



**Universitat
Autònoma
de Barcelona**

3D Reconstruction of the Coronary Tree using Biplane Snakes

A dissertation submitted by **Cristina Cañero
Morales** at Universitat Autònoma de Barcelona
to fulfil the degree of **Doctor en Informàtica**.

Bellaterra, November 11, 2002

Director: **Dra. Petia Radeva Ivanova**
Universitat Autònoma de Barcelona
Dept. Informàtica & Computer Vision Center



This document was typeset by the author using L^AT_EX 2_ε.

The research described in this book was carried out at the Computer Vision Center, Universitat Autònoma de Barcelona.

Copyright © 2002 by Cristina Cañero Morales. All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the author.

ISBN 84-932156-4-3

Printed by Ediciones Gráficas Rey, S.L.

a Dani

Amb el suport del Departament de Universitats, Recerca
i Societat de la Informació de la Generalitat de Catalunya.
Beca 1999FI 00753 APTIND CVC.

Agradecimientos

Hacer una tesis doctoral es una aventura repleta de obstáculos, tanto externos como internos. Externos, porque, por lo menos para mí, una tesis requiere dedicación casi exclusiva, y eso es difícil de conseguir. Internos, porque también requiere constancia, tenacidad y sobretodo un gran esfuerzo para confiar en uno mismo. Creo que sin toda la ayuda que he recibido no hubiera podido superar estos obstáculos, y ahora no estaría escribiendo estas líneas de agradecimiento. Las escribo en un día lluvioso, así que perdonadme la cursilería que pudiera perpetrar.

La principal responsable de que haya terminado esta tesis es mi directora, Petia Radeva. A ella no sólo le agradezco que me ayudase a conseguir una beca predoctoral, sin la cual no hubiera podido dedicarme a la tesis tanto como lo he hecho, sino que también le agradezco todo su apoyo, tanto científico como moral. Gracias Petia, no sólo he aprendido mucho de ti, además me has servido de guía a la hora de orientar mi carrera científica, que espero sólo acabe de comenzar.

Por supuesto, no habría empezado esta tesis en el Centro de Visión por Computador si Juanjo Villanueva no hubiera "visto" algo en mí para invitarme a unirme a su grupo de investigación. Le estoy muy agradecida por ello, pero también por varias conversaciones de café que tuvimos y que me fueron muy útiles.

¡Y cómo iba a olvidarme del culpable de que me metiese en este lío que es una tesis doctoral, Jaime López-Krahe! Mi estancia con él en París fue determinante para que me decidiese a orientar mi carrera hacia la ciencia. Además, hizo que me sintiera muy acogida en un país distinto y un grupo de gente nueva, y eso te lo agradezco sinceramente.

Esta tesis se ha realizado utilizando imágenes reales de un hospital, y eso quiere decir que varias personas han dedicado parte de su precioso tiempo a tomar imágenes. Gracias Fina, Eduard, Oriol y el resto del equipo de la unidad coronaria del Hospital Universitari "Germans Trias i Pujol" por vuestra buena disposición, tan difícil de encontrar. Pero también, gracias a todos los pacientes que se prestaron a aguantar la respiración y a estarse lo más quietos posible mientras se adquirían las imágenes. Esta tesis también va dedicada a ellos.

Mis agradecimientos también van dirigidos a mis "compañeros de viaje", aquéllos que empezaron la tesis conmigo y con los que he compartido los buenos y los malos momentos: Oriol, Débora, Juanma, Xevi, Poal y Marc. Otros compañeros, como David Guillamet, Fernando, Marco, Ramón Baldrich, Josep Lladós, Xavi Varona, Albert Pujol, Sergio, Xavier Otazu, Misael, Jaume Amores, Ernest Estruga, David Rotger, Felipe Lumbreras, Andrés Solé y, muy especialmente, Robert Benavente,

merecen también mi agradecimiento. Por sonreír los lunes, por hablar los martes, por saludar los miércoles, por ayudarme los jueves, por reír conmigo los viernes y por acompañarme los fines de semana de trabajo.

La primera sonrisa de la mañana al entrar en el centro me la regalasteis vosotros: Ana Celia, Mari Carmen y Pedro. También fuisteis siempre amables y atentos conmigo, así como lo fueron Montse, Pilar, Raquel i Joan Masoliver. Gracias por vuestra amistad, y, cómo no, por resolverme muchos problemas. Gracias también a Loli, por ser siempre tan amable y cariñosa conmigo, y por algunos cafés que compartimos a las cinco de la mañana.

No me puedo olvidar de grandes amigas, a las que veo poco, pero que siempre me perdonan, y con las que he pasado, paso y espero seguir pasando muy buenos momentos. Muchas gracias a todas. A Menchu, "la fée clochettes", por ser tan especial y dulce ¡Qué suerte tiene los ingleses de tenerte allí! A Ana Serrano, por ser tan creativa y divertida. Espero que por fin logres encontrar tu camino en el mundo de la comunicación audiovisual. A Manoli, mi psicóloga particular, que siempre me apoya y me anima, aunque sólo sea porque tomemos un café juntas. A Antonia, por ser siempre capaz sorprenderme, y de reírse y hacer reír a los demás de todas las cosas que nos depara la vida. Y a Anuska, compañera de carrera y amiga, con la que he compartido "pecera" muchas veces, pero también grandes tertulias sobre la vida.

Mi más cariñoso y sincero agradecimiento va a mis padres, a mi hermana y a mi familia más próxima, que siempre han sido mis mayores fans, y siempre consiguen animarme. Sobretudo a mis padres, porque desde pequeña me habéis apoyado siempre en mis decisiones, y habéis confiado muchísimo más que yo en mí. De vosotros he aprendido el valor del esfuerzo y del trabajo, del afán de auto superación, y lo importante que es que cada uno trace su propio camino. Y todo esto lo aprendí siguiendo vuestro ejemplo. Además, también he recibido mucho amor y cariño, lo que me ha convertido en la persona que soy. Por todo ello, muchas gracias, os quiero mucho.

Finalmente, me gustaría agradecer a mi compañero, Dani, todos los momentos que hemos pasado juntos hasta ahora. Por tu compañía, por tu cariño, por tu amor, porque siempre sabes cuándo necesito un abrazo. Pero también por tus consejos y tu ayuda en el terreno científico. He tenido en ti un gran apoyo en todo lo que he hecho, o me he propuesto hacer. Incluso cuando esto sea motivo de que nos dejemos de ver por algún tiempo, siempre me has ayudado a perseguir mis sueños. T'estimo i espero que continuem caminant junts endavant!

Cristina Cañero

Abstract

This thesis explores all problems involved in the 3D reconstruction of coronary vessels from X-ray angiographies: calibration, vessel extraction in X-ray images and 3D reconstruction of the vessel in space.

We divide the calibration into two procedures: The first procedure is devoted to the correction of the geometrical distortion, and the second one is focused on the estimation of the extrinsic and intrinsic parameters of the X-ray acquisition system. We correct the **geometrical distortion** introduced by the image intensifier by defining a novel polynomial model for the distortion. The main advantage of our approach is that the distortion can be corrected for any view, and thus eliminates the need of distortion calibration for each acquisition. Regarding the estimation of the **acquisition geometry**, we show that the assumptions made by other researchers to estimate the extrinsic parameters are only valid in very restricted cases and thus introduce high degree of error. Therefore, we propose a new, more general and flexible model to calibrate the parameters of the C-arm. Using our proposed model, we can obtain a highly accurate estimation of the acquisition geometry.

Once calibrated, next steps of the thesis refer to vessel detection in images and 3D reconstruction. Due to the low signal-to-noise ratio of X-ray angiographies, the **vessel segmentation** procedure is not a trivial procedure. We develop a new anisotropic diffusion filter that enhances the vessels on the X-ray images in order to improve the vessel segmentation result. The proposed filter automatically chooses the scale of the diffusion tensor to be applied for each pixel, and achieves selective enhancement of the vessels enhancing the contrast-to-noise ratio. Given the detected vessel points in the angiography, 3D reconstruction is applied applying techniques similar to stereo vision avoiding the problem of determining the **corresponding points** of the object in both images. We define the biplane snakes that are deformable models consisting on a three-dimensional elastic curve that deforms in space to adapt its projections to the vessels in the images. The proposed approach inherently solves the point correspondence problem and has a good behaviour when the acquisition conditions are known with low precision.

The proposed methods are extensively validated with imaged phantoms and real angiographic sequences acquired with a monoplane angiographic system and the valuable help of the physicians of Hospital Universitari "Germans Trias i Pujol".

Resumen

Esta tesis explora los problemas de la reconstrucción 3D de los vasos coronarios a partir de angiografías: calibración, extracción de los vasos a partir de las imágenes, y reconstrucción 3D del vaso.

La calibración se divide en dos procedimientos: El primer procedimiento corrige la distorsión geométrica, y el segundo se concentra en la estimación de los parámetros extrínsecos e intrínsecos del sistema de adquisición. La **distorsión geométrica** introducida por el intensificador de imagen se corrige mediante la definición de un nuevo modelo de polinomios para la distorsión. La principal ventaja de esta aproximación es que la distorsión se puede corregir para cualquier vista, eliminando así la necesidad de calibrar para cada adquisición. Respecto la estimación de la **geometría de adquisición**, mostramos que las asunciones hechas por otros investigadores a la hora de estimar los parámetros extrínsecos sólo son válidas en casos muy restringidos, y por tanto introducen un alto grado de error. Por ello, proponemos un modelo nuevo, más general y flexible, para calibrar los parámetros del brazo C. Gracias a este modelo, podemos obtener una estimación precisa de la geometría de adquisición.

Después de calibrar, los siguientes pasos de esta tesis se refieren a la detección de los vasos y a la reconstrucción 3D. Debido a la baja relación señal-ruido de las imágenes de angiografías, la **segmentación de los vasos** no es un problema trivial de resolver. Nosotros desarrollamos un filtro de difusión anisotrópica para mejorar el resultado de la segmentación de los vasos. Este filtro escoge automáticamente la escala del tensor de difusión para cada píxel y consigue un realce selectivo de los vasos. A partir de los puntos detectados del vaso, la reconstrucción 3D se consigue aplicando técnicas parecidas a las de la visión estéreo, pero evitando el problema de la determinación de las **correspondencias de puntos** del objeto en las dos imágenes. Para ello, definimos los snakes biplanos, que son modelos deformables consistentes en una curva tridimensional que se deforma en el espacio para adaptar sus proyecciones a los vasos en las imágenes. Esta técnica resuelve el problema de las correspondencias de puntos de forma inherente, y tiene un buen comportamiento cuando las condiciones de adquisición no se conocen con gran precisión.

Los métodos propuestos son extensamente validados usando *phantoms* y secuencias reales de angiografías adquiridas con un sistema monoplaneo de angiografías y la inestimable ayuda de los médicos del Hospital Universitari "Germans Trias i Pujol".

Contents

1	Introduction	1
1.1	Improvements using Computers	1
1.2	Coronary Angiography	3
1.3	3D Reconstruction from Coronary Angiography	6
1.3.1	Geometrical Distortion Correction	6
1.3.2	Determination of the Acquisition Geometry	8
1.3.3	Automatic Vessel Detection	11
1.3.4	Corresponding Points Determination	13
1.4	Previous Publications	13
1.5	Thesis Organization	14
2	Background	15
2.1	X-Ray Angiography	15
2.1.1	X-Rays History	15
2.1.2	The Development of Cardiac Angiography	17
2.1.3	The Cardiac Angiography Acquisition System	18
2.2	3D Reconstruction	23
2.2.1	Cameras and Angiographers are very Similar	23
2.2.2	Three-Dimensional Reconstruction of a Point	25
2.2.3	Reconstruction error	27
2.2.4	Camera Calibration using the Zhang's Method	27
2.2.5	Determination of Corresponding Points	31
2.2.6	Computation of the Epipolar Line	31
2.3	Image Enhancement Techniques	34
2.3.1	Anisotropic Diffusion	34
2.3.2	Frangi's Filter	35
2.4	Deformable Models	36
2.4.1	Deformable Models Overview	37
2.4.2	Parametric Deformable Models	37
2.4.3	Internal Energy	39
2.4.4	External Energy	40
2.4.5	Numerical Representation of the Snake	44
3	Calibration	49

3.1	Distortion Correction	50
3.1.1	Geometrical Distortion Estimation	51
3.1.2	Modelling Geometrical Distortion	54
3.2	Extrinsic Parameters Calibration	57
3.2.1	Using Zhang's Method in the Angiographic Frame.	57
3.2.2	Modelling the Movement of the C-arm	58
3.2.3	Calibrating the C-arm's Movement Model	63
4	3D Reconstruction using Biplane Snakes	69
4.1	X-Ray Angiography Acquisition	69
4.2	Vessel Detection	71
4.2.1	Vessel Enhancement Diffusion	71
4.3	Biplane Snakes	74
4.3.1	Biplane Snakes Initialization	74
4.3.2	Biplane Snakes Deformation	77
5	Experimental results	81
5.1	Geometrical (Un)Distortion Model	82
5.1.1	Polynomial Suitability to Model Geometrical Distortion	83
5.1.2	Focal Variation Influence	85
5.1.3	Optimal Polynomial Degree for $\delta_{\alpha,\beta}^u(u, v)$, $\delta_{\alpha,\beta}^v(u, v)$, $\delta_{\alpha,\beta}^{\hat{u}}(\hat{u}, \hat{v})$ and $\delta_{\alpha,\beta}^{\hat{v}}(\hat{u}, \hat{v})$	86
5.1.4	Determination of the Optimal Polynomial Degree for $\{\omega^{u,i}(\alpha, \beta)\}$, $\{\omega^{v,i}(\alpha, \beta)\}$, $\{\omega^{\hat{u},i}(\alpha, \beta)\}$, and $\{\omega^{\hat{v},i}(\alpha, \beta)\}$	88
5.2	Model for the Acquisition Geometry	93
5.2.1	Zhang's Calibration on a SIEMENS C-Arm	93
5.2.2	Accuracy Evaluation for M0 , M1 , M2 , M3 and M4	94
5.3	Biplane Snakes	95
5.3.1	Experiments with Synthetic Images	96
5.3.2	Tests on Real Images	99
5.4	Vessel Enhancement Diffusion	106
5.4.1	Quantitative Evaluation on Synthetic Images	108
5.4.2	Tests on Real Images	113
6	Conclusion	117
6.1	Geometrical Distortion	117
6.2	Determination of the Acquisition Geometry	118
6.3	Automatic Vessel Detection	119
6.4	Automatic Corresponding Points Determination	119
6.5	Applications	120
6.6	Suggesting New Directions of Study	121
A	Matrix Representation for Deformable Models	123
A.1	One-Dimensional Deformable Models (B-Snakes)	123
A.1.1	B-Spline Inner product	124
A.1.2	<i>Stiffness</i> Matrix	124

A.1.3	Energy-Minimization Algorithm	126
A.2	Two-Dimensional Deformable Models (B-Sheets)	126
A.2.1	Energy Minimization Algorithm	127
B	Crossing points detection	129
B.1	Accuracy on Crossing Points Detection	130
C	Bivariate polynomial fitting	133
	Bibliography	135
	Publications	143
	Curriculum Vitae	147

List of Tables

5.1	Error of the predicted intersection points for different models	95
5.2	Error of the predicted 3D curve for different models	101

List of Figures

1.1	An IVUS and Angiogram fusion application. The user can know for each IVUS image the corresponding location of the catheter on the angiographic image.	3
1.2	Coronary Angiography	4
1.3	Perspective Effect	4
1.4	Foreshortening Effect	5
1.5	IVUS image	5
1.6	IVUS provide local cross-sectional information of the vessel.	6
2.1	W.C. Roentgen	16
2.2	The first radiograph	16
2.3	X-ray generation	17
2.4	First angiography	17
2.5	Angiographer	19
2.6	Image intensifier	20
2.7	Geometrical distortion	20
2.8	Rotation, denoted by α	21
2.9	Angulation, denoted by β	22
2.10	Varying the SID. The X-ray source does not move.	22
2.11	Angiographer vs Camera	24
2.12	3D Reconstruction	26
2.13	The world coordinate system $\{O, X, Y, Z\}$ is defined so that the calibration object is placed on $Z = 0$	28
2.14	Epipolar constraint	31
2.15	Epipolar lines for automatic correspondence determination.	33
2.16	Epipolar line on X-ray angiography.	33
2.17	<i>Thin-plate</i> energy discourage the blending of the shape.	39
2.18	Two approaches to obtain a potential: Gaussian convolution and distance map.	42
2.19	Distance map gradient and GGVF	44
2.20	B-Splines have local control.	45
2.21	Obtained B-Spline interpolation with degree $d = 3$ for different knot multiplicities (μ) at point A	45

2.22	B-spline blending functions for $d = \{1, 2, 3\}$. The open curve begins and ends at the extrema control points.	47
3.1	Anatomical angles	50
3.2	Calibration grid placed against the II screen	51
3.3	Relationship between grid coordinates and ideal image coordinates when the grid is attached to the image intensifier screen.	52
3.4	Dumay's Model	59
3.5	Model M1	61
3.6	Non intersecting rotation axes.	62
4.1	The ECG peaks correspond to diastole state of the heart.	70
4.2	Multiscale dominant orientation computation	73
4.3	Biplane Snake adaptation	74
4.4	The usefulness of additional initialization points	75
4.5	External forces for the biplane snake	78
5.1	Error vs polynomial degree of the model	84
5.2	Evolution of the detected crossing points of a grid fixed to the Image Intensifier screen while varying the focal distance.	85
5.3	Focal variation influence	87
5.4	Distortion influence color-map	89
5.5	Error vs polynomial degree of the orientation-dependent model	90
5.6	Error vs polynomial degree of the predictive model	92
5.7	Anatomical angles	94
5.8	Mean reconstruction error for computational phantom.	96
5.9	Concave shapes capture performance	98
5.10	Distortion influence on 3D reconstruction	100
5.11	3D reconstruction of a wire phantom from two views.	101
5.12	Reconstructed 3D curve on different views	102
5.13	3D reconstruction of a phantom with contrast	104
5.14	3D reconstruction of two wires	105
5.15	Vessel centerline features	106
5.16	3D reconstruction on real cardiac images	107
5.17	Vessel model used to create the synthetic images.	109
5.18	Synthetic Vessels	109
5.19	Contrast-to-Noise Ratio computed on the gray-scale image for different values η of noise strength (isolated vessels).	110
5.20	Contrast-to-Noise Ratio computed on the vesselness image for different values η of noise strength (isolated vessels).	111
5.21	Contrast-to-Noise Ratio computed on the gray-scale image for different values η of noise strength (bifurcated vessels).	112
5.22	Contrast-to-Noise Ratio computed on the vesselness image for different values η of noise strength (bifurcated vessels).	112
5.23	Contrast-to-Noise Ratio computed on the gray-scale image for different values η of noise strength (bifurcated vessels + background structures).	113

5.24 Contrast-to-Noise Ratio computed on the vesselness image for different values η of noise strength (bifurcated vessels + background structures). 114

5.25 Results on a ROI of one of the images. 115

5.26 Results on a ROI of another of the images. 115

B.1 Crossing points detection 130

B.2 Function fitted to the projection of the intersection areas 131

B.3 Crossing points detection error for each method 131

Chapter 1

Introduction

The rapid development and proliferation of medical imaging technologies is revolutionizing medicine. Medical imaging allows scientists and physicians to glean potentially lifesaving information by peering into the human body. The role of medical imaging has expanded beyond the simple visualization and inspection of anatomic structures. It has become a tool for surgical planning and simulation, intra-operative navigation, radiotherapy planning and for tracking the progress of disease.

The analysis of coronary arteries is of particular importance in medical diagnosis of heart disease. It has been well known that cardiologic injury is presently the single greatest health problem in most developed countries. The most frequent pathology affecting the heart arteries or *coronaries* is the partial or total obstruction due to lipid accumulation in their interior walls. This pathology is known as *stenosis*. An image technique is needed for its diagnosis and evaluation.

At this moment, the current most used technique is coronary angiography. Another useful technique is the IntraVascular UltraSound (IVUS), although it is not so widely used. Both techniques complement each other. Recently, a new technique has been developed, the High-speed Magnetic Resonance Coronary Angiography, that provides an image volume of the heart, where the vessels can be appreciated. However, this technique is not yet available in most catheterism laboratories, and, moreover, the acquisition cost is very high.

1.1 Improvements using Computers

As showed by McInerney et al. in [57], with medical imaging playing an increasingly prominent role in the diagnosis and treatment of diseases, it becomes nearly necessary to extract, with the assistance of computers, clinically useful information about anatomic structures imaged through X-Ray Angiography, IntraVascular UltraSound (IVUS), Computed Tomography (CT), Magnetic Resonance (MR), Positron

Emission Tomography (PET), and other modalities. Although these modern imaging devices provide exceptional views of internal anatomy, the use of computers to quantify and analyze the embedded structures with accuracy and efficiency is limited. Accurate, repeatable, quantitative data must be efficiently extracted in order to support the diagnosis. Moreover, and beyond providing quantitative data, computers can compensate some of the lacks that have the current medical imaging techniques.

In the framework of coronary vessels examination, much work has been done to improve the diagnosis potential of imaging techniques:

3D Lesion Length Measurement. Using a pair of views of the lesion, a 3D reconstruction of the vessel segment containing the lesion is reconstructed. From this reconstruction, one can obtain a more reliable measurement of the lesion length. An example of this application can be found in [8].

Fusion of X-Ray Coronary Angiography and IVUS. X-ray coronary angiography presents the lack of the foreshortening effect when evaluating the degree of a stenosis, while IVUS has the problem of an excessive locality of the information that they provide. An approach to alleviate these problems is to perform a fusion between the information provided by the angiography (the 3D catheter path) and the one provided by the IVUS (the cross-sectional transmural data). The result obtained allows a global assessment of the vessel lesion, by means of volumetric measurements and 3D visualizations. This approach has been presented in [73, 93, 78]. Figure 1.1 shows a snapshot of the application developed by Rotger in [77] to relate and fuse data from angiography and IVUS.

Determination of Optimal Views. The main idea of this procedure is to acquire two different views of the vessel to be analysed and, from this image pair, to deduce which should be the two optimal views to acquire the vessel of interest. The optimal views shall be two perpendicular ones that are both parallel to the main axis of the vessel. From these views, the perspective effect should be minimized. Dumay et al. presented in [22] a method to obtain such views.

Case Retrieval. Using computers, the individual cases can be stored with all the information related to them: diagnosis, treatment, patient evolution, etc. Hence, given a new case, the physician can perform a search on the database to obtain similar past cases. By examining the effects of the different treatments applied to each case, he/she can decide the treatment to be applied. At this moment, a research project is being carried out in our group devoted to the development of a case retrieval application for IVUS images using elastic matching.

Vessel Tagging. Given the complexity of the coronary tree and the wide range of views from which it can be acquired, it is often necessary to get some help on identifying the different coronary vessels from a coronary angiography. The main applications of vessel tagging are to improve inter-physician communication and to help the case characterisation for other purposes, as case retrieval. An approach to obtain this tagging is to adapt a model of the coronary tree containing the vessel tags to the image [30].

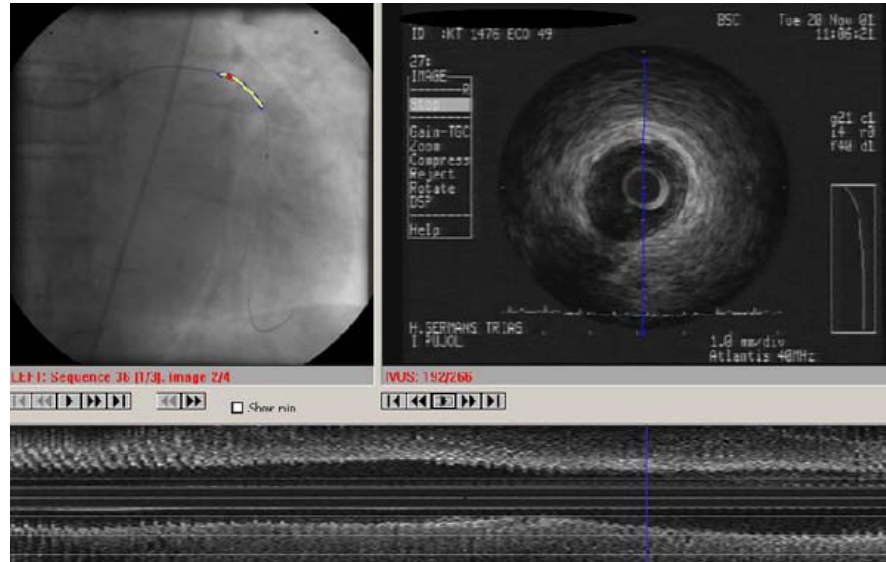


Figure 1.1: An IVUS and Angiogram fusion application. The user can know for each IVUS image the corresponding location of the catheter on the angiographic image.

Among these improvements, there are several, specially lesion length measurement and the fusion of X-ray coronary angiography and IVUS, that need a three-dimensional reconstruction of the vessel from angiographic image pairs. Hence, it is a key point on the development of these techniques. That is why in this thesis we focus the three-dimensional reconstruction of the vessels from angiographies and the problems associated with it.

1.2 Coronary Angiography

Coronary angiography consists of a vessel radiography with radio-opaque fluid injected to improve contrast (see figure 1.2 for an example). This technique is specially used to detect stenosis and other lesions on the coronary arteries. Also, it is used to determine the length of the stent to be introduced in order to treat the detected stenosis and to guide the stent to the desired location during its introduction.

However, it presents some drawbacks. First, when measuring the length of a lesion, the measurement has to be obtained by comparing the size of the lesion to the width of the catheter appearance in the image, disregarding the imprecision due to the perspective projection. As shown in figure 1.3, unless we acquire the lesion from the appropriate angulation, that is, parallel to the vessel, we will obtain a length measure shorter than the real one due to the perspective effect.

Another drawback of this technique is that the degree of stenosis can not be

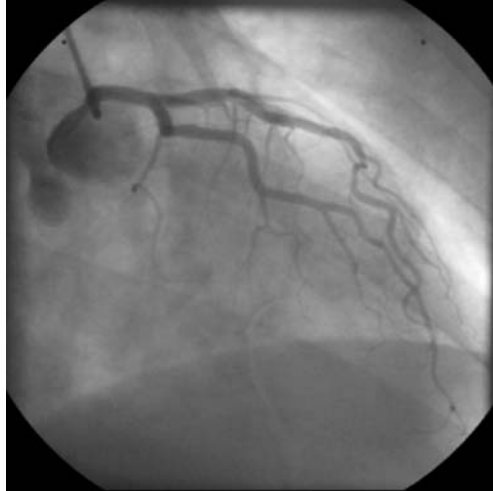


Figure 1.2: A coronary angiography image.

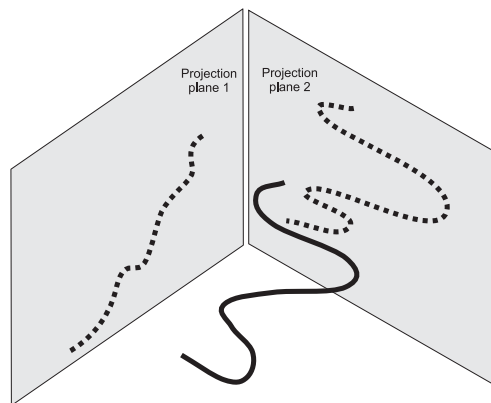


Figure 1.3: The perspective projection problem. Unless we acquire the lesion from the appropriate angulation, that is, parallel to the vessel, we will obtain a length measure shorter than the real one.

evaluated due to the foreshortening effect: depending on the view angle of the lesion, the degree of stenosis can be under-estimated. For an illustration of this effect, see figure 1.4. Moreover, even when this evaluation is performed from an optimal view, the precision obtained is not very high due to the low signal-to-noise ratio of this kind of images and the size of the vessels on the image.

Finally, vessel overlapping is another shortcoming of this technique. The desired lesion could be hidden by another vessel in front of it from the acquisition point of view.

A technique that can complement X-ray angiography is Intravascular Ultrasound,

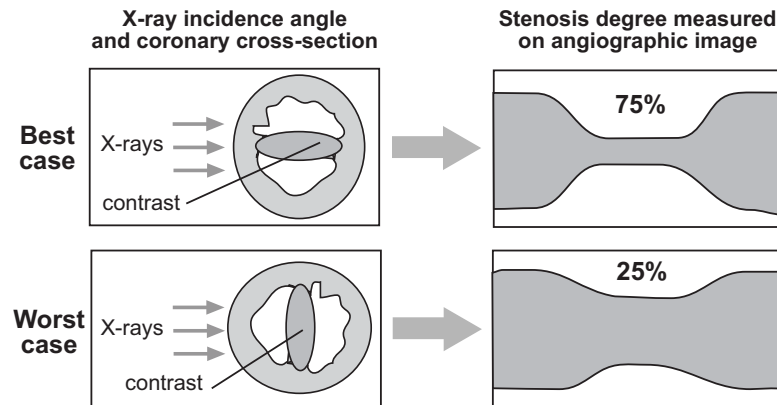


Figure 1.4: The foreshortening effect. Depending on the acquisition angle, the degree of stenosis can be under-estimated using X-ray angiography.

which is obtained by introducing a catheter into the vessel. This catheter has a sensor on its extreme that, using ultrasound, provides transmural cross-sectional information of the vessel at the position of the sensor (see figure 1.5).

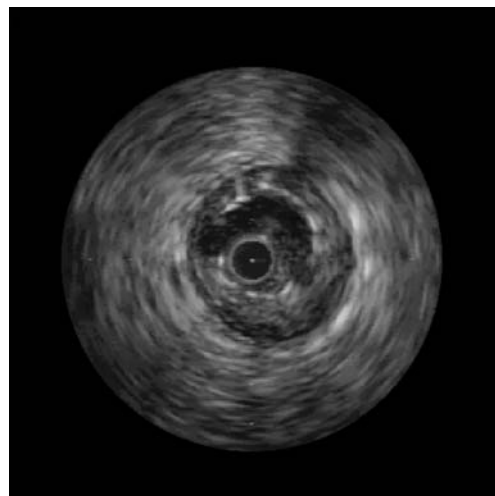


Figure 1.5: An IntraVascular UltraSound image.

This technique, although invasive, allows in vivo a more comprehensive assessment of the plaque, as it provides transmural cross-sectional imaging of coronary arteries and allows diameter and area measurements of both lumen and atherosclerotic plaque. To examine a section of a vessel, the catheter is introduced until the beginning of the desired section and then a motorized system slowly extracts the catheter while obtaining cross-sectional images of the vessel at each position (see figure 1.6). This procedure is called *pull-back*.

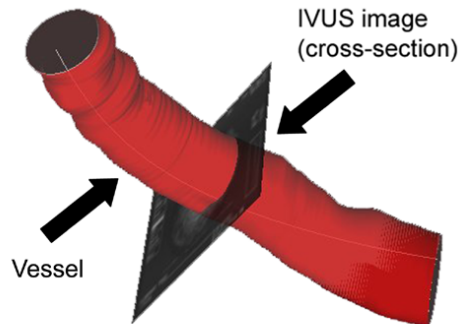


Figure 1.6: IVUS provide local cross-sectional information of the vessel.

The main drawback of this technique is that each image only provides local information of the vessel, i.e. the physician does not know exactly which location of the vessel each frame refers to. By examining the pull-back sequence, he/she has to imagine how the global situation of the vessel is. This mental exercise is of course not very easy to achieve, due to the lack of references in the images. Moreover, due to the movement of the heart, each frame is rotated from the previous one, complicating even more the task to the physician.

1.3 3D Reconstruction from Coronary Angiography

As stated before, obtaining an accurate three-dimensional reconstruction is key for several applications. However, we have found several difficulties when approaching the 3D reconstruction of the coronary vessels from angiography. The main difficulties that we have found are the correction of the geometrical distortion introduced by the image intensifier, the acquisition geometry determination, the automatic vessel detection and the corresponding points determination. In this thesis, we analyse these difficulties and propose an approach to alleviate them.

1.3.1 Geometrical Distortion Correction

When approaching the three-dimensional reconstruction of the coronary vessels, we find that angiographic images suffer from geometrical distortion, which is introduced by the image intensifier. The main difficulty is that, unlike the distortion introduced by the optics in the camera frame, the distortion introduced by the image intensifier varies from one view to another.

This distortion is generally described as the combination of two effects from different sources: First, the peripheral concavity of the Image Intensifier (II) and the non-linearity of the electronic devices that causes pincushion and other geometrical

distortions. Second, the influence of the earth’s magnetic field is quite perceptible and the distortion depends on the orientation of the detector; hence, the distortion changes depending on the orientation of the II [29, 17]. Other non-fixed distortion sources may include fields emitted by nearby monitors, or even non-linear distortions with a warming-up of the image intensifier. Since in our case the monitors remain at a fixed position, we assume that the distortion introduced by them can also be assumed as depending on the orientation of the detector. Regarding the warming-up of the II, authors in [17] claim that the distortion remains quite constant even during several months, always after a warming-up of a minimum of 2h from the switching on of the II.

In routine clinical practice, where a Field-Of-View (FOV) of 17–23cm is commonly used for the acquisition of coronary vessels, this distortion introduces a positional error of up to 7 pixels for an image matrix size of 512×512 and a FOV of 17cm. This error increases with the size of the FOV. Geometrical distortions have a significant effect on the validity of the 3D reconstruction of vessels from these images. Hence, in order to obtain an accurate three-dimensional reconstruction of vessels from angiography, X-ray image distortion must be taken into account [21].

Some investigators use a few small lead markers attached to the II screen to estimate the distortion [26, 67, 63]. Since the geometrical distortion is complex, in most cases a large number of lead markers [17] or a calibration grid attached to the II screen, is necessary [27, 2, 76, 29, 17, 38, 28, 25, 21]. Applying this procedure for each acquisition may be inconvenient in clinical practice; hence, some researchers have suggested to calibrate the system only once [28]. Thus, they correct only the radial or pincushion distortion. However, the influence of the II orientation on the geometrical distortion [29] and its effect on the 3D reconstruction is too large to be neglected [17, 21]. Other authors propose to predict the distortion [29, 25] from the anatomical angles (rotation and angulation) of the II by interpolation, obtaining a mean error of about 1.5 pixels and a maximum of about three pixels, for an image matrix size of 512×512 pixels [25]. Even though these approaches are a good step forward, the precision obtained could be insufficient for some applications. Therefore, we conclude that although much work has been done, the problem of correcting X-ray image distortion remains open.

Our proposal

In this thesis, we propose a model of geometrical distortion based on anatomical angles, since the geometrical distortion varies with the extrinsic parameters. First, we prove that the distortion can be accurately modeled using a polynomial for each view. Also, we show that the estimated polynomial is independent of focal length, but not of changes in anatomical angles, as the II is influenced by the earth’s magnetic field. Thus, we decompose the polynomial into two components, namely the steady and the orientation-dependent component. This fact simplifies the estimation of the orientation-dependent polynomial, since the number of polynomial coefficients to be estimated is lower.

For the development of the distortion model, we have been inspired by the work of Kerrien et al. in [25]. In the method proposed by Kerrien et al. [25], the position of the grid nodes are estimated using a bivariate polynomial of α and β (the anatomical angles, which describe the orientation of the C-arm). Using the estimated grid nodes, a vector field containing the image displacement due to distortion is computed for the whole image. The main differences with this work are: First, we estimate a global model of the distortion consisting on a bi-variate polynomial of the image coordinates, instead of estimating the positions of the grid nodes on the image. This polynomial is selected to have the optimal number of coefficients to describe the distortion, and this simplifies the estimation of these coefficients from the anatomical angles α and β . Second, we decompose the distortion model into two components, namely the steady and the orientation-dependent component, thus reducing the number of coefficients to be estimated from α and β , since the orientation-dependent distortion is less complex than the steady one. This should simplify the model and its calibration. And third, we use a bi-variate polynomial of α and β of optimal degree for the estimation of the orientation-dependent distortion component.

1.3.2 Determination of the Acquisition Geometry

In order to obtain an accurate 3D reconstruction of a point or curve, the acquisition geometry must be known with a high precision. Hence, camera calibration is a necessary step in order to extract metric information from 2D images. Much work has been done, starting in the photogrammetry community (see [6, 31] to cite a few), and more recently in computer vision [35, 60, 91, 96, 56].

We can classify the different calibration techniques in the literature into two categories: photogrammetric calibration and self-calibration[103]:

Photogrammetric calibration. Camera calibration is performed by observing a calibration object whose geometry in 3D space is known with very good precision. Calibration can be very efficiently done [32]. The calibration object usually consists of two or three planes orthogonal to each other. Sometimes a plane undergoing a precisely known translation is also used [91]. These approaches require an expensive calibration apparatus and an elaborated setup.

Self-calibration. Techniques in this category do not use any calibration object. Just by moving a camera in a static scene, the rigidity of the scene provides in general two constraints [56, 52] on the cameras internal parameters from the camera displacement by using image information alone. Therefore, if images are taken by the same camera with fixed internal parameters, correspondences between three images are sufficient to recover both the internal and external parameters which allow to reconstruct 3D structure up to a similarity [53, 40]. While this approach is very flexible, it is not yet mature [5]. Because there are many parameters to estimate, we cannot always obtain reliable results.

Other techniques exist: vanishing points for orthogonal directions [11, 47], and calibration from pure rotation [40, 86].

Thus, a lot of work on the estimation of the camera projection matrix has been done, and hence one can find a lot of methods on the Computer Vision literature. However, in the angiographic frame we have some difficulties that are harder to overcome than in the camera calibration frame and that limitate the applicability of some methods in the angiographic frame.

The first difficulty is the calibration object manufacture. Since we are not working with visible light, but with X-rays, we have to accurately choose the materials, this object is manufactured of. Hence, radio-opaque materials, as iron and other metals, should be used only as calibration markers, while radio-transparent materials, as expanded polystyrene (better known as *porexpan*) or polymethyl methacrylate, should be kept for supporting the markers. Working with this kind of materials is more difficult than simply printing a calibration pattern on a paper and sticking it to a rigid planar surface. Thus, the calibration method should require a calibration object as easy to construct as possible. Therefore, we will discard the calibration methods which require a complicate calibration object, as most methods coming from Photogrammetric Calibration do. Nevertheless, we can find methods that use a Photogrammetric approach in the angiographic frame, as the one presented by Hoffmann et al. in [28]. This method needs a calibration object consisting of a polymethyl methacrylate cube with 12 lead marks attached to it at known 3D positions. The calibration is done by repeating the acquisition but acquiring a calibration object instead of the patient. Although this method obtains good results, the calibration object is hard to manufacture and hence we have not followed this approach.

Another difficulty is that some methods require a precisely known translation [91] or rotation[39, 86]. In the camera calibration frame, this is already a difficulty, since it requires an elaborated setup to perform such a movement. In the angiographic frame, a movement of this kind (as a rotation around the camera focus) is simply impossible, or at least very hard to obtain. Moreover, for the sake of the clinical applicability of the reconstruction method, the calibration procedure should be as simple to perform as possible. Thus, we have also rejected all the calibration methods that needed such movements.

Finally, self-calibration techniques were also discarded for this thesis. This approach has been already explored in the angiographic frame [62, 13]. The main idea is to define point correspondences on at least two different projections of the heart, typically corresponding to vessel bifurcations. From these correspondences, these methods are able to recover the intrinsic and extrinsic parameters up to a similarity. This means that knowledge about the size of the imaged objects should be known with high precision to obtain good results. In practice, this approach presents two main problems: first, knowing the size of the imaged objects, or, alternatively, the distance between the two X-ray sources is difficult to achieve. Moreover, the accuracy of these techniques highly relies on the precision obtained when detecting the corresponding features on the images. Since an automatic method to detect corresponding vessel bifurcations on two images with high precision is very hard to achieve, this is an additional obstacle for the practical application of this technique on routine clinical practice.

Hence, we are looking for an approach as flexible as self-calibration but which does not need a knowledge hard to obtain (as the distance between the two X-ray sources for each acquisition pair) nor a complicated calibration object. Recently, two methods that fulfill these requirements have appeared in the Computer Vision literature [87, 103]. Although both methods were independently developed, they are very similar. They only require the camera (or angiography acquisition system) to observe a planar pattern shown at a few (at least two) different orientations. Either the camera or the planar pattern can be freely moved. The motion need not be known. By assuming that the model plane is on $Z = 0$ of the world coordinate system, two constraints are defined, provided that the intrinsic parameters of the camera (as the focal length) do not change. The position and orientation of the camera for each acquisition can be derived from these constraints, and also the camera intrinsic parameters.

However, all these methods require performing a calibration at each acquisition, which becomes a time-consuming and tedious task in clinical practice. Hence, another common approach in the angiographic frame is to use a model of the movement of the C-arm, as the proposed by Dumay et al. in [22], which is widely used to obtain an estimation of the acquisition geometry. Nevertheless, authors in [23] state that the classical isocentric model could not satisfy the accuracy required, and thus propose to use a variable isoaxis instead of a variable isocenter, and to refine the acquisition geometry using image correspondences at each acquisition. Chen et al. in [13] propose also this refining for the same reason. As we have already discussed, automatically detecting corresponding features on cardiac images with high precision is hard to achieve. Moreover, the first difficulty that one finds when aiming to use a model is the estimation of the model parameters which best fit a particular angiographic system.

Our proposal

Our objective is to develop an improved model of the movement of the C-arm inspired from the model proposed by Dumay et al. in [22], which imposes less constraints and then is closer to the reality. By extending the model, we soften the requirement of a refinement at each acquisition. The calibration procedure should be performed only once (or at least only after a long period of time from the last calibration). This should facilitate the final clinical application of the method.

Also, we aim to formulate a flexible calibration method to estimate the parameters of this model, which will depend on each particular system. The calibration method should only require a calibration object easy to construct and a short acquisition time. In this thesis, we have chosen the method proposed by Zhang in [103] as starting point for the calibration method.

1.3.3 Automatic Vessel Detection

After correcting the image distortion and determining the acquisition geometry, we approach the three-dimensional reconstruction of the vessels. One of the most difficult problems when approaching this task is the automatic vessel extraction. A gray-scale image thresholding approach to segment the vessels is not applicable here for two reasons: First, the X-rays incidence is not homogeneous on the image, and therefore the image is brighter near the center of the image. And second, as opposite to the optical camera frame, the images obtained with X-rays are of an accumulative nature, and therefore the vessels will be darker or brighter depending on the background structures behind or in front of them. Hence, only a local contrast measure is applicable here. Moreover, current angiographic systems store the acquired sequences on a digital support in DICOM format. Even if this support allows applications which were nearly impossible with the film support, storing the images in DICOM format currently means an image size of 512×512 pixels and only 256 gray levels. This fact really makes difficult vessel segmentation, specially for small and/or poor contrasted vessels.

We can find many methods devoted to the vessel centerline extraction in the literature (see [50, 37] for a review). Some of them are based on specially designed algorithms, others on mathematical morphology and others on general valley detectors tuned to the particular problem. Nevertheless, in this thesis we do not focus the study of the different methods to detect the vessel centerline. Instead, we have chosen the most available for us, since it has been developed in our group, as starting point to develop our vessel centerline feature extraction procedure.

This method was presented by Lopez et al. in [50]. It consists on a scanning method, which is based on the computation of a robust creaseness measure named multilocal level curve curvature, a geometric entity defined by the same authors. Hence, the vessel centerlines are obtained by thresholding this creaseness measure and afterwards applying a thinning algorithm to the obtained binary image to obtain the centerlines. In order to eliminate spurious background structures, the authors in [50] propose to eliminate the centerlines of short length. Although this actually improves the result, it can have the undesired effect of eliminating real vessel centerlines when the centerline appears broken.

Hence, a problem that we have found when applying this method is that some background structures are also detected as vessel centerlines. In our opinion, the reason is that the creaseness measure does not take into account the image contrast at the creases, and hence poorly contrasted creases obtain the same score as the strong contrasted ones, which are more likely to correspond to a vessel.

Digital Subtraction Angiography (DSA) is a common approach to alleviate the problems originated from background structures and the heterogeneous X-ray intensity distribution on the image: an image acquired prior to contrast injection is subtracted from the contrasted image, thus removing the background and obtaining the vessels. For still images the results of this technique are so good that no further processing is necessary. However, in cardiac applications the coronary vessels are

in constant movement, and a raw DSA approach will create motion artifacts which could be detected as vessels. In [59], several methods for the reduction of motion artifacts are presented. Nevertheless, a conclusion of this paper is that in coronary angiography these techniques are hard to apply.

On the other hand, we can find specific multiscale approaches for vessel enhancement without subtracting the background in [81, 49, 34]. In particular, the filter proposed by Frangi et al. in [34] gives good noise and background suppression results. However, we have noticed some problems to segment the vessels from two-dimensional X-ray cardiac angiographies. We think that one of the reasons is that this method does not consider the coherence, that is if a given pixel is considered to be part of a vessel, the neighbour pixels in the vessel direction are also very likely to correspond to a vessel.

By analysing the literature, we can find filters that aim increasing the coherence of the features in the images. That is the case of the diffusion processes based on the structure tensor to enhance linear structures [95] and creases [85]. Since the vessels can be regarded as creases and Solé's filter [85] aims enhancing the coherence of the creases on the image, this filter seems a priori appropriate for enhancing the image prior the vessel segmentation. However, the ridges/valleys measure used does not take into account the image contrast, and therefore, when applying it to regularize the vessels, some undesired background structures are also enhanced. Also, the vessels appear at different scales, and therefore a multiscale approach is necessary [34].

Our proposal

In this thesis, we aim to develop a method to enhance vessel structures to make easier vessel segmentation. A vessel enhancement improves vessel delineation and reduces noise and background objects. This facilitates visual analysis and improves the results of vessel centerline segmentation using any of the techniques present in the literature, as for example the one presented by Lopez et al. in [50].

To this aim, we propose a new anisotropic diffusion filter inspired in that of Solé et al. [85], which enhances the creases on the image. The main difference is that our filter enhances the creases on the image with a diffusion strength determined by the Frangi's vesselness measure [34]. Moreover, for each pixel the scale of the diffusion tensor to be applied is chosen using the Frangi's vesselness measure.

By applying this pre-processing before the vessel centerline detection procedure of Lopez et al. [50], we obtain two benefits: First, since the filtering increases the crease coherence on the image, the centerlines corresponding to a vessel are less likely to be broken after thresholding the creaseness measure. Second, we can use the vesselness measure, which will be now improved after applying the anisotropic vesselness enhancement diffusion filter, to eliminate the detected centerlines corresponding to the background while preserving those corresponding to vessels.

1.3.4 Corresponding Points Determination

Once that we have detected the points on the image corresponding to the vessel centerline, we will be interested in obtaining the three-dimensional centerline curve. To this aim, the usual approach consists of reconstructing corresponding points in both X-ray images, and then obtaining a 3D curve by interpolation between the obtained 3D points. This is the approach used in [84, 36, 22, 94, 23, 12, 13, 28, 98, 21].

However, this bottom-up approach has many shortcomings. First, automatic corresponding points determination is in general computationally expensive, mainly when we do not have in general two similar views but two different ones as angulated as possible. While this fact improves the accuracy of the three-dimensional reconstruction, the two views will be very different one from another. Moreover, there is a big degree of ambiguity present on this kind of images, since all the vessels are similar. The usual approach is therefore asking the user to manually mark corresponding points. This task is also non-trivial for the user, and can become quite tedious if high accuracy is needed. Second, even when the user is helped by the epipolar line [22] to match points in different views, inaccuracies on the biplane geometry estimation may make fail the epipolarity constraint, and therefore the user can be driven to mark incorrect matches. And third, the curve is directly interpolated among marked points, and therefore the user must mark a lot of corresponding points to accurately follow the shape of the vessel, which in many cases can be quite wavy.

Our proposal

In this thesis we propose to reduce both problems by using biplane snakes, which are derived from the physics-based model of snakes [42]. The snake initialization is done by a user interaction that determines two pairs of approximately corresponding points near the beginning and the end of the part of the vessel to be reconstructed. Then, an initial 3D curve is obtained from this correspondence. Such a curve deforms in order to adapt its projections to the images, inherently solving the point correspondence problem.

Preliminary results of this approach were presented in [64] for orthogonal views in the frame of catheter path reconstruction. We aim to develop an improved version of the technique and apply it to the three-dimensional reconstruction of vessel centerlines from arbitrary views.

1.4 Previous Publications

Parts of this thesis have been previously published. The model for the movement of the C-arm and the calibration techniques were published in [10]. The distortion/undistortion model and the associated calibration were published in [7] and in journal version in [9]. Preliminary results on the use of biplane snake to reconstruct the vessels were presented in [15, 66, 8]. A journal version of biplane snakes using

GGVF was published in [9].

Although not described in this thesis, we have applied the results obtained in this thesis to other applications. In particular, results for IVUS and angiogram data fusion were published in [74, 55, 79].

1.5 Thesis Organization

The remaining of this thesis is organized as follows. In Chapter 2 we provide the background material about X-ray angiography, 3D Reconstruction, Deformable Models and Image Enhancement techniques. In Chapter 3, we develop our proposal for the geometrical distortion correction, and the model of the movement of the C-arm. We also propose calibration methods to obtain the parameters of the proposed models. In Chapter 4, we focus on the three-dimensional reconstruction of the coronary vessels using snakes. We propose a method to improve the image quality in order to obtain a better segmentation of the vessels and develop the biplane snakes, extended to arbitrary views, to improve the performance against wavy vessels and to take into account image distortion. In Chapter 5, we present experimental results of the methods proposed in Chapters 3 and 4 using the angiographic images provided by the Hospital Universitary "Germans Trias i Pujol". Finally, we conclude this thesis in Chapter 6 with a summary and a discussion of future research areas.

Chapter 2

Background

2.1 X-Ray Angiography

In this section, we aim to provide the basics on digital X-ray angiography. First, we relate the X-Rays discovery and describe the properties of this type of radiation. Next, we concentrate on the development of coronary angiography. Finally, we describe a coronary angiographer, how it works, and the problems that could arise when working with it.

2.1.1 X-Rays History

Wilhelm Conrad Roentgen (figure 2.1) discovered the X-rays while studying cathode rays. In order to facilitate his observations, he wrapped the gas discharge tube in black paper and darked the room. Thus, he noticed that when he operated his gas discharge tube, a fluorescent screen of barium platinocyanide, which was accidentally near the tube, lighted up.

This is the first documented observation of X-rays, and occurred on November 8, 1895. Right away he recognised the importance of his discovery and locked him up in his laboratory for seven weeks. Within these weeks, he discovered and analyzed practically all properties of X-rays relevant for imaging and even produced some radiographs of hands on photographic plates, the first known X-ray images showing human anatomy (see figure 2.2).

During his early experiments with X-rays, Roentgen identified as the source of the radiation that area of the glass tube which was hit by the cathode rays, i.e. accelerated free electrons (see figure 2.3). The cloud of electrons was generated by the small voltage, and accelerated by the big one to the cathode. When they hit the glass at very high speed, they were able to liberate electrons from the higher energy layers of the glass atoms, and hence an electromagnetic radiation of very high energy (i.e. frequency) was produced.

Thus, the mysterious X-rays discovered by Roentgen were just another type of electromagnetic radiation like radio waves, infra-red, visible light, or ultra-violet, however, of a very much higher frequency or energy than e.g. ultra-violet and, therefore,



Figure 2.1: Wilhelm Conrad Roentgen, discoverer of the X-rays.



Figure 2.2: The first radiograph performed by W.C. Roentgen, probably Mrs. Roentgen's hand.

drastically different behavior in its interaction with matter. In particular, X-rays have the ability to penetrate solid objects, and the degree of penetration depends on

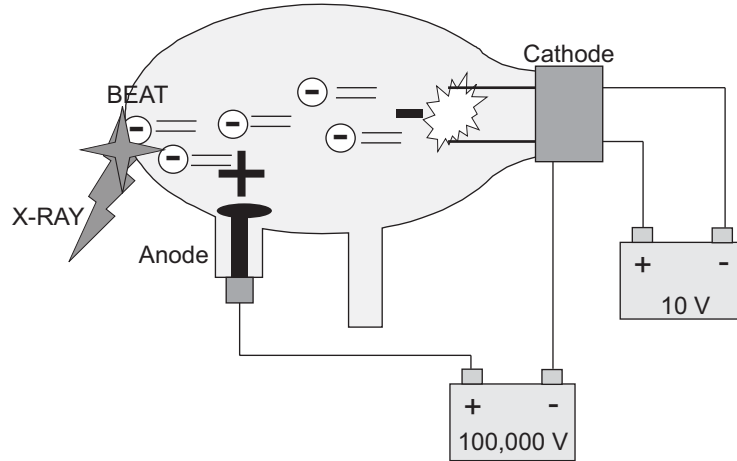


Figure 2.3: X-ray generation on Roentgen's glass tube.

the energy of the X-ray quanta but also on the thickness and density of the object. Therefore, X-rays have become an invaluable tool for medical imaging and industrial non-destructive testing.

2.1.2 The Development of Cardiac Angiography

The first angiography was obtained in January 1896, by post-mortem injection of mercury compounds on the vessels of a hand (see figure 2.4). This compound was necessary in order to enhance the contrast of the vessels, since blood is nearly radio-transparent. Since mercury is highly radio-opaque, it would be a good contrast compound if a mercury injection were not so mortal. Fortunately, for the medical application of angiography, innocuous compounds for the enhancement of the contrast of vessels were posteriorly developed. Now the most commonly injected contrast compound is a iodine solution at exactly 36.5°C .

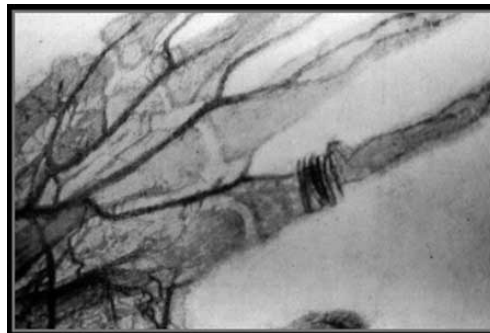


Figure 2.4: Post-mortem injection of mercury compounds, January 1896.

However, coronary angiography appeared much later, since this technique required direct contrast injection on the vessels to be imaged. A student of the University Hospital of Berlin, called Werner Forssmann, played a key role on the development of this technique. In 1929, he used a urology catheter of 60cm to reach the heart by inserting it through the arteries. He was so convinced of the feasibility of this that he decided to test the technique with himself. Hence, he carried himself out a cut on his forearm, and opened an artery to introduce the catheter into it. After introducing the catheter into his body, he decided that he had to test whether he had succeeded by imaging the catheter with X-rays, and therefore, he went to the upper floor through the stairs to the X-ray laboratory with the catheter introduced into his body. There he realized that he could lead the catheter to his heart: cardiac catheterism was born. Although he was dismissed due to the riskiness of his experiments, in 1956 he shared with André Cournand and Dickinson Richards the Nobel Prize for his contribution in the development of cardiac catheterism.

Nowadays, it is not necessary to perform such a cut to the patient, since it is enough to inflict a puncture. The usual location of this puncture is on the femoral artery, although at this moment new techniques allow to perform this puncture on the forearm of the patient (radial puncture). This is much more comfortable to the patient, since after the exploration the patient can stand up and walk, which is not the case for the femoral puncture.

Another great improvement has been the introduction of the image intensifier, and the development of new techniques to produce X-rays. Thanks to these improvements, now it is possible to image coronary vessels in real time and with high resolution, without undergoing the patient to a big amount of radiation.

2.1.3 The Cardiac Angiography Acquisition System

A typical cardiac X-ray angiography acquisition system is composed by the following: A X-ray generator tube or X-ray source, a table onto which the patient can be placed, an Image Intensifier which converts the X-ray signal into a visible image, a granting system to move the source-intensifier pair in order to obtain different views of the patient's anatomy and an image acquisition system. Figure 2.5 shows the system available at Hospital Universitari "Germans Trias i Pujol" of Badalona, with whom we collaborate.

The X-Ray Source

Whereas the applications as computerized tomography use a fine X-ray beam to obtain high-resolution images, X-ray cardiac angiography must find an equilibrium between a high spatial resolution and a high temporal resolution.

On the one hand, high spatial resolution is needed, since coronary vessels must be imaged with enough accuracy to determine the severity of an arterial obstruction or stenosis. On the other hand, a high temporal resolution is needed to deal with the movement of the heart and avoid blurred images.

The problem is that the X-ray tube has its limitations. To give a maximal resolution, either spatial or temporal, angiographic systems use fan-shaped X-ray sources,

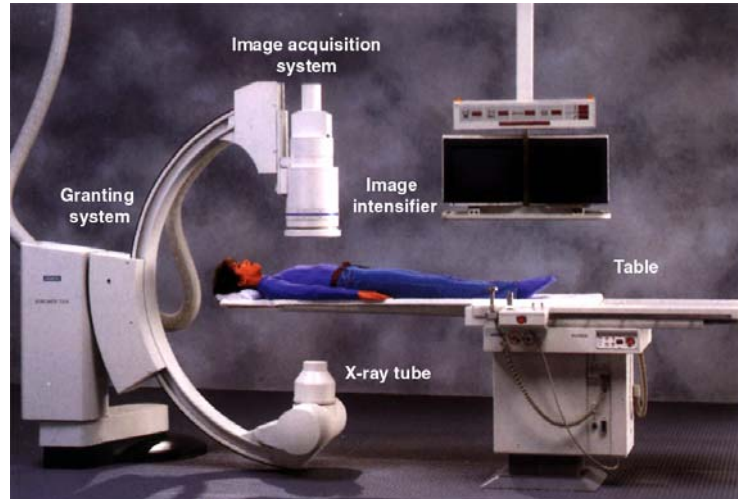


Figure 2.5: An example of angiographer.

which allow a fast image acquisition at the expense of a small distortion on the image. Nevertheless, recently new techniques for high-speed angiographic system have been developed, which provide a higher resolution, both spatial and temporal[92].

The Image Intensifier

The image acquisition system is based on a property called *fluorescence*. This property consists of the capability to transform an incident radiation in another radiation of a different wave length (a more detailed description of this phenomenon can be found in any elemental physics book). Thus, a screen covered by a substance which absorbs X-ray and emits them as visible light allows to obtain an image¹. However, in order to obtain an image directly visible by the human eye, the incident radiation must be intense.

The problem is that a big amount of X-ray radiation can bring about tumours, as first X-ray investigators as Mihran Kassabian (1870-1910), who lost its two hands, unfortunately ought to note. Therefore, in medical applications the minimization of the amount of radiation received by the patient is a must. So, an alternative way for image intensification should be explored, and this leads to the application of the image intensifier.

Figure 2.6 shows a sketch of the image intensifier, which consists of a fluorescent input screen mounted on a thin transparent base on a tube, in which vacuum has been made.

X-ray collides with the fluorescent input screen, thus leading to the emission of visible light. This light is then transformed in an electron beam by means of the fotocathode. An anode placed at the other extreme of the tube accelerates the electrons

¹For instance, the screen of barium platinocyanide present in Roentgen's laboratory was fluorescent.

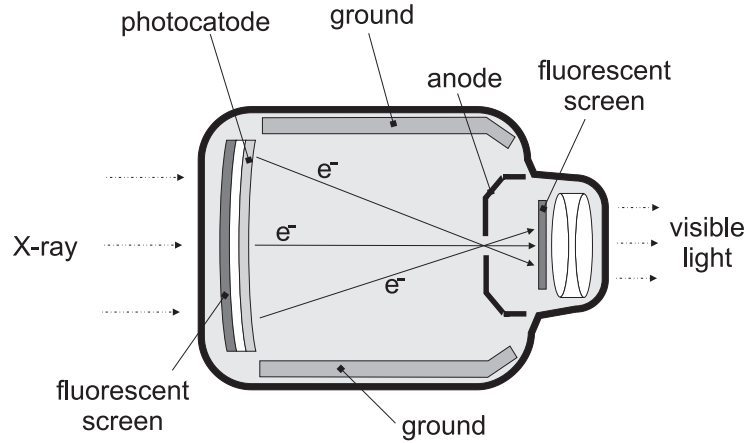


Figure 2.6: A image intensifier.

and concentrates them into a groove. The electrons cross the groove and collide with a second fluorescent screen which transform them again in visible light, but around 10,000 times brighter than the original [92].

The main drawback of the image intensifier is that it introduces geometrical distortion on the image. This distortion is generally described as the combination of two effects from different sources: First, the peripheral concavity of input screen of the Image Intensifier and the non-linearity of the electronic devices cause pincushion and geometrical S distortions (see figure 2.7 for an illustration of these distortions). Second, the influence of the earth's magnetic field is quite perceptible and the distortion depends on the orientation of the detector, provoking something similar to an image rotation. These effects are quite perceptible. For instance, in clinical practice, where a Field-Of-View(FOV) of about 17cm is commonly used, we have found that the positional error is up to 7 pixels for a 512×512 image. This error increases with the size of the FOV.

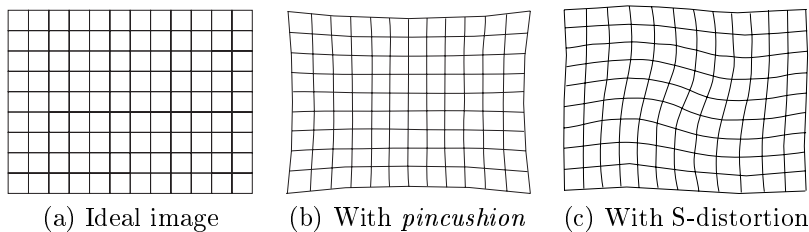


Figure 2.7: Ideal image and two types of geometrical distortion.

The *Granting* System

There exist two types of granting system: L-arm and C-arm. In this work, we will concentrate on the description of the C-arm, which is the most common and also the one which we work with.

The C-arm configuration owes its name to the its C-shape. The image intensifier and the image acquisition system are mounted on the top extreme of the C, and the X-ray tube is mounted on the bottom extreme.

In clinical practice, it is very useful to change the point of view depending on the vessel to be imaged. Instead of rotating and moving the patient, the granting system is able to rotate with two degrees of freedom (*rotation* and *angulation*), which are denoted by the so called *anatomical angles* α and β , respectively. Let us analyze the two movements that the C-arm can carry out:

Rotation: It is determined by the anatomical angle α . As shown in figure 2.8, when $\alpha > 0$, we have a Right Anterior Oblique (RAO) view, and for $\alpha < 0$ we obtain a Left Anterior Oblique (LAO) view.

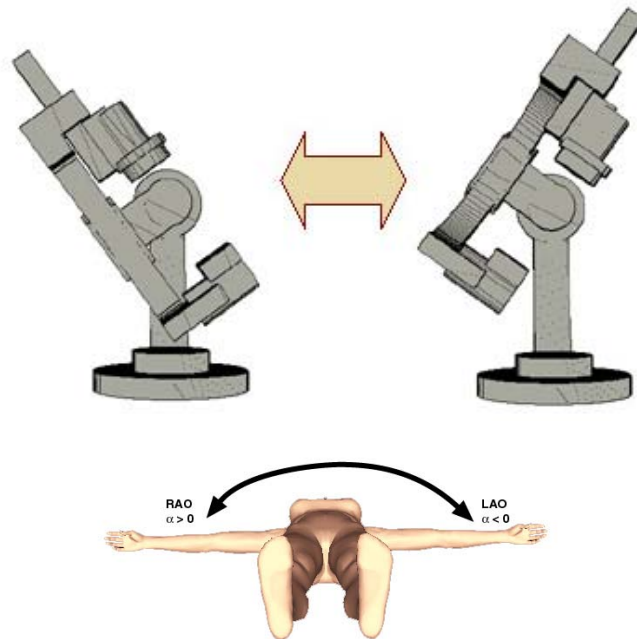


Figure 2.8: Rotation, denoted by α .

Angulation: Described by the anatomical angle β . When $\beta > 0$ the view is classified as Cranial (CRA), and when $\beta < 0$ we talk about a Caudal (CAU) view, as illustrated in figure 2.9.

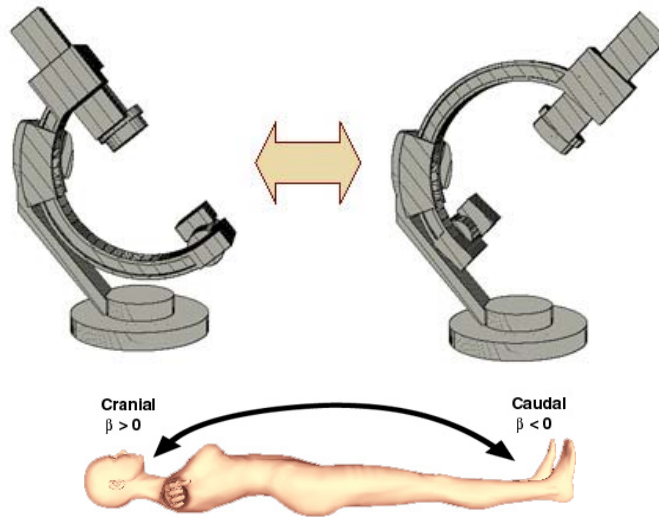


Figure 2.9: Angulation, denoted by β .

Other named views are the one corresponding to $\alpha = 0$ and $\beta = 0$ configuration, called Anterior-Posterior (AP), which provides a frontal view of the heart, and this of $\alpha = -90$ and $\beta = 0$, which is called Lateral 90 (LAT90) and is mainly used to image the right coronary.

The physician has also the possibility of varying the X-ray source and image intensifier distance (SID). As shown in figure 2.10, the image intensifier moves up and down, whereas the X-ray source remains at a constant position. Hence, by moving the image intensifier away from the patient, the objects appear magnified, since the focal length is longer.

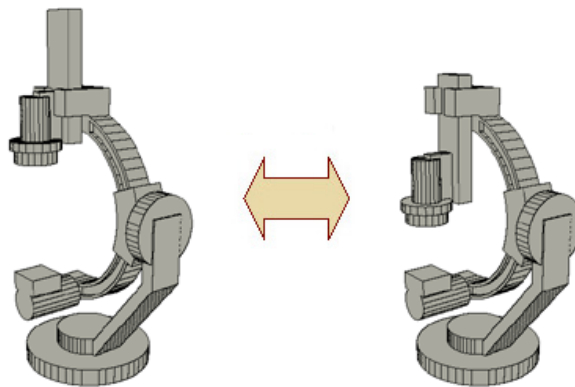


Figure 2.10: Varying the SID. The X-ray source does not move.

Image Acquisition System

In a traditional angiographic acquisition system, a *beam-splitter* is placed in front of the output screen, in order that both the video camera and the 35cm cinefilm can pick the image. This configuration allows the real-time visualization of the angiography, but also to save the sequence on a high-resolution support which is easy to transport.

Modern angiographic systems, on the other hand, have replaced both supports on a unique one: the digital one. Images can also be examined in real time, but at the end of the exploration they are stored in a CD-ROM using the DICOM[92] protocol. This means that we have the sequences stored with a resolution of 512×512 pixels, 256 gray levels and typically 12.5 frames per second. Although the resolution is not so high as the provided by the 35cm cinefilm, the numerous advantages of the digital support, as the possibility of storing all the information about the acquisition conditions (rotation and angulation angle, type of contrast, physician which performed the exploration, etc), about the patient (name, sex, age, weight, etc.) and above all the ECG, fully justify the use of the digital support, since it opens the possibility of a wide variety of applications. For the application of this thesis, where a digitalization is required in order to work with the images on a computer, a digital system avoids a lot of problems with the 35cm cinefilm digitalization procedure. Moreover, the new DICOM standard will include the possibility of storing the images at a resolution of 1024×1024 pixels with 12 bits of gray-level depth.

Although we started this thesis with a non-digital, traditional angiographic system, very soon we had available digital one. The results presented are all for a digital acquisition system.

2.2 3D Reconstruction

In this section, we aim to provide the reader with a minimal background on three-dimensional reconstruction to understand the difficulties found in this thesis. Hence, what is presented here is not an exhaustive review on three-dimensional, but a collection of concepts which will be used to calibrate the X-ray angiographer for 3D reconstruction purposes. To this aim, we first illustrate the relationship existing between the camera model used in Computer Vision applications and that of the X-ray angiographer. Then, we give some insights on 3D reconstruction of a point and describe two feasible reconstruction error measures. After, we describe a calibration approach that fits our requirements and that is used in this thesis. Finally, we describe a tool- namely, the epipolar line-, which is used to determine corresponding points in two views.

2.2.1 Cameras and Angiographers are very Similar

There is a strong relationship between the camera model used in Computer Vision applications and the angiography acquisition system. Let us consider a pinhole camera model and note the differences with the angiography model. Plane \mathcal{F} , called the *focal plane*, is placed in front of the *image plane* \mathcal{I} . Distance f is the *focal length*. Let

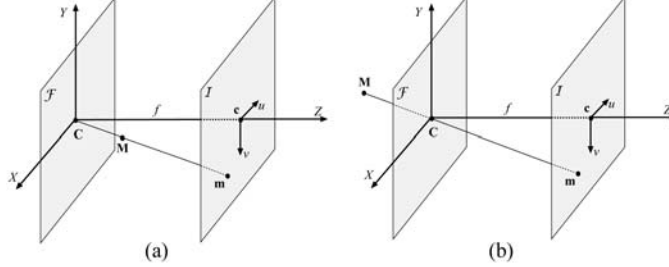


Figure 2.11: The angiography acquisition model (a) and the pinhole camera model (b)

point \mathbf{C} be the *optical center*, then the *optical axis* is the line passing through \mathbf{C} perpendicular to \mathcal{F} . The optical axis intersects \mathcal{I} at point \mathbf{c} , called the *principal point*. Figure 2.11(a) illustrates the angiography acquisition model and figure 2.11(b) shows the pinhole camera model (see [32, 102] for more details). It should be noted that, from a geometric viewpoint, in both models we deal with the perspective projection. The main difference between them is that the object in the X-ray image is magnified and not inverted compared to the pinhole camera image.

The transformation from 3D world coordinates to camera pixel coordinates is modeled in [91] as a process of four stages. By changing some details, we obtain an equivalent procedure.

First, we perform a linear transformation from the three-dimensional world coordinate system (X_w, Y_w, Z_w) to the three-dimensional camera coordinate system (X, Y, Z) :

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = [\mathbf{R} \quad \mathbf{t}] \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}, \quad (2.1)$$

where matrix \mathbf{R} is a 3×3 rotation matrix and \mathbf{t} is a translation vector. \mathbf{R}, \mathbf{t} constitute the *extrinsic parameters*.

Second, a perspective projection is applied, following the pinhole camera model, which is also applicable in the angiographic context:

$$\begin{bmatrix} U \\ V \\ S \end{bmatrix} = \mathbf{A} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}, \quad (2.2)$$

$$u = \frac{U}{S}, \quad v = \frac{V}{S} \quad \text{when } S \neq 0$$

where matrix \mathbf{A} embodies the *intrinsic parameters*. A standard expression for the matrix \mathbf{A} is:

$$\mathbf{A} = \begin{bmatrix} fk & 0 & u_0 \\ 0 & fk & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

where k is the inverse pixel size and (u_0, v_0) are the coordinates of the principal point c on the image matrix. This step differs from Tsai's approach [91], where matrix \mathbf{A} depends only on f . There, the pixel aspect ratio is addressed at a fourth step by including the aspect factor s_x , and expressing f in pixels/cm. We do not introduce any aspect factor, as we consider it as a part of image distortion. Another difference with the Tsai's approach is that in [91], the image center is supposed to be at $(0, 0)$, which is defined at the center of the image matrix. Instead, we include the parameters u_0, v_0 .

Equations (2.1) and (2.2) can be combined as follows:

$$\begin{bmatrix} U \\ V \\ S \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}, \quad (2.3)$$

$$u = \frac{U}{S}, \quad v = \frac{V}{S} \quad \text{when } S \neq 0$$

However, ignoring distortion is unacceptable when doing 3D measurement [91]. This is also true in the context of angiographic projection, since the pinhole camera model is not accurate. Therefore, a third step to model geometrical distortion is needed. Thus, we perform a non-linear transformation using functions ρ^u and ρ^v to obtain *distorted pixel coordinates* (\hat{u}, \hat{v}) :

$$\begin{aligned} \hat{u} &= u + \rho^u, \\ \hat{v} &= v + \rho^v, \end{aligned}$$

This is the model to be applied in order to obtain a unwarped image. The reason is that for each coordinate of the target undistorted image, we obtain a unique corresponding distorted coordinate on the image to be unwarped. If we use instead a undistortion model to unwarp images, we will ought deal with image holes or multiple pixel values.

Nevertheless, image unwarping introduces noise in the image, and is computationally expensive. So, it is preferable to avoid it. Instead, we propose to estimate also a model to obtain the undistorted coordinates:

$$\begin{aligned} u &= \hat{u} + \rho^{\hat{u}}, \\ v &= \hat{v} + \rho^{\hat{v}}, \end{aligned}$$

Thus, to project a 3D point on the distorted image (as we will see later) we use the distortion model, and to reconstruct a 3D point from a pair of points on the images, we use the undistortion model. Thus, no image unwarping is needed.

2.2.2 Three-Dimensional Reconstruction of a Point

Let $\mathbf{C}_1, \mathbf{C}_2$ be the focus position of the X-ray beams in two views, and $\mathbf{m}_1 = (u_1, v_1)$ and $\mathbf{m}_2 = (u_2, v_2)$ the 2D projections of a three-dimensional point \mathbf{M} , which we aim to obtain.

Theoretically, the 3D reconstruction of a point \mathbf{M} from two projections $\mathbf{m}_1, \mathbf{m}_2$ corresponds to the intersection of projective rays $\mathbf{m}_1^1\mathbf{C}^1$ and $\mathbf{m}_2^2\mathbf{C}^2$. However in practice they may fail to intersect. Dumay et al. [22] propose as an approximated reconstruction (also called *retroprojection*) the point \mathbf{M}' , which minimizes the distance to both projective lines (see figure 2.12).

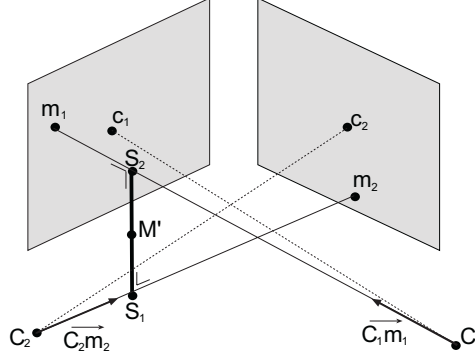


Figure 2.12: The minimum distance three-dimensional reconstruction of corresponding points \mathbf{m}_1 and \mathbf{m}_2 .

This point is situated upon the segment $\mathbf{S}_1\mathbf{S}_2$, which is perpendicular to both projective lines. The vectorial representation of this segment is as follows:

$$\mathbf{S}_1\mathbf{S}_2 = (\mathbf{C}_2 + \sigma\mathbf{C}_2\mathbf{m}_2) - (\mathbf{C}_1 + \tau\mathbf{C}_1\mathbf{m}_1) \quad (2.4)$$

where τ, σ are determined from the perpendicular constraints:

$$\mathbf{S}_1\mathbf{S}_2 \cdot \mathbf{C}_1\mathbf{m}_1 = \mathbf{S}_1\mathbf{S}_2 \cdot \mathbf{C}_2\mathbf{m}_2 = 0$$

The 3D point corresponding to both projections of $\mathbf{m}_1, \mathbf{m}_2$ (which is called the *backprojection* φ^{-1} of \mathbf{m}_1 and \mathbf{m}_2) is therefore expressed as follows [22]:

$$\mathbf{M}' = \varphi^{-1}(\mathbf{m}_1, \mathbf{m}_2) = \mathbf{C}_1 + \tau\mathbf{C}_1\mathbf{m}_1 + \frac{1}{2}\mathbf{S}_1\mathbf{S}_2$$

However, this backprojection method does not necessarily minimize the distances $\|\mathbf{m}_1 - \mathbf{m}'_1\|$ and $\|\mathbf{m}_2 - \mathbf{m}'_2\|$, where \mathbf{m}'_1 and \mathbf{m}'_2 are the projections of \mathbf{M}' . Moreover, it is not clear how to apply it when more than two views are available. Therefore, we use an alternative method to compute \mathbf{M}' .

Let $\mathbf{P}_i = \mathbf{A}_i \begin{bmatrix} \mathbf{R}_i & \mathbf{t}_i \end{bmatrix}$ be the projection matrix for view i . Using equation (2.3), we can derive the following constraint for each view[102]:

$$\begin{bmatrix} (\mathbf{a}_1^i - u_i\mathbf{a}_3^i)^T \mathbf{R}_i \\ (\mathbf{a}_2^i - v_i\mathbf{a}_3^i)^T \mathbf{R}_i \end{bmatrix} \mathbf{M}' = \begin{bmatrix} (u_i\mathbf{a}_3^i - \mathbf{a}_1^i)^T \mathbf{t}_i \\ (v_i\mathbf{a}_3^i - \mathbf{a}_2^i)^T \mathbf{t}_i \end{bmatrix}$$

where \mathbf{a}^i_j corresponds to the j^{th} row of matrix \mathbf{A}_i . For any two views, we have 4 equations and 3 unknowns. Any linear least-squares technique can be used to solve this problem.

However, the previous approach minimizes a criterion that does not have a good physical interpretation (see [102] for details). We use the obtained solution as starting point to minimize the following functional:

$$\sum_i \|\mathbf{m}_i - \mathbf{m}_i'\|^2 = \sum_i \left[\left(u_i - \frac{\mathbf{p}_1^i \mathbf{T} \mathbf{M}'}{\mathbf{p}_3^i \mathbf{T} \mathbf{M}'} \right)^2 + \left(v_i - \frac{\mathbf{p}_2^i \mathbf{T} \mathbf{M}'}{\mathbf{p}_3^i \mathbf{T} \mathbf{M}'} \right)^2 \right]$$

where \mathbf{p}_j^i corresponds to the j^{th} row of matrix \mathbf{P}_i .

2.2.3 Reconstruction error

Now that we know how to obtain the backprojection \mathbf{M}' from two projections $\mathbf{m}_1, \mathbf{m}_2$, we are interested in measuring with how much precision it can be done.

The problem is that in most cases we will not have the exact geometry of the object to be reconstructed, and therefore, we need to define some measure which provides an idea of the amount of error committed when reconstructing. A possible measure of the reconstruction error, which is commonly used, is the distance between projective lines calculated from (2.4):

$$\varepsilon_1(\mathbf{m}_1, \mathbf{m}_2) = \|S_1 S_2\| \quad (2.5)$$

Alternatively, considering $\mathbf{m}_1', \mathbf{m}_2'$ as the projections of \mathbf{M}' in both planes, we can define the reconstruction error as:

$$\varepsilon_2(\mathbf{m}_1, \mathbf{m}_2) = \|\mathbf{m}_1 \mathbf{m}_1'\| + \|\mathbf{m}_2 \mathbf{m}_2'\| \quad (2.6)$$

The use of the first or the second error measure will depend on the application, although both are clearly related. It is easy to see that by reducing (2.6) we reduce (2.5). Error (2.6) is in the image plane units (typically in pixels), while error (2.6) is in world units (typically, millimeters or centimeters).

2.2.4 Camera Calibration using the Zhang's Method

The calibration technique proposed by Zhang in [103] only requires the camera (or angiography acquisition system) to observe a planar pattern shown at a few (at least two) different orientations. Either the camera or the planar pattern can be freely moved. The motion need not be known.

This approach, which uses 2D metric information, lies between the photogrammetric calibration, which uses explicit 3D model, and self-calibration, which uses motion rigidity or equivalently implicit 3D information. Compared to classical techniques, this technique is considerably more flexible: Anyone can make a calibration pattern and, in the angiographic frame, we can use the calibration grid which is commonly available in the catheterism laboratories to estimate the geometrical distortion. Compared with self-calibration, it gains a considerable degree of robustness. Moreover, we do not need any additional knowledge besides the spacing of the grid, which is well known.

Let $\tilde{\mathbf{m}} = [u, v, 1]^T$ be an augmented 2D point. A 3D point is denoted by $\tilde{\mathbf{M}} = [X, Y, Z, 1]^T$. A camera is then modeled by the usual pinhole:

$$s\tilde{\mathbf{m}} = \mathbf{A} \begin{bmatrix} \mathbf{R} & \mathbf{t} \end{bmatrix} \tilde{\mathbf{M}} \quad \text{with} \quad \mathbf{A} = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.7)$$

where s is an arbitrary scale factor, (\mathbf{R}, \mathbf{t}) the rotation and translation which relates world coordinates with camera coordinates, and \mathbf{A} is the camera intrinsic matrix, with (u_0, v_0) the coordinates of the principal point, α and β the scale factors and γ the skew factor.

In this method, the point is to assume that the model plane is on $Z = 0$ of the world coordinate system, so we can denote $\tilde{\mathbf{M}} = [X, Y, 1]^T$. Figure 2.13 illustrates the concept.

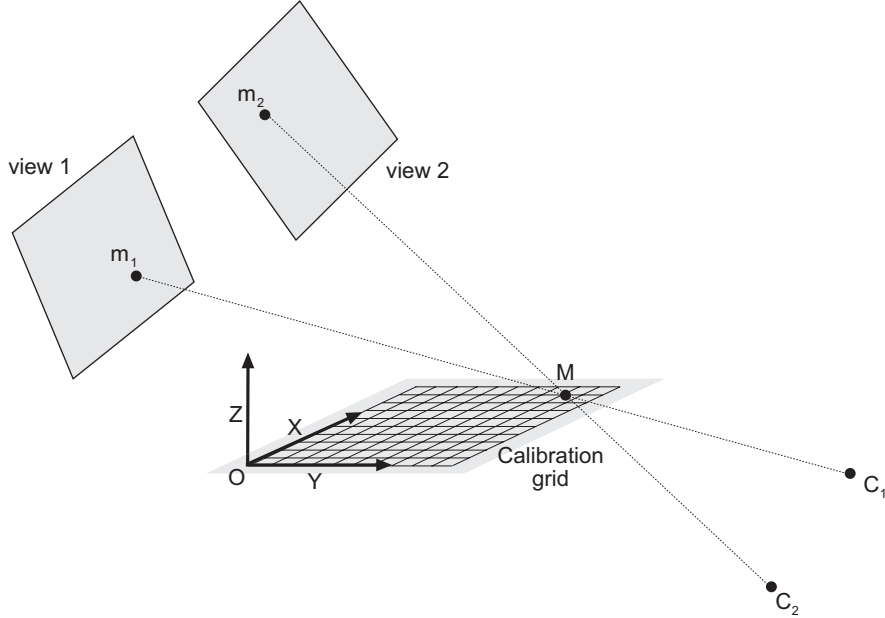


Figure 2.13: The world coordinate system $\{O, X, Y, Z\}$ is defined so that the calibration object is placed on $Z = 0$.

From equation (2.7) we have:

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{r}_3 & \mathbf{t} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}$$

Therefore, a model point on world coordinates can be expressed in image coordinates by applying the homography \mathbf{H} :

$$s\tilde{\mathbf{m}} = \mathbf{H}\tilde{\mathbf{M}} \quad \text{with} \quad \mathbf{H} = \mathbf{A} \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 & \mathbf{t} \end{bmatrix} \quad (2.8)$$

where $\mathbf{r}_1, \mathbf{r}_2$ are the first and second columns of \mathbf{R} . \mathbf{H} is defined up to a scale factor and can be estimated for each view.

There are many ways to estimate the homography between the model plane and its image. Here, we present a technique based on a maximum-likelihood criterion, which is the one proposed by Zhang in [103]. Let \mathbf{M}_i and \mathbf{m}_i be the model and image points, respectively. Ideally, they should satisfy (2.8). In practice, they do not because of noise in the extracted image points. Let us assume that m_i is corrupted by Gaussian noise with mean 0 and covariance matrix $\Lambda_{\mathbf{m}_i}$. Then, the maximum-likelihood estimation of \mathbf{H} is obtained by minimizing the following functional:

$$\sum_i (\mathbf{m}_i - \hat{\mathbf{m}}_i)^T \Lambda_{\mathbf{m}_i}^{-1} (\mathbf{m}_i - \hat{\mathbf{m}}_i)$$

where

$$\hat{\mathbf{m}}_i = \frac{1}{\bar{\mathbf{h}}_3^T \mathbf{M}_i} \begin{bmatrix} \bar{\mathbf{h}}_1^T \mathbf{M}_i \\ \bar{\mathbf{h}}_2^T \mathbf{M}_i \end{bmatrix} \quad \text{with } \bar{\mathbf{h}}_i, \text{ the } i\text{-th row of } \mathbf{H}.$$

In practice, it is simply assumed $\Lambda_{\mathbf{m}_i} = \sigma^2 \mathbf{I}$ for all i . This is reasonable if points are extracted independently with the same procedure. In this case, the above problem becomes a nonlinear least-squares one, i.e., $\min_{\mathbf{H}} \sum_i \|\mathbf{m}_i - \hat{\mathbf{m}}_i\|^2$. The nonlinear minimization can be carried out using any non-linear least squares procedure, as the Levenberg-Marquardt Algorithm. Since this requires an initial guess, Zhang in [103] proposes to obtain it as follows:

Let $\mathbf{x} = [\bar{\mathbf{h}}_1^T, \bar{\mathbf{h}}_2^T, \bar{\mathbf{h}}_3^T]$. Then, equation (2.8) can be rewritten as follows:

$$\begin{bmatrix} \tilde{\mathbf{M}}^T & \mathbf{0}^T & -u\tilde{\mathbf{M}}^T \\ \mathbf{0}^T & \tilde{\mathbf{M}}^T & -v\tilde{\mathbf{M}}^T \end{bmatrix} \mathbf{x} = \mathbf{0}.$$

Given n points, one obtains n above equations, which can be written in matrix equation as $\mathbf{L}\mathbf{x} = \mathbf{0}$, where \mathbf{L} is a $2n \times 9$ matrix. As \mathbf{x} is defined up to a scale factor, the solution is well-known to be the right singular vector of \mathbf{L} associated with the smallest singular value (or equivalently, the eigenvector of $\mathbf{L}^T \mathbf{L}$ associated with the smallest eigenvalue)².

Using the obtained homographies, Zhang defines in [103] the following constraints for each view:

$$\mathbf{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_2 = 0 \quad (2.9)$$

$$\mathbf{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_1 = \mathbf{h}_2^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_2 \quad (2.10)$$

where \mathbf{h}_i is the i -th column vector of matrix \mathbf{H} .

These constraints are derived from equation (2.8), from which $\mathbf{h}_1 = \mathbf{A}\mathbf{r}_1$ and $\mathbf{h}_2 = \mathbf{A}\mathbf{r}_2$. Hence, using the knowledge that \mathbf{r}_1 and \mathbf{r}_2 are orthonormal and therefore $\mathbf{r}_1^T \mathbf{r}_2 = 0$ and $\mathbf{r}_1^T \mathbf{r}_1 = \mathbf{r}_2^T \mathbf{r}_2$, Zhang defines the constraints (2.9) and (2.10). For a geometrical interpretation of these constraints, see [103].

²In \mathbf{L} , some elements are constant 1, some are in pixels, some are in world coordinates, and some are multiplication of both. This makes \mathbf{L} poorly conditioned numerically. Much better results can be obtained by performing a simple data normalization prior to running the above procedure [103].

Let $\mathbf{B} = \mathbf{A}^{-T} \mathbf{A}^{-1}$. If the skewless constraint $\gamma = 0$ is imposed, from (2.7) the following expression for matrix \mathbf{B} is obtained:

$$\mathbf{B} = \mathbf{A}^{-T} \mathbf{A}^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix} = \begin{bmatrix} \frac{1}{\alpha^2} & 0 & -\frac{u_0}{\alpha^2} \\ 0 & \frac{1}{\beta^2} & -\frac{v_0}{\beta^2} \\ -\frac{u_0}{\alpha^2} & -\frac{v_0}{\beta^2} & \frac{u_0^2}{\alpha^2} + \frac{v_0^2}{\beta^2} + 1 \end{bmatrix} \quad (2.11)$$

Since \mathbf{B} is symmetric, it is defined by the 6D vector \mathbf{b} :

$$\mathbf{b} = [B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33}]^T \quad (2.12)$$

Let the i -th column vector of matrix \mathbf{H} be $\mathbf{h}_i = [h_{i1}, h_{i2}, h_{i3}]^T$. Then the two constraints (2.9) and (2.10) are expressed as two homogeneous equations in matrix form as follows [103]:

$$\begin{bmatrix} \mathbf{v}_{12}^T \\ (\mathbf{v}_{11} - \mathbf{v}_{22})^T \end{bmatrix} \mathbf{b} = \mathbf{0} \quad (2.13)$$

where

$$\mathbf{v}_{ij} = [h_{i1}h_{j1}, h_{i1}h_{j2} + h_{i2}h_{j1}, h_{i2}h_{j2}, h_{i1}h_{j3} + h_{i3}h_{j1}, h_{i2}h_{j3} + h_{i3}h_{j2}, h_{i3}h_{j3}]^T$$

For n images of the model plane, one can stack equations of (2.13) as:

$$\mathbf{V}\mathbf{b} = \mathbf{0} \quad (2.14)$$

Since the skewless constraint $\gamma = 0$, i.e., $[0, 1, 0, 0, 0, 0]\mathbf{b} = 0$ has been imposed³, an additional equation to (2.14) can be added. This allows to obtain the remaining parameters by only using 2 views. The solution to (2.14) is the eigenvector of $\mathbf{v}^T \mathbf{v}$ associated with the smallest eigenvalue.

The matrix \mathbf{B} is estimated up to a scale factor, i.e., $\mathbf{B} = \lambda \mathbf{A}^{-T} \mathbf{A}^{-1}$ with λ an arbitrary scale. Without difficulty, the intrinsic parameters can be extracted by inspecting equation (2.11):

$$\begin{aligned} \lambda &= B_{33} - \frac{B_{13}^2 B_{22} + B_{23}^2 B_{11}}{B_{22} B_{11}} \\ u_0 &= -\frac{B_{13}}{B_{11}} & v_0 &= -\frac{B_{23}}{B_{22}} \\ \alpha &= \sqrt{\frac{\lambda}{B_{11}}} & \beta &= \sqrt{\frac{\lambda}{B_{22}}} \end{aligned}$$

Once \mathbf{A} is known, the extrinsic parameters for each view are readily computed from (2.7) as follows:

$$\begin{aligned} \mathbf{r}_1 &= \frac{\mathbf{A}^{-1} \mathbf{H}_1}{\|\mathbf{A}^{-1} \mathbf{H}_1\|} & \mathbf{r}_2 &= \frac{\mathbf{A}^{-1} \mathbf{H}_2}{\|\mathbf{A}^{-1} \mathbf{H}_2\|} \\ \mathbf{r}_3 &= \mathbf{r}_1 \times \mathbf{r}_2 & \mathbf{t} &= \frac{\mathbf{A}^{-1} \mathbf{H}_3}{\|\mathbf{A}^{-1} \mathbf{H}_1\|} = \frac{\mathbf{A}^{-1} \mathbf{H}_3}{\|\mathbf{A}^{-1} \mathbf{H}_2\|} \end{aligned}$$

For further details, see [103].

³In practice, we impose $[0, 1000, 0, 0, 0, 0]\mathbf{b} = 0$, since it makes somewhat stronger this constraint.

2.2.5 Determination of Corresponding Points

Once the acquisition geometry is known, the next problem to solve is determining the corresponding points on an image pair of the objects to be reconstructed. This task has been deeply studied, and we can find a lot of methods in the Computer Vision literature (see [104] for an example of one of these methods). We will not describe here these methods, but a common characteristic among them is that they use the epipolar line.

The epipolar line is the intersection of the image plane with the epipolar plane. This plane is defined by two projective rays, and therefore contains the focal points \mathbf{F}_1 , \mathbf{F}_2 and the three-dimensional point \mathbf{M} to be reconstructed. Figure 2.14 illustrates the concept for the angiographic case.

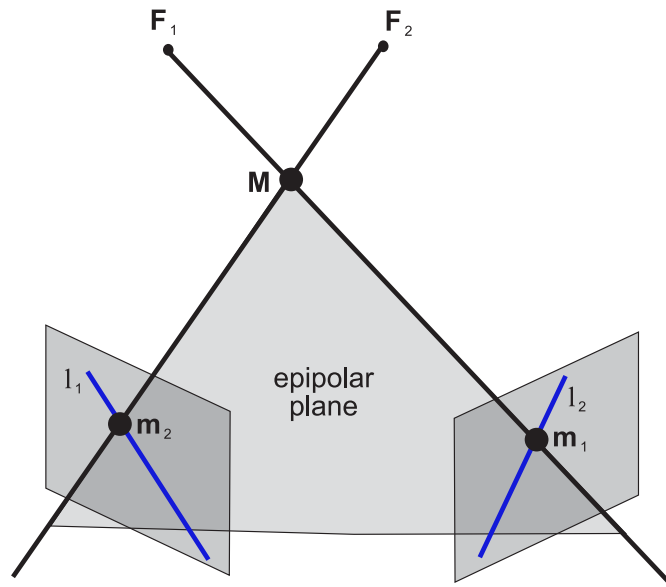


Figure 2.14: The epipolar plane for the angiographic case define the epipolar lines l_1 and l_2 .

Note that projections \mathbf{m}_1 and \mathbf{m}_2 of the three-dimensional point \mathbf{M} are on the epipolar lines. Provided that we know the acquisition geometry, and although the three-dimensional point \mathbf{M} is unknown, by determining the projection of point \mathbf{M} on one view we can obtain the corresponding epipolar line on the other view.

2.2.6 Computation of the Epipolar Line

Let $\mathbf{R}_1, \mathbf{t}_1$ be the extrinsic parameters of the first camera, and $\mathbf{R}_2, \mathbf{t}_2$ the extrinsic parameters of the second one. The transformation from the camera coordinates of the first camera to the camera coordinates of the second one is then described by the rigid transform (\mathbf{R}, \mathbf{t}) , where $\mathbf{R} = \mathbf{R}_2 \mathbf{R}_1^T$ and $\mathbf{t} = \mathbf{t}_2 - \mathbf{R}_2 \mathbf{R}_1^T \mathbf{t}_1$.

The *Essential Matrix* \mathbf{E} is defined from \mathbf{R}, \mathbf{t} as follows [32]:

$$\mathbf{E} = \mathbf{TR}$$

where

$$\mathbf{T} = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} \quad \text{with } \mathbf{t} = [t_1, t_2, t_3]^T$$

This matrix has a number of important properties, based on the fact that it is the algebraic representation of epipolar geometry (see [32] for details). Then, using the calibration matrices $\mathbf{A}_1, \mathbf{A}_2$, that relate an image point in pixels to the same point in camera coordinates, the *fundamental matrix* is defined as follows:

$$\mathbf{F} = \mathbf{A}_2^{-T} \mathbf{E} \mathbf{A}_1^{-1}$$

Given two projections $\mathbf{m}_1, \mathbf{m}_2$ from the three-dimensional point \mathbf{M} , the epipolar constraint can be written using the fundamental matrix \mathbf{F} as follows:

$$\mathbf{m}_2^T \mathbf{F} \mathbf{m}_1 = 0$$

Using the fundamental matrix \mathbf{F} , the equation of the epipolar line is very easy to obtain. Given the point \mathbf{m}_1 on the first image in pixel coordinates, the equation of the epipolar line on the second image $a_2u + b_2v + c_2 = 0$ is obtained as follows:

$$\begin{bmatrix} a_2 & b_2 & c_2 \end{bmatrix}^T = \mathbf{F} \mathbf{m}_1$$

On the other hand, given the point \mathbf{m}_2 on the second image, the equation of the epipolar line on the first image $a_1u + b_1v + c_1 = 0$ is derived using the following:

$$\begin{bmatrix} a_1 & b_1 & c_1 \end{bmatrix} = \mathbf{m}_2^T \mathbf{F}$$

As we stated before, in the Computer Vision frame, the epipolar line is therefore used to automatically determine the corresponding points. Given the projection of a point on one view, the corresponding point is searched only along the epipolar line defined on the other view. Figure 2.15 shows this concept for a stereo pair. The user has manually marked some points, which determine the corresponding epipolar lines on the other view. It is clear that given the nature of these images, a simple correlation procedure shall perform quite well.

However, in the angiographic frame the automatic correspondence determination presents two main difficulties. First, we do not have in general two similar views but two different views as angulated as possible. While this fact improves the accuracy of the three-dimensional reconstruction, the two views will be very different one from another. The second difficulty is the big degree of ambiguity present on this kind of images, since all the vessels are similar. Figure 2.16 shows an example of this situation. The epipolar line crosses several vessels, and it is difficult to determine which is the correct one without additional knowledge.

That is the reason why most researchers propose to obtain the three-dimensional reconstruction of the vessel using corresponding points, which are manually marked by the user.

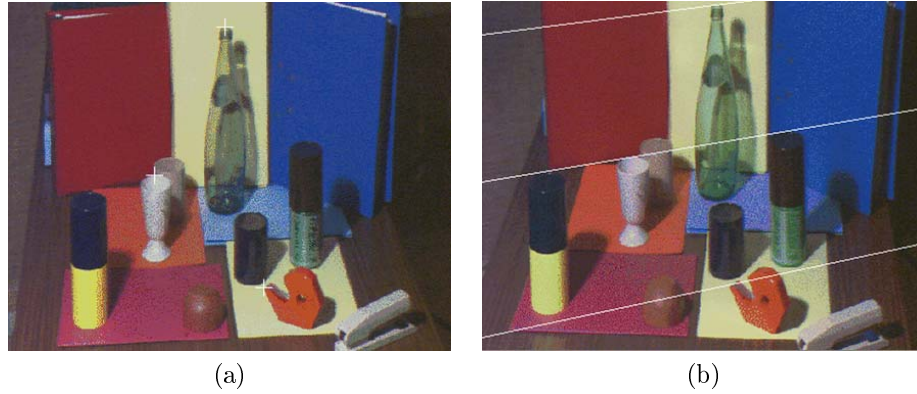


Figure 2.15: The points marked by the user on image (a) determine epipolar lines on image (b). These lines are useful to constraint the search of corresponding points.

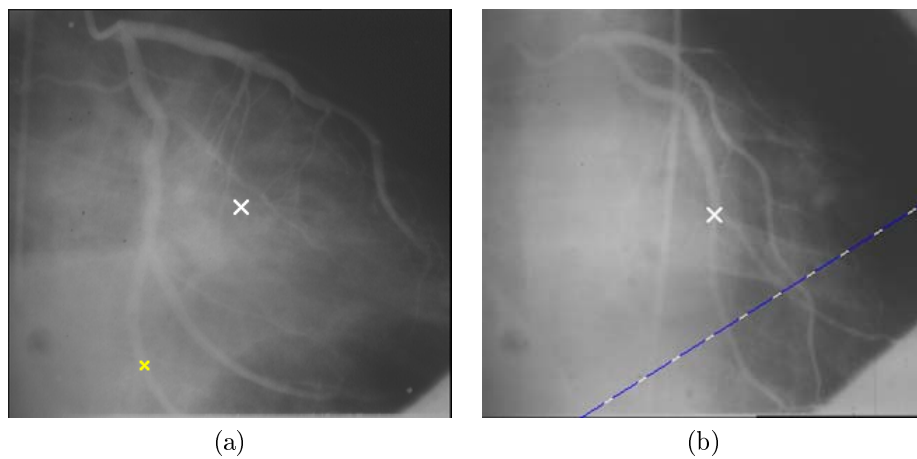


Figure 2.16: Epipolar line on X-ray angiography is not enough to determine corresponding points due to the ambiguities.

However, as stated in Chapter 1, even when the user is helped by the epipolar line [22] to match points in different views, this is a quite tedious task if high accuracy is needed. Moreover, the curve is directly interpolated among marked points, and therefore the user must mark a lot of corresponding points to accurately follow the shape of the vessel, which in many cases can be quite wavy. Hence, we propose the use of biplane snakes to perform this task, which are derived from the physics-based model of snakes [42].

2.3 Image Enhancement Techniques

As we stated in Chapter 1, one of the most difficult problems that we find when approaching the 3D reconstruction of the vessels is the automatic vessel segmentation. The main difficulty relies on the fact that X-ray cardiac angiographies are very noisy, and small vessels with low contrast are hard to detect.

In this section, we describe two techniques that are applicable in order to improve the quality of the X-ray images of the vessels. First we give some insight on anisotropic diffusion techniques. Special attention is paid to a crease enhancement diffusion process, since the vessels can be described as valleys on the image. Second, we describe a specific technique developed to enhance the vessels on cardiac images.

2.3.1 Anisotropic Diffusion

Diffusion processes derive from Fick's law [33] and the continuity equation. Fick's law expresses that a gradient concentration leads to a flow which compensates for it. If we include Fick's law in the continuity equation, which expresses that mass is only transported but can be neither created nor destroyed, we obtain the diffusion equation:

$$\frac{\partial \mathbf{u}}{\partial t} = \operatorname{div}(\mathbf{D} \cdot \nabla \mathbf{u}) \quad (2.15)$$

where \mathbf{D} is a diffusion tensor and \mathbf{u} corresponds to the mass concentration (for images, to the gray level at each pixel). The diffusion tensor defines the diffusion process; for example, if we choose \mathbf{D} as the identity matrix we will obtain a filtering equivalent to the convolution of the function \mathbf{u} by a Gaussian kernel [46, 97]. Another type of filtering, proposed by Perona and Malik (see [72]), can be seen as a diffusion filtering:

$$\frac{\partial \mathbf{u}}{\partial t} = \operatorname{div} \left(\frac{1}{1 + \frac{\|\nabla \mathbf{u}\|^2}{K^2}} \cdot \nabla \mathbf{u} \right)$$

where K controls the diffusion strength. This filtering avoids the blurring at the edges.

In all these processes, the diffusion direction is always collinear to the image gradient $\nabla \mathbf{u}$ and its perpendicular $\nabla \mathbf{u}^\perp$. Weickert proposed in [95] the tuning of the diffusion directions according to the dominant orientation of each image pixel, in order to enhance linear structures. This direction can be obtained through structure tensor analysis [41]. Then, Weickert constructed the diffusion tensor which eigendirections coincide with the eigenvectors of the structure tensor but have different eigenvalues; namely:

$$\begin{aligned} \lambda_1 &= \alpha \\ \lambda_2 &= \alpha + (1 - \alpha) e^{\frac{-C}{(\mu_1 - \mu_2)^{2m}}}, \end{aligned}$$

where μ_1 and μ_2 are the eigenvalues of the structure tensor and m , C , and α are parameters controlling the diffusion strength in the noncoherence direction. However, Sole et al. state in [85] that with this filtering many junctions disappear and original structures are deformed. Therefore, they propose another diffusion tensor. First, they compute the *normalized Hessian* by first regularizing the initial image \mathbf{u} to obtain

robust derivatives (derivation step) and then regularizing the tensor field (integration step) as follows:

$$\mathbf{H}_{\sigma,\rho} = \mathbf{G}_\rho * \left[\frac{1}{\sqrt{1 + \|\nabla \mathbf{u}_\sigma\|^2}} \begin{pmatrix} \frac{\partial^2 \mathbf{u}_\sigma}{\partial x^2} & \frac{\partial^2 \mathbf{u}_\sigma}{\partial x \partial y} \\ \frac{\partial^2 \mathbf{u}_\sigma}{\partial x \partial y} & \frac{\partial^2 \mathbf{u}_\sigma}{\partial y^2} \end{pmatrix} \right] \quad \text{with } \mathbf{u}_\sigma = \mathbf{G}_\sigma * \mathbf{u}, \quad (2.16)$$

where \mathbf{G}_ρ is a Gaussian kernel of size ρ (*integration scale*) and σ is the *differentiation scale* that controls the size of the neighbourhood in which an orientation is dominant. This tensor provides a multilocal version of the eigendirections and normalized eigenvalues of the Hessian. Let v_1 be the first eigendirection corresponding to the highest eigenvalue in absolute value ($\tilde{\kappa}_1 = |\kappa_1|$) and v_2 be the first eigendirection corresponding to the lowest one. The *creaseness diffusion tensor* is then constructed to have the same eigenvectors v_1 and v_2 and the associated eigenvalues:

$$\begin{aligned} \lambda_1 &= \epsilon & \epsilon &\in (0, 1) \\ \lambda_2 &= \alpha\mu_r + \beta\mu_v & \alpha, \beta &\in [0, 1], \end{aligned}$$

where α and β are parameters controlling the diffusion strenght in the presence of ridges and valleys, respectively. The value ϵ ensures the semidefinite propoperty of the diffusion tensor, and should be small. μ_r and μ_v reach their highest values in the presence of ridges and valleys, respectively, and have the following expression:

$$\mu_r = \begin{cases} \frac{\tilde{\kappa}_1 - \tilde{\kappa}_2}{\tilde{\kappa}_1 + \tilde{\kappa}_2} & \text{if } \kappa_1 < 0 \\ 0 & \text{if } \kappa_1 \geq 0 \end{cases} \quad \mu_v = \begin{cases} \frac{\tilde{\kappa}_1 - \tilde{\kappa}_2}{\tilde{\kappa}_1 + \tilde{\kappa}_2} & \text{if } \kappa_1 > 0 \\ 0 & \text{if } \kappa_1 \leq 0 \end{cases}$$

Since this diffusion process enhances the creases on the image, and since the vessels can be regarded as creases, it seems appropriate for enhancing the image prior the vessel segmentation. However, the ridge/valley measure does not take into account the image contrast, and therefore, when applying it to regularize the vessels, some undesired background structures are also enhanced. Also, the vessels appear at different scales, and therefore a multiscale approach is necessary.

2.3.2 Frangi's Filter

Although the filter proposed by Frangi et al. in [34] is applicable both for 3D and 2D images, in this thesis we will focus on the two-dimensional case.

The main idea is that a vessel can be regarded as a tubular structure. Since vessels appear in different sizes, a measurement scale which varies within a certain range is introduced. To analyze the local behaviour of an image, denoted by \mathbf{u} , we can use the information provided by the Hessian computed at a given scale σ as follows:

$$\mathbf{H}_\sigma(\mathbf{u}) = \sigma^2 \begin{pmatrix} \frac{\partial^2 \mathbf{u}_\sigma}{\partial x^2} & \frac{\partial^2 \mathbf{u}_\sigma}{\partial x \partial y} \\ \frac{\partial^2 \mathbf{u}_\sigma}{\partial x \partial y} & \frac{\partial^2 \mathbf{u}_\sigma}{\partial y^2} \end{pmatrix} \quad (2.17)$$

where \mathbf{u}_σ is the initial image \mathbf{u} convolved by a Gaussian kernel of size σ . Note that the Hessian is normalized by multiplying it by σ^2 in order to obtain a fair comparison of the response of differential operators at multiple scales [48].

By analysing the eigenvectors and eigenvalues of the Hessian, Frangi et al. propose two descriptors of the local structure useful for vessel detection. Let h_k be the eigenvalue with the k -th smallest magnitude ($|h_1| \leq |h_2|$) and \mathbf{v}_k the corresponding eigenvector.

First, a *blobness* measure \mathcal{R}_β is defined in 2D as follows:

$$\mathcal{R}_\beta = \left| \frac{h_1}{h_2} \right| \quad (2.18)$$

Note that \mathcal{R}_β is near to 0 for tubular structures and near to 1 when no privileged direction is present. This ratio is grey-level invariant (i.e. it remains constant under intensity re-scalings) and therefore only capture the local geometric information of the image. However, vessel structures are darker than the background and occupy a (relatively) small volume of the dataset. To quantify this, Frangi et al. propose the following *contrast* descriptor:

$$\mathcal{S} = \|\mathbf{H}_\sigma\|_F = \sqrt{h_1^2 + h_2^2} \quad (2.19)$$

which will be low in the background where no structure is present and in regions with high contrast will become larger, since at least one of the eigenvalues will be large.

Using these two descriptors, a vesselness function is defined for each scale σ as follows:

$$\mathcal{V}_0(\sigma) = \begin{cases} 0 & \text{if } h_2 < 0 \text{ (brighter structure),} \\ \left(1 - e^{-\frac{\mathcal{S}^2}{2c^2}}\right) e^{-\frac{\mathcal{R}_\beta^2}{2\beta^2}} & \text{otherwise} \end{cases} \quad (2.20)$$

where β and c are parameters to control the sensitivity of the filter to the measures \mathcal{R}_β and \mathcal{S} . Frangi et al. propose a value of 0.5 for β and half the maximum Hessian norm for c . These are the values that we use in this work.

The vesselness measure is analyzed at different scales σ . Since the response will be maximum at a scale that approximately matches the size of the vessel to detect, the integrated filter response is computed as follows:

$$\mathcal{V}_0 = \max_{\sigma_{min} \leq \sigma \leq \sigma_{max}} \mathcal{V}_0(\sigma) \quad (2.21)$$

This filter gives good noise and background suppression results, but we have noticed problems to segment the vessels from two-dimensional X-ray cardiac angiographies. One of the reasons is that this method does not consider the coherence, that is the fact that if a given pixel is considered to be part of a vessel, the neighbour pixels in the vessel direction are also very likely to correspond to a vessel.

2.4 Deformable Models

Medical images are often corrupted by noise and sampling artifacts. Even if we use image enhancement techniques, classical segmentation techniques either fail completely or need a post-processing step when dealing with this type of images. To

address these difficulties, *deformable models* have been extensively studied and widely used, with promising results. Hence, in this thesis we use a deformable model technique to segment the vessels on the images. Moreover, we use an special type of deformable models, which we called biplane snakes, that inherently solves the corresponding points determination problem.

In this section, we give some background knowledge about one-dimensional deformable models or *snakes*, which are used to develop biplane snakes in Chapter 4.

2.4.1 Deformable Models Overview

The term *deformable models* first appeared in the work of Terzopolous and his collaborators in [89, 42]. Since the publication of deformable models in [42], they have grown to be one of the most active and successful research areas in image segmentation and tracking. Other names, as snakes, active contours or surfaces, ballons and deformable contours or surfaces, have been used in the literature to refer the deformable models.

The main idea of deformable models is to adapt a shape on an image domain. Different approaches are present in the literature, by changing the representation of the shape to be adapted or the method to adapt it to the image. Basically, there are two types of deformable models: *parametric deformable models* and *geometric deformable models*. Parametric deformable models are curves or surfaces explicit in their parametric forms during deformation. This representation allows direct interaction with the model and can lead to a compact representation for fast real-time implementation. However, it is difficult to adapt model topology, such as splitting or merging parts during deformation. Geometrical deformable models, on the other hand, can handle topological changes in a natural way. These models are based on the theory of curve evolution [80, 65, 43, 44], and the level set method [68, 82], and the curves or surfaces are represented implicitly as a level set of a higher-dimensional scalar function. Hence, their parametrizations are computed only after complete deformation, and thus facilitating topological adaptivity.

We have chosen parametric deformable models in this thesis for two reasons: first, the particular application of this thesis does not require topology adaptivity. Moreover, we do not have control on the topology of the final result. And second, due to the definition of geometric deformable models, the obtained shapes, as level set curves are closed curves instead of open curves, which are the ones that we aim to reconstruct. Therefore, in the remainder of this section we focus on the basics of the parametric deformable models. For a wider and more detailed description on deformable models, both parametric and geometric, see [99].

2.4.2 Parametric Deformable Models

Deformable models, as physics-based models, are defined in terms of energies. The deformable model tries to minimize its external energy and in this way approaches the image features, where this energy is minimum, whereas internal forces keep the smooth shape. Therefore, the total energy of the deformable model is given by the

following functional:

$$E_{total}(\mathbf{Q}(\mathbf{u})) = E_{int}(\mathbf{Q}(\mathbf{u})) + E_{ext}(\mathbf{Q}(\mathbf{u})) \quad (2.22)$$

where $\mathbf{Q}(\mathbf{u})$ is the deformable model shape and $\mathbf{u} = (u_1, u_2, \dots, u_n)$ is the internal parameter of the deformable model. For curves, we will have an one-dimensional internal parameter, and for surfaces, we need a two-dimensional internal parameter. Therefore, the deformable models consisting of a curve are one-dimensional deformable models, and the deformable surfaces are two-dimensional deformable models. Note that this notion of dimensionality is independent of the output space. In the remaining of this thesis we will focus on one-dimensional deformable models and hence parameter \mathbf{u} will be uni-dimensional and we will denote it as u .

A method to minimize the energy expression (2.22) of the deformable model was presented in [42] for one-dimensional snakes. The main idea is that minimizing the energy functional equals to solve the following equation:

$$\frac{\partial E_{int}}{\partial u} + \frac{\partial E_{ext}}{\partial u} = 0 \quad (2.23)$$

To gain some insight about the physical behaviour of deformable models, we can view equation (2.23) as a force balance equation:

$$\mathbf{F}_{int}(\mathbf{Q}(u)) + \mathbf{F}_{ext}(\mathbf{Q}(u)) = 0$$

Hence, a solution of the energy minimization procedure can be viewed as one which achieves an equilibrium between the internal and external forces in the energy equation and reaches a minimum of the total energy.

To find the shape $\mathbf{Q}(u)$ which satisfies equation (2.23), an $O(n)$ iterative technique is proposed in [42]. The proposed approach uses sparse matrix methods, where n is the number of data points which represent the snake. Each iteration effectively takes implicit Euler steps with respect to the internal energy and explicit Euler steps with respect to the external energy. Hence, an expression of $\mathbf{Q}_t(u)$ in terms of $\mathbf{Q}_{t-1}(u)$ can be obtained, where $\mathbf{Q}_0(u)$ is the initial shape of the deformable model.

According to the Newton's second law, the dynamics of $\mathbf{Q}(t, u)$ must satisfy the following:

$$\mu \frac{\partial^2 \mathbf{Q}}{\partial t^2} = \mathbf{F}_{damp}(\mathbf{Q}) + \mathbf{F}_{int}(\mathbf{Q}) + \mathbf{F}_{ext}(\mathbf{Q})$$

where μ is a coefficient that has a mass unit and \mathbf{F}_{damp} is the damping (viscous) force, defined as $-\gamma \frac{\partial \mathbf{Q}}{\partial t}$, with γ being the damping coefficient. In image segmentation, the mass coefficient μ in front of the inertial term is usually set to zero, since the inertial term may cause the contour to pass over the weak edges [99]. Therefore, the dynamics of a deformable model becomes:

$$\gamma \frac{\partial \mathbf{Q}}{\partial t} = \mathbf{F}_{int}(\mathbf{Q}) + \mathbf{F}_{ext}(\mathbf{Q})$$

Using this formulation, external forces can be broken down into more manageable terms. For example, one could define the external force as the sum of distance-map potential forces and pressure forces, as we will see later.

The way we define each of these forces (internal and external) will determine the behaviour of the deformable model and therefore the final segmentation result.

2.4.3 Internal Energy

The deformable models are typically modeled to have restrictions to the stretching and bending of the model. This fact should prevent the model from being caught by spurious features corresponding to noise. A common approach is to define the internal energy associated to the model to decrease when the magnitude of first and second derivatives reduce. With this purpose, two components of the internal energy are defined by means of a sum of the *membrane* energy and *thin-plate* energy[42] as follows:

$$E_{int}(\mathbf{Q}) = E_{membrane}(\mathbf{Q}) + E_{thin-plate}(\mathbf{Q}) \quad (2.24)$$

Each of these energies have a different effect on the snake shape:

- The *membrane* energy is given by the first derivative, and keeps the continuity between pixels:

$$E_{membrane}(\mathbf{Q}) = \int \alpha(u) \left(\frac{\partial \mathbf{Q}}{\partial u} \right)^2 du \quad (2.25)$$

This energy has the side effect of stretching the shape in absence of external forces. While this property is often desirable when deforming a closed shape, when using open curves or sheets, parameter α is normally set to zero.

- The *thin-plate* energy is computed from the second derivative and tries to keep a low curvature, avoiding the bending of the model:

$$E_{thin-plate}(\mathbf{Q}) = \int \beta(u) \left(\frac{\partial^2 \mathbf{Q}}{\partial u^2} \right)^2 du \quad (2.26)$$

The influence of the *thin-plate* energy is illustrated in figure 2.17. A twisted tube under the influence of thin-plate energy tends to become a cylinder. Note that it is not stretched, since the *membrane* parameter has been set to zero.

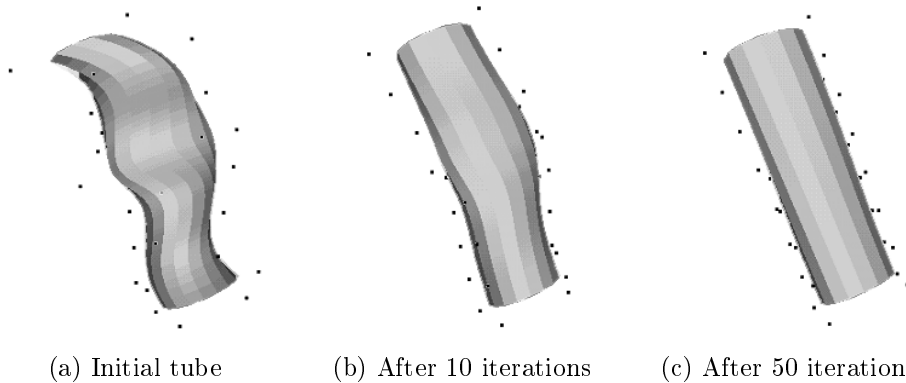


Figure 2.17: *Thin-plate* energy discourage the bending of the shape.

Hence, from equations (2.24),(2.25) and (2.26) we can obtain the following expression for the internal energy:

$$E_{int}(\mathbf{Q}) = \int \alpha(u) \left(\frac{\partial \mathbf{Q}}{\partial u} \right)^2 du + \int \beta(u) \left(\frac{\partial^2 \mathbf{Q}}{\partial u^2} \right)^2 du$$

where $\alpha(u), \beta(u)$ are often assumed constant to simplify the energy minimization procedure and hence the following expression is used:

$$E_{int}(\mathbf{Q}) = \alpha \int \left(\frac{\partial \mathbf{Q}}{\partial u} \right)^2 du + \beta \int \left(\frac{\partial^2 \mathbf{Q}}{\partial u^2} \right)^2 du \quad (2.27)$$

Although there are a lot of alternative ways of expressing the internal energy, this is one of the most popular ways, due to its simplicity and possibility to reduce the energy-minimizing procedure to linear equations. That is the reason to use expression (2.27) in our work.

2.4.4 External Energy

Here we will describe several kinds of external forces for deformable models. First, we will describe the potential forces, which are derived from the image, generally from the output of a feature detection algorithm. This kind of forces guides the deformable model to the desired objects. However, the different types of potential forces have the problem of convergence into object concavities. To solve this problem, several approaches have been presented in the literature. Here we will describe two of them: Pressure forces and Generalized Gradient Vector Flow. Finally, we will describe the interactive forces, which are derived from the user's interaction to help the model to converge to the desired location, or to avoid a region on the image.

Potential Forces

Although there are some subtle differences, the general expression of potential forces is as follows:

$$\mathbf{F}_{pot}(\mathbf{Q}) = -\nabla P(\mathbf{Q})$$

where P is a potential derived from the output of a feature detection algorithm which is applied to the image. The feature detection method to be applied depends on the application. For instance, if we aim to segment the contour of a shape, we typically run an edge detection algorithm, as a convolution with a Sobel kernel. In this thesis, we aim to segment the vessel centerline, and therefore we need a more sophisticated feature detector. Anyway, once the features have been detected, the potential is computed in the same way, no matter which feature detector is used.

All the techniques to construct a potential field are characterized by the fact that the potential is continuous, and that the valleys of potential P are at the location of the extracted image features. Elsewhere the potential forces push the snake towards the valleys of the potential field following its slope as if the snake was under the effect of gravity.

The different types of potential forces are derived from the method to compute the potential from the features. Here we explain three different types of potential forces: gaussian potential force, multiscale gaussian potential force, and distance potential force.

Gaussian potential field. The potential is computed as the output result of the features detector $F(x, y)$ convolved by a Gaussian function $G_\sigma(x, y)$ as follows:

$$P(x, y) = -G_\sigma(x, y) * F(x, y)$$

For the case of edge detection, this type of potential is computed as [42]:

$$P(x, y) = -|\nabla (G_\sigma(x, y) * I(x, y))|^2$$

where $I(x, y)$ is the original image.

This type of force has the problem of the attraction range: while increasing σ extends the attraction range, larger σ can also cause a shift on the feature location, resulting in a less accurate result.

Multiscale Gaussian potential field. This approach was proposed in [42] to remedy the problems of the Gaussian potential force. The idea is to first use a large σ to create the potential force. When the contour reaches an equilibrium, the value of σ is reduced to allow a more accurate result. Although this approach effectively extends the attraction range without compromising accuracy, there is no established theorem for how to schedule changes in σ .

Distance potential field. The potential is defined as a function of the distance map from the image features using Euclidean distance[18] or Chamfer distance[4]:

$$P(x, y) = d(x, y) \text{ or alternatively, } P(x, y) = -e^{-d(x, y)^2}$$

This procedure has the advantage, compared to the Gaussian forces, that it provides influence of distant features and at the same time does not shift the features. However, weak features attract the snake with the same force as strong features.

Multi-thresholded Distance potential force. This approach is used in [75] to surmount the above problem. Instead of applying a multiple threshold to the detected features and computing the distance map from the resulting binary image, several distance maps are computed for different threshold values. Hence, the snake first deforms on a distance map computed from the stronger features. Once it has stabilized, the potential force is replaced by a distance map computed from weaker features. Using this approach we prevent the snake of being attracted by weak edges at the beginning of the deformation, although we take them into account to obtain a finer fitting.

Figure 2.18 illustrates the differences between the Gaussian force and the distance potential force. For the sake of a better visualization, brighter values correspond to low potential values, and darker, to high potential values.

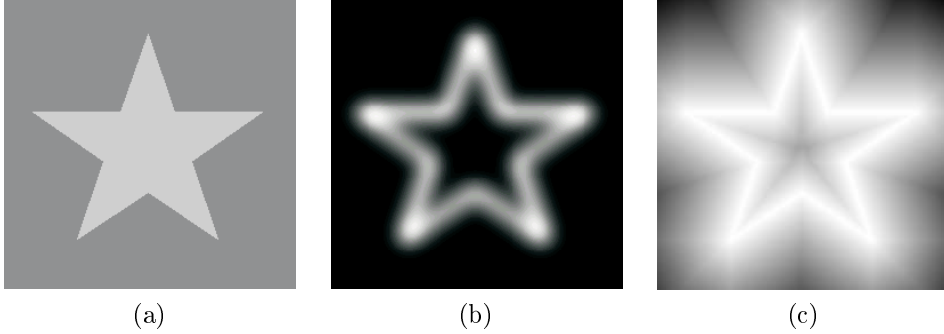


Figure 2.18: Two approaches to obtain a potential. (a) is the original image, from which we have detected the image contours, (b) is the output result of the features detector convolved by a Gaussian function and (c) is a function of the distance map from the image features (in this case, the identity function).

Although multiscale Gaussian potential force and especially distance potential force solve the problem of feature capture range, when reconstructing wavy shapes the snake has problems to converge into concavities. This kind of shapes is very common on X-ray angiography, since the coronary vessels are very likely to be wavy.

Here we present three approaches that have been proposed in the literature to solve this problem: Pressure Forces [14], Dynamic Distance Force[19, 54] and Generalized Gradient Vector Flow[100].

Pressure Force

Cohen proposed in [14] to use the pressure force together with the potential force computed from the image. Since this force can either inflate or deflate the model, a deformable model that uses a pressure force is also known as a balloon.

The pressure force is defined as:

$$\mathbf{F}_p(\mathbf{Q}) = w_p \mathbf{N}(\mathbf{Q})$$

where $\mathbf{N}(\mathbf{Q})$ is the outward unit normal of the model \mathbf{Q} and w_p is a constant weighting parameter. The sign of w_p determines whether to inflate or deflate the model and is typically chosen by the user. The value of w_p determines the strength of the pressure force. It must be carefully selected so that the pressure force is slightly smaller than the potential force at significant features, but large enough to pass through weak or spurious features. When the model deforms, the pressure force keeps inflating or deflating the model until it is stopped by the potential force. This has two advantages: it removes the requirement to initialize the model near the desired object and the pressure forces can push the model into object concavities. The disadvantage in using pressure forces is that they may cause the deformable model to cross itself and form loops[88], although this fact should be alleviated by the internal forces.

However, in the context of vessel centerline segmentation, we are working with open curves, and hence this approach is not appropriate. Therefore, we need another method to force the snake to converge into concavities present on wavy vessels.

Dynamic Distance Force

The dynamic distance force [19, 54] is derived by computing a signed distance $D(\mathbf{Q})$ at each point on the deformable contour or surface. This signed distance is calculated by determining the closest image feature along the model normal direction. Each time the model is deformed, the distance values are recomputed. To avoid confusion with outliers and to reduce the computation time, the maximum search distance is specified using distance threshold D_{max} .

Hence, the dynamic distance force has the following expression:

$$\mathbf{F}_D(\mathbf{Q}) = w_D \frac{D(\mathbf{Q})}{D_{max}} \mathbf{N}(\mathbf{Q})$$

where w_D is a constant weighting parameter which determines the strength of this force and $\mathbf{N}(\mathbf{Q})$ corresponds to the outward unit normal.

This force has the weakness that the relatively time-consuming 1D search along the normal direction must be performed each time the model deforms. Although lowering the distance threshold can reduce the computation time, the attraction range also decreases [99].

Generalized Gradient Vector Flow (GGVF)

Xu and Prince presented in [100] a regularization method for the potential gradient in order to improve active contour convergence into indentations. This method is called Generalized Gradient Vector Flow (GGVF). In this work, we have used GGVF as external force to improve the convergence to wavy vessels.

GGVF has many desirable properties as an external force for snakes, since it improves active contour convergence into long, thin boundary indentations, has an extended capture range and excellent performance on noisy and real medical images [100]. The main idea is to replace the external force term $\mathbf{F}_{ext} = -\nabla \mathbf{V}(\mathbf{Q})$ by $\nu(\mathbf{Q})$, where ν is the GGVF defined as the equilibrium solution of the following system of partial differential equations:

$$\begin{aligned} \nu_t &= g(|\nabla f|) \nabla^2 \nu - h(|\nabla f|) (\nu - \nabla f) \\ \nu_0 &= \nabla f \end{aligned}$$

where f corresponds typically to a Gaussian smoothing of the detected target features of the image.

The first term on the right $\nabla^2 \nu$ is referred to as the *smoothing term* since this term will produce a smoothly varying vector field. The second term $(\nu - \nabla f)$ is referred to as the *data term* since it encourages the vector field ν to be close to ∇f computed from the data. The weighting functions $g(\cdot)$ and $h(\cdot)$ apply to the smoothing and data terms, respectively. Xu and Prince use in [100] the following weighting functions for GGVF:

$$\begin{aligned} g(|\nabla f|) &= e^{-\frac{|\nabla f|}{k}}, \\ h(|\nabla f|) &= 1 - e^{-\frac{|\nabla f|}{k}} \end{aligned}$$

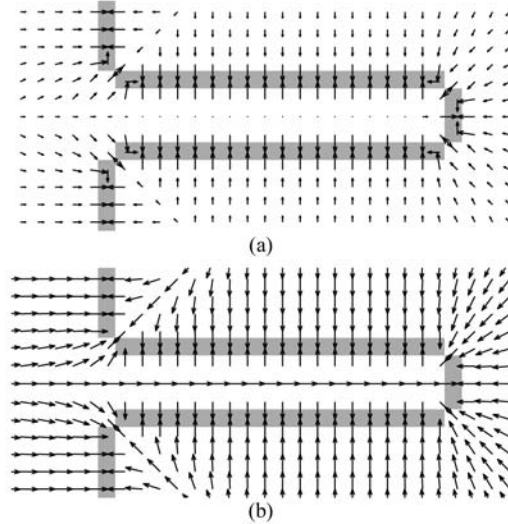


Figure 2.19: Difference between distance map gradient (a) and GGVF external forces (b) when dealing with thin concavities.

The GGVF field computed using this pair of weighting functions will conform to the distance map gradient near the features, but will vary smoothly away from them. The specification of K determines to some extent the degree of tradeoff between field smoothness and gradient conformity. For numerical implementation details of GGVF, see [101, 100].

Figure 2.19 illustrates the differences between the gradient of the distance map and GGVF when thin concavities are present. Figure 2.19(a) shows the distance map gradient. In this case, the distance map gradient is zero into the concavity, and thus the snake will fail to capture the contour, since none external force attracts the snake into the concavity. On the other hand, it is clear that the GGVF on figure 2.19(b) is able to attract the snake to the target contour.

2.4.5 Numerical Representation of the Snake

Until here we have described the different forces that can guide the deformation of the snake and the different methods to implement this deformation. However, it still remains to discuss the different numerical representations for curve $\mathbf{Q}(u)$.

Among the numerous numerical representations used to implement deformable models, the more representative are two: point-based [42] deformable models and B-Spline based [61] deformable models.

The B-Spline based representation has several important advantages [61]:

Local control The B-Spline blending functions have local control. This means that the modification of one control point only has local effect on the curve (see figure 2.20).

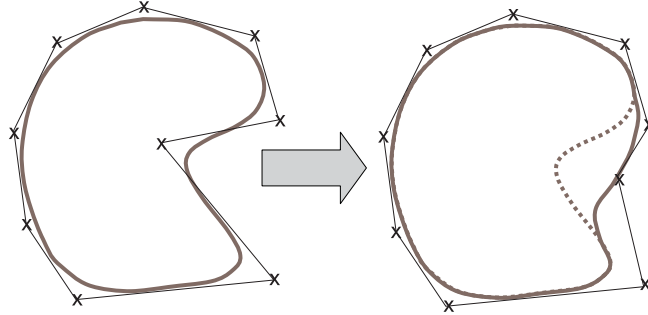


Figure 2.20: B-Splines have local control.

Controlled continuity The B-Splines with degree d have continuity C^{d-1} . This is of interest, since first and second derivatives are needed in order to compute the internal energy of the deformable model. Nevertheless, one can modify this continuity by introducing multiple knots. Thus, if μ is the multiplicity of a knot, the continuity in this knot will be $C^{d-1-\mu}$. Figure 2.21 illustrates this concept. Menet et al. use knot multiplicity in order to detect corners [61].

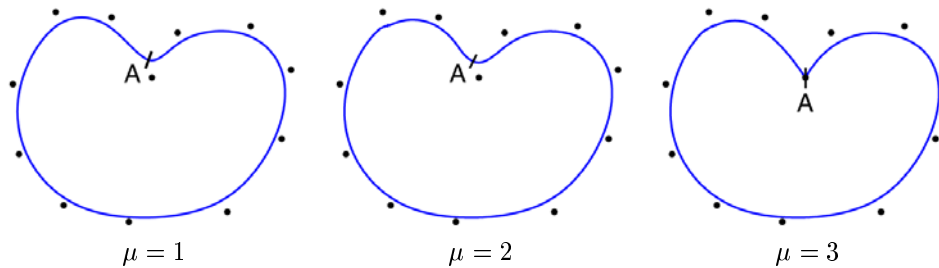


Figure 2.21: Obtained B-Spline interpolation with degree $d = 3$ for different knot multiplicities (μ) at point A .

Compact representation The B-Spline based model is represented only by a small number of control points, compared to the number of points needed in a point-based representation. This is very important from the point of view of computational efficiency.

Note that a point-based representation is equivalent to consider a linear uniform B-Spline whose control points coincide with the points of this representation. At this point, we can conclude that in this work we will use the B-Spline based representation.

B-Spline Representation

Given a set of control points \mathbf{V}_i with $i = 1 \dots n$, the B-spline curve $\mathbf{Q}(u)$ is computed as follows:

$$\mathbf{Q}(u) = \sum \mathbf{V}_i B_i(u)$$

where $B_i(u)$ are B-spline blending functions. There is a blending function for each control point \mathbf{V}_i . These functions are recursive piecewise polynomials with finite support.

The blending splines are defined by a sequence of real values k_j called *knots* defined in increasing order, i.e. $k_j \leq k_l$ for $j < l$. For a closed curve, we will need n knots to define the blending functions, while for open curves, we will use $n + d$ knots. Given a knot sequence, the blending splines of degree one are defined as:

$$B_{j,1}(u) = \begin{cases} 1 & \text{if } k_j \leq u < k_{j+1} \\ 0 & \text{otherwise} \end{cases}$$

and higher-order splines are recursively computed by blending lower-order splines as follows:

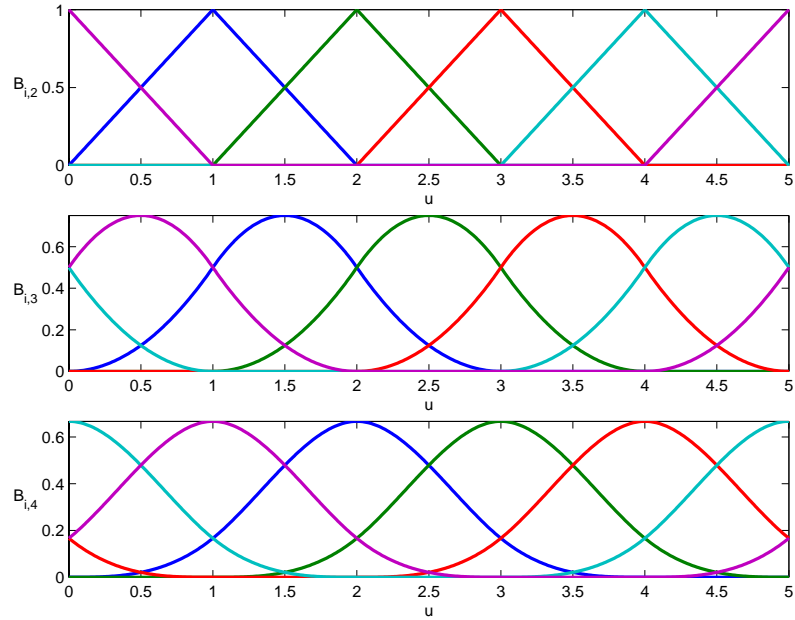
$$B_{j,d+1}(u) = \frac{u - k_j}{k_{j+d} - k_j} B_{j,d}(u) + \frac{k_{j+d+1} - u}{k_{j+d+1} - k_{j+1}} B_{j+1,d}(u)$$

The blending functions $B_{j,d+1}$ are therefore polynomials of degree d with finite support. They are globally of continuity C^{d-1} . A C^0 curve is simply a continuous curve. A C^1 curve is a curve whose derivative is C^0 . In fact, a C^n curve is derivable n times. As stated before, this continuity can however be modified by repeating successive knot values.

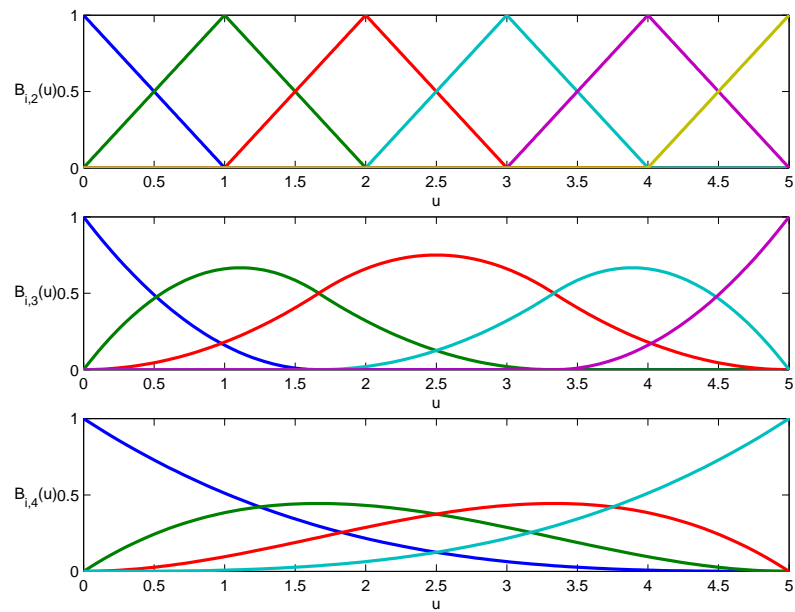
Figure 2.22 illustrates the B-spline blending functions for different values of d . For the open case, the knot multiplicity is $d + 1$ at the beginning and the end of the curve to force the curve to interpolate the first and last control points.

In appendix A, we show how to implement an one-dimensional B-spline based deformable model using a matrix representation. This approach is very simple to implement on a matrix manipulator as Matlab, or using any standard matrix library.

As stated before, we use an special type of B-spline based one-dimensional deformable models, called *biplane snakes*, to segment the vessels on the X-ray images and to obtain its three-dimensional reconstruction. However, before approaching the three-dimensional reconstruction, we first need to calibrate the acquisition system.



(a) Closed curve.



(b) Open curve.

Figure 2.22: B-spline blending functions for $d = \{1, 2, 3\}$. The open curve begins and ends at the extrema control points.

Chapter 3

Calibration

As stated in section 2.2.4, in order to obtain an accurate 3D reconstruction of a point \mathbf{M} from two projections \mathbf{m}_1 and \mathbf{m}_2 , we need to accurately know the intrinsic and extrinsic parameters of the cameras, which are condensed on its projection matrices $\mathbf{P}_i = \mathbf{A}_i [\mathbf{R}_i \quad \mathbf{t}_i]$. Moreover, since X-ray angiographies suffer from pincushion and other geometrical distortions (see section 2.1.3), we need to accurately estimate these distortions in order to perform an accurate correction and hence obtain a precise 3D reconstruction.

In this chapter, we present a new method to obtain an accurate estimate of the intrinsic and extrinsic parameters of a X-ray angiographer, as well as of the geometrical distortion, from the parameters provided by the system. They are derived using the values of the rotation angle (α), the angulation angle (β) and the distance from the X-ray source to the image intensifier (*SID*).

To this aim, we develop a model of the X-ray angiographer, and propose a flexible calibration method to estimate the parameters of this model, which will depend on each particular system. The calibration method only needs a radio-opaque calibration grid and around ten minutes of acquisition time. Moreover, the calibration procedure must be performed only once (or at least after long periods of time), since it remains valid even three months after the calibration. Provided that most calibration methods for X-ray angiography need to be performed by each acquisition, this is a strong point for our method. Also, the calibration phantom is commonly available at the catheterism laboratories, and, anyway, it is easier to construct than others proposed in the literature.

In section 3.1, we explain how we model and predict the geometrical distortion¹. Then, in section 3.2, we propose a new model of the movement of the C-arm, and a method to calibrate its parameters. Experimental results of the methods proposed in this chapter are presented in Chapter 5.

¹Section 3.1 has been extracted from the article: "Predictive (un)distortion model and 3D Reconstruction by Biplane Snakes", by C. Cañero et al., published in *IEEE Transactions on Medical Imaging* (see reference [9] for details).

3.1 Distortion Correction

As stated in section 2.2.1, image distortion can be described with the additive functions $\rho^u(u, v)$ and $\rho^v(u, v)$. Thus, the distorted coordinates can be computed using the following expression:

$$\begin{aligned}\hat{u} &= u + \rho^u(u, v), \\ \hat{v} &= v + \rho^v(u, v),\end{aligned}$$

where u and v are the undistorted pixel coordinates and \hat{u} and \hat{v} are the distorted ones. To obtain the undistorted pixel coordinates, on the other hand, we will use the additive functions $\rho^{\hat{u}}(\hat{u}, \hat{v})$ and $\rho^{\hat{v}}(\hat{u}, \hat{v})$ as follows:

$$\begin{aligned}u &= \hat{u} + \rho^{\hat{u}}(\hat{u}, \hat{v}), \\ v &= \hat{v} + \rho^{\hat{v}}(\hat{u}, \hat{v}),\end{aligned}$$

Depending on the application, we will be interested on the estimation of the distortion functions $\rho^u(u, v)$, $\rho^v(u, v)$ or on the undistortion functions $\rho^{\hat{u}}(\hat{u}, \hat{v})$, $\rho^{\hat{v}}(\hat{u}, \hat{v})$.

In Computer Vision, for real cameras, the geometrical distortion is introduced by the imperfect lens shape [102], and $\rho^{\hat{u}}(\hat{u}, \hat{v})$ and $\rho^{\hat{v}}(\hat{u}, \hat{v})$ can be approximated as polynomials of \hat{u} , \hat{v} [91, 32, 20, 102].

However, in the case of digital angiography, ρ^u and ρ^v change depending on the orientation of the Image Intensifier [29, 17], since it is affected by the Earth's magnetic field. Hence, we need a more elaborated model.

Our aim is to propose a model to estimate ρ^u , ρ^v , $\rho^{\hat{u}}$ and $\rho^{\hat{v}}$ from a given orientation of the detector, which is determined by a rotation angle α and an angulation angle β . Figure 3.1 illustrates the rotation and angulation angles in physical terms. These angles are directly available from the X-ray sequence DICOM file with a precision of 0.1° .

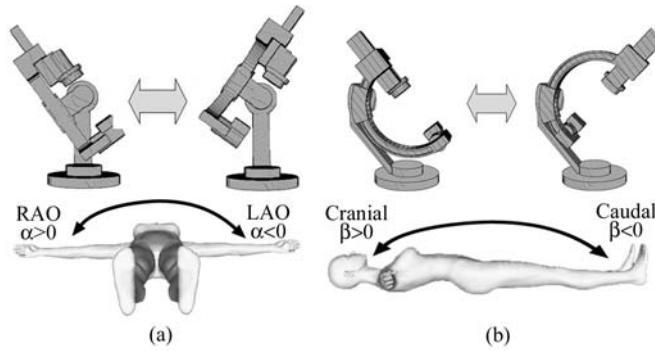


Figure 3.1: Anatomical angles, which define the orientation of the detector. (a) illustrates the rotation, denoted by α , and (b) the angulation, denoted by β .

3.1.1 Geometrical Distortion Estimation

To estimate the distortion, a calibration grid with a spacing of 1.0 cm is placed against the Image Intensifier screen, as shown in figure 3.2.

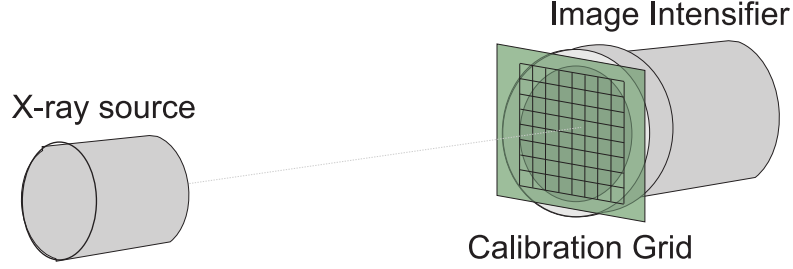


Figure 3.2: The calibration grid is placed against the Image Intensifier screen to estimate geometrical distortion.

We acquire the grid for different rotation and angulation angles of the acquisition system. Thus, for each image k , we obtain subpixel estimates of crossing points ($\hat{\mathbf{u}}^k, \hat{\mathbf{v}}^k$) and their corresponding physical position ($\mathbf{x}^k, \mathbf{y}^k$) (see appendix B for details).

Our aim is to determine the *ideal pixel coordinates* ($\mathbf{u}^k, \mathbf{v}^k$) of the grid for each image. The main drawback is that ($\mathbf{x}^k, \mathbf{y}^k$) may have a different coordinate origin for each image. To solve this problem, the authors in [2] use the fact that near the center of the image the distortion is low. They propose to track along the images a crossing point placed near the image center and to assign the origin of coordinates to it.

We can define the world coordinate system as placed on the calibration grid. Thus, the world coordinates of the crossing points on the grid will be:

$$\mathbf{X}_w^k = \mathbf{x}^k, \quad \mathbf{Y}_w^k = \mathbf{y}^k, \quad \mathbf{Z}_w^k = 0 \quad (3.1)$$

Since we place the calibration grid on the II screen, and thus parallel to the image plane, we can assume that the external parameters \mathbf{R}, \mathbf{t} are :

$$\mathbf{R} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \mathbf{t} = \begin{bmatrix} t_x \\ t_y \\ f - d \end{bmatrix} \quad (3.2)$$

where d is the distance between the II screen and the image plane, θ describes a rotation around the Z axis, and (t_x, t_y) is the displacement parallel to the image plane between the coordinate origin of the grid and the one of the image. Figure 3.3 illustrates this situation.

From (2.3), (3.1) and (3.2) we obtain the following ²:

$$\begin{aligned} \mathbf{u}^k &= \frac{fk}{f-d} \cos \theta \mathbf{x}^k - \frac{fk}{f-d} \sin \theta \mathbf{y}^k + \frac{fk}{f-d} t_x + u_0 \\ \mathbf{v}^k &= \frac{fk}{f-d} \sin \theta \mathbf{x}^k + \frac{fk}{f-d} \cos \theta \mathbf{y}^k + \frac{fk}{f-d} t_y + v_0 \end{aligned} \quad (3.3)$$

²Please, note the difference between pixel size k and view index \mathbf{k} .

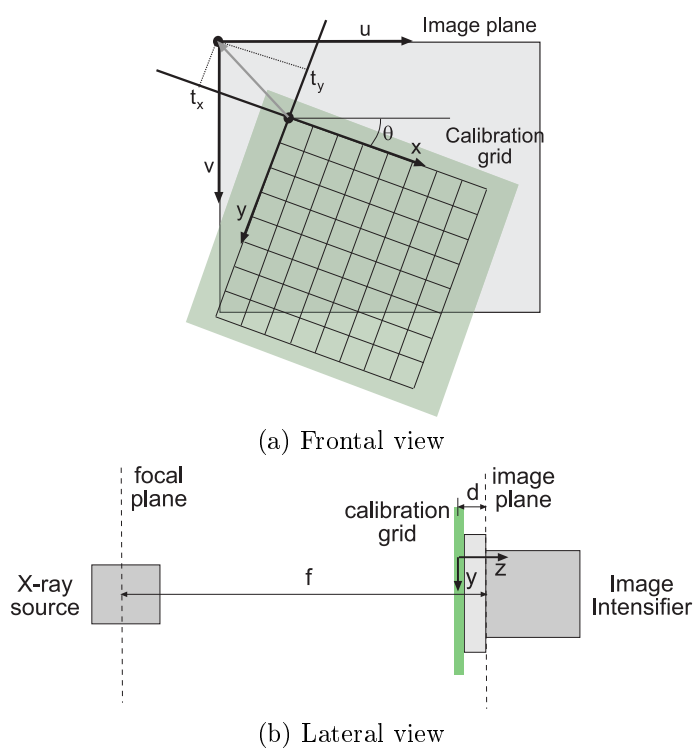


Figure 3.3: Relationship between grid coordinates and ideal image coordinates when the grid is attached to the image intensifier screen.

Let $s = \frac{fk}{f-d}$, $t_u = \frac{fk}{f-d}t_x + u_0$ and $t_v = \frac{fk}{f-d}t_y + v_0$, then we can re-write equation (3.3) as follows:

$$\begin{bmatrix} \mathbf{u}^k \\ \mathbf{v}^k \end{bmatrix} = \begin{bmatrix} s \cos \theta & -s \sin \theta & t_u \\ s \sin \theta & s \cos \theta & t_v \end{bmatrix} \begin{bmatrix} \mathbf{x}^k \\ \mathbf{y}^k \\ \mathbf{1} \end{bmatrix}$$

Therefore, we can compute the ideal undistorted coordinates $(\mathbf{u}^k, \mathbf{v}^k)$ from parameters s , θ , t_u and t_v . These parameters must minimize the following expression:

$$\sum_k \sum_i [(\hat{u}_i^k - u_i^k)^2 + (\hat{v}_i^k - v_i^k)^2]$$

where $(\hat{u}_i^k, \hat{v}_i^k)$ is the i th detected grid node in view k . To minimize this expression, parameters a_u and a_v are introduced to linearise the problem as follows [16]:

$$a_u = s \cos \theta, \quad a_v = s \sin \theta$$

The least squares approach leads to a set of four linear equations:

$$\begin{bmatrix} A & -B & 1 & 0 \\ B & A & 0 & 1 \\ C & 0 & A & B \\ 0 & C & -B & A \end{bmatrix} \begin{bmatrix} a_u \\ a_v \\ t_u \\ t_v \end{bmatrix} = \begin{bmatrix} U \\ V \\ D_1 \\ D_2 \end{bmatrix}$$

where

$$A = \frac{1}{N} \sum_k \sum_i \mathbf{x}_i^k \quad B = \frac{1}{N} \sum_k \sum_i \mathbf{y}_i^k$$

$$C = \frac{1}{N} \sum_k \sum_i [(\mathbf{x}_i^k)^2 + (\mathbf{y}_i^k)^2]$$

$$U = \frac{1}{N} \sum_k \sum_i \hat{\mathbf{u}}_i^k \quad V = \frac{1}{N} \sum_k \sum_i \hat{\mathbf{v}}_i^k$$

$$D_1 = \frac{1}{N} \sum_k \sum_i (\hat{\mathbf{u}}_i^k \mathbf{x}_i^k + \hat{\mathbf{v}}_i^k \mathbf{y}_i^k)$$

$$D_2 = \frac{1}{N} \sum_k \sum_i (\hat{\mathbf{v}}_i^k \mathbf{x}_i^k - \hat{\mathbf{u}}_i^k \mathbf{y}_i^k)$$

and N is the total number of detected crossing points in all images. When there is high distortion, we can constrain the estimate of these parameters to the crossing points at the central part of the image. For a FOV of 17cm, all the crossing points can be used.

3.1.2 Modelling Geometrical Distortion

Once the ideal pixel coordinates are obtained, we can model $\rho^u(u, v)$ (and, analogously, $\rho^v(u, v)$) as the polynomial ³:

$$\rho^u(u, v) = \sum_{n=0}^N \sum_{m=0}^n a_j^u u^{n-m} v^m$$

where $j = \frac{n(n+1)}{2} + m$ and N is the polynomial degree. The coefficients $\{a_j^u\}$ are chosen to minimize the following expression⁴:

$$\sum_i [\hat{u}_i - u_i - \rho^u(u_i, v_i)]^2 = \sum_i [\hat{u}_i - u_i - \sum_{n=0}^N \sum_{m=0}^n a_j^u u_i^{n-m} v_i^m]^2$$

where (\hat{u}_i, \hat{v}_i) are the distorted coordinates of each of the detected nodes of the grid and (u_i, v_i) the corresponding undistorted ones. The optimal polynomial degree N can be empirically determined (see Chapter 5 for details).

Analogously, the polynomial $\rho^{\hat{u}}(\hat{u}, \hat{v})$ (as $\rho^{\hat{v}}(\hat{u}, \hat{v})$) is given by:

$$\rho^{\hat{u}}(\hat{u}, \hat{v}) = \sum_{n=0}^N \sum_{m=0}^n a_j^{\hat{u}} \hat{u}^{n-m} \hat{v}^m$$

where the coefficients $\{a_j^{\hat{u}}\}$ are chosen to minimize the expression:

$$\sum_i [u_i - \hat{u}_i - \rho^{\hat{u}}(\hat{u}_i, \hat{v}_i)]^2 = \sum_i [u_i - \hat{u}_i - \sum_{n=0}^N \sum_{m=0}^n a_j^{\hat{u}} \hat{u}_i^{n-m} \hat{v}_i^m]^2$$

We have chosen polynomials for $\rho^u(u, v)$, $\rho^v(u, v)$, $\rho^{\hat{u}}(\hat{u}, \hat{v})$ and $\rho^{\hat{v}}(\hat{u}, \hat{v})$ because most distortion / undistortion models present in Computer Vision literature can be reduced to a bi-variate polynomial [91, 32, 20, 102, 103]. Moreover, the authors in [38] pay special attention to the suitability of a bi-variate polynomial for the modelling of the X-ray angiography geometrical distortion. The degrees of these polynomials have been empirically estimated (see section 5.1.1).

However, the estimated distortion is not always the same, since it is a combination of two components: one is fixed and the other depends on the orientation. We expect to model the distortion variation by a low degree polynomial. Thus, we re-define the polynomials $\rho^u(u, v)$, $\rho^v(u, v)$, $\rho^{\hat{u}}(\hat{u}, \hat{v})$ and $\rho^{\hat{v}}(\hat{u}, \hat{v})$ as:

$$\begin{aligned} \rho^u(u, v, \alpha, \beta) &= \bar{\rho}^u(u, v) + \delta^u(u, v, \alpha, \beta) \\ \rho^v(u, v, \alpha, \beta) &= \bar{\rho}^v(u, v) + \delta^v(u, v, \alpha, \beta) \\ \rho^{\hat{u}}(\hat{u}, \hat{v}, \alpha, \beta) &= \bar{\rho}^{\hat{u}}(\hat{u}, \hat{v}) + \delta^{\hat{u}}(\hat{u}, \hat{v}, \alpha, \beta) \\ \rho^{\hat{v}}(\hat{u}, \hat{v}, \alpha, \beta) &= \bar{\rho}^{\hat{v}}(\hat{u}, \hat{v}) + \delta^{\hat{v}}(\hat{u}, \hat{v}, \alpha, \beta) \end{aligned}$$

where $\bar{\rho}$ corresponds to the static component of distortion and δ to the orientation-dependent one.

³For simplicity, we omit here the k superindex, but this process is applied for each image.

⁴See appendix C for details on bi-variate polynomial fitting

The polynomial $\bar{\rho}$ corresponds to the previous ρ , but now estimated to fit the distortion for all orientations of the C-arm (see section 5.1.3 for details).

For instance, the coefficients $\{\bar{a}_j^u\}$ of the polynomial

$$\bar{\rho}^u(u, v) = \sum_{n=0}^N \sum_{m=0}^n \bar{a}_j^u u^{n-m} v^m$$

are chosen to minimize the following expression⁵:

$$\sum_k \sum_i [\hat{u}_i^k - u_i^k - \bar{\rho}^u(u_i^k, v_i^k)]^2 = \sum_k \sum_i [\hat{u}_i^k - u_i^k - \sum_{n=0}^N \sum_{m=0}^n \bar{a}_j^u (u_i^k)^{n-m} (v_i^k)^m]^2$$

We follow an analogous procedure to estimate the coefficients of the polynomials $\bar{\rho}^v(u, v)$, $\bar{\rho}^{\hat{u}}(\hat{u}, \hat{v})$ and $\bar{\rho}^{\hat{v}}(\hat{u}, \hat{v})$.

The residual distortion which can not be modelled by $\bar{\rho}$, i.e., the orientation-dependent distortion, is therefore modelled using δ . To this aim, the polynomials $\delta^u(u, v, \alpha, \beta)$, are defined as bi-variate polynomials of (u, v) , namely $\delta_{\alpha, \beta}^u(u, v)$, whose coefficients are bi-variate polynomials of (α, β) as follows:

$$\begin{aligned} \delta^u(u, v, \alpha, \beta) &= \sum_{n=0}^N \sum_{m=0}^n \omega^{u, j}(\alpha, \beta) u^{n-m} v^m = \delta_{\alpha, \beta}^u(u, v) \\ \delta^v(u, v, \alpha, \beta) &= \sum_{n=0}^N \sum_{m=0}^n \omega^{v, j}(\alpha, \beta) u^{n-m} v^m = \delta_{\alpha, \beta}^v(u, v) \end{aligned}$$

with $j = \frac{n(n+1)}{2} + m$ and N is the polynomial degree of $\delta_{\alpha, \beta}^u(u, v)$ and $\delta_{\alpha, \beta}^v(u, v)$

Therefore, the polynomial $\omega^{u, j}(\alpha, \beta)$ represents the j th coefficient of the polynomial $\delta_{\alpha, \beta}^u(u, v)$. The polynomials $\delta^{\hat{u}}(\hat{u}, \hat{v}, \alpha, \beta)$ and $\delta^{\hat{v}}(\hat{u}, \hat{v}, \alpha, \beta)$ are constructed analogously:

$$\begin{aligned} \delta^{\hat{u}}(\hat{u}, \hat{v}, \alpha, \beta) &= \sum_{n=0}^N \sum_{m=0}^n \omega^{\hat{u}, j}(\alpha, \beta) \hat{u}^{n-m} \hat{v}^m = \delta_{\alpha, \beta}^{\hat{u}}(\hat{u}, \hat{v}) \\ \delta^{\hat{v}}(\hat{u}, \hat{v}, \alpha, \beta) &= \sum_{n=0}^N \sum_{m=0}^n \omega^{\hat{v}, j}(\alpha, \beta) \hat{u}^{n-m} \hat{v}^m = \delta_{\alpha, \beta}^{\hat{v}}(\hat{u}, \hat{v}) \end{aligned}$$

Now our aim is the determination of the polynomial degree N of $\delta_{\alpha, \beta}^u(u, v)$, $\delta_{\alpha, \beta}^v(u, v)$, $\delta_{\alpha, \beta}^{\hat{u}}(\hat{u}, \hat{v})$ and $\delta_{\alpha, \beta}^{\hat{v}}(\hat{u}, \hat{v})$. Let us define the polynomials describing the

⁵Please, note that here k is the view index and not an exponent.

orientation-dependent component of distortion for a given view k as:

$$\begin{aligned}\delta^u(u, v, \alpha_k, \beta_k) &= \delta_k^u(u, v) = \sum_{n=0}^N \sum_{m=0}^n c_j^{u,k} u^{n-m} v^m \\ \delta^v(u, v, \alpha_k, \beta_k) &= \delta_k^v(u, v) = \sum_{n=0}^N \sum_{m=0}^n c_j^{v,k} u^{n-m} v^m \\ \delta^u(\hat{u}, \hat{v}, \alpha_k, \beta_k) &= \delta_k^{\hat{u}}(\hat{u}, \hat{v}) = \sum_{n=0}^N \sum_{m=0}^n c_j^{\hat{u},k} \hat{u}^{n-m} \hat{v}^m \\ \delta^v(\hat{u}, \hat{v}, \alpha_k, \beta_k) &= \delta_k^{\hat{v}}(\hat{u}, \hat{v}) = \sum_{n=0}^N \sum_{m=0}^n c_j^{\hat{v},k} \hat{u}^{n-m} \hat{v}^m\end{aligned}$$

where α_k is the rotation angle and β_k is the angulation angle for view k , $j = \frac{n(n+1)}{2} + m$, N is the polynomial degree and $\{c_j^{u,k}\}$ are the coefficients we aim to determine.

For a given polynomial degree N , we can estimate the coefficients $\{c_j^{u,k}\}$ of $\delta_k^u(u, v)$ for each view k by minimizing the expression⁶:

$$\begin{aligned}\sum_i [\hat{u}_i^k - u_i^k - \bar{\rho}^u(u_i^k, v_i^k) - \delta_k^u(u_i^k, v_i^k)]^2 &= \\ = \sum_i \left[\hat{u}_i^k - u_i^k - \bar{\rho}^u(u_i^k, v_i^k) - \sum_{n=0}^N \sum_{m=0}^n c_j^{u,k} (u_i^k)^{n-m} (v_i^k)^m \right]^2\end{aligned}$$

As with $\bar{\rho}^u(u, v)$, the optimal polynomial degree for $\delta_k^u(u, v)$ (and therefore, for $\delta_{\alpha,\beta}^u(u, v)$) can be empirically determined (see section 5.1.3).

The same procedure can be applied to estimate the coefficients and the optimal polynomial degree for the polynomials $\delta_{\alpha,\beta}^v(u, v)$, $\delta_{\alpha,\beta}^{\hat{u}}(\hat{u}, \hat{v})$ and $\delta_{\alpha,\beta}^{\hat{v}}(\hat{u}, \hat{v})$, by minimizing the expressions:

$$\sum_i (\hat{v}_i^k - v_i^k - \bar{\rho}^v(u_i^k, v_i^k) - \delta_k^v(u_i^k, v_i^k))^2,$$

$$\sum_i (u_i^k - \hat{u}_i^k - \bar{\rho}^{\hat{u}}(\hat{u}_i^k, \hat{v}_i^k) - \delta_k^{\hat{u}}(\hat{u}_i^k, \hat{v}_i^k))^2$$

and

$$\sum_i (v_i^k - \hat{v}_i^k - \bar{\rho}^{\hat{v}}(\hat{u}_i^k, \hat{v}_i^k) - \delta_k^{\hat{v}}(\hat{u}_i^k, \hat{v}_i^k))^2,$$

respectively.

Now our aim is being able to estimate the polynomials $\delta^u(u, v, \alpha, \beta)$, $\delta^v(u, v, \alpha, \beta)$, $\delta^{\hat{u}}(\hat{u}, \hat{v}, \alpha, \beta)$ and $\delta^{\hat{v}}(\hat{u}, \hat{v}, \alpha, \beta)$ for any rotation angle α and angulation angle β . This means that we need to estimate the coefficients of the polynomials $\{\omega_j^u(\alpha, \beta)\}$, $\{\omega_j^v(\alpha, \beta)\}$, $\{\omega_j^{\hat{u}}(\alpha, \beta)\}$ and $\{\omega_j^{\hat{v}}(\alpha, \beta)\}$.

⁶Note that k is not an exponent, but a view index.

Note that for each view k , we have fitted the polynomials $\delta_k^u(u, v)$, $\delta_k^v(u, v)$, $\delta_k^{\hat{u}}(\hat{u}, \hat{v})$ and $\delta_k^{\hat{v}}(\hat{u}, \hat{v})$, represented by the coefficients $\{c_j^{u,k}\}$, $\{c_j^{v,k}\}$, $\{c_j^{\hat{u},k}\}$ and $\{c_j^{\hat{v},k}\}$, respectively. Each of these coefficients has its own associated polynomial. For instance, the coefficient $c_j^{u,k}$ is estimated using the polynomial $\omega_j^u(\alpha, \beta)$ with $\alpha = \alpha_k$ and $\beta = \beta_k$, where $\omega_j^u(\alpha, \beta)$ has the following expression:

$$\omega_j^u(\alpha, \beta) = \sum_{n=0}^N \sum_{m=0}^n d_i^{u,j} \alpha^{n-m} \beta^m$$

Hence, the coefficients $\{d_i^{u,j}\}$ of polynomial $\omega_j^u(\alpha, \beta)$ are chosen to minimize:

$$\sum_k [c_j^{u,k} - \omega_j^u(\alpha_k, \beta_k)]^2 = \sum_k [c_j^{u,k} - \sum_{n=0}^N \sum_{m=0}^n d_i^{u,j} \alpha_k^{n-m} \beta_k^m]^2$$

using the coefficients $\{c_j^{u,k}\}$ estimated before.

We follow an equivalent procedure for the other polynomials $\{\omega_j^v(\alpha, \beta)\}$, $\{\omega_j^{\hat{u}}(\alpha, \beta)\}$ and $\{\omega_j^{\hat{v}}(\alpha, \beta)\}$. The degree of these polynomials is also empirically determined (see section 5.1.4 for details).

Once geometrical distortion is calibrated, the biplane imaging geometry can be determined by any of the methods proposed in the Computer Vision or Medical Imaging literature [22, 23, 13, 28]. In the remainder of this chapter, we present a method to calibrate the extrinsic and intrinsic parameters and a model to estimate the acquisition geometry given the rotation angle α and the angulation angle β .

3.2 Extrinsic Parameters Calibration

Given the general problem of pinhole camera calibration, Zhang in [103] proposes a technique to calibrate both the intrinsic and extrinsic parameters by using a planar pattern, provided that the intrinsic parameters are constant for all views. This method can also be applied on the angiographic frame when the distance from the Image Intensifier to the X-ray source remains unchanged for all views. For a more detailed explanation of this method, see section 2.2.4.

3.2.1 Using Zhang's Method in the Angiographic Frame.

The calibration pattern required by Zhang's method can be easily constructed by printing a checker-box or a grid onto a paper and then fixing it to a rigid planar surface, for instance to a hard-covered book. Although this approach is not applicable when working with X-rays, we can use as calibration object the grid we used in section 3.1.1, when calibrating the distortion introduced by the Image Intensifier. Thus, we propose to place the grid on the examination table, and acquire it by changing the anatomical angles for each view. Zhang's method [103] can then be applied to obtain the intrinsic parameters, represented by matrix \mathbf{A} , and the extrinsic parameters, corresponding for each view i to matrices $\hat{\mathbf{R}}_i$ and $\hat{\mathbf{t}}_i$.

Nevertheless, this calibration is only valid for the acquired views, and we are interested in being able to compute the extrinsic parameters $\hat{\mathbf{R}}$, $\hat{\mathbf{t}}$ for any given configuration of the C-arm, determined by the anatomical angles α and β . To do that, we define a model of the movement of the C-arm, as the one proposed by Dumay et al. in [22], and then we estimate the parameters of this model through the obtained $\hat{\mathbf{R}}_i$ and $\hat{\mathbf{t}}_i$ using the Zhang's method.

3.2.2 Modelling the Movement of the C-arm

Modelling the movement of the C-arm means defining how to compute the extrinsic parameters \mathbf{R} , \mathbf{t} for any given configuration of the C-arm, determined by the anatomical angles α and β .

Dumay et al. defined in [22] the characteristic movements of a ceiling-mounted C-arm. First, we analyze this model and then introduce some modifications to improve its accuracy.

Dumay's Model (M0).

In the model proposed by Dumay et al. in [22], the fixed axis of rotation is directed horizontally and in longitudinal direction to the table. This axis is always projected 'vertically' in the images, i.e. its projection coincides with the v-axis on the image matrix. The projection of this fixed axis is therefore chosen as vector \mathbf{l} on the image plane. The axis perpendicular to \mathbf{l} on the image plane is denoted as \mathbf{k} (see figure 3.4) and is defined as:

$$\mathbf{k} = [0, -\cos \alpha, \sin \alpha]^T$$

The position of the plane, $\mathbf{C} = \|C\|\mathbf{c}$, where \mathbf{c} can be computed as:

$$\mathbf{c} = [\sin \beta, \sin \alpha \cos \beta, \cos \alpha \cos \beta]^T$$

The local reference system is chosen as a left-hand oriented orthogonal system, and thus the remaining axis \mathbf{l} is solved from:

$$\mathbf{l} = \mathbf{c} \times \mathbf{k}$$

From this model, we can define the extrinsic parameters \mathbf{R}' , \mathbf{t}' as:

$$\mathbf{R} = \bar{\mathbf{R}} \quad \mathbf{t} = \bar{\mathbf{t}} \quad (3.4)$$

where

$$\bar{\mathbf{R}} = [\mathbf{k} \quad \mathbf{l} \quad \mathbf{c}]^T$$

and

$$\bar{\mathbf{t}} = [0, 0, \|C\|]^T$$

where $\|C\|$ is defined as the distance from the isocenter of rotation to the entrance screen of the image intensifier. Figure 3.4 illustrates this situation.

However, the model proposed by Dumay et al. makes the following assumptions (see fig. 3.4):

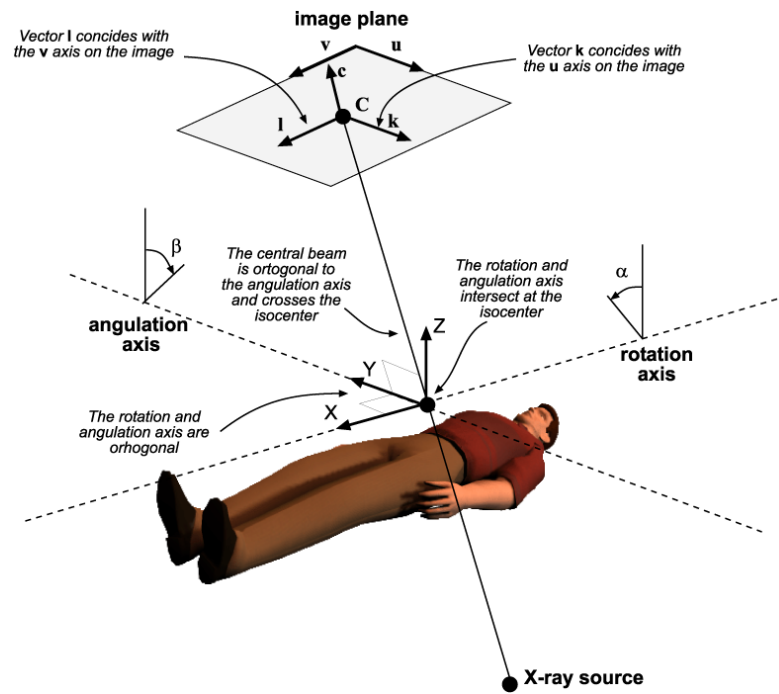


Figure 3.4: Dumay's model for the movement of the C-arm.

- Vector \mathbf{l} coincides with the v axis on the image, and \mathbf{k} coincides with the u axis on the image.
- The rotation and angulation axes do intersect, i.e. there is an isocenter.
- The imaging equipment is properly aligned and thus the central beams intersect at the isocenter of rotation.
- The rotation and angulation axes are perfectly orthogonal.
- The central beam is orthogonal to the angulation axis.

We will analyse each assumption and, if necessary, we will define a model for each case.

Model with Image Rotation (M1).

Vector \mathbf{k} and \mathbf{l} may not coincide with the u axis and v axis on the image, respectively, since the Image Intensifier introduces a distortion with a rotational effect, which depends on the rotation angle and the angulation angle. Our undistortion method presented in section 3.1.1 fixed image rotation and other distortions, even the orientation-dependent ones, by estimating the distortion model from the anatomical angles. Nevertheless, and since the ground truth orientation of the image is unknown, the resulting undistorted image will be in general rotated from the ideal undistorted one. Moreover, rotation and angulation axis of the isocentric model could not be aligned with the ideal image.

Therefore, image rotation can stem both from Image Intensifier distortion and from image misalignment. For simplicity, we will combine both rotations on a unique rotation matrix $\mathbf{R}_d = R_z(\theta)$, which will be applied before applying the perspective projection defined by the matrix \mathbf{A} . Since this rotation is not in general performed on the origin of coordinates, we should add a displacement vector $\mathbf{t}_d = [d_k, d_l, 0]^T$ after rotating⁷. Thus, the extrinsic parameters defined by this model are:

$$\mathbf{R} = \mathbf{R}_d \bar{\mathbf{R}} \quad \mathbf{t} = \mathbf{R}_d \bar{\mathbf{t}} + \mathbf{t}_d$$

Since $\bar{\mathbf{t}}$ is a displacement among z axis, matrix \mathbf{R}_d has no effect on it, and thus:

$$\mathbf{R} = \mathbf{R}_d \bar{\mathbf{R}} \quad \mathbf{t} = \bar{\mathbf{t}} + \mathbf{t}_d \quad (3.5)$$

Figure 3.5 illustrates this new model.

Non Intersecting Central Beams.

If the imaging equipment is not properly aligned, the central beam does not intersect the isocenter of rotation, i.e., the isocenter is also displaced from the central beam in the plane perpendicular to it. We assume that this displacement is constant for all views. This can be modelled with the previous model, since parameters d_k , d_l can compensate this effect. Therefore, we will not define a model for this situation.

⁷Although the rotation related to angulation and rotation angles misalignment is performed before projecting and the one related to II distortion is performed after, it is very easy to show that we can do this simplification.

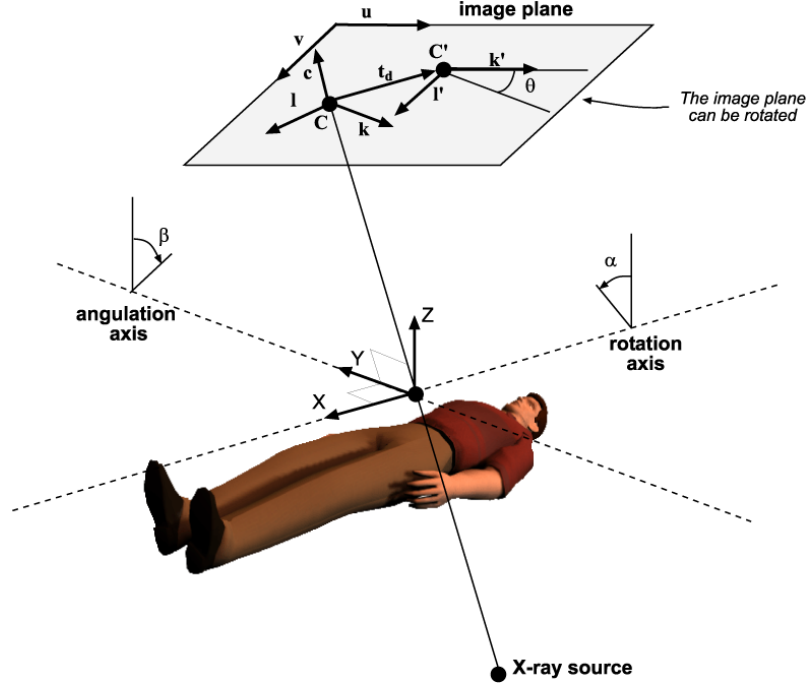


Figure 3.5: Model M1 for the movement of the C-arm.

Non Intersecting Rotation Axis (M2).

If both rotation axis do not intersect, the isocentric point model is not valid. Instead, we should define a model with a displacement between the two rotation axis and perpendicular to both. Since rotation, defined by angle α , is performed around the y axis and angulation around the x axis, this displacement will be along the z axis after rotation. Figure 3.6 illustrates this situation.

To define the external parameters of the model, we will decompose the rotation matrix \mathbf{R} as:

$$\mathbf{R} = \mathbf{R}_d R_x(\beta) R_y(\alpha)$$

where $R_y(\alpha)$ corresponds to the rotation and $R_x(\beta)$ to the angulation⁸.

Let ϵ be the displacement between the rotation and the angulation axis. Hence, this displacement must be computed after applying the rotation $R_y(\alpha)$ but before computing the angulation $R_x(\beta)$. Since this displacement is along the z axis, we add

⁸Note that if we compute $R_x(\beta)R_y(\alpha)$ we will not obtain $\bar{\mathbf{R}} = [\mathbf{k} \ \mathbf{l} \ \mathbf{c}]$, but $\bar{\mathbf{R}}\mathbf{R}_a$, where

$$\mathbf{R}_a = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

This is not important to obtain the three-dimensional reconstruction, neither has effect when calibrating.

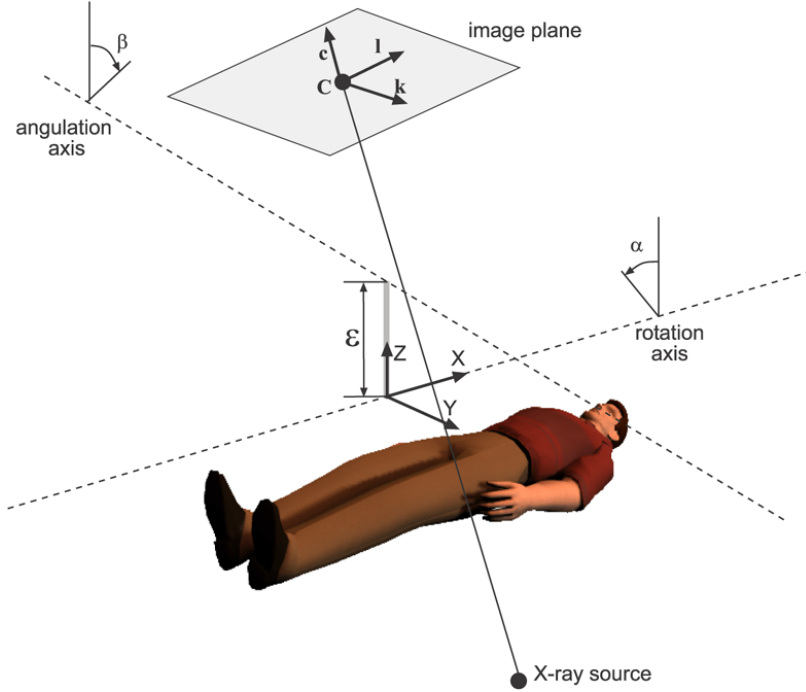


Figure 3.6: Non intersecting rotation axes.

the following translation term:

$$\mathbf{R}_d R_x(\beta) [0, 0, \epsilon]^T$$

where the rotation matrix $R_y(\alpha)$ does not appear, since the rotation movement is performed before translating by $[0, 0, \epsilon]^T$.

Hence, the external parameters for model **M2** are computed as follows:

$$\begin{aligned} \mathbf{R} &= \mathbf{R}_d R_x(\beta) R_y(\alpha) \\ \mathbf{t} &= \bar{\mathbf{t}} + \mathbf{t}_d + \mathbf{R}_d R_x(\beta) \begin{bmatrix} 0 \\ 0 \\ \epsilon \end{bmatrix} \end{aligned} \quad (3.6)$$

Non Orthogonal Rotation Axes (M3).

Although the rotation axes are expected to be orthogonal, perfect orthogonality is hard to achieve, and thus the angulation axis will be slightly rotated from its ideal position. Moreover, the anatomical angles could be measured as $\alpha' = \alpha + \Delta\alpha$ and $\beta' = \beta + \Delta\beta$. This fact would have an equivalent effect on the extrinsic parameters as the non orthogonality of the rotation axes.

Let matrix \mathbf{R}_p the matrix describing the rotation between the rotation and the

angulation axis. Thus, matrix \mathbf{R} is re-defined as:

$$\mathbf{R} = \mathbf{R}_d R_x(\beta) \mathbf{R}_p R_y(\alpha)$$

where matrix \mathbf{R}_p has been introduced between matrices $R_x(\beta)$ and $R_y(\alpha)$, which describe the angulation and rotation movements, respectively.

The translation vector has also to be re-defined as follows:

$$\mathbf{t} = \bar{\mathbf{t}} + \mathbf{t}_d + \mathbf{R}_d R_x(\beta) \mathbf{R}_p \begin{bmatrix} 0 \\ 0 \\ \epsilon \end{bmatrix}$$

where matrix \mathbf{R}_p has been after the term $[0, 0, \epsilon]^T$, because this translation is performed before angulation. The terms $\bar{\mathbf{t}}$ and \mathbf{t}_d are not modified because these transformations are done after angulation.

Hence, the extrinsics parameters for model (M3) are computed as follows:

$$\begin{aligned} \mathbf{R} &= \mathbf{R}_d R_x(\beta) \mathbf{R}_p R_y(\alpha) \\ \mathbf{t} &= \bar{\mathbf{t}} + \mathbf{t}_d + \mathbf{R}_d R_x(\beta) \mathbf{R}_p \begin{bmatrix} 0 \\ 0 \\ \epsilon \end{bmatrix} \end{aligned} \quad (3.7)$$

Central Beam Non Orthogonal to Angulation Axis (M4).

In theory, the central beam should be perpendicular to the angulation axis. Nevertheless, this may not be true, and hence we introduce matrix \mathbf{R}_b in the model. This matrix will describe the rotation between the central beam and the ideal one, which is perpendicular to the angulation axis. Since matrix \mathbf{R}_b describes any rotation, the application of matrix \mathbf{R}_d , describing a rotation of the image plane around the z -axis is no more necessary. In fact, matrix \mathbf{R}_b can be understood as a generalization of matrix \mathbf{R}_d , and therefore is defined as M3, but replacing \mathbf{R}_d by \mathbf{R}_b .

Hence, we obtain the following expression for the extrinsic parameters of model (M4):

$$\begin{aligned} \mathbf{R} &= \mathbf{R}_b R_x(\beta) \mathbf{R}_p R_y(\alpha) \\ \mathbf{t} &= \bar{\mathbf{t}} + \mathbf{t}_d + \mathbf{R}_b R_x(\beta) \mathbf{R}_p \begin{bmatrix} 0 \\ 0 \\ \epsilon \end{bmatrix} \end{aligned} \quad (3.8)$$

3.2.3 Calibrating the C-arm's Movement Model

Zhang's approach assumes that the world coordinate system is placed on the calibration pattern. The axes of this coordinate system are not in general parallel to the rotation and angulation axes, and the origin of coordinates on the calibration pattern does not coincide with the isocenter, if any. Thus, there is a rigid transform, represented by \mathbf{R}_w , \mathbf{t}_w between the two coordinate systems. If the model is exact and well calibrated, the extrinsic parameters \mathbf{R}_i , \mathbf{t}_i for a given view i predicted by the

model and the matrices $\hat{\mathbf{R}}_i$, $\hat{\mathbf{t}}_i$, estimated for view i using Zhang's approach, should fulfill the following:

$$\hat{\mathbf{R}}_i = \mathbf{R}_i \mathbf{R}_w \quad \hat{\mathbf{t}}_i = \mathbf{R}_i \mathbf{t}_w + \mathbf{t}_i \quad (3.9)$$

Our aim is thus to estimate the parameters of each model and the rigid transform \mathbf{R}_w , \mathbf{t}_w , which best fits the estimated matrices $\hat{\mathbf{R}}_i$, $\hat{\mathbf{t}}_i$ using the Zhang's approach to the ideal matrices \mathbf{R}_i , \mathbf{t}_i of each model.

M0 Calibration.

The parameters to be estimated are \mathbf{R}_w , \mathbf{t}_w and $\|C\|$. From equation (3.4) we have:

$$\mathbf{R}_i = \bar{\mathbf{R}} \quad \mathbf{t}_i = \bar{\mathbf{t}}$$

Hence, using equation (3.9), we can state the following:

$$\hat{\mathbf{R}}_i = \bar{\mathbf{R}}_i \mathbf{R}_w \quad \hat{\mathbf{t}}_i = \bar{\mathbf{R}}_i \mathbf{t}_w + \bar{\mathbf{t}}_i$$

Note that, since $\bar{\mathbf{R}}_i$ is known, we can compute for each view i an estimation of \mathbf{R}_w as:

$$\mathbf{R}_w = \bar{\mathbf{R}}_i^T \hat{\mathbf{R}}_i$$

To obtain a better estimation of \mathbf{R}_w , we can compute the mean of the rotation matrices $\{\bar{\mathbf{R}}_i^T \hat{\mathbf{R}}_i\}$. This can be done by using the approach proposed in [71].

On the other hand, for each view we can define the following set of homogeneous linear constraints:

$$\begin{bmatrix} \bar{\mathbf{R}}_i & \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} & -\hat{\mathbf{t}}_i \end{bmatrix} \begin{bmatrix} \mathbf{t}_w \\ \|C\| \\ 1 \end{bmatrix} = \mathbf{0}$$

These constraints can be stacked to construct the following overdetermined linear system of equations:

$$\mathbf{U} \begin{bmatrix} \mathbf{t}_w \\ \|C\| \\ 1 \end{bmatrix} = \mathbf{0}.$$

Since there are 4 independent unknown variables - namely, the three components of \mathbf{t}_w and the distance value $\|C\|$ -, we need at least 2 views. The solution to this system is well-known to be the eigenvector $\mathbf{v} = [v_1, \dots, v_5]^T$ of $\mathbf{U}^T \mathbf{U}$ with the associated smallest eigenvalue. Using the obtained \mathbf{v} , we can compute the desired parameters as follows:

$$\begin{aligned} \mathbf{t}_w &= \begin{bmatrix} \frac{v_1}{v_5}, \frac{v_2}{v_5}, \frac{v_3}{v_5} \end{bmatrix} \\ \|C\| &= \frac{v_4}{v_5} \end{aligned}$$

M1 Calibration.

From equation (3.5), we have the following:

$$\begin{aligned}\mathbf{R}_i &= \mathbf{R}_d \bar{\mathbf{R}}_i \\ \mathbf{t}_i &= \bar{\mathbf{t}}_i + \mathbf{t}_d\end{aligned}$$

Hence, using equation (3.9), we obtain:

$$\begin{aligned}\hat{\mathbf{R}}_i &= \mathbf{R}_d \bar{\mathbf{R}}_i \mathbf{R}_w \\ \hat{\mathbf{t}}_i &= \mathbf{R}_d \bar{\mathbf{R}}_i \mathbf{t}_w + \bar{\mathbf{t}}_i + \mathbf{t}_d\end{aligned}\quad (3.10)$$

Note that \mathbf{R}_w , \mathbf{t}_w , \mathbf{R}_d , \mathbf{t}_d and $\bar{\mathbf{t}}_i = [0, 0, \|C\|]^T$ should be constant for all views.

First, we need to estimate \mathbf{R}_d . For each two views i, j with $i \neq j$, we can compute the following:

$$\hat{\mathbf{R}}_i \hat{\mathbf{R}}_j^T = \mathbf{R}_d \mathbf{R}_i \mathbf{R}_w \mathbf{R}_w^T \mathbf{R}_j^T \mathbf{R}_d^T$$

and using that \mathbf{R}_w is a rotation matrix and thus $\mathbf{R}_w \mathbf{R}_w^T = \mathbf{I}$, we can state that:

$$\hat{\mathbf{R}}_i \hat{\mathbf{R}}_j^T = \mathbf{R}_d \mathbf{R}_i \mathbf{R}_j^T \mathbf{R}_d^T$$

We can stack these constraints to obtain a non-linear system. Since \mathbf{R}_d is uniquely determined by the angle θ , this non-linear system is very easy to solve by applying any standard non-linear least squares approach. In this work, we have applied the Gauss-Newton method to estimate θ . To construct the error function, we have used the metric proposed in [71] to compute the distance between two rotations.

Once we have an estimate of \mathbf{R}_d , we can proceed to estimate the other parameters.

First, using equation (3.10), we can compute an estimate of \mathbf{R}_w for each view i as follows:

$$\mathbf{R}_w = \bar{\mathbf{R}}_i^T \mathbf{R}_d^T \hat{\mathbf{R}}_i$$

As with the previous model, we can obtain a better estimation of \mathbf{R}_w by computing the mean of the rotation matrices $\{\bar{\mathbf{R}}_i^T \mathbf{R}_d^T \hat{\mathbf{R}}_i\}$.

Second, for each view i we can define the following set of linear constraints using equation (3.10):

$$\begin{bmatrix} \mathbf{R}_d \mathbf{R}_i & \mathbf{I} & -\hat{\mathbf{t}}_i \end{bmatrix} \begin{bmatrix} \mathbf{t}_w \\ \mathbf{t}_i + \mathbf{t}_d \\ 1 \end{bmatrix} = \mathbf{0}$$

where

$$\mathbf{t}_i + \mathbf{t}_d = \begin{bmatrix} d_k \\ d_l \\ \|C\| \end{bmatrix}$$

These constraints can be stacked to construct the following overdetermined linear system of equations:

$$\mathbf{U} \begin{bmatrix} \mathbf{t}_w \\ d_k \\ d_l \\ \|C\| \\ 1 \end{bmatrix} = \mathbf{0}.$$

Since there are 6 independent unknown variables - namely, the three components of \mathbf{t}_w , d_k , d_l and $\|C\|$ -, we need at least 2 views. The solution to this system is well-known to be the eigenvector $\mathbf{v} = [v_1, \dots, v_7]^T$ of $\mathbf{U}^T \mathbf{U}$ with the associated smallest eigenvalue. Using the obtained \mathbf{v} , we can compute the desired parameters as follows:

$$\begin{aligned} \mathbf{t}_w &= \left[\frac{v_1}{v_7}, \frac{v_2}{v_7}, \frac{v_3}{v_7} \right] \\ d_k &= \frac{v_4}{v_7} \\ d_l &= \frac{v_5}{v_7} \\ \|C\| &= \frac{v_6}{v_7} \end{aligned}$$

The obtained parameters θ , \mathbf{R}_w , \mathbf{t}_w , d_k , d_l and $\|C\|$ can then be refined by using a non-linear least squares minimization method using as error function the sum of the squared distances between the detected intersection points on the images and the predicted ones using the model.

M2 Calibration.

Using equations (3.6) and (3.9), we obtain the following relation:

$$\begin{aligned} \hat{\mathbf{R}}_i &= \mathbf{R}_d R_x(\beta_i) R_y(\alpha_i) \mathbf{R}_w \\ \hat{\mathbf{t}}_i &= \mathbf{R}_d R_x(\beta_i) R_y(\alpha_i) \mathbf{t}_w + \bar{\mathbf{t}}_i + \mathbf{t}_d + \mathbf{R}_d R_x(\beta_i) \begin{bmatrix} 0 \\ 0 \\ \epsilon \end{bmatrix} \end{aligned}$$

To estimate the parameters of the model, we first estimate θ by using the non-linear approach used to calibrate the previous model, replacing $\hat{\mathbf{R}}_i$ by $R_x(\beta_i) R_y(\alpha_i)$. Once we have an estimate of θ , and thus of \mathbf{R}_d , we can define the following set of linear constraints for each view i :

$$\left[\begin{array}{cc} \mathbf{R}_d R_x(\beta_i) R_y(\alpha_i), & \mathbf{R}_d R_x(\beta_i) \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, & \mathbf{I}, & -\hat{\mathbf{t}}_i \end{array} \right] \begin{bmatrix} \mathbf{t}_w \\ \epsilon \\ d_k \\ d_l \\ \|C\| \\ 1 \end{bmatrix} = \mathbf{0}$$

These constraints can be stacked to construct the following overdetermined linear homogeneous system of equations:

$$\mathbf{U} \begin{bmatrix} \mathbf{t}_w \\ \epsilon \\ d_k \\ d_l \\ \|C\| \\ 1 \end{bmatrix} = \mathbf{0}.$$

Since there are 7 independent unknown variables - namely, the three components of \mathbf{t}_w , ϵ , d_k , d_l and $\|C\|$ -, we need at least 3 views. The solution to this system is well-known to be the eigenvector $\mathbf{v} = [v_1, \dots, v_8]^T$ of $\mathbf{U}^T \mathbf{U}$ with the associated smallest eigenvalue. Using the obtained \mathbf{v} , we can compute the desired parameters as follows:

$$\begin{aligned} \mathbf{t}_w &= \begin{bmatrix} \frac{v_1}{v_8}, \frac{v_2}{v_8}, \frac{v_3}{v_8} \\ \frac{v_4}{v_8} \\ \frac{v_5}{v_8} \\ \frac{v_6}{v_8} \\ \frac{v_7}{v_8} \end{bmatrix} \\ \epsilon &= \frac{v_4}{v_8} \\ d_k &= \frac{v_5}{v_8} \\ d_l &= \frac{v_6}{v_8} \\ \|C\| &= \frac{v_7}{v_8} \end{aligned}$$

Finally, the obtained parameters θ , \mathbf{R}_w , \mathbf{t}_w , ϵ , d_k , d_l and $\|C\|$ can be refined by using a non-linear least squares minimization method to fit the predicted positions of the intersection points of the grid to the detected ones on the images, as we have done with model **M1**.

M3 Calibration.

Using equations (3.7) and (3.9), we can state the following for each view i :

$$\begin{aligned} \hat{\mathbf{R}}_i &= \mathbf{R}_d R_x(\beta) \mathbf{R}_p R_y(\alpha) \mathbf{R}_w \\ \hat{\mathbf{t}}_i &= \mathbf{R}_d R_x(\beta) \mathbf{R}_p R_y(\alpha) \mathbf{t}_w + \bar{\mathbf{t}}_i + \mathbf{t}_d + \mathbf{R}_d R_x(\beta) \mathbf{R}_p \begin{bmatrix} 0 \\ 0 \\ \epsilon \end{bmatrix} \end{aligned}$$

Since we have not found a closed form solution to estimate the parameters of this model, we propose to fit model **M2** and then use the obtained parameters θ , \mathbf{R}_w , \mathbf{t}_w , ϵ , d_k , d_l and $\|C\|$ and the added $\mathbf{R}_p = \mathbf{I}$ as starting guess for a non-linear minimization procedure. As before, the error functional to be minimized will be the squared distance between the predicted positions of the intersections on the grid and the detected ones on the images.

M4 Calibration.

Using equations (3.8) and (3.9), we can obtain the following expression for matrices $\hat{\mathbf{R}}_i$, $\hat{\mathbf{t}}_i$:

$$\begin{aligned} \hat{\mathbf{R}}_i &= \mathbf{R}_b R_x(\beta) \mathbf{R}_p R_y(\alpha) \mathbf{R}_w \\ \hat{\mathbf{t}}_i &= \mathbf{R}_b R_x(\beta) \mathbf{R}_p R_y(\alpha) \mathbf{t}_w + \bar{\mathbf{t}}_i + \mathbf{t}_d + \mathbf{R}_b R_x(\beta) \mathbf{R}_p \begin{bmatrix} 0 \\ 0 \\ \epsilon \end{bmatrix} \end{aligned}$$

To estimate the parameters of this model, we can start with the parameter estimate for model **M3** as initial guess (with $\mathbf{R}_b = \mathbf{R}_d$), and run a non-linear minimization procedure. Again, the error function will be the squared distance between the positions of the intersections on the grid predicted by the model and the detected ones on the images.

In section 5.2 we present experimental results for each of the proposed models.

Chapter 4

3D Reconstruction using Biplane Snakes

In Chapter 3, we proposed a model of the acquisition geometry for a X-ray angiographic system. Also, we proposed a model of the geometrical distortion introduced by the Image Intensifier. Hence, we are now able to obtain an accurate three-dimensional reconstruction of the vessels.

In this chapter, we focus on the three-dimensional reconstruction of the vessels from two X-ray angiographic views using Biplane Snakes. First, in section 4.1, we discuss the acquisition and digitalization of the X-ray angiographic images. Then, in section 4.2, we present a new method to enhance the vessels in the angiographic images in order to obtain a more robust segmentation of the vessels. In section 4.3, we define Biplane Snakes and propose their use to solve the point correspondence problem. We describe how the initial shape can be obtained and define the Biplane Snakes deformation procedure. Finally, we show how Biplane Snakes can be combined with the GGVF in order to deal with wavy vessels, and how geometrical distortion can be included into the Biplane Snakes formulation and hence avoid image unwarping. Experimental results of the 3D reconstruction on synthetic images and real cardiac images are not presented in this chapter but in section 5.3.

4.1 X-Ray Angiography Acquisition

We started this thesis working with a non digital biplane angiographic system. This kind of systems can provide two almost simultaneous views of the coronary vessels. Hence, no synchronization is needed. However, this system was not a digital one. Since in such a system the images are stored on a film support, no additional information about the acquisition parameters relating the exact angulation parameters is available, and has to be noted down by the physician for each view from the data provided by the monitors. Anyway, the main problem was the digitalization of these images. The hospital had available a cinefilm to convert the films into video sequences, and hence we digitalized the video signal. The point was that no frame synchronism signal was available, and hence we ought to acquire the films (of about

100 frames) frame by frame. Moreover, each time we loaded a new film, the new film could be acquired displaced and with a different scale from the previous one. Hence, despite the clear advantage of acquiring simultaneous different views of the heart, the many shortcomings introduced by the non digital nature of the system made very difficult the obtention of a reliable three-dimensional reconstruction.

Fortunately, the hospital purchased a new system, a monoplane digital one. This kind of system provides direct digital images, which can be stored in a CD-ROM in DICOM format. Hence, a lot of information about the acquisition conditions is also stored with each image sequence.

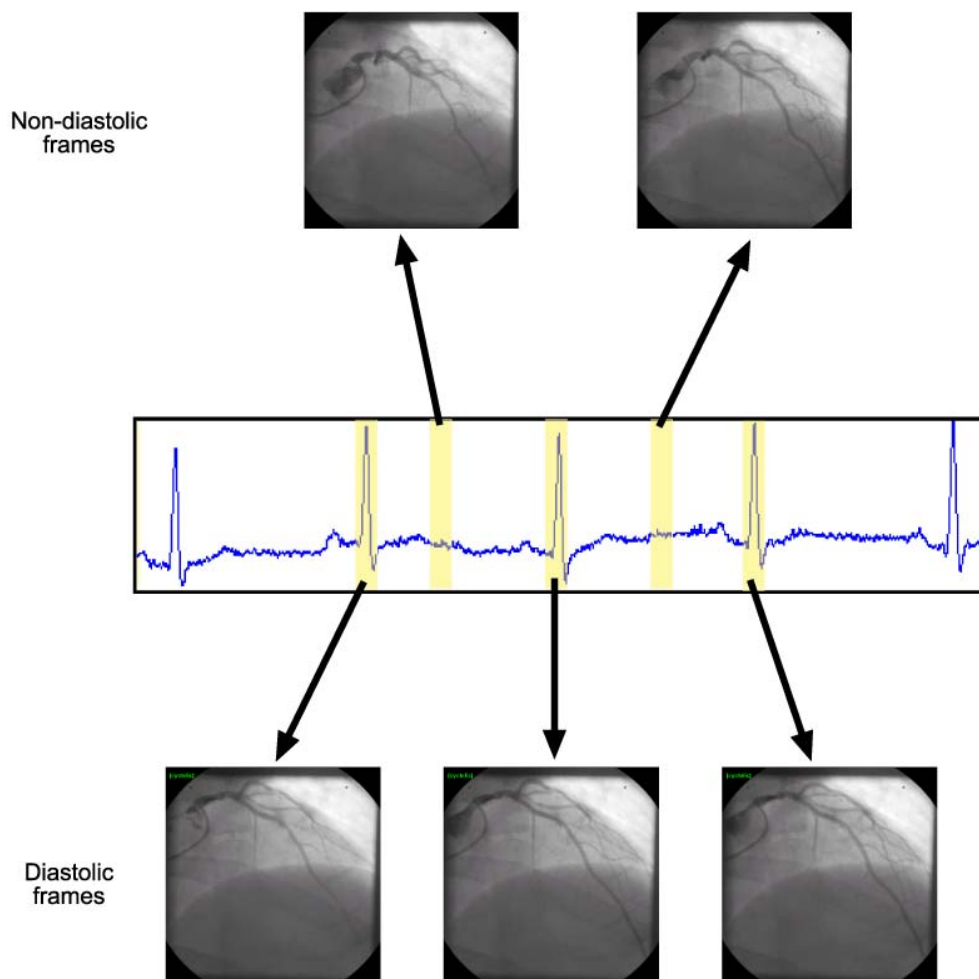


Figure 4.1: The ECG peaks correspond to diastole state of the heart.

Although the ideal situation is to have available a biplane digital system, we can alleviate the shortcomings of having a monoplane system by using the information of the electrocardiogram (ECG), which is stored in the DICOM file for each image

sequence. The ECG curve provides information about the movement of the heart in the sense that from this curve we can deduce which frame corresponds to a heart in diastole (see figure 4.1). The only difficulty is that the breathing of the patient and hence the movements of its lungs move also the heart. To overcome this issue, we ask the patient to exhale all the air and retain breathing before each acquisition.

4.2 Vessel Detection

As stated in the Chapter 1, in this thesis we aim to develop a method to enhance vessel structures to make easier vessel segmentation. A vessel enhancement improves vessel delination and reduces noise and background objects. This facilitates visual analysis and improves the results of vessel centerline segmentation using any of the techniques present in the literature, as for example the one presented by Lopez et al. in [50].

To this aim, we propose a new anisotropic diffusion filter inspired in that of Solé et al. [85], which enhances the creases on the image. The main difference is that our filter enhances the creases on the image with a diffusion strenght determined by the Frangi's vesselness measure[34]. Moreover, for each pixel the scale of the diffusion tensor to be applied is chosen using the Frangi's vesselness measure.

4.2.1 Vessel Enhancement Diffusion

We can see the image $\mathbf{u}(u, v)$ as a surface corresponding to the mass concentration (the gray level). From Fick's law [33] and the continuity equation, we obtain the diffusion equation:

$$\frac{\partial \mathbf{u}}{\partial t} = \text{div}(\mathbf{D} \cdot \nabla \mathbf{u}),$$

where \mathbf{D} is a diffusion tensor, which defines the diffusion process. The filters defined using this formulation are named diffusion filters.

We have thus defined our vessel enhancement diffusion filter iteratively. At each iteration, we compute the Frangi's vesselness measure $\mathcal{V}_o(\sigma)$ for each scale $\sigma = \sigma_{min} \dots \sigma_{max}$ as follows [34]:

$$\mathcal{V}_o(\sigma) = \begin{cases} 0 & \text{if } h_2 < 0, \\ \left(1 - e^{-\frac{1}{2\sigma^2}(h_1^2+h_2^2)}\right) e^{-\frac{1}{2\sigma^2}\left(\frac{h_1}{h_2}\right)^2} & \text{otherwise} \end{cases}$$

where h_1, h_2 are the eigenvalues of the Hessian \mathbf{H}_σ computed at scale σ , chosen to fulfill that $|h_1| \leq |h_2|$. The Hessian is computed as follows:

$$\mathbf{H}_\sigma = \sigma^2 \begin{pmatrix} \frac{\partial^2 \mathbf{u}_\sigma}{\partial u^2} & \frac{\partial^2 \mathbf{u}_\sigma}{\partial u \partial v} \\ \frac{\partial^2 \mathbf{u}_\sigma}{\partial u \partial v} & \frac{\partial^2 \mathbf{u}_\sigma}{\partial v^2} \end{pmatrix}$$

where \mathbf{u}_σ is the initial image \mathbf{u} convolved by a Gaussian kernel of size σ . Note that the Hessian is normalized by multiplying it by σ^2 in order to obtain a fair comparison of the response of differential operators at multiple scales [48].

Given a pixel corresponding to a vessel, the vesselness measure shall be maximal at a scale that approximately matches the size of the vessel to detect. If the pixel does not correspond to a vessel, then the maximal vesselness shall be low.

Let $\sigma_f(u, v)$ be the scale at which the vesselness measure is maximal for each pixel (u, v) , that is:

$$\sigma_f(u, v) = \underset{\sigma = \sigma_{min}, \dots, \sigma_{max}}{\arg \max} \mathcal{V}_o(\sigma).$$

Then, we compute the *normalised Hessian* at scale $\sigma_f(u, v)$ for each pixel (u, v) as follows:

$$\hat{\mathbf{H}}_{\sigma_f, \rho} = \mathbf{G}_\rho * \left[\frac{1}{\sqrt{1 + \|\nabla \mathbf{u}_{\sigma_f}\|^2}} \begin{pmatrix} \frac{\partial^2 \mathbf{u}_{\sigma_f}}{\partial u^2} & \frac{\partial^2 \mathbf{u}_{\sigma_f}}{\partial u \partial v} \\ \frac{\partial^2 \mathbf{u}_{\sigma_f}}{\partial u \partial v} & \frac{\partial^2 \mathbf{u}_{\sigma_f}}{\partial v^2} \end{pmatrix} \right] \quad \text{with } \mathbf{u}_{\sigma_f} = \mathbf{G}_{\sigma_f} * \mathbf{u},$$

where \mathbf{G}_ρ is a Gaussian kernel of size ρ (*integration scale*) and $\sigma_f(u, v)$ is the *differentiation scale* that controls the size of the neighbourhood in which an orientation is dominant. Note that here we are using different differentiation scales for each pixel, determined by the vesselness measure.

The normalized Hessian provides information about the dominant orientation of each image pixel, which corresponds to the eigenvector corresponding to highest eigenvalue in absolute value. Thus, we extract the eigenvectors and eigenvalues of $\hat{\mathbf{H}}_{\sigma_f, \rho}$. Let \mathbf{v}_1 be the first eigenvector corresponding to highest eigenvalue in absolute value ($\tilde{\kappa}_1 = |\kappa_1|$) and \mathbf{v}_2 the eigenvector corresponding to the lowest eigenvalue. Note that the eigendirections $\mathbf{v}_1, \mathbf{v}_2$ providing information about the dominant orientation are computed at the scale of the vessel at each pixel, given by $\sigma_f(u, v)$. This situation is illustrated in figure 4.2.

We define the *vessel enhancement diffusion tensor* that is constructed to have eigenvectors \mathbf{v}_1 and \mathbf{v}_2 , and the following associated eigenvalues:

$$\lambda_1 = \epsilon$$

$$\lambda_2 = \begin{cases} \eta & \text{if } \mathcal{V}_o(\sigma_f) \geq \theta \\ 0 & \text{otherwise} \end{cases}$$

where $\epsilon \in (0, 1)$ ensures the semidefinite property of the diffusion tensor and should be a small value. Parameter θ is a threshold parameter to decide whether a pixel corresponds to a vessel and η is the creaseness measure defined by Solé et al. in [85]:

$$\eta = \frac{\tilde{\kappa}_1 - \tilde{\kappa}_2}{\tilde{\kappa}_1 + \tilde{\kappa}_2}$$

In this work, we use $\theta = 0.5$ and $\epsilon = 0.01$. Nevertheless, θ can be automatically selected by using the Otsu's method [69].

Hence, λ_2 reaches a high value in the presence of creases corresponding to a vessel and is zero for pixels corresponding to background. Therefore, the final expression for the vesselness enhancement diffusion tensor is:

$$\mathbf{D} = [\mathbf{v}_1 \quad \mathbf{v}_2] \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} [\mathbf{v}_1 \quad \mathbf{v}_2]^T$$

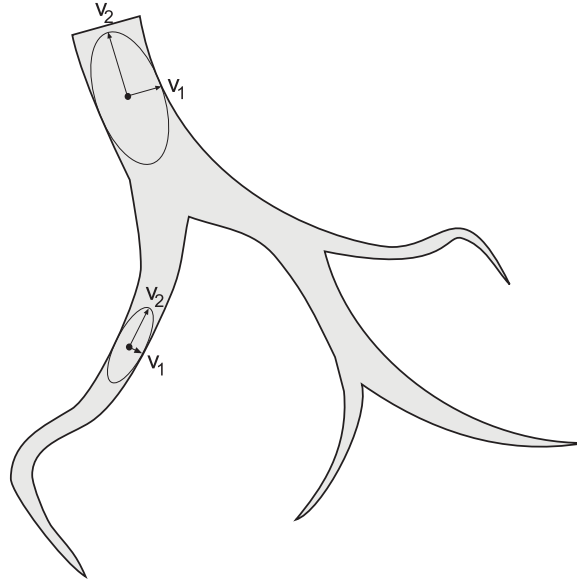


Figure 4.2: The dominant orientation of the vessel is computed at the scale of the vessel at each pixel, given by $\sigma_f(u, v)$.

When the pixel corresponds to a vessel, the eigenvector \mathbf{v}_2 indicates the local orientation of the vessel, and hence the vessel is smoothed mainly along its main axis, and not across, since $|\lambda_1| \ll |\lambda_2|$. On the other hand, when the pixel does not correspond to a vessel, λ_2 is zero and therefore the smoothing is performed across the crease, attenuating it. Hence, the diffusion tensor is anisotropic, since the smoothing is performed along a dominant direction, which is not necessarily collinear to the image gradient $\nabla \mathbf{u}$.

By applying this pre-processing before the vessel centerline detection procedure of Lopez et al. [50], we obtain two benefits: First, and since the filtering increases the crease coherence on the image, the centerlines corresponding to a vessel are less likely to be broken after thresholding the creaseness measure. Second, we can use the vesselness measure, which is now improved after applying the anisotropic vesselness enhancement diffusion filter, to eliminate the detected centerlines corresponding to the background while preserving those corresponding to vessels.

In Chapter 5 we present some results of the proposed vessel enhancement method. This method is applied in the next section to improve the detection result of the image features corresponding to vessel centerlines, which are used to compute the external force of the Biplane Snakes when approaching the three-dimensional reconstruction of the vessels.

4.3 Biplane Snakes

Biplane snakes are one-dimensional deformable models consisting on a three-dimensional curve, which deforms in space to minimize the distance (external energy) between its projections and the vessels from two (or more) image planes. The external energy is defined to be low when the projection of the snake is near the desired features on the image planes. Given that the object to be reconstructed is also curvilinear, the external forces guide the snake to the desired shape, since the snake deforms to adapt its projections to the two-dimensional views of the part of the vessel to be reconstructed. To obtain the initial curve, the user manually determines two pairs of corresponding points near the beginning and the end of the object to be reconstructed. From these corresponding points, we obtain the initial three-dimensional shape of the snake. Figure 4.3 illustrates the whole process. Note the small importance of near initialization.

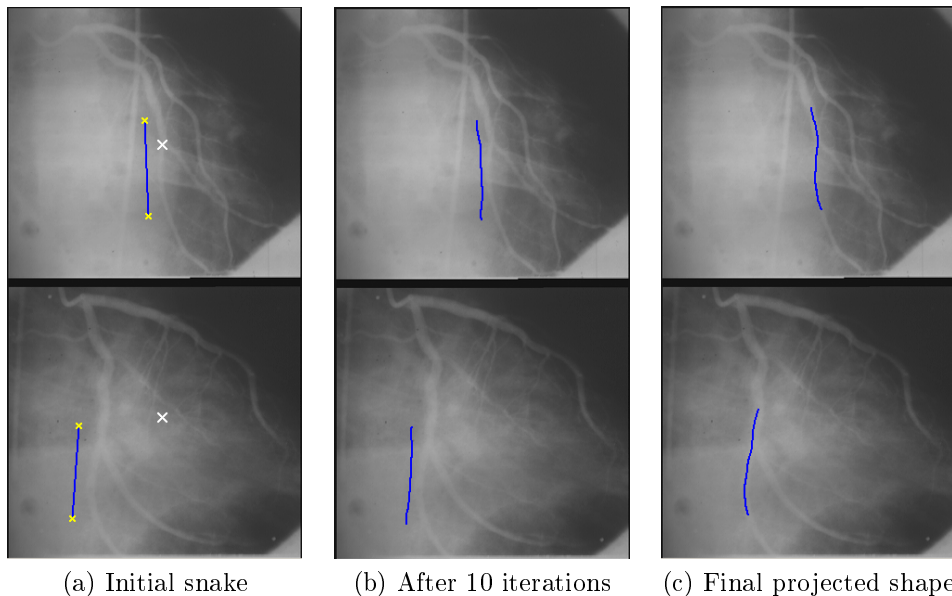


Figure 4.3: The user marks two pairs of corresponding points near the vessel to be reconstructed, from which the initial three-dimensional curve is obtained. This curve deforms in space in order to adapt its projections to the vessels on the image.

4.3.1 Biplane Snakes Initialization

As stated above, the initial snake shape is introduced by the user manually determining two corresponding points among two views. In order to help the user to determine these correspondences, we use the epipolar line. For details on the computation of the epipolar line, see section 2.2.5.

Although the initialization usually consists in only two pairs of corresponding

points near the beginning and the end of the object to be reconstructed, it is in some cases very useful to introduce some additional points to overcome some difficulties.

For instance, when another vessel intersects the path that the projections of the snake should follow to the desired part of the vessel, the snake will not be able to reach the correct position. This situation is illustrated in figure 4.4(a). To overcome this difficulty, the user can introduce some additional initialization points, as shown in figure 4.4(b). Another especial situation is when the projection of the vessel intersects itself for a given view (see figure 4.4(c)). In this case, the introduction of additional initialization points is also very useful, as shown in figure 4.4(d). In Chapter 5, we illustrate these cases using real images.

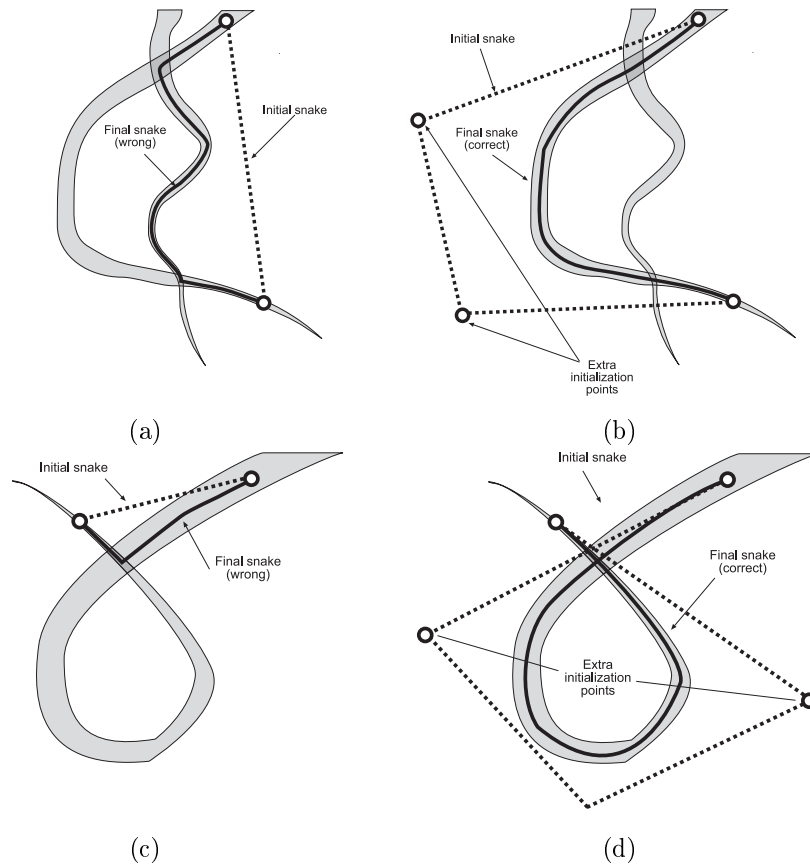


Figure 4.4: Additional initialization points can be very useful when a vessel intersects the path that should follow the snake projection to the vessel part to be reconstructed (a) and (b), or when the projection of the vessel part to be reconstructed intersects itself (c) and (d).

In this thesis, we implement biplane snakes using a B-spline numerical representation. Therefore, we have to obtain the control points of the B-spline curve that passes through the initialization points provided by the user to obtain the initial B-snake

shape.

To assure that the obtained B-spline curve interpolates the initialization points when projected on the image, we first obtain the three-dimensional reconstruction of the initialization points, which will be used to obtain the three-dimensional control points of the curve. Note that if we first compute the control points for each image and afterwards we reconstruct them, we can not assure that the projection of the B-spline curve interpolates the initialization points on the image due to the properties of the B-spline.

Interpolation Using B-Splines

Hence, given m three-dimensional initialization points $\mathbf{P}_j = (x_j, y_j, z_j)$ $j = 1, \dots, m$, we aim finding $n = m - d + 3$ control points \mathbf{V}_i , $i = 1, \dots, n$ such that at each knot k_j in the range $[k_0, k_{m-1}]$ the curve attains a specified point. This means that we want to compute \mathbf{V}_i that fulfill the following constraints:

$$\mathbf{Q}(k_j) = \sum_{i=1}^n \mathbf{V}_i B_{i,k}(k_j) = \mathbf{P}_j \quad \forall j = 1, \dots, m.$$

Although in our case \mathbf{V}_i and \mathbf{P}_i are three-dimensional points, since B-splines can be expressed in terms of individual components we develop the method for only one component. Hence, we rewrite the above constraints using x_j for \mathbf{P}_j and v_i for \mathbf{V}_i as follows:

$$Q(k_j) = \sum_{i=0}^n v_i B_{i,k}(k_j) = x_j \quad \forall j = 1, \dots, m.$$

The problem is that we only have m equations and $n = m - d + 3$ unknowns. Hence, we need $d - 1$ additional constraints. For cubic B-splines, which are the ones used in this thesis, $d = 3$ and therefore we need two more constraints. For instance, we can add the following two extra equations:

$$\sum_{i=1}^n x_i B_{i,4}^{(2)}(k_0) = 0$$

and

$$\sum_{i=1}^n x_i B_{i,4}^{(2)}(k_{m-1}) = 0$$

Although there are plenty of other possibilities, as discussed in [1], we use this approach to obtain the control points of the initial B-snake.

Hence, this method provides $m - d + 3$ control points from a given set of m initialization points. Nevertheless, in order to capture the shape of wavy vessels we could need more points than the initially provided by the user. To cope with this issue, the approach that we follow first introduces extra initialization points using linear interpolation and then applies the above B-spline interpolation technique to obtain the snake initial control points.

4.3.2 Biplane Snakes Deformation

Let $\mathbf{Q}(u)$ be the three-dimensional curve, which has been obtained following the above procedure. The snake curve $\mathbf{Q}(u)$ is therefore defined by the control points \mathbf{V}_i , with $i = 1, \dots, n$ as follows:

$$\mathbf{Q}(u) = \sum_{i=1}^n B_i(u) \mathbf{V}_i$$

This equation can be expressed in matrix form as:

$$\mathbf{Q}_t = \mathbf{B} \mathbf{V}_t$$

where \mathbf{B} is a $m \times n$ matrix containing the blending functions sampled at m values u_1, \dots, u_m of parameter u , that is, $\mathbf{B}_{ij} = B_i(u_j)$. Due to the properties of the B-spline blending functions, this matrix is sparse and therefore the matrix multiplication required to obtain \mathbf{Q} can be computed very efficiently.

The snake deformation is guided by an energy minimization procedure. The general form of the energy to be minimized for a snake has the following expression:

$$E(\mathbf{Q}) = \int E_{int}(\mathbf{Q}) + E_{ext}(\mathbf{Q}) du$$

where $E_{int}(\mathbf{Q})$ is the internal energy and preserves smoothness, and $E_{ext}(\mathbf{Q})$ is the external energy and attracts the snake to the features in the image, computed following the method proposed in section 4.2.

To minimize the snake energy, we use an iterative procedure. Let \mathbf{V}_t be the control points of the B-spline curve that defines the snake shape at step t . These control points are obtained from the previous \mathbf{V}_{t-1} as follows:

$$\mathbf{V}_t = (\mathcal{H} + \gamma \mathbf{I})^{-1} (\gamma \mathbf{V}_{t-1} - \mathbf{g}(\mathbf{Q}_{t-1})),$$

Parameter γ controls the deformation speed and is usually set to 1, and matrix \mathcal{H} is the stiffness matrix, which controls the snake shape. For details on the computation of matrix \mathcal{H} and the blending functions $B_i(u_j)$, see appendix A. The term $\mathbf{g}(\mathbf{Q})$ is a $n \times 1$ vector, which is derived from the external force as:

$$[\mathbf{g}(\mathbf{Q})]_i = \int B_i(u) \mathbf{F}_{ext}(\mathbf{Q}(u)) du \quad i = 1 \dots n$$

In practice, we compute an approximation to $\mathbf{g}(\mathbf{Q})$ as follows:

$$\mathbf{g}(\mathbf{Q}) = \mathbf{B}^T \mathbf{F}_{ext}(\mathbf{Q})$$

We aim obtaining the three-dimensional reconstruction of the curve so that its projections coincide with the vessels in the X-ray images. To this end, we define the external force $\mathbf{F}_{ext}(\mathbf{Q}) = -\nabla \mathbf{V}(\mathbf{Q})$ of the 3D snake curve as a function of distances of the projected snake to the image features:

$$\mathbf{F}_{ext}(\mathbf{Q}) = \varphi^{-1}(\mathbf{q}_1 - \nabla \mathbf{V}_1(\mathbf{q}_1), \mathbf{q}_2 - \nabla \mathbf{V}_2(\mathbf{q}_2)) - \mathbf{Q} \quad (4.1)$$

where $\mathbf{q}_1, \mathbf{q}_2$ are the projections of \mathbf{Q} on view 1 and 2, respectively, $\mathbf{V}_1, \mathbf{V}_2$ are the feature distance maps for each view and φ^{-1} is the retroprojection operator. See figure 4.5 for a geometrical interpretation of formula (4.1). Guided by this force, the 3D snake deforms to adjust its projections to the image vessels. Note that the external energy stems from two images, and thus we name our 3D deformable model as *biplane snake*.

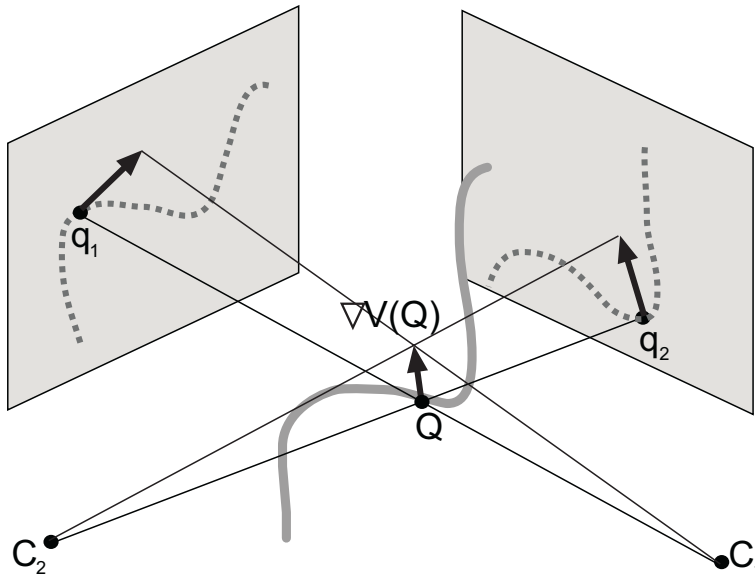


Figure 4.5: External forces deform the snake in order to adapt its projections to the images.

Since the curve deforms in order to adapt its projections to the images, the point correspondence problem is inherently solved. On the other hand, the precision obtained is at least as good as a user would obtain (see the experiment in section 5.3.1). Moreover, since we are using deformable models, feature extraction is not so crucial, since snakes can deal with noise.

Dealing with Concavities

Despite the numerous advantages of using biplane snakes, we observed that when reconstructing wavy vessels the snake has difficulties converging into concave shapes. The problem of active contours convergence to concave objects is addressed in [101] by defining a new kind of external force, called the Gradient Vector Flow (GVF). Later, in [100] the authors presented an improved version called the Generalized Gradient Vector Flow which overcomes a number of problems with the GVF. Therefore, we extend the biplane snakes to deform on a GGVF to allow the 3D reconstruction of complex shapes of coronary vessel centerlines.

First, we compute the Generalized Gradient Vector Flow ν_i for each view, using as features (f) the detected centerline points (see section 4.2). The deformation

procedure is the same, but now we re-define the external force from equation (4.1) of the biplane snake as:

$$F_{ext}(\mathbf{Q}) = \varphi^{-1}(\mathbf{q}_1 + \nu_1(\mathbf{q}_1), \mathbf{q}_2 + \nu_2(\mathbf{q}_2)) - \mathbf{Q} \quad (4.2)$$

Dealing with Distortion

The external force defined in (4.2) has a good convergence behaviour in this application. However, it does not take into account the distortion, since $\mathbf{q}_1, \mathbf{q}_2$ are the projection of \mathbf{Q} on the ideal undistorted coordinates.

A first approach to take into account the distortion is to unwarped the original image, compute the features on the undistorted image and then obtain the undistorted GGVF from these features. Nevertheless, this approach has the problem that unwarped deteriorates the quality of the image, and thus feature extraction becomes harder.

On the other hand, one can compute the target features directly from the original image and compute the distorted GGVF from the distorted features. Then one can obtain $\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2$ on distorted coordinates by applying the corresponding distortion model for each view to $\mathbf{q}_1, \mathbf{q}_2$, which is estimated by means of the model proposed in Chapter 3. We have chosen this last approach, since no unwarped is needed. Moreover, the distortion model has only to be applied at the coordinates of the curve, and not to all the image pixels.

The final external force expression is then as follows:

$$F_{ext}(\mathbf{Q}) = \hat{\varphi}^{-1}(\hat{\mathbf{q}}_1 + \nu_1(\hat{\mathbf{q}}_1), \hat{\mathbf{q}}_2 + \nu_2(\hat{\mathbf{q}}_2)) - \mathbf{Q}$$

where $\hat{\varphi}^{-1}$ is the retroprojection operator including a prior undistortion, i.e.:

$$\hat{\varphi}^{-1}(\hat{\mathbf{q}}_1, \hat{\mathbf{q}}_2) = \varphi^{-1}(\mathbf{q}_1, \mathbf{q}_2)$$

In Chapter 5, we present experimental results of the behaviour of the Biplane Snakes for the three-dimensional reconstruction of coronary vessels using a monoplane system.

