0 & \text{if } y > 0. \end{cases}$$

### 3.3 Experiments and Results

The 3D-2D registration method described in this paper, was evaluated on a phantom vessel, simulated phantom X-rays and clinical datasets. Patient data and phantom images were acquired with Phillips Brilliance 64 channel CT Scanner, Phillips Allura Xper FD 10 C-arm device and kindly provided by our partner physicians.

#### 3.3.1 Rotation Recovery

First measure, Fourier Magnitude Correlation was used in this study with exception of few cases as explained in section 3.4. This measure’s mean and standard deviation for matched results of 15 patient and 5 phantom registrations are  $0.8647 \pm 0.0369$  and  $0.9833 \pm 0.0045$  respectively.

#### 3.3.2 Validation on Phantom Dataset

A synthetic solid 3D vessel sized  $11cm \cdot 10cm \cdot 6cm$ , made of plastic material was printed by ProJetT HD 3000 Professional 3D Printer. The ends of the vessels were left open for discharging support material (wax) after printing and filling it with the contrast agent material of diluted Ultravist 370. Eleven radio-opaque localizer

markers, i.e. metal balls of 2.5 mm diameter, were taped on the phantom vessel. It was then placed in an open box and tightly attached to it with adhesive tapes as depicted in Figure 3.5. One CTA scan and several C-arm X-ray images were acquired at our clinical partner site, Yeditepe University hospital, without changing the position of markers and the vessel within the box. Phantom CTA and X-ray images were segmented by thresholding, hole filling and region growing. The 2D landmark errors were measured after the translation registration had been completed. This error metric is referred as mean Projection Distance Error (mPDE) or in-plane error [59, 60]. Although several other studies on 3D-2D registration approaches have reported results in terms of mean Target Registration Errors (mTREs) [23, 61, 24, 62, 63], this metric is not suitable for single view registration that lacks depth information [59]. Therefore, we opted to show our results by means of mPDEs and difference of actual and recovered angles. Most of the markers were not visible in a single X-ray frame since they were occluded by the contrast agent. In the phantom experiment, the DICOM header primary, secondary angles are considered as the ground truth because the C-arm device was calibrated and there was no intensifier motion during acquisition. Source to patient, source to detector distances and pixel spacings were also obtained from the DICOM headers as ground truth.

In this dataset, the gantry angle range was set to 30 degrees and sample size was set to 12 per rotation axis of the intensifier. The volume rotation range was 15 degrees and the sample size was 6. Therefore  $12 \cdot 12 \cdot 6 = 864$  candidate DRRs were produced for experiment.

Figure 3.6 depicts registration results in five different views of the phantom dataset where successful marker alignments are clearly visible. Results in Table 3.1 indicate that actual rotation angles are also close to the recovered rotation angles from frequency domain registration.

The average mPDEs given in Table 3.2 fall in the interval of [2.3, 4.2] mm and average to 2.898 mm. Metz et al. [21] and Rivest-Henault et al. [22] have also reported mPDEs for their coronary artery registrations. Rivest Henault et al. [22] have performed different optimizers for Rigid Registration. The average mPDE of this study is lower than the lowest error achieved by Cobyala optimizer (3.43 mm). Our average mPDE is slightly higher than the mPDEs reported in Metz et al.'s [21]

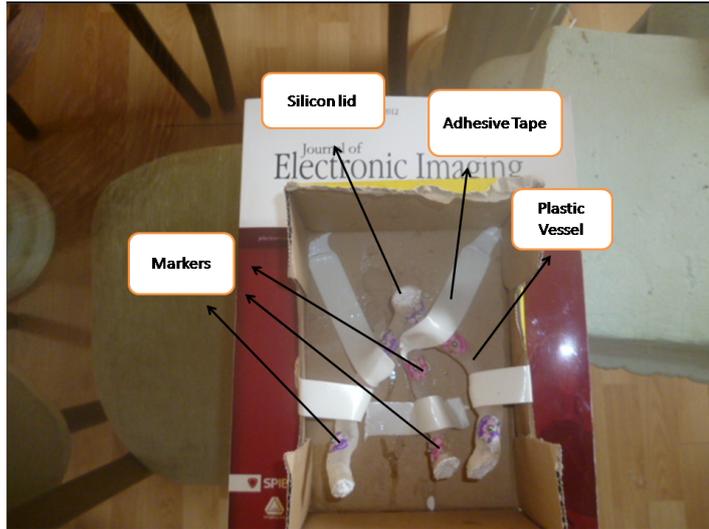


Figure 3.5: Phantom Vessel

work which fall in the range of [1.3, 5.1] mm. However experiments in Metz et al. [21] have started from low offsets set manually before registrations. Our method does not require close initialization of rotation and only rough manual translation initialization therefore our method can perform in more difficult scenarios in lack of close initializations. Recovered rotation angles are close to ground truth rotation angles obtained from the DICOM header.

To give an idea about the energy function value in the non-registered scenarios, in Table 1, for the phantom X-ray index 3 (row 3), the initial values before registration are as follows:

Energy function value: 14.1367, False positive: 77.80, False negative: 77.60, Overlap ratio: 22.22.

### 3.3.3 Validation on Synthetic Projections

Ten random rigid transformations in angle range  $[-90, 90]$  and translation range  $[-40, 40]$  were generated and DRRs of the phantom vessel were produced with them. Figure 3.7 shows the produced DRRs and difference images at the end of the registration. The initial translation offset before registration for all DRRs was  $(0,0,0)$ . In this dataset, the gantry angle range and volume rotation range were set to 25 and 12.5 degrees respectively. Sample size is same as in the phantom dataset, therefore  $12 \cdot 12 \cdot 6 = 864$  candidate DRRs were produced for each experiment which are

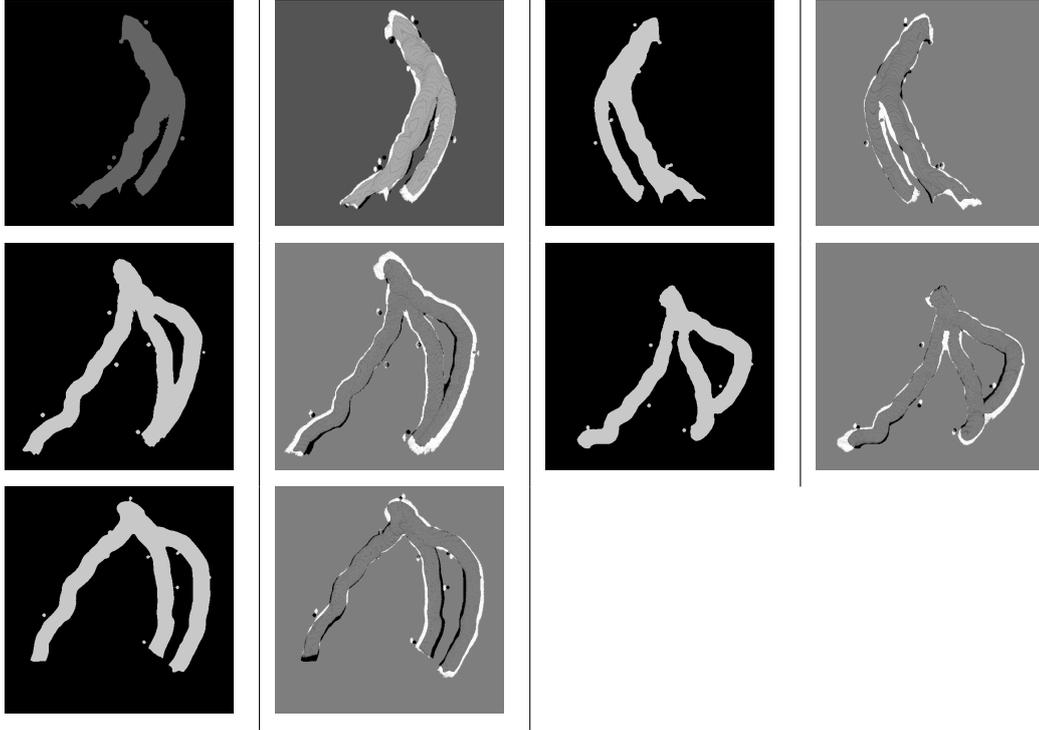


Figure 3.6: Phantom Image Registration Results: X-ray Images and Difference Images

reported in Table 3.3.

The average 2D landmark matching errors (mPDEs) for phantom DRR registration given in Table 3.4 fall in the interval of  $[0.0488, 1.9109]$  mm and average to 0.8046 mm, which is lower than all rigid registration mPDEs reported in Metz et al. [21] and Rivest-Henault et al. [22]. This dataset did not require any manual initialization before registration. Actual and recovered primary and secondary angles are very close in each experiment which are reported in Table 3.3. Recovered and actual translation values differ significantly in many cases. The high errors in translation results from the difficulties the optimizer faces in finding depth.

### 3.3.4 Validation on Deformed Phantom Dataset

To simulate deformed vessel situations and to observe how this rigid registration algorithm is affected by deformations in the vessel, the phantom CTA image was deformed by a vector field whose  $x$ ,  $y$ ,  $z$  components are Gaussian functions with mean as the center of the volume. Three translation registrations were performed on three phantom X-rays at the subsection 3.3.3. Figure 3.8 shows the deformed CTA

Phantom X-ray Index	Actual Primary and Secondary Angles	Primary and Secondary Angles Found	False Positive, False Negative (%)	Translation Values Found (mm)	Energy Function Value
1	0, 90	0, 90	17.81, 4.02	-20.1434, 2.8437, -17.8150	-1.4627
2	0, -90	-7, -89	26.02, 25.37	19.4380, 4.8225, -18.9541	-1.4885
3	-2, 44.7	-2, 47	11.95, 9.89	35.5469, 0.0592, 23.5023	-1.1820
4	44.7, 0	39, -5	16.25, 5.58	11.4881, -19.5214, 2.7578	-1.5170
5	30.2, 29.7	33, 30	12.99, 5.27	9.6186, 26.9833, 0.9414	-1.5032

Table 3.1: Numerical Results of Phantom Registration

Phantom X-ray Index	Number of Visible Landmarks	Average and SD mPDE (mm)
1	5	$4.2458 \pm 1.1651$
2	6	$2.6713 \pm 2.4338$
3	6	$2.2750 \pm 1.7103$
4	5	$2.3470 \pm 0.8921$
5	7	$3.0546 \pm 1.5492$

Table 3.2: Landmark Errors of Phantom Registration

and difference images at the end of the registration. Results in Table 3.5 show the numerical results of deformed registration with higher false positive and negatives as expected however, as observed in Figure 3.8, the rigid registration results are good with a qualitative visual evaluation.

Phantom DRR Index	Actual Primary and Secondary Angles	Primary and Secondary Angles Found	False Positive, False Negative (%)	Actual Translation (mm)	Translation Values	Translation Values Found (mm)	Energy Function Value
1	50.75, -12.25	50.75, 10.17	1.83, 6.26	6.83408	-4.60730	9.3975 -5.0618 -11.6186	-1.8851
2	-50.24, 12.97	-50.24, 12.97	1.90, 4.54	14.7015	-10.3839	5.5144 -12.9882	-1.9090
3	-78.89, 28.48	-78.89, 26.20	3.90, 6.02	14.2919	3.4667	3.3292, 9.5084, -6.8852	-1.8630
4	56.31, 37.67	56.31, 37.67	0.42, 4.95	1.73199	-6.8314	3.1123 -5.3409	-1.9626
5	41.0, -53.13	41.0, -53.13	0.40, 5.86	2.4041	-6.3340	2.777 -6.9845 -15.4042	-1.9166
6	71.68, -84.70	71.68, -82.9	3.30, 10.15	9.0220	-21.1613	6.690, -15.7458 -6.0045	-1.9349
7	84.62, 53.04	82.53, 49.80	8.08, 11.67	-13.1714	11.5716	-11.3300, 12.7954, -22.3988	-1.7769
8	-66.33, -59.9	-66.33, -57.8	4.29, 6.94	11.6863	-6.6731	10.1385, -9.0360, -9.9316	-1.8177
9	34.29, 66.11	34.29, 68.20	6.13, 7.49	6.8724	-24.5034	10.9970, -7.4626, 11.7311	-1.7423
10	24.89, -44.47	24.89, -44.47	4.23, 5.23	6.2224, 22.6618	9.4860, -	11.3815, -3.7706, -9.9220	-1.8542

Table 3.3: Numerical Results of Phantom DRR Registration

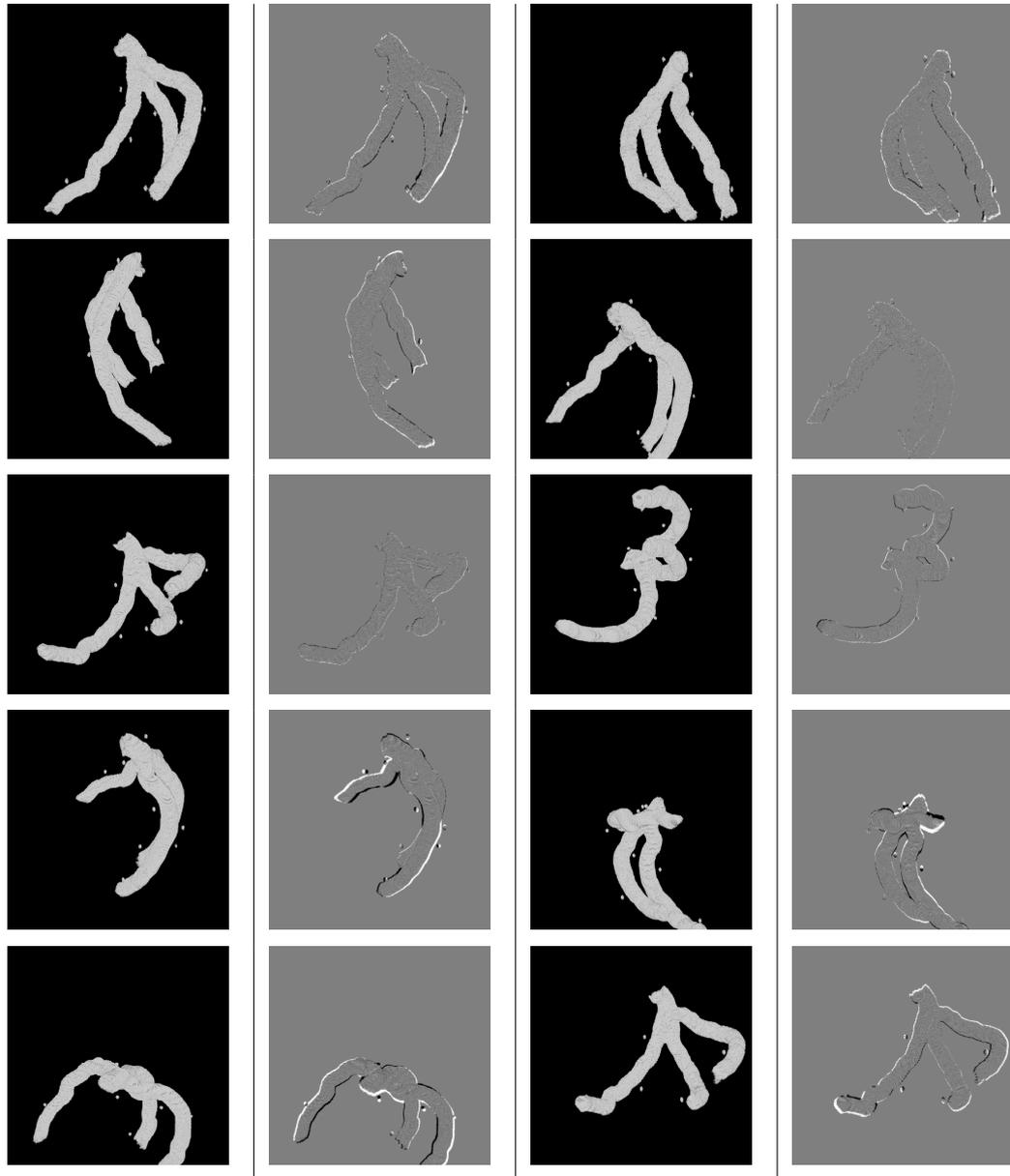


Figure 3.7: Phantom DRR Image Registration Results: DRR Images and Difference Images after Registration

Phantom DRR Index	Number of Visible Landmarks	Average and SD mPDE (mm)
1	6	$1.0789 \pm 0.7829$
2	7	$0.7775 \pm 0.7350$
3	3	$0.9942 \pm 0.5720$
4	7	$0.0488 \pm 0.1292$
5	5	$0.4868 \pm 0.5882$
6	6	$0.4449 \pm 0.4275$
7	7	$1.1804 \pm 1.0008$
8	7	$0.8013 \pm 0.8195$
9	5	$1.9109 \pm 1.2158$
10	4	$1.1313 \pm 0.5278$

Table 3.4: Landmark Errors of Phantom DRR Registration

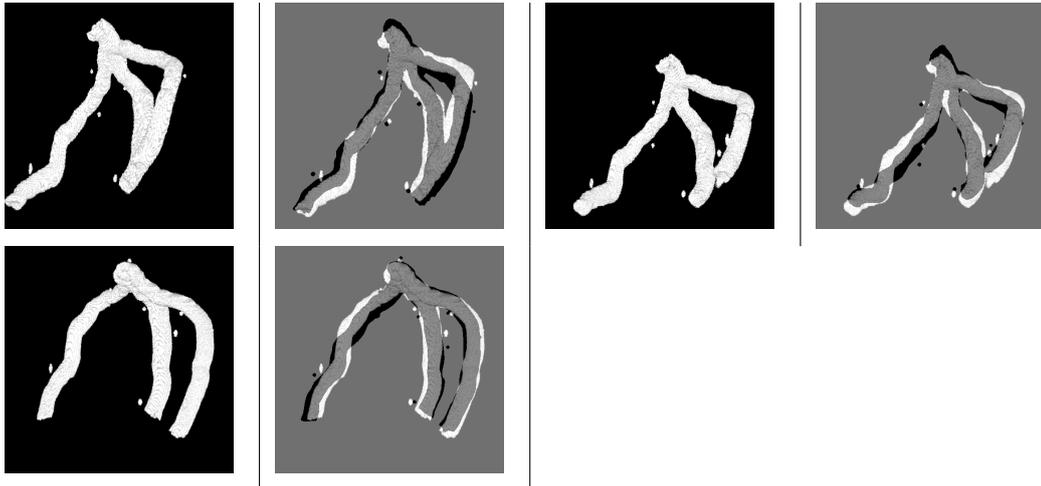


Figure 3.8: Deformed Phantom Image Registration Results: Deformed CTA and Difference Images

Phantom X-ray Index	False Positive, False Negative (%)	Translation Values Found (mm)	Energy Function Value
3	25.80, 21.32	5.8786, -3.3130, -9.1146	0.130
4	26.55, 16.71	4.5709 -13.2017 -13.1773	-0.002
5	20.62, 15.37	-4.0607, -14.9186, -12.4765	-0.5509

Table 3.5: Numerical Results of Deformed Phantom Registration

### 3.3.5 Validation on Clinical Dataset

Clinical experiments were performed on 10 X-ray images of 8 different patients. The angle ranges and sample size is same as the phantom dataset.

X-ray Image Index	DICOM Primary and Sec- ondary Angles	Primary and Sec- ondary Angles Found	Translation Parameters Found (mm)	Energy Function Value
1	-3.4, 44.1	3, 45	-31.6189, 22.0681, -26.8139	5.5919
2	-40.3, 26.8	-40, 31	-15.8008, -7.7006, 6.8409	6.2534
3	39, 38	54, 35	-17.0615, 13.5423, 13.5953	7.7289
4	-26.2, 43	-33, 55	-22.493, 0.89, 16.788	3.5326
5	-29.0, -11.7	-39, -4	-5.8977, -10.4569, -2.1360	4.9570
6	41.6, 35.7	41, 40	-6.9413, 5.9905, 23.1248	5.7021
7	-40.3, 26.8	-40, 31	-14.883, -17.8963, -13.3422	3.3520
6	-3.4, 44.1	3, 45	-31.6189, 22.0681, -26.8139	5.5919
7	-3.301, 34.4	1, 44	0.6029, -7.984, 12.145	4.320
8	-40.3, 26.8	-40, 31	-14.883, -17.8963, -13.3422	3.3520
9	-8.7, 43.1	3, 45	-14.5022, 0.8477, -3.6518	1.7182
10	-34.3, -20.1	-36, -22	-96.4062, -57.8730, 91.1996	3.7451

Table 3.6: Numerical Results for Patient Registration

As there is no ground truth data available for patient X-ray images, the visual inspection and cost function values are the only indicators of success. The clinical data parameters as well as the cost function values in Eq 3.8 are reported in Table 3.6. In the second column, Primary and Secondary Angles from the DICOM headers are the initial values used in our registration method. The estimated angles and translation parameters are reported in the third and fourth columns of the table. Because the energy function measures the average pixel distances and overlap ratio, the energy function values of different datasets are directly comparable. The minimum value that the cost function can take is -2. and the values obtained from the patient data experiments are reported in the last column of Table 3.6. The registration results are visualized in Figure 3.9, which show successful overlaps. Our team’s clinician partners qualitatively evaluated the difference images and expressed that

the unconformities in alignments are result of non-rigid deformations mainly due to heart and/or breathing motions.

### 3.3.6 Effect of $\lambda$ on Registration Results

The registration results have been evaluated for different values of regularizer  $\lambda$  in the energy function. The box whisker plots of average CTA distance, average X-ray distance and overlap ratios for different lambda values are given in Figures 3.10 and 3.11 for 5 phantom and 3 patient registrations. Average CTA distance is the average distance of all pixels on DRR vessel from X-ray vessel and average X-ray distance is the average distance of all pixels on X-ray vessel from the DRR vessel.

Plots show that  $\lambda$  values have higher impact on patient data than phantom data. For both datasets, lambda value 5 yields the largest overlap ratios and smallest distances between vessels. The plots indicate that setting lambda to a value in the vicinity of 5 would yield the optimal results. In our experiments lambda was set to 2 to bring the overlap ratio in the same scale with other terms.

### 3.3.7 Computation Time

Currently, the algorithms are implemented on CPU with a single thread yielding to average computation times of 4 minutes for template generation, 6-7 minutes for rotation recovery, and 5-6 minutes for translation recovery. Employing code parallelization strategies, however, these algorithms can run much faster. Each template is generated by a DRR computation, which takes 1-2 seconds on average. Since the templates are unrelated to each other, they can be produced independently. Rotation recovery essentially consists of a comparison of each template with a given X-ray image, which takes 1-2 seconds and is again independent. Here, parallel threads would reduce performance time significantly. Translation recovery is an inherently sequential optimization search. MinIP is performed several times in every step for cost function evaluation and DRR is called for visualization. Both MinIP and DRR are ray casting methods, that can be implemented on GPU for several fold speed improvement [32].

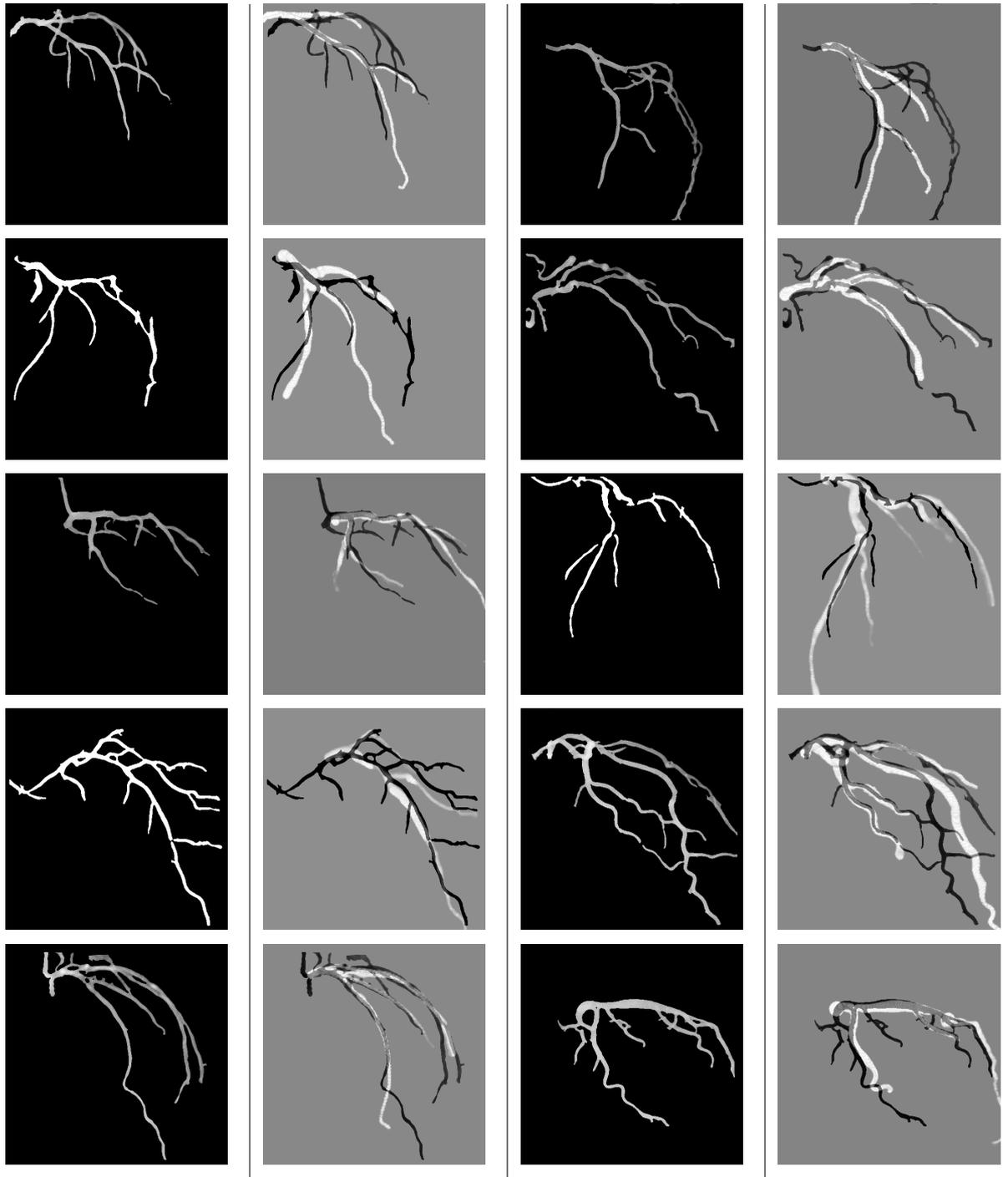


Figure 3.9: Patient X-ray Registration Results: X-ray Image and Difference Image

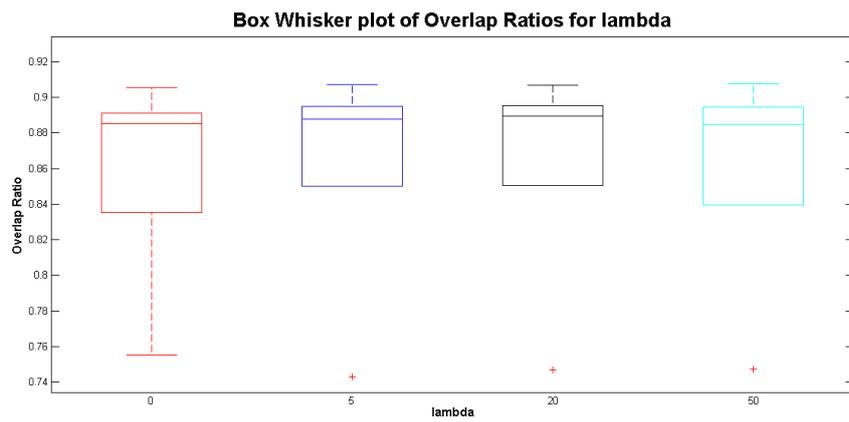
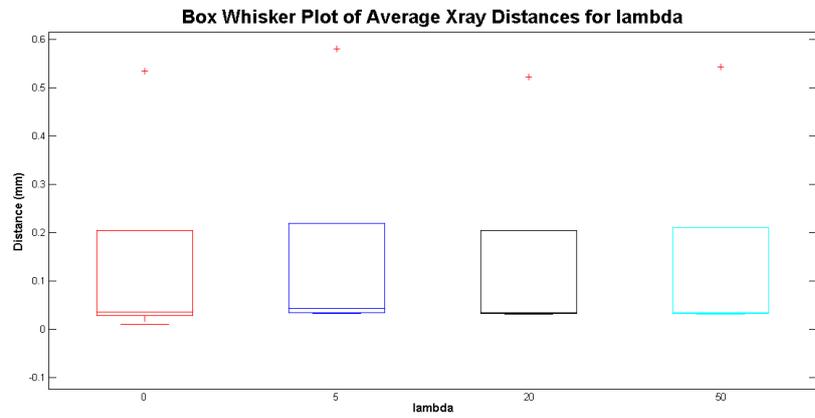
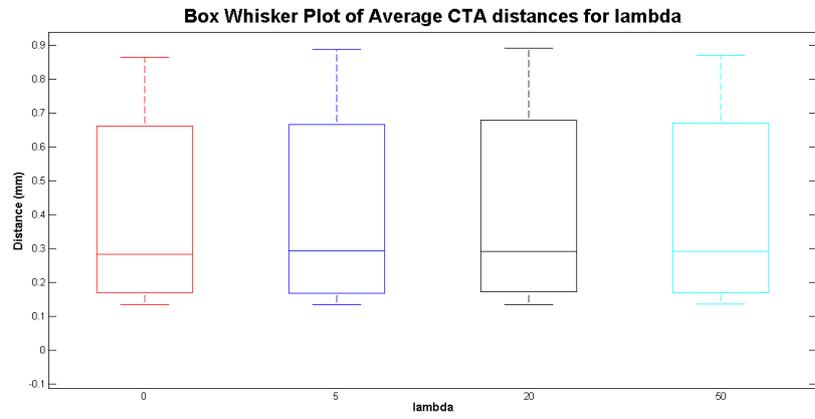


Figure 3.10: Box Whisker Plots for Phantom X-ray

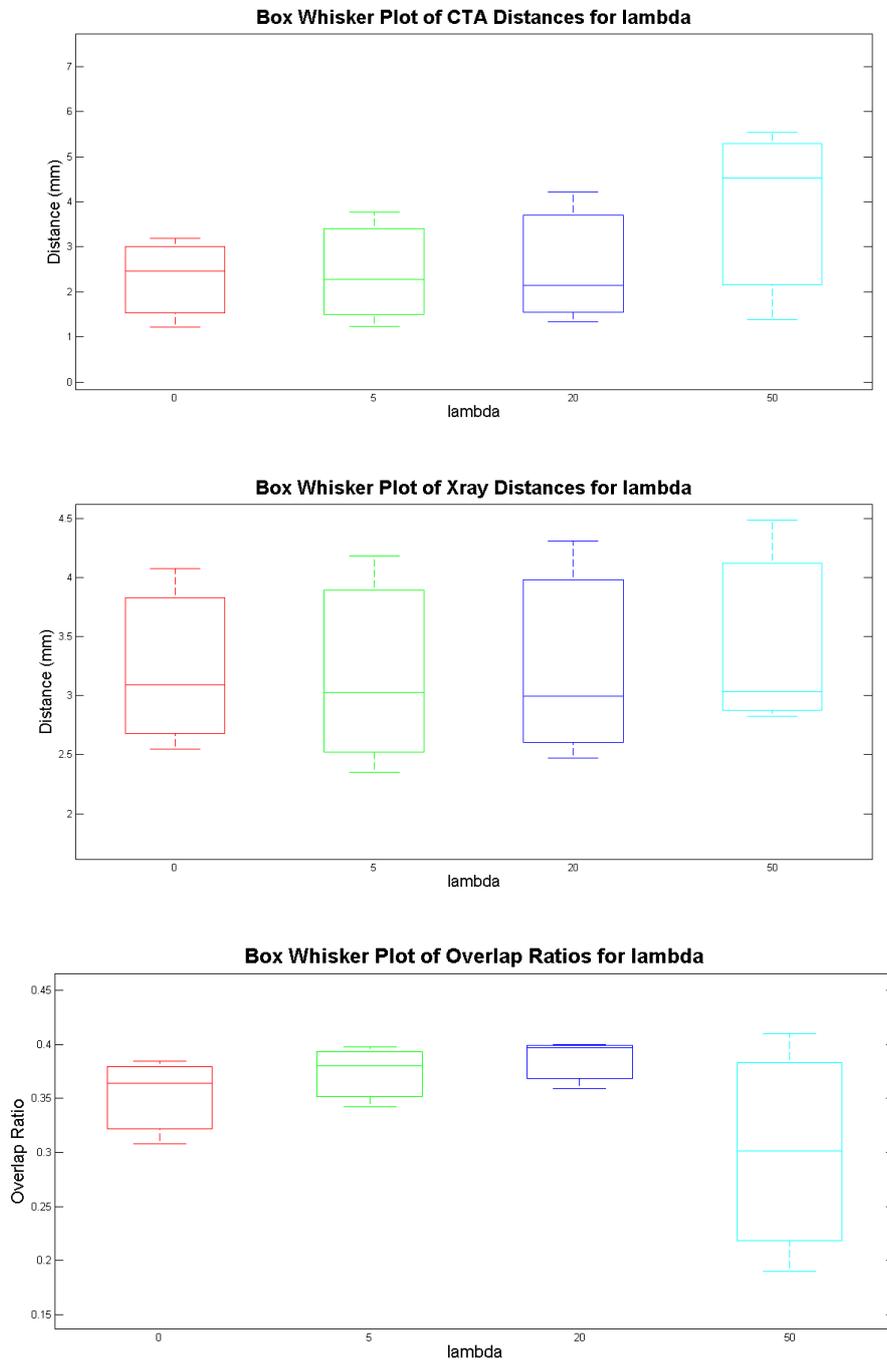


Figure 3.11: Box Whisker Plots for Patient X-ray

### 3.4 Discussion and Conclusions

In this chapter, a novel 3D-2D registration scheme for IGCI, in particular matching of 3D coronary artery tree to 2D coronary arteries model is presented. Instead of recovering all 9 free parameters at the same time, this approach reduces the dimensionality of the parameter space to 6 by including prior knowledge and effectively separates the recovery of rotation from translation parameters. This way, rotation can be estimated in Fourier domain, which is faster, more accurate and robust than existing methods. Translation is eventually estimated in spatial domain via distance maps. This method was evaluated on phantom, synthetic and patient data.

Among similarity measures for rotation recovery, results showed that the first measure always returned the closest pose if the main vessels in X-ray images are not occluded or clipped to a great extent. If a large part of the CTA vessels is not visible in X-ray, then the second measure yielded the closest result. Inverse Fourier phase correlation measure yielded the least similar candidates.

CTA and X-ray segmentations of patient arteries may be different for two reasons. First only a part of the whole vessel structure in segmented CTA may be visible in X-ray image due to vessels not covered by field of view, the contrast agent not penetrating to all branches of interest, or rare occlusion by other tissues. It is also possible that CTA segmentation may not segment all vessels especially small ones that are visible in the X-ray. Because the rotation pose estimation takes place in the frequency domain and depend on statistical correlation measures, they are robust against segmentation differences to a certain degree. When large differences between both images exist, scale ratio can not be computed accurately. In such cases, rotation signature measure should be used. In this study, the vessels in the X-ray image were mostly visible and occlusion was not frequent. In experiments, the second measure was used only twice due to severe occlusion cases, which would be problematic for most intrinsic registration methods. In the other datasets, the first measure was used. A source of error in rotation angles is discretization where the true angle may fall between the sampled angles. Translation errors are due to inability of the optimization algorithm to estimate the depth accurately.

In the phantom and synthetic dataset, the frequency domain method recovered the actual primary and secondary angles with small deviations and distance-map

based BFGS algorithm aligned the shapes successfully. Landmark errors are comparatively small with respect to the studies in the literature as reported in Section 3.3.2.

In the clinical dataset, rigid alignment is successful when judged qualitatively. Segmentation differences may affect translation recovery if there are large detached structures not present on the other map causing distance map values to vary significantly. In such cases, the undesired structures must be removed by connected components operation during segmentation. Another reason for imperfect alignments is that vessels undergo non-rigid deformation due to heart beat motion. Because patient coronary arterial deformations in C-arm may be different than in CTA volume, the rigid transform cannot account for all the motion. Equalization of heart phases can eliminate such deformations.

## Chapter 4

# Monoplane 3D-2D Rigid Registration based on Stratified Parameter Optimization Method

1

### 4.1 INTRODUCTION

Most tasks during Image Guided Interventions (IGI) are carried out under a monoplane X-ray. The registration of a 3D volume to a monoplane X-ray is ill-posed, because out-of-plane translation or differences in depth are difficult to detect in a monoplane X-ray. The bulk of previously published methods for 3D-2D image registration consider the use of biplane, or even multiplane, X-rays. Although methods exist for 3D to monoplane 2D image registration [26, 64], they are typically devised for a specific application in IGI. A general solution that adequately addresses the challenges of 3D to monoplane 2D image registration still does not seem to exist.

Modern facilities for IGI are equipped with an X-ray imaging device on a robotic C-arm system, which can acquire the X-ray projection from almost any arbitrary view. It is based on several X-rays acquired during a semicircular rotation of the

---

<sup>1</sup>This chapter's contents are in preparation for submission as a journal article in collaboration with Dr. Ziga Spielin and Prof.Dr. Franjo Pernus of University of Ljubljana, Slovenia.

C-arm, and can also generate a 3D cone beam CT (CBCT). Although machine-based 3D-2D registration [24] is possible on C-arm imaging systems, usually an initialization of the pose of the 3D image is required, since large alignment errors may arise from patient or C-arm movement during IGI. When the 3D and 2D imaging systems are different, the initial pose of the 3D image is roughly established based on standardized patient position during imaging. However, the initial alignment errors of the 3D image in the coordinate frame of 2D imaging system can be very large. Hence, a method for 3D-2D image registration is required, which is capable of overcoming large initial alignment errors.

Intensity, feature and gradient based methods reviewed in Chapter 2 have drawbacks for IGI applications for the following reasons. Most of the intensity-based image similarity measures have poor sensitivity with respect to translation of 3D image along the direction of projection (out-of-plane translation). Furthermore, generating the projection of raw 3D image is computationally demanding and may represent a bottleneck for clinical application despite speed improvements in recent GPU implementations [65]. The accuracy and robustness of feature-based methods are heavily dependent on the quality of segmentations and extracted features, which must be devised for each image modality and target anatomy of interest. As varying conditions are usually encountered during IGI, these methods may be unreliable, or may require case-by-case tuning, and thus seem less suitable for practical use. When gradient-based methods are used for registration to monoplane X-ray, the similarity measures employed in these methods, similarly to the intensity-based methods, exhibit poor sensitivity to out-of-plane translations. This can be addressed by carefully devising the strategy to optimize the similarity measure.

Usually, the 3D-2D image registration is performed by an iterative optimization of similarity measures [66], which require a good initial guess of transformation parameters. Ruijters et al. [24] compared Powell's and stochastic optimizers and found that the latter was less likely to get stuck in local optimum. Otake et al. [20] employed Covariance Matrix Adaptation Evolution Strategy (CMAES), a global optimization method, which typically requires several thousand similarity measure evaluations. To further improve the robustness to large initial alignment errors, they used CMAES with multiple start and re-start strategy [65], which additionally in-

creased the number of similarity measure evaluations. Florin et al. [67] proposed sequential Monte-Carlo sampling to obtain hypothetical rigid-body transforms, which were coupled to particle filters to get improved transformation. Zikic et al. [68] employed Markov random fields and discrete optimization, in which the values of each parameter were sampled from a uniform discrete value range, while during optimization, they iteratively refined the value range to achieve sub-millimeter registration accuracy. These optimization methods primarily address large initial alignment errors, which are common in 3D-2D registration for IGIs due to the use of different 3D and 2D imaging systems or uncalibrated X-ray imaging systems. Furthermore, these methods generally require a large number of similarity measure evaluations, which may also vary because the number of iterations to convergence will vary between different pairs of 3D and 2D images. Hence, the iterative optimization methods may easily extend the average computational times beyond what is acceptable for clinical application in IGIs.

Optimization strategies that stratify the 3D rigid-body parameters into subsets and perform a consecutive estimation of the subsets of parameters may be employed to increase the sensitivity of any similarity measure to out-of-plane translation of the 3D image. Namely, the sensitivity is generally the highest when all other rigid-body parameters are close to their optimal values. Because the dimension of search space is reduced in the individual consecutive steps, the registration process is usually faster, while exhaustive search can be applied to overcome large initial alignment errors.

Using the stratified parameter estimation, several researchers first determined the in-plane translation parameters by exhaustive grid search [69, 70] or frequency domain methods [71]. Kerrien et al. [69] found the in-plane translations by optimizing normalized cross correlation (NCC) over a fixed grid, followed by optimization of all rigid-body parameters. In a two-stage method, Kita et al. [71] first determined the in-plane translations by optimizing NCC in the frequency domain. By optimizing NCC over fixed grids in a three-stage approach, Hentschke and Tönnies [70] first determined the in-plane translations, then the out-of-plane translation and in-plane rotation and, finally, the two remaining rotations. Kubias et al. [72] determined the in-plane prior to the out-of-plane parameters in a multi-resolution and multi-stage

optimization strategy using low- and high-resolution images in consecutively applied global and local optimizers, respectively. To recover the rotations and scale, Van der Bom et al. [73] used projection-slice theorem, followed by phase correlation to recover the in-plane translations, however, the method resulted in high alignment errors around 20 mm. Aksoy et al. [64] matched rotation templates (set of DRRs of segmented 3D image in a discrete set of rotations) to the segmented X-ray image by a scale and translation invariant measure computed in the frequency domain. After selecting the optimal rotations, all rigid-body parameters were iteratively optimized through minimization of the overlap between the best matching pair of the segmented DRR and the X-ray. Nevertheless, the estimation of the out-of-plane translation in monoplane 3D-2D registration was not addressed adequately in any previous methods.

#### 4.1.1 Contributions

In this chapter, an intrinsic registration method for rigid alignment of a 3D image to a monoplane 2D projection based on a novel stratified parameter optimization is proposed. For efficiency, the method employs a small set of high-magnitude intensity gradients in 3D and 2D. The stratified parameter optimization is based on previously proposed rotation templates [64] to recover 3D rigid-body rotations, while here *we introduce depth templates to recover the out-of-plane translation*. The in-plane translations are found by efficient exhaustive search based on gradient phase correlation [46]. For objective evaluation of the proposed method we used twenty pairs of 3D and 2D images of ten patients undergoing cerebral endovascular IGI [4]. Experiments involved 3D to monoplane 2D registrations in a multi-resolution framework with starting poses that had initial mean target registration error up to 100 mm. The registration accuracy was below the image sampling step and the execution time was constant and below 10 s, while the obtained high success rate of more than 90% shows a high potential of the proposed method for clinical applications.

## 4.2 METHODOLOGY

### 4.2.1 Geometry of 3D-2D Registration

Let the 3D image  $V$  be positioned in 3D by a rigid-body transformation  $\mathbf{T} = \mathbf{T}(\mathbf{q})$  that has six parameters  $\mathbf{q} = (t_x, t_y, t_z, \omega_x, \omega_y, \omega_z)$  and is expressed by a  $4 \times 4$  matrix as

$$\mathbf{T}(\mathbf{q}) = \left[ \begin{array}{ccc|c} & & & t_u \\ \mathbf{R}_{\omega_x} & \mathbf{R}_{\omega_y} & \mathbf{R}_{\omega_z} & t_v \\ & & & t_w \\ 0 & 0 & 0 & 1 \end{array} \right]. \quad (4.1)$$

where  $\mathbf{R}_{\omega_x}$ ,  $\mathbf{R}_{\omega_y}$  and  $\mathbf{R}_{\omega_z}$  are  $3 \times 3$  rotation matrices. For convenience, translations  $t_x, t_y, t_z$  of the 3D volume in world axes  $xyz$  will be expressed in the coordinate system of 2D detector with axes  $uvw$  (Fig. 4.1) as

$$\mathbf{t}_V = [t_u \ t_v \ t_w]^T = [\mathbf{e}_u \ \mathbf{e}_v \ \mathbf{e}_w] [t_x \ t_y \ t_z]^T \quad (4.2)$$

where  $\mathbf{e}_u$ ,  $\mathbf{e}_v$  and  $\mathbf{e}_w$  are  $3 \times 1$  unit vectors of axes  $uvw$  in world coordinate system. The benefit of using this notation is that translation  $t_w$  expresses the out-of-plane, while  $t_u$  and  $t_v$  express the in-plane translations. Our aim is to recover patient's rigid motion between the 3D image  $V$  and the X-ray image  $I$  represented by rotation

$$\mathbf{R}_V = \mathbf{R}_{\omega_x} \cdot \mathbf{R}_{\omega_y} \cdot \mathbf{R}_{\omega_z} \quad (4.3)$$

and translation  $\mathbf{t}_V$  (Eq. (4.2)).

Geometric setup of a C-arm is shown in Fig. 4.1. The projection of 3D points into the 2D detector plane is determined by calibration of the 2D imaging system, such as the C-arm in Fig. 4.1. The output of calibration is a  $3 \times 4$  projection matrix  $\mathbf{P}$ , defined as [74]:

$$\mathbf{P} = \begin{bmatrix} -SID & 0 & -u_0 \\ 0 & -SID & -v_0 \\ 0 & 0 & 1 \end{bmatrix} \left[ \begin{array}{cc|c} & & 0 \\ \mathbf{R}_{PA} & \mathbf{R}_{SA} & 0 \\ & & -SOD \end{array} \right], \quad (4.4)$$

where  $SID$  and  $SOD$  are the respective source-to-detector and source-to-object distances,  $(u_0, v_0)$  the coordinates of the principle point on the 2D detector and  $\mathbf{R}_{PA}$  and  $\mathbf{R}_{SA}$  the  $3 \times 3$  rotation matrices that describe the C-arm gantry rotation w.r.t.

primary ( $PA$ ) and secondary ( $SA$ ) angles (Fig. 4.1), respectively. The principle point ( $u_0, v_0$ ) is generally in the center of the 2D detector (intensifier) and is thus set to  $IS/2$ , where  $IS$  is the intensifier size. The values of parameters  $SID$ ,  $SOD$ ,  $PA$ ,  $SA$  and  $IS$  are accessible from the DICOM header of the 2D images.

Given the projection matrix  $\mathbf{P}$  and rigid-body parameters  $\mathbf{q}$ , any point  $\mathbf{p} = [x\ y\ z]^T$  in 3D has its corresponding point  $\mathbf{t} = [u\ v\ f]$  in the 2D detector (normalization  $\mathbf{t}/f$  yields homogeneous 2D coordinates) that is calculated by:

$$\mathbf{t} = \mathbf{P} \cdot \mathbf{T}(\mathbf{q}) \cdot \mathbf{p}. \quad (4.5)$$

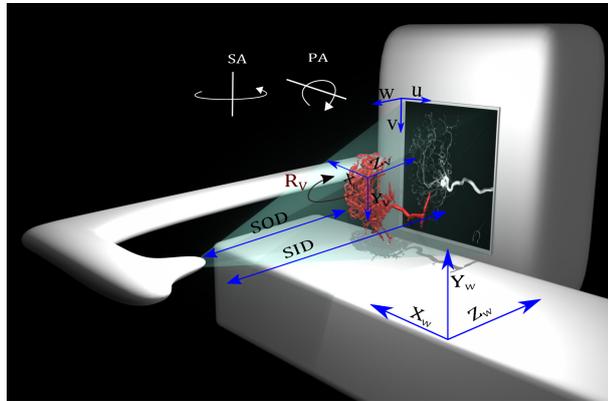


Figure 4.1: Illustration of a C-arm imaging system with indicated geometric parameters in (4.4) and axes  $x_w, y_w, z_w$  of the world coordinate system, axes  $x_v, y_v, z_v$  of the 3D image and the axes  $u, v, w$  of the 2D detector.

## 4.2.2 3D-2D Registration Method

Here we describe a gradient-based 3D-2D rigid registration based on a novel optimization process, which stratifies the 3D rigid-body motion estimation into two consecutive steps: 1) estimation of rotation  $\mathbf{R}_V$  (Eq.(4.2)) and out-of-plane translation  $t_w$  and 2) estimation of in-plane translations  $t_u, t_v$  (Eq.(4.1)). The 3D rigid-body motion estimation is performed in a multi-resolution framework for a discrete set of image resolutions  $r_k; k = 1, \dots, K$ . The methodology is outlined by the flowchart in Figure 4.2. In the following we describe the representation and computation of gradient images, the estimation of rotations and out-of-plane translations, and the estimation of in-plane translations.

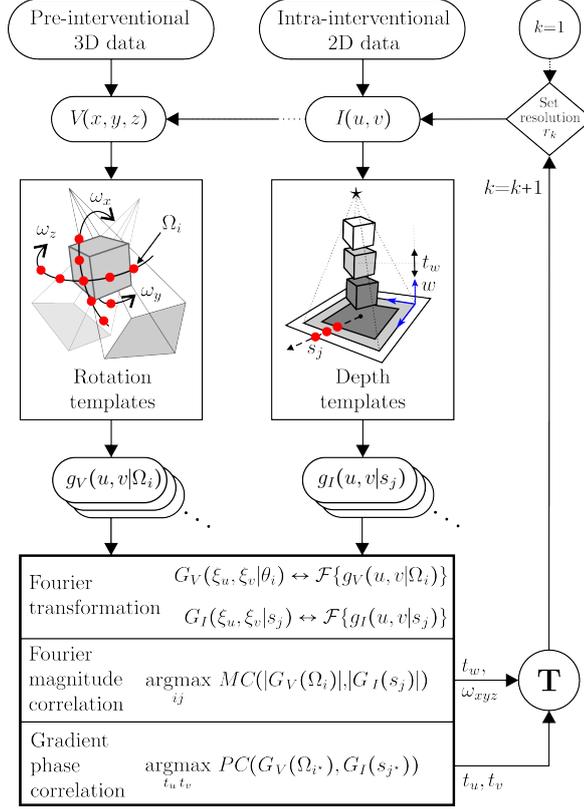


Figure 4.2: Flow chart of the 3D-2D registration method.

### 4.2.3 Gradient Images

Gradients provide the advantage of extracting the edges of salient, high-contrast structures and diminishing the contribution of faint, low-contrast structures such as low frequency noise and blur. Using gradients for image registration may cancel the need to accurately segment the structures of interest. The 2D intensity gradient  $\nabla I$  of the 2D image  $I$  is represented by a complex value [46] as

$$\begin{aligned}
 g_I(u, v) &= \mathbf{e}_u^T \nabla I(u, v) + j \mathbf{e}_v^T \nabla I(u, v) \\
 &= \text{Re}(g_I) + j \text{Im}(g_I),
 \end{aligned} \tag{4.6}$$

where, following a Gaussian smoothing of the 2D image, the gradient components were computed based on finite differences. We base the registration on a small subset of high-magnitude intensity gradients

$$g_I(u, v) = \begin{cases} g_I(u, v), & \text{if } |g_I(u, v)| \geq M_t \\ 0, & \text{if } |g_I(u, v)| < M_t \end{cases}, \tag{4.7}$$

where  $M_t$  is a threshold on gradient magnitude  $|g_I(u, v)|$ . The projection of intensity gradients  $\nabla V$  of the 3D image into 2D is defined as [34]

$$g_V(u, v) \sim \left[ \mathbf{e}_u^T + j \mathbf{e}_v^T \right] \cdot \int_{\lambda} \lambda \nabla V(\mathbf{p}(\lambda)) d\lambda, \quad (4.8)$$

where  $\mathbf{p}(\lambda)$  are points on a ray from the X-ray source to the 2D detector. For efficiency, we compute the projected 3D intensity gradients only in those points  $(u, v)$ , where  $|g_I(u, v)| \geq M_t$ .

#### 4.2.4 Invariance to In-plane Translation

We employ the rotation templates, generated from the 3D image, and depth templates, generated from the 2D image, to simultaneously estimate both the rotations and the out-of-plane translation. Since the in-plane translations are not known, and generally cannot be reliably estimated without knowing the rotations and out-of-plane translation, the method comparing rotation and the depth templates has to be invariant to in-plane translation.

McGuire [43] transformed the images into Fourier Magnitude domain, which is invariant to translation, in order to recover a similarity transform. The similarity spatial transform between 2D complex gradient maps  $g_I(u, v)$  and  $g_V(u, v)$  can be expressed as

$$g_I(u, v) = g_v(t_u + s(u \cos\theta - v \sin\theta), \quad (4.9)$$

$$t_v + s(u \cos\theta + v \sin\theta))$$

where  $t_u, t_v$  are translations,  $\theta$  a rotation, and  $s$  the scale.

By transforming the gradient maps into Fourier domain as  $G_I(\xi_u, \xi_v) = \mathcal{F}\{g_I(u, v)\}$  and  $G_V(\xi_u, \xi_v) = \mathcal{F}\{g_V(u, v)\}$ , the relation in Eq. (4.9) becomes

$$G_I(\xi_u, \xi_v) = s^2 e^{j2\pi(\xi_u t_u + \xi_v t_v)/s}$$

$$G_V((\xi_u \cos\theta + \xi_v \sin\theta)/s, \quad (4.10)$$

$$(\xi_u \sin\theta - \xi_v \cos\theta)/s).$$

When the 2D gradient map with the same scale  $s$  as the gradient map of 3D image is compared, the scale difference is eliminated from the equation above and taking its Fourier Magnitude  $|\cdot|$  relates the two sides as:

$$|G_I(\xi_u, \xi_v)| = |G_V(\xi_u \cos\theta + \xi_v \sin\theta, \quad (4.11)$$

$$\xi_u \sin\theta - \xi_v \cos\theta)|.$$

As seen in Eq. (4.11) the Fourier magnitude of two gradient maps differ only by rotation when scales are equalized.

## 4.2.5 Estimation of Rigid-body Rotations and Out-of-plane Translation

For the purpose of 3D-2D image registration the estimates of rotations and out-of-plane translation are found by maximizing the correlation coefficient of Fourier magnitudes ( $MC$ ) between *rotation and depth templates*. Rotation templates are obtained by rotating 3D intensity gradients  $\nabla V$  for a discrete set of rotations  $\Omega_i = \{(\omega_x, \omega_y, \omega_z)_i\}; i = 1, \dots, N$  and projecting into 2D to yield gradient maps  $g_V(u, v)$ .

Motion of the 3D image along the direction of projection is reflected as a change in scale of the 2D projection image (Fig. 4.2). In order to determine the out-of-plane translation, the 2D gradient maps  $g_I(u, v)$  are scaled, for some expected depth range, in a discrete set of scales  $s_j; j = 1, \dots, M$ , which represent the *depth templates*.

The optimal rotations  $\Omega_{i^*}$  and scale  $s_{j^*}$  are obtained by maximizing  $MC$  as

$$(i^*, j^*) = \underset{ij}{\operatorname{argmax}} MC(|G_V(\xi_u, \xi_v | \Omega_i)|, |G_I(\xi_u, \xi_v | s_j)|), \quad (4.12)$$

Given the optimal scale  $s_{j^*}$ , the corresponding out-of-plane translation  $t_w$  can be derived based on the projection matrix  $\mathbf{P}$  of the original 2D image ( $s = 1$ ). Assuming that we select the optimal rotation ( $\mathbf{R}_V \approx \mathbf{Id}$ ) and, without loss of generality, that the world and 2D detector's coordinate systems are aligned ( $\mathbf{R}_{PA}\mathbf{R}_{SA} \approx \mathbf{Id}$ ,  $xyz \leftrightarrow uvw$ ), then using (4.4) the 3D-2D correspondence becomes:

$$\mathbf{t} = \begin{bmatrix} -SID & 0 & -u_0 & u_0 SOD \\ 0 & -SID & -v_0 & v_0 SOD \\ 0 & 0 & 1 & -SOD \end{bmatrix} \mathbf{p}. \quad (4.13)$$

Let  $\mathbf{p} = [u' v' w']^T$ . We will also assume that the 2D image is isotropically scaled, thus we can consider only one of the two 2D coordinates ( $u, v$ ). For the original 2D image and coordinate  $u$  we have:

$$u = \frac{-SID(u' + t_u) - u_0(w' + t_w - SOD)}{w' + t_w - SOD}, \quad (4.14)$$

which simplifies to

$$u + u_0 = \frac{-SID(u' + t_u)}{w' + t_w - SOD}. \quad (4.15)$$

For the scaled 2D image at optimal scale  $s_{j^*}$ , the out-of-plane translation is required to match (in scale) the structures in 3D and 2D, hence  $t_w = 0$ , and (4.15) becomes:

$$s_{j^*}(u + u_0) = \frac{-SID(u' + t_u)}{w' - SOD}. \quad (4.16)$$

By dividing (4.16) and (4.15) we get the  $t_w$  for central point of 3D volume as:

$$t_w = (w' - SOD) \cdot (s_{j^*} - 1). \quad (4.17)$$

Note that  $w' - SOD$  represents the distance from the X-ray source to the position of center, or any other reference point, of the 3D image. In the axes  $xyz$  of the world coordinate system, the out-of-plane translation is expressed as

$$\begin{bmatrix} t_x & t_y & t_z \end{bmatrix}^T = t_w \mathbf{e}_w^T \cdot \begin{bmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \end{bmatrix}. \quad (4.18)$$

The discrete value of the optimal rigid-body rotations  $\Omega_{i^*}$  obtained at maximal  $MC$  is further refined by a linear interpolation scheme. Note that  $\Omega_{i^*}$  is a 3-vector from a discrete set of rotations, i.e.  $[\omega_x \ \omega_y \ \omega_z]_{i^*}^T$ . Let  $\Omega$  represent a 3-vector with arbitrary values and let  $\Delta\omega$  denote the rotation step used to sample rotation along all three rotation axes. A refined rotation estimate  $\Omega^*$  is obtained by computing a matrix of linear weights as:

$$W_{lmn} = MC(\Omega_{i^*} + [l, m, n]^T \cdot \Delta\omega), \quad (4.19)$$

for  $l, m, n = \{-1, 0, 1\}$ . Matrix  $W$  is normalized as:

$$\tilde{W}_{lmn} = W_{lmn} - \min_{l,m,n}\{W_{lmn}\}; \quad \forall l, m, n \quad (4.20)$$

Then, the refined rotations are obtained as:

$$\Omega^* = \Omega_{i^*} + \Delta\omega \cdot \frac{\sum_{l,m,n=-1}^1 \tilde{W}_{lmn} \cdot [l, m, n]^T}{\sum_{l,m,n=-1}^1 \tilde{W}_{lmn}}. \quad (4.21)$$

The discrete value of optimal scale  $s_{j^*}$  is further refined based on the maximal  $MC$  and neighboring values with the following a linear regression model:

$$s_{err} = \beta_1 d_1 + \beta_2 d_2 + \beta_3 \quad (4.22)$$

where  $s_{err} = s_t - s_{j^*}$  is the true scale and  $d_1, d_2$  are the differences the MC value at neighboring discrete scales, i.e.  $d_1 = MC_{i^*(j^*-1)} - MC_{i^*j^*}$  and  $d_2 = MC_{i^*(j^*+1)} - MC_{i^*j^*}$ .

The coefficients  $\beta_i; i = 1, 2, 3$  of the linear regression model are determined in a training phase, which involved the creation of several projected gradient maps  $g_V(t_w)$  at different out-of-plane translations  $t_w$ . The MC values of these gradient maps were computed w.r.t. depth templates, created from a projected gradient map  $g_V(0)$  at the current position of the 3D image. For the purpose of training, the out-of-plane translations  $t_w$  were randomly sampled in the expected range of out-of-plane motion, thereby knowing the actual true scale  $s_t$ .

The aim of the regression model is to capture the error  $s_{err}$  of discrete scale estimate  $s_{j^*}$ , from which a refined "continuous" estimate is obtained by  $s_t = s_{err} + s_{j^*}$ .

#### 4.2.6 Estimation of In-plane Translation

In-plane translation between the gradient maps  $g_V(u, v | \Omega_{i^*})$  and  $g_I(u, v | s_{j^*})$  is obtained by phase correlation (*PC*). When the gradient map  $g_V$  is obtained from  $g_I$  by in-plane translation  $(t_u, t_v)$ , their power spectra remains the same ( $|G_I \overline{G_I}| = |G_V \overline{G_V}|$ ), hence, computing normalized correlation in Fourier domain as  $G_I \overline{G_V} / |G_I \overline{G_V}|$  would yield the phase information as:

$$\begin{aligned} PC(\xi_u, \xi_v) &= \frac{G_I(\xi_u, \xi_v) \overline{G_I(\xi_u, \xi_v)} \cdot e^{j2\pi(\xi_u t_u + \xi_v t_v)}}{|G_I(\xi_u, \xi_v) \overline{G_I(\xi_u, \xi_v)} \cdot e^{j2\pi(\xi_u t_u + \xi_v t_v)}} \\ &= e^{j2\pi(\xi_u t_u + \xi_v t_v)} \end{aligned} \quad (4.23)$$

where  $G_V \doteq G_I(w_x, w_y) \cdot e^{j2\pi(\xi_u t_u + \xi_v t_v)}$ . As a result, the inverse Fourier transform of  $PC(\xi_u, \xi_v)$  is a Dirac impulse translated by corresponding in-plane translation  $(t_u, t_v)$  [47]. The use of gradient maps in *PC* is robust against occlusions and image noise, while registration performance is stable even if there is small overlap between the two images [46].

The obtained optimal translations  $(t_u, t_v)$  in the 2D detector plane can be mapped into the in-plane translations in world coordinate system as follows

$$\begin{bmatrix} t_x & t_y & t_z \end{bmatrix}^T = -\frac{SOD}{SID} (t_u \mathbf{e}_u^T + t_v \mathbf{e}_v^T) \cdot \begin{bmatrix} \mathbf{e}_x & \mathbf{e}_y & \mathbf{e}_z \end{bmatrix}. \quad (4.24)$$

## 4.3 Experiments and Results

### 4.3.1 Experiment Description

The experiments involved 10 datasets of cerebral angiograms [4]<sup>2</sup>. Objective evaluation of the proposed 3D-2D registration method was performed between 3D digitally subtracted angiograms (3D-DSAs) and 2D-DSA images. These images contained two common vascular pathologies, i.e. aneurysm and arteriovascular malformation (AVM). The 2D-DSAs were acquired in the anterior-posterior (AP) and lateral (LAT) projection views, thus forming 20 pairs of 3D and 2D images. For each 3D-2D image pair a reference or gold standard registration was established by co-registration of fiducial markers [4].

The experiments involved 3D and 2D image registration by the proposed method and a recent iterative method based on matching gradient covariances [75]. A set of starting positions for 3D-2D registration, i.e. the initial rigid-body parameters of the 3D image, were defined w.r.t. the gold standard registration as described in the next section. The proposed and the iterative methods were tested as follows: 1) the proposed method was executed with the 3D image in one of the starting positions; 2) the iterative method was executed with the 3D image in one of the starting positions; and, finally, 3) the proposed method was executed with the 3D image in one of the starting positions, and the obtained registered position was further improved by the iterative method.

The three experimental setups are referred to as the

1. *proposed*,
2. *iterative* and
3. *combined*

registrations.

The proposed method was executed in three resolution stages ( $r_1 = 1.5$  mm,  $r_2 = 1$  mm and  $r_3 = 0.5$  mm), in each stage improving the rigid-body parameters of the 3D image over the parameter values obtained at previous stage. In each stage,

---

<sup>2</sup>URL: <http://lit.fe.uni-lj.si/tools.php?lang=eng>

the discrete set of scales and rotations ( $\Omega_i$  and  $s_j$ ) was updated to a smaller range of rotations and scales, and to smaller discrete sampling steps  $\Delta\omega$  and  $\Delta s$ . Rotation and depth templates were created by using the following rotation and scale ranges ( $\Omega_i$  and  $s_j$ , respectively):

1.  $\Omega_i = -10$  to  $10^\circ$  with  $\Delta\omega = 5^\circ$  (125 tmpls.)  
and  $s_j = 0.75$  to  $1.15$  with  $\Delta s = 0.05$  (9 tmpls.)
2.  $\Omega_i = -5$  to  $5^\circ$  with  $\Delta\omega = 2.5^\circ$  (125 tmpls.)  
and  $s_j = 0.90$  to  $1.10$  with  $\Delta s = 0.025$  (9 tmpls.)
3.  $\Omega_i = -1$  to  $1^\circ$  with  $\Delta\omega = 1^\circ$  (27 tmpls.)  
and  $s_j = 0.95$  to  $1.05$  with  $\Delta s = 0.025$  (5 tmpls.)

### 4.3.2 Registration Evaluation

Registration metrics that underlie the evaluation of 3D-2D registration methods were based on measuring the quality of initial and final alignment of 3D targets w.r.t. their gold standard position. The 3D targets were the vessels' centerlines, extracted from 3D-DSA in each of the 10 datasets[4]. Initial positions of 3D images were defined in terms of mean target registration error (mTRE), generated in the range 0–100 mm mTRE w.r.t. the gold standard position by randomly sampling rigid body transformations. For each 3D-2D image pair, one set of rigid-body parameters was generated in each 1 mm subinterval of mTRE, hence, in total 100 per dataset.

Translations were randomly sampled in the range  $[-100, 100]$  mm, while rotations were sampled in the range  $[-5, 5]$  degrees, such that the desired initial mTRE was achieved. The following ranges of translations and rotations correspond to the initial misalignment of the pre-operative 3D and intra-operative 2D images that is expected in a typical interventional C-arm suite [76, 77] and may be due to the use of uncalibrated C-arm pose, due to C-arm table movement (no external tracking system) or due to possible patient movement. Since there were 20 pairs of 3D- and 2D-DSAs and 100 initial rigid-body parameters per each image pair, and three registration tests were executed, we altogether performed 6000 3D-2D registration trials.

After executing 3D-2D registration, accuracy of alignment was measured by mean reprojection distance (mRPD). The mRPD is the mean of minimal distances between the 3D target points in the goldstandard position and lines connecting the X-ray source and the 3D target points in the registered position. The final registration was considered a successful 3D-2D registration if mRPD was less than 2 mm. Overall accuracy of the 3D-2D registration was defined as  $\text{MEAN} \pm \text{STD}$  of mRPD of all the corresponding successful registrations. Overall success rate (SR) was defined as the percentage of successful registrations.

Cumulative success rate (cSR) was computed w.r.t. 2 mm subintervals of the initial mTRE, which consisted of 20 registration trials, such that the cumulative number of successful registration trials was divided by the number of all registration trials. Capture range (CR) was defined as the first mRPD subinterval of length 2 mm, in which more than one of the 20 corresponding registrations failed. This setting corresponds to a 95% confidence level for the CR estimate.

### 4.3.3 Registration Results

The obtained overall results for the proposed, iterative and combined registration are reported in Table 4.1 for the experiments on clinical image datasets of cerebral angiograms involving 20 pairs 3D- and 2D-DSAs, with 10 2D-DSAs in AP and 10 2D-DSAs in LAT projection view. Results per each dataset are shown in Fig. 4.3.

The overall results for registrations by the proposed method involving the AP projection views were: registration accuracy as mean mRPD of  $0.99 \pm 0.48$ , SR of 64.8%, and CR of 4 mm. The results improved significantly with use of combined registration, i.e. mean mRPD of  $0.39 \pm 0.32$  mm, SR of 96% and CR of 18 mm. The differences in registration accuracy per each dataset and overall were significant (p-values  $\ll 0.05$ ) according to Mann-Whitney-Wilcoxon non parametric test for the two unpaired groups (Proposed vs. Combined Registration). On the other hand, using solely the iterative registration [75] resulted in lower mean mRPD compared to the proposed method, however, the SR of iterative registration almost halved compared to the proposed registration (Table 4.1). Fig. 4.4 *left* shows the cumulative SR, which is linearly decreasing from 100% to 95% across the 0–30 mm initial mTRE range and is then nearly constant at slightly above 95%.

The overall registration results on image pairs involving LAT projection views were similar to the results on image pairs involving AP projection views. The proposed method achieved registration accuracy with mean mRPD of  $1.03 \pm 0.47$  at SR of 75.2% and CR of 4 mm, while the respective results for the combined registration were  $0.40 \pm 0.18$  mm, 98.6% and 100 mm. Again, the iterative registration was not robust to high initial misalignment, as it had a almost twice lower SR. The registration accuracy per each dataset and overall improved significantly by the use of combined registration compared to the proposed registration with p-values  $\ll 0.05$ . Fig. 4.4 *left* shows the cumulative SR, which is around 99% across the 0–100 mm range of mTRE.

Clearly, the combined registration yields a very accurate and robust 3D-2D image registration in the range from 0–100 mm of initial mTRE. The proposed method seems very robust for initializing the 3D-2D image registration, based on which accurate alignment of 3D and 2D images can be achieved by the use of the iterative registration [75].

Table 4.2 reports computational times of the proposed method w.r.t. three resolution stages, generation of rotation and depth templates and registration execution time for 3D-2D registration to AP and LAT views. The mean registration times of the iterative registration were 1.80 s, hence, the overall mean times of the combined registration were around 12 seconds.

## 4.4 DISCUSSION

We proposed an intrinsic image-based registration method for rigid alignment of a 3D image to a monoplane 2D projection based on a novel stratified parameter optimization. The dimensional correspondence between 3D and 2D images is achieved through projection, while the method employs, as registration features, a small set of high-magnitude intensity gradients in 3D and 2D. In this way, only a small subset of 3D intensity gradients needs to be projected to 2D, which makes the projection very efficient. Further, we introduce a novel parameter optimization approach, which stratifies the 3D rigid-body motion estimation into two consecutive steps; first, the estimation of rigid-body rotations and out-of-plane translation and, second, the esti-

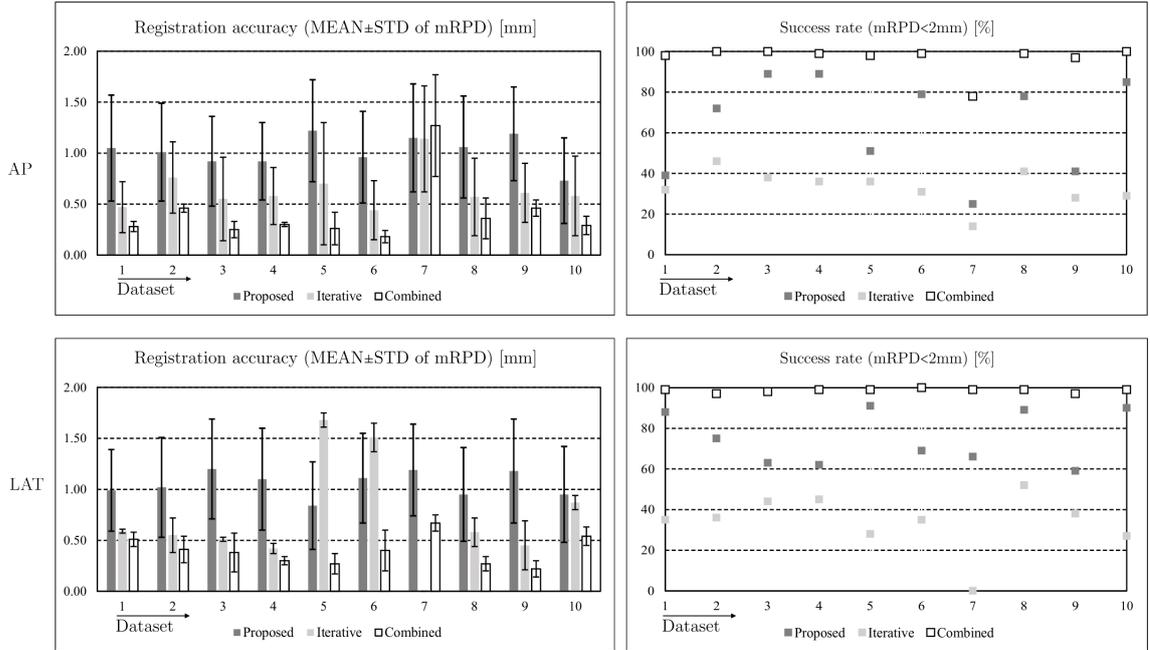


Figure 4.3: Results for the proposed, iterative and combined registration on each of 10 clinical image datasets of cerebral angiograms, reported for 3D- and 2D-DSA image pairs involving AP (*top*) and LAT (*bottom*) projection views. Registration accuracy as mRPD is shown on the *left* and success rate on the *right*.

mation of in-plane translations. The 3D rigid-body motion estimation is performed in a multi-resolution framework for a discrete set of image resolutions.

The purpose of the stratification of 3D rigid-body parameters into two subsets and their consecutive estimation was to increase the robustness of parameter estimation. The method employs two gradient-based similarity measures (SMs), both computed in the frequency domain. The first SM is the correlation coefficient between Fourier magnitudes (MC) of complex gradient maps, which is used to estimate the rigid-body rotations and out-of-plane translation. In the second step, in-plane translations are estimated by gradient phase correlation. By decoupling the estimation of 3D rigid-body parameters into two consecutive steps, the estimation of out-of-plane translation, which may otherwise hamper the joint estimation of 3D rigid-body parameters via state-of-the-art Similarity Measures [30, 20, 17, 31, 32, 6], was better posed. The main reason is that the employed MC is invariant to the in-plane translations, to which the state-of-the-art SMs are generally most sensitive.

Registration of 3D to monoplanar 2D projection is difficult because the lack of

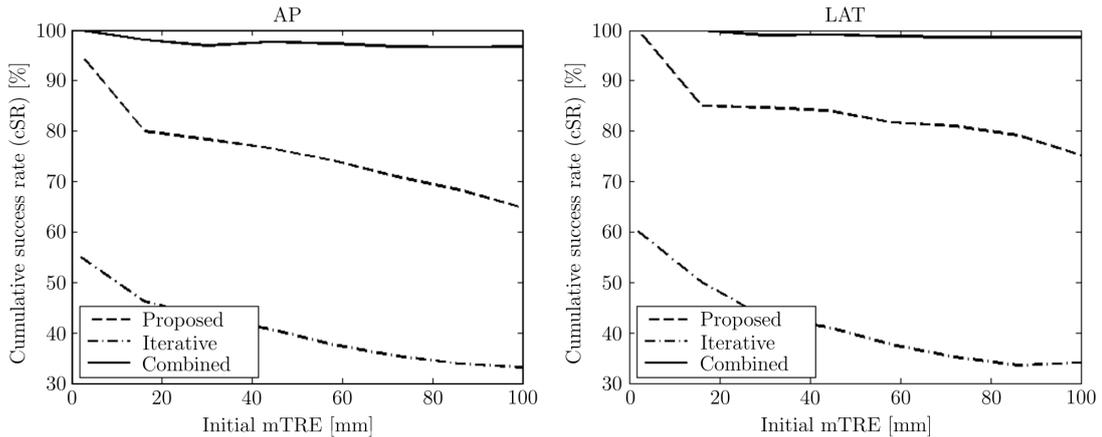


Figure 4.4: Cumulative success rate (cSR) as fraction of registration trials, which had  $mRPD < 2$  mm, with respect to initial mTRE. The cSR is computed over 10 clinical image datasets that involved registrations of 3D-DSA to 2D-DSA in either AP (*left*) or LAT (*right*) projection views.

depth information in the 2D image renders the estimation of out-of-plane translation ill-posed. The main reason is that most of the intensity-based image similarity measures have poor sensitivity with respect to the out-of-plane translation. However, the sensitivity of (any) similarity measure to out-of-plane translation is generally the highest when all rigid-body parameters, are close to their optimal values. Hence, the proposed method can be used to improve the performance of existing, state-of-the-art registration methods. This was validated by executing the iterative registration method [75] with or without prior rigid registration with the proposed method (Table 4.1).

Most of the state-of-the-art 3D-2D image registration methods, for which error rates were reported smaller than for the proposed method, require that the initial alignment of 3D and 2D images is very close to the optimal alignment. For this purpose, manual user interaction with the image-guidance system was generally required [19], which, however, is not suitable for routine clinical use. The comparison of the combined method (proposed and iterative) with the iterative method also showed that the stratified parameter optimization, which is based on a rather limited sets of rotation and depth templates, can be applied to overcome usually large initial translation errors encountered with the use of C-arm imaging systems [78].

The registration times of the proposed method compare favorably to state-of-

Table 4.1: The overall results for registrations on 10 patient datasets of cerebral angiograms [4] between 3D- and 2D-DSA in either AP or LAT projection views. Registration accuracy is reported as MEAN $\pm$ STD of mRPD of successful registrations (mRPD $<$ 2 mm), and success rates (SR) computed for registrations over 1000 registrations of 3D and 2D image pairs. Mann-Whitney-Wilcoxon non parametric test was used to evaluate the difference mRPD between of successful registration trials.

View	Method	MEAN $\pm$ STD [mm]	SR [%]
AP	Proposed	0.99 $\pm$ 0.48	64.8
	Iterative	0.62 $\pm$ 0.41	33
	<b>Combined</b>	<b>0.39 <math>\pm</math> 0.32<sup>†</sup></b>	<b>96.8</b>
LAT	Proposed	1.03 $\pm$ 0.47	75.2
	Iterative	0.74 $\pm$ 0.44	34
	<b>Combined</b>	<b>0.40 <math>\pm</math> 0.18<sup>†</sup></b>	<b>98.6</b>

<sup>†</sup> $p < 10^{-20}$

the-art GPU implementations of DRR generation by Otake et al. [65], who reported high DRR rates up to 22 000 s<sup>-1</sup>, however, the registration times were more than 6 seconds due to the use of global CMAES optimization methods, including multi-start and re-start strategies. Such a fast DRR generation was obtained by increasing ray step length to 3 voxels, while we used 0.5 voxels in our implementation of gradient projection. Increasing the step length is reasonable for MIP/DRR generation from images containing relatively large anatomical structures such as vertebrae considered in [65], however, it may completely obscure tiny structures such as vessels, which we considered in this work. Since vertebrae are much smaller in size compared to cerebral angiograms, given equal spatial image sampling, the clinical image of cerebral angiograms are much larger and, thus, require more computations for obtaining a projection image. Opposed to state-of-the-art methods that use general-purpose optimization techniques, which have variable number of iterations, and thus variable convergence times, due to varying initial misalignment and image quality conditions,

Table 4.2: Execution times (MEAN  $\pm$  STD) of the proposed method in each of three resolution stages for the generation of rotation and depth templates and registration execution time computed over all 10 datasets for registration to AP and LAT projection views.

Stage	Template generation		Registration times [s]
	Rotation [s]	Depth [s]	
AP-1	0.22 $\pm$ 0.08	0.08 $\pm$ 0.11	0.46 $\pm$ 0.07
AP-2	1.16 $\pm$ 0.25	0.14 $\pm$ 0.20	2.80 $\pm$ 0.40
AP-3	1.38 $\pm$ 0.39	0.14 $\pm$ 0.21	2.89 $\pm$ 0.42
LAT-1	0.18 $\pm$ 0.07	0.06 $\pm$ 0.15	0.36 $\pm$ 0.10
LAT-2	0.97 $\pm$ 0.36	0.11 $\pm$ 0.30	2.21 $\pm$ 0.61
LAT-3	1.21 $\pm$ 0.41	0.11 $\pm$ 0.32	2.29 $\pm$ 0.67

the proposed method has a fixed number of computations and thus provides image alignment in nearly constant time.

In conclusion, the main advantages of the proposed method are: (i) improved estimation of depth or out-of-plane translation; (ii) robustness against image noise and occlusions; (iii) robustness to high initial alignment errors; (iv) fixed number of computations. Furthermore, the proposed method does not depend on segmentation or extraction of features, thereby its performance is less dependent on the quality of 3D and, especially, 2D images. A subset of high-magnitude gradients can be easily extracted using heuristic techniques, while the projection of a subset of 3D image gradients into 2D is also very efficient compared to MIP/DRR generation. These characteristics increase the potential of the proposed method, and of the 3D-2D image registration technology in general, for a successful translation into routine use in clinical applications.

## Chapter 5

# Conclusions and Outlook

### 5.1 Summary and Conclusions

This thesis presents two novel methods for 3D-2D registration of vessels from a single view for Image Guided Interventions (IGI).

Chapter 3 proposes decoupling of rotation and translation in estimation of rigid transformation parameters. Rotation estimation is based on matching rotational DRR templates of segmented CTA to the segmented X-ray in the frequency domain. Translation is estimated by minimization of distances between two segmented vessels by an optimizer. An initial alignment is not needed since DRR templates and distance based optimization avoid convergence to local minima. The errors are measured by distances on the projection plane. The mean error is low, however because a gold standard is missing three dimensional misalignment distances cannot be measured. This metric is necessary for validation of IGI applications. The main strengths of this method are: robustness against initial alignment errors; noise; and occlusions in the image. The sources of error for this method are: its dependence on the segmentation success; the discrete estimation of rotation angles; and poor estimation of depth by the iterative optimization algorithm.

Chapter 4 builds on the method of Chapter 3, and adds the crucial idea of separation of depth estimation from in-plane translation, hence dividing the estimation of rigid parameters into three stages. A second difference is the replacement of the error prone segmentation by gradients of DSA images. In this way, the full structure of vessels in both modalities are preserved. Both gradients and frequency

domain methods provide robustness against partial occlusion and noise in the images. Scale space of the X-ray DSA images are compared to rotated template DRRs of Rotational DSA in Fourier domain to recover depth and rotation. The in-plane translation is recovered by Fourier Phase correlation. The discrete estimates of scale and rotation are interpolated by linear regression and linear interpolation respectively. The results are reported in mean Reprojection Distance (mRPD) metric. The results show low alignment errors and indicate that this method can be employed safely as an initialization for other local registration algorithms. Although mean mRPD of the proposed method is not high, the success rates can be further improved by a local optimizer or another stage in the registration process. The computation time compares favorably to other methods implemented on GPU therefore the proposed method is clinically feasible for IGIs. The main shortcoming of the method at present is relatively low success rates in mRPD metric. The main advantages of the method are: (i) improved depth estimation; (ii) robustness against noise and occlusions; (iii) robustness to high initial alignment offsets. Particularly, the latter advantage is very important. The state-of-the-art techniques in the literature, which obtained smaller error rates than the proposed method, all required very close initialization conditions, which was typically carried out interactively. Furthermore, the method involves fast and fixed number of computations, as opposed to optimization-based techniques in the literature, which have variable convergence times due to varying initial conditions and image quality.

## 5.2 Outlook

The future areas in 3D-2D registration from a single view that need improvement are the ability to function under changing conditions, large initial alignment errors. 3D-2D registration methods based on a single view, which are not affected by initial alignment errors and have high final success rates in three dimensional distance metrics are still open research targets. Some of the intrinsic methods reporting high success rates and low alignment errors require close initialization. Some of the methods with low final error rates employ more than one view which may not be realistic as monoplane interventions are still common. Another deficiency in this

area is poor generalization to other applications. Most intensity and feature based registration methods are designed for specific scenarios and may not be reliable if conditions and image characteristics change significantly. A single 3D-2D registration algorithm reliable enough to be employed in all IGI scenarios concerning an anatomy is still not available. Moreover some of the proposed methods require complex manual interactions from the surgeon during the intervention which may not be desirable under real-time constraints.

This thesis, particularly the second method proposed in Chapter 4 presented solutions to some of these issues: (i) automatic correction of large initialization errors; (ii) decreased sensitivity to image noise; (iii) fast and deterministic computation times.

Taking a template-based approach, as well as focusing on improvement of the depth (out-of-plane translation) estimation as proposed in this thesis could be the basis of future research directions in this field. Ideally a robust 3D-2D registration algorithm should not be sensitive to image noise and artifacts. In order for the method to perform reliably under variable circumstances, the errors must be low and predictable. Another bottleneck for IGI applications is complexity. The real time constraints require the algorithm complexities to be tractable and converge to optimal solution in few number of steps. Iterative optimization methods have the drawbacks of possibility of converging to local minima and uncertainty in number of steps required. A time-wise efficient algorithm should have a bounded worst case running time, which supports our hypothesis that template-based methods are more suitable to the solution of the 3D-2D registration problem.

# Bibliography

- [1] (2015). [Online]. Available: <http://www.uchospitals.edu/news/2005/20050401-64CT.html>
- [2] (2015). [Online]. Available: <http://www.healthcare.siemens.com/surgical-c-arms-and-navigation/hybrid-or/artis-zee-biplane>
- [3] M. Groher, “2d 3d registration of vascular images towards 3d-guided catheter interventions,” Ph.D. dissertation, Technical University of Munich, 2007.
- [4] U. Mitrovic, Z. Spiclin, B. Likar, and F. Pernus, “Method for 3d-2d registration of vascular images: Application to 3d contrast agent flow visualization,” in *Clinical Image-Based Procedures. From Planning to Intervention*, ser. Lecture Notes in Computer Science, K. Drechsler, M. Erdt, M. Linguraru, C. Oyarzun Laura, K. Sharma, R. Shekhar, and S. Wesarg, Eds. Springer Berlin Heidelberg, 2013, vol. 7761, pp. 50–58.
- [5] M. V. Wyawahare, D. P. M. Patil, and H. K. Abhyankar, “Image Registration Techniques: An overview,” *International Journal of Signal Processing, Image Processing and Pattern Recognition*, vol. 2, no. 3, 2009.
- [6] J. B. A. Maintz and M. A. Viergever, “A survey of medical image registration,” *Medical Image Analysis*, vol. 2, no. 1, pp. 1–36, 1998.
- [7] I. N. Bankman, Ed., *Handbook of Medical Image Processing and Analysis*, 2nd ed. Academic Press, 2008.
- [8] A. Sotiras, C. Davatzikos, and N. Paragios, “Deformable medical image registration: A survey,” *Medical Imaging, IEEE Transactions on*, vol. 32, no. 7, pp. 1153–1190, July 2013.

- [9] B. Zitova and J. Flusser, “Registration methods: a survey,” *Image and Vision Computing*, vol. 21, pp. 977–1000, 2003.
- [10] T. Lindeberg, “Scale Invariant Feature Transform,” vol. 7, no. 5, p. 10491, 2012.
- [11] H. Bay, T. Tuytelaars, and L. Van Gool, “Surf: Speeded up robust features,” in *Computer Vision ECCV 2006*, ser. Lecture Notes in Computer Science, A. Leonardis, H. Bischof, and A. Pinz, Eds. Springer Berlin Heidelberg, 2006, vol. 3951, pp. 404–417.
- [12] M. Nixon and A. S. Aguado, *Feature Extraction & Image Processing, Second Edition*, 2nd ed. Academic Press, 2008.
- [13] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006.
- [14] T. Makela, P. Clarysse, O. Sipila, N. Pauna, Q. C. Pham, T. Katila, and I. E. Magnin, “A Review of Cardiac Image Registration Methods,” *IEEE Trans. on Medical Imaging*, vol. 21, no. 9, pp. 1011 – 1021, 2002.
- [15] M.-K. Hu, “Visual pattern recognition by moment invariants,” *Information Theory, IRE Transactions on*, vol. 8, no. 2, pp. 179–187, February 1962.
- [16] J. A. M. H., M. Santos, and J. de Lope, “Orthogonal variant moments features in image analysis,” *Journal of Information Sciences*, vol. 180, pp. 846 – 860, 2010.
- [17] G. P. Penney, J. Weese, J. A. Little, P. Desmedt, D. L. G. Hill, and D. J. Hawkes, “A comparison of similarity measures for use in 2d 3d medical image registration,” *IEEE Trans. on Medical Imaging*, vol. 17, no. 4, pp. 586–595, 1998.
- [18] P. Markelj, D. Tomazevic, B. Likar, and F. Pernus, “A review of 3d-2d registration methods for image-guided interventions,” *Medical Image Analysis*, vol. 16, pp. 642–661, 2012.

- [19] A. Varnavas, T. Carrell, and G. Penney, “Increasing the automation of a 2d-3d registration system,” *Medical Imaging, IEEE Transactions on*, vol. 32, no. 2, pp. 387–399, 2013.
- [20] Y. Otake, M. Armand, R. S. Armiger, M. Kutzer, E. Basafa, P. Kazanzides, and R. Taylor, “Intraoperative image-based multiview 2d/3d registration for image-guided orthopaedic surgery: Incorporation of fiducial-based c-arm tracking and gpu-acceleration,” *Medical Imaging, IEEE Transactions on*, vol. 31, no. 4, pp. 948–962, April 2012.
- [21] C. Metz, M. Schaap, S. Klein, P. Rijnbeek, L. Neefjes, N. Mollet, C. Schultz, P. Serruys, W. Niessen, and T. van Walsum, “Alignment of 4d coronary cta with monoplane x-ray angiography,” in *6th International Workshop, AE-CAI 2011, Held in Conjunction with MICCAI 2011*, 2011, pp. 106–116.
- [22] D. Rivest-Henault, H. Sundar, and M. Cheriet, “Nonrigid 2d/3d registration of coronary artery models with live fluoroscopy for guidance of cardiac interventions,” *Medical Imaging, IEEE Transactions on*, vol. 31, no. 8, pp. 1557–1572, Aug 2012.
- [23] G. A. Turgeon, G. G. Glen Lehmann, M. Drangova, D. Holdsworth, and T. Peters, “2D-3D registration of coronary angiograms for cardiac procedure planning and guidance,” *Medical Physics*, vol. 32, no. 12, pp. 3737 – 3749, 2005.
- [24] D. Ruijters, B. M. ter Haar Romeny, and P. Suetens, “Vesselness-based 2d-3d registration of the coronary arteries,” *International Journal of Computer Assisted Radiology and Surgery*, vol. 4, no. 4, pp. 391 – 397, 2009.
- [25] M. Groher, N. Padoy, T. F. Jakobs, and N. Navab, “New cta protocol and 2d-3d registration method for liver catheterization,” in *Proc. Medical Image Computing and Computer Assisted Interventions*, 2006, pp. 873 – 881.
- [26] M. Groher, D. Zikic, and N. Navab, “Deformable 2d-3d registration of vascular structures in a one view scenario,” *IEEE Trans. on Medical Imaging*, vol. 28, pp. 847–860, 2009.

- [27] N. Baka, C. Metz, C. Schultz, L. Neefjes, R. van Geuns, B. Lelieveldt, W. Niessen, T. van Walsum, and M. de Bruijne, “Statistical coronary motion models for 2d+t/3d registration of x-ray coronary angiography and cta,” *Medical Image Analysis*, vol. 17, no. 6, pp. 698 – 709, 2013.
- [28] C. Metz, M. Schaap, S. Klein, N. Baka, L. Neefjes, J. Schultz, W. Niessen, and T. van Walsum, “Registration of 3d+t coronary cta and monoplane 2d+t x-ray angiography,” *Medical Imaging, IEEE Transactions on*, vol. 32, no. 5, pp. 919–931, May 2013.
- [29] N. Baka, C. Metz, C. Schultz, R.-J. van Geuns, W. Niessen, and T. van Walsum, “Oriented gaussian mixture models for nonrigid 2d/3d coronary artery registration,” *Medical Imaging, IEEE Transactions on*, vol. 33, no. 5, pp. 1023–1034, May 2014.
- [30] L. Zollei, E. Grimson, A. Norbash, and W. Wells, “2D-3D rigid registration of X-ray fluoroscopy and CT images using mutual information and sparsely sampled histogram estimators,” in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*. IEEE Computer Society, 2001.
- [31] L. M. G. Brown and T. E. Boult, “Registration of planar film radiographs with computed tomography,” in *Workshop on Mathematical Methods in Biomedical Image Analysis*. IEEE Computer Society, 1996, pp. 42–51.
- [32] S. Demirci, M. Baust, O. Kutter, F. Manstad-Hulaas, H.-H. Eckstein, and N. Navab, “Disocclusion-based 2d-3d registration for angiographic interventions,” *Computers in Biology and Medicine*, vol. 43, pp. 312–322, 2013.
- [33] S. Dong, J. Kettenbach, I. Hinterleitner, H. Bergmann, and W. Birkfellner, “The Zernike Expansion An Example of a Merit Function for 2D-3D Registration Based on Orthogonal Functions,” in *Proc. Medical Image Computing and Computer Assisted Interventions*, 2008, pp. 964–971.
- [34] H. Livyatan, Z. Yaniv, and L. Joskowicz, “Gradient-based 2-d/3-d rigid registration of fluoroscopic x-ray to CT,” *IEEE Trans. on Medical Imaging*, vol. 22, no. 11, pp. 1395–1406, 2003.

- [35] D. Tomaevic, B. Likar, T. Slivnik, and F. Pernu, “3-d/2-d registration of CT and MR to x-ray images,” *IEEE Trans. on Medical Imaging*, vol. 22, no. 11, pp. 1407–1416, 2003.
- [36] P. Markelj, D. Tomazevic, F. Pernu, and B. Likar, “Robust gradient-based 3-d/2-d registration of CT and MR to x-ray images,” *IEEE Trans. on Medical Imaging*, vol. 27, no. 12, pp. 1704–1714, 2008.
- [37] D. Tomazevic, B. Likar, and F. Pernus, “3D-2D image registration: The impact of Xray views and their number,” *Proc. Medical Image Computing and Computer Assisted Interventions*, vol. 4791, pp. 450–457, 2007.
- [38] M. Prummer, J. Hornegger, M. Pfister, and A. Dorfler, “Multi-modal 2d3d nonrigid registration,” in *Medical Imaging 2006: Image Processing, SPIE*, vol. 6144, 2006.
- [39] E. Serradell, A. Romero, R. Leta, C. Gatta, and F. Moreno-Noguer, “Simultaneous correspondence and non-rigid 3d reconstruction of the coronary tree from single x-ray images,” in *Computer Vision (ICCV), 2011 IEEE International Conference on*, Nov 2011, pp. 850–857.
- [40] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, “Multiscale vessel enhancement filtering,” in *Proc. Medical Image Computing and Computer Assisted Interventions*, 1998, pp. 130–137.
- [41] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006.
- [42] J. Kremer, “Recovering deformable 3d coronary arteries from a monoplane angiographic sequence,” Master’s thesis, Technical University of Munich, 2013.
- [43] M. McGuire, “An image registration technique for recovering rotation, scale and translation parameters,” *NEC Res. Inst. Tech. Rep., TR*, pp. 98–018, 1998.
- [44] G. Wolberg and S. Zokai, “Robust image registration using log polar transform,” in *Proc. IEEE Int. Conf. on Image Processing*, 2000.

- [45] J. Nocedal and S. Wright, *Numerical Optimization, 2nd edition*, ser. Springer Series in Operations Research. Springer, 2006.
- [46] G. Tzimiropoulos, V. Argyriou, S. Zafeiriou, and T. Stathaki, “Robust fft-based scale-invariant image registration with image gradients,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 10, pp. 1899–1906, Oct 2010.
- [47] J. N. Sarvaiya, S. Patnaik, and K. Kothari, “Image registration using log polar transform and phase correlation to recover higher scale,” *Journal of Pattern Recognition Research*, vol. 7, pp. 90 – 105, 2012.
- [48] Redwood, Curzen, and Thomas, *Oxford Textbook of Interventional Cardiology*. Oxford, UK: Oxford University Press.
- [49] J. Prince and J. Links, *Medical Imaging Signals and Systems*. Pearson Education, Limited, 2014. [Online]. Available: <http://books.google.com.tr/books?id=HHrangEACAAJ>
- [50] S. Demirci, “New approaches to computer assistance for endovascular aortic repairs,” Ph.D. dissertation, Technical University of Munich, 2011.
- [51] N. Strobel, O. Meissner, J. Boese, T. Brunner, B. Heigl, M. Hoheisel, G. Lauritsch, M. Nagel, M. Pfister, E.-P. Rhrnschopf, B. Scholz, B. Schreiber, M. Spahn, M. Zellerhoff, and K. Klingenbeck-Regn, “3d imaging with flat-detector c-arm systems,” in *Multislice CT*, ser. Medical Radiology, M. F. Reiser, C. Becker, K. Nikolaou, and G. Glazer, Eds. Springer Berlin Heidelberg, 2009, pp. 33–51.
- [52] S. Cetin, A. Demir, A. J. Yezzi, M. Degertekin, and G. Unal, “Vessel tractography using an intensity based tensor model with branch detection,” in *IEEE Trans. on Medical Imaging*, 2013, pp. 348–363.
- [53] J. Weickert, *Anisotropic Diffusion in Image Processing*. B. G. Teubner (Stuttgart), 2008.
- [54] R. Hartley and A. Zisserman, *Multiple View Geometry*. Cambridge University Press, 2003.

- [55] Nokia, “Qt development framework 4.8.0,” 2012.
- [56] Kitware, “Visualization toolkit 5.8.0,” 2011.
- [57] ———, “Insight segmentation and registration toolkit 4.1.0,” 2012.
- [58] J. N. Naoaki Okazaki, “liblbfgs: a library of l-bfgs,” 2010.
- [59] E. B. van de Kraats, G. P. Penney, D. Tomazevic, T. van Walsum, and W. J. Niessen, “Standardized evaluation methodology for 2-d3-d registration,” *IEEE Trans. on Medical Imaging*, vol. 24, no. 9, pp. 1177 – 1189, 2005.
- [60] S. A. Pawiro, P. Markelj, F. Pernus, C. Gendrin, M. Figl, C. Weber, F. Kainberger, I. Nobauer-Huhmann, H. Bergmeister, M. Stock, D. Georg, H. Bergmann, and W. Birkfellner, “Validation for 2d/3d registration i: A new gold standard data set,” *Medical Physics*, vol. 38, no. 3, pp. 1481–1490, 2011.
- [61] D. Knaan and L. Joskowicz, “Effective intensity-based 2d-3d rigid registration between fluoroscopic x-ray and ct,” in *Proc. Medical Image Computing and Computer Assisted Interventions*, 2003, pp. 351 – 358.
- [62] D. Tomazevic, B. Likar, and F. Pernus, “Gradient-based registration of 3d mr and 2d x-ray images,” in *International Congress Series 1230*, 2001, pp. 338 – 345.
- [63] P. Markelj, B. Likar, and F. Pernus, “Standardized evaluation methodology for 3d/2d registration based on the visible human data set,” *Medical Physics*, vol. 37, no. 9, pp. 4643–4647, 2010.
- [64] T. Aksoy, G. Unal, S. Demirci, N. Navab, and M. Degertekin, “Template-based cta to x-ray angio rigid registration of coronary arteries in frequency domain with automatic x-ray segmentation,” *Medical Physics*, vol. 40, no. 10, 2013.
- [65] Y. Otake, A. S. Wang, J. W. Stayman, A. Uneri, G. Kleinszig, S. Vogt, A. J. Khanna, Z. L. Gokaslan, and J. H. Siewerdsen, “Robust 3d2d image registration: application to spine interventions and vertebral labeling in the presence of anatomical deformation,” *Physics in Medicine and Biology*, vol. 58, no. 23, p. 8535, 2013.

- [66] M. J. van der Bom, S. Klein, M. Staring, R. Homan, L. W. Bartels, and J. P. W. Pluim, “Evaluation of optimization methods for intensity-based 2d-3d registration in x-ray guided interventions,” in *SPIE Medical Imaging 2011: Image Processing*, vol. 7962, Orlando, FL, USA, 2011, pp. 796 223–796 223–15.
- [67] C. Florin, J. Williams, A. Khamene, and N. Paragios, “Registration of 3d angiographic and x-ray images using sequential monte carlo sampling,” in *Computer Vision for Biomedical Image Applications*, ser. Lecture Notes in Computer Science, Y. Liu, T. Jiang, and C. Zhang, Eds. Springer Berlin Heidelberg, 2005, vol. 3765, pp. 427–436.
- [68] D. Zikic, B. Glocker, O. Kutter, M. Groher, N. Komodakis, A. Kamen, N. Paragios, and N. Navab, “Linear intensity-based image registration by markov random fields and discrete optimization,” *Medical Image Analysis*, vol. 14, no. 4, pp. 550–562, 2010.
- [69] E. Kerrien, M. Berger, E. Maurincomme, L. Launay, R. Vaillant, and L. Picard, “Fully automatic 3d-2d subtracted angiography registration,” in *Proceedings of the Second International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 1679, 1999, pp. 664 –671.
- [70] C. M. Hentschke and K. D. Tönnies, “Automatic 2d/3d-registration of cerebral dsa data sets.” in *Bildverarbeitung für die Medizin*. Citeseer, 2010, pp. 162–166.
- [71] Y. Kita, D. L. Wilson, and A. Noble, “Real-time registration of 3D cerebral vessels to x-ray angiograms,” in *Medical Image Computing and Computer-Assisted Intervention - MICCAI 1998*. Berlin, Heidelberg: Springer, 1998, vol. 1496, pp. 1125–1133.
- [72] A. Kubias, F. Deinzer, T. Feldmann, and D. Paulus, “Extended global optimization strategy for rigid 2d/3d image registration,” in *Computer Analysis of Images and Patterns*, ser. Lecture Notes in Computer Science, W. Kropatsch, M. Kampel, and A. Hanbury, Eds. Springer Berlin Heidelberg, 2007, vol. 4673, pp. 759–767.
- [73] M. J. van der Bom, L. W. Bartels, M. J. Gounis, R. Homan, J. Timmer, M. A. Viergever, and J. P. W. Pluim, “Robust initialization of 2d-3d image

- registration using the projection-slice theorem and phase correlation,” *Medical physics*, vol. 37, no. 4, pp. 1884–1892, 2010.
- [74] G. Shechter, B. Shechter, J. Resar, and R. Beyar, “Prospective motion correction of x-ray images for coronary interventions,” *Medical Imaging, IEEE Transactions on*, vol. 24, no. 4, pp. 441–450, April 2005.
- [75] Z. Špiclin, B. Likar, and F. Pernuš, “Fast and robust 3D to 2D image registration by backprojection of gradient covariances,” in *Biomedical Image Registration*, ser. Lecture Notes in Computer Science, S. Ourselin and M. Modat, Eds. Springer International Publishing, 2014, no. 8545, pp. 124–133.
- [76] J. H. Siewerdsen, D. J. Moseley, S. Burch, S. K. Bisland, A. Bogaards, B. C. Wilson, and D. A. Jaffray, “Volume CT with a flat-panel detector on a mobile, isocentric c-arm: pre-clinical investigation in guidance of minimally invasive surgery,” *Medical Physics*, vol. 32, no. 1, pp. 241–254, 2005.
- [77] M. J. Daly, J. H. Siewerdsen, Y. B. Cho, D. A. Jaffray, and J. C. Irish, “Geometric calibration of a mobile c-arm for intraoperative cone-beam CT,” *Medical Physics*, vol. 35, no. 5, pp. 2124–2136, 2008.
- [78] G. Warmerdam, P. Steininger, M. Neuner, G. Sharp, and B. Winey, “Influence of imaging source and panel position uncertainties on the accuracy of 2d/3d image registration of cranial images,” *Medical Physics*, vol. 39, no. 9, pp. 5547–5556, 2012.