UNIVERSIDADE DA BEIRA INTERIOR
Ciências da Saúde

Medical Image Registration by Neural Networks

A Regression-Based Registration Approach

Ana Isabel Antunes Dias Rodrigues Gouveia

Tese para obtenção do Grau de Doutor em
Biomedicina
(3º ciclo de estudos)

Orientador: Prof. Doutor Luís Freire
Co-orientador: Prof. Doutor Pedro Almeida
Co-orientador: Prof. Doutor João Queiroz

Covilhã, outubro de 2014

To Ricardo.

To Eduarda, Henrique. and Lourenço

Acknowledgments

In these last years I've lived some of the most important experiences in my life. Personally speaking, some of the best but for sure I've also lived the worst moments. On work I must say it was the hardest experience, which is maybe an evidence for the majority of the ex-PhD students. But I believe each case is unique and so it is my experience. I had to leave my comfort zone many times but it was worth it. If what I learned and achieved was represented by some data, it could be modeled by a flat curve in the beginning and it would turned into an exponential at some point. The slope of that curve would definitely change at my first stay in Biomedical Imaging Group Rotterdam at the end of 2009. That made the difference!

I hope this work will be useful for someone, it means something. Ultimately, I hope my thesis brings something to imaging registration framework. It definitively changed ME! And mainly because of people I had the chance to meet. I am grateful to so many people!

First, my supervisor Prof. Luís Freire. If today my research topic is this one, it is because of you. After college it was not expectable (at all!) to work on this, but today I do thank you to have introduced me to it. First image registration, and it has always been dilatable to hear you talking about this topic. Later was machine learning, a new unknown world to me, which you opened it with your contagious enthusiasm about Neural Networks. You brought me to this and that makes you co-responsible to where I am now. Thank you for your support through these years, and your many answers to my institutional demands. Finally, I must enhance your admirable innovator spirit, so important to point out new solutions, which I admire.

To my co-supervisor Prof. Pedro Almeida. What a good coincidence your arrival from France at the moment I needed a final project!! That was the beginning of a long collaboration and, although with some kilometers between us, a good friendship. You have always been there, sometimes when no one else was. I have to thank your efficiency and your care. Your enthusiasm by merging the physics and engineering to the medicine and biology, together with the casual and nice environment, represent the best of IBEB. Your will to find new projects, to open new doors, to create synergies is admirable. I am grateful to have you as a co-supervisor and also as a friend.

To my co-supervisor Prof. João Queiroz. You were so true when you told me the first year would be so low productive! It is a pity I took some more time to find my way, but I did it! I am grateful for the institutional support you gave me, and to have kept me under your supervision even when so many changes happened in your career.

To Doctor Stefan Klein. I do not know how to thank everything you have done for me. In the first place, my gratitude is for our common friends Patrick Andersson and Sandra Tecelão; I still remember his words “Ana must meet Stefan!”. Indeed! You have been my mentor: your determination, your crucial support, your opportune and timely feedback, your guidance through a difficult way whenever I needed you. You also have been an example to me: your pragmatism, your integrity and your efficiency as a researcher and as an individual. And even after so many emails, chats and discussions, you were able to surprise me with your care and encouraging comments. From you, I always had such great motivation!

Then, Doctor Coert Metz. Your collaboration was so important! I am really grateful for all the data, the computations, the inputs, the comments you have provided to my work. Your help made this Thesis possible! And I also have to thank you for your patience and tolerance to my deadlines, in spite of your own deadlines.

A word to Prof. Wiro J. Niessen, who always made me feel welcome in Biomedical Imaging Group Rotterdam of the Erasmus MC University Medical Center. I really enjoyed living some months in that environment and to share the daily life routine of such a high standard research center. Certainly your example will be fruitful to me!

To the deans of Faculty of Health Sciences, I would like to thank all the institutional support through these years. A special thanks to Prof. Miguel Castelo-Branco, who provided the main work tool of this project and supported me with great sensitivity.

To University of Beira Interior, which provided me time to work on my PhD project. To IBEB, which always has the door open for me; for its financial support to attend the scientific conferences.

To UBI staff for its technical support. To Rui Costa, Ricardo Relvas and, in a particular way, to Pedro Sousa by spending so many hours solving my computers’ problems and by the useful and practical tips. Thank you for being around!

To my colleagues from UBI. The collaboration and understanding of some colleagues created better conditions to accomplish this work. A special thanks to Graça Baltazar, Fernando Ferreira, José Amoreira, and also Luís Alexandre for his insight in Neural Networks. And being, in some way, an outsider in my own faculty, the support and care of some was very important. Luiza Granadeiro, Carla Fonseca and Isabel Neto: good to have you nearby!

To people from BIGR that I had the luck to meet or to work with. To Theo van Walsum for the data and, together with Nóra Baka, for the help on the development of the gold standard of this Thesis. To Ana Afonso, the “other portuguese Ana”. You were a good and supportive surprise!

To my colleagues and ex-colleagues from IBEB. Even if there is time and space between us, I know you are there. A particular thanks to Beatriz Lampreia.

Now my good friends who shared with me this journey. Friends from *Covilhã 11*, you are like a family to me. To Ana Teresa, the coffees we shared just before a hardworking and lonely day were very cheering. To Arminda, which friendship has grown side by side to my PhD project. Rita Nunes, Ana Sofia, Ângela Brandão, Celina, Helga, Pe. Henrique, Hugo and Catarina, thank you for your care, support and trust. A special thanks to Sandra, it is good to share my thoughts with you, and maybe because we have so much in common we rely on each other. And Claudia, you are the sister I would choose! Even if you so often do not hear the phone or read emails, I know you are always there. You have listened, put up with so much, and always with cheering words.

E à minha família, a Gouveia e a Rodrigues, e aos meus queridos sobrinhos. E agora em português. Ouviram as minhas lamentações mas sobretudo suportaram as minhas más disposições, silêncios e ausências. Aos meus irmãos, e aos cunhados Gouveia, em particular à Mónica; apesar das diferenças vocês são o meu porto seguro: *wherever*, *whenever*! Um obrigada especial à minha irmã João e aos meus sobrinhos João e José, que me acolheram tantas e tantas vezes em sua casa. Aos meus cunhados e cunhadas Rodrigues que, pela sua amizade, generosidade e proximidade, tantas vezes me ajudaram. Agradeço especialmente à minha mãe pela sua constante disponibilidade e preciosa ajuda. Aos meus sogros, que me acolhem como uma filha e nos apoiam incondicionalmente. E não posso deixar de mencionar o meu pai, que acredito vela por mim, e o meu querido Pe. Sanches. Estão sempre comigo!

Esta tese é dedicada aqueles que mais do que fazer parte de mim, se fundem em mim. E a eles lhes agradeço do fundo do meu coração. Primeiro os meus filhos. O meu lindo e doce bebé, o Lourenço. Que me mima tanto com os seus sorrisos e a sua ternura; e que me lembra como a felicidade pode estar nas pequeninas coisas. O meu menino, desde bebé um pequeno grande lutador, o Henrique. Cuja alegria me contagia e cuja exigência me fez manter ligada à vida, mesmo nos momentos mais complicados. A minha querida menina, a Eduarda, que foi crescendo e tomando consciência do que a rodeava. Quiseste estar ao meu lado, saber o que me preocupava, mesmo que isso te tenha angustiado e obrigado a crescer. Foste a minha companheira e uma grande motivação. Finalmente, o meu marido, companheiro, amigo. E até orientador em tantos momentos onde o meu pouco pragmatismo me prejudicou. Foste a minha força, sempre; a minha motivação, nos momentos mais difíceis; e foste uma inspiração como investigador e como pessoa. Mostraste-me o que é central, o que realmente importa. Focaste-me tantas vezes no essencial! Este trabalho também é teu. E é-te dedicado, com todo o meu amor.

Resumo

O alinhamento de imagens é o processo que permite estabelecer uma correspondência espacial entre o mesmo elemento de tecido representado em duas ou mais imagens, em diferentes momentos, de diferentes perspectivas, e/ou de diferentes modalidades. O processo computacional do alinhamento implica a determinação de uma transformação geométrica entre sistemas de coordenadas.

Esta transformação geométrica é calculada pela optimização de um critério de alinhamento, usualmente através de uma procura iterativa na qual uma estimativa inicial da transformação é gradualmente refinada por um processo numérico de optimização (Hill, Batchelor, Holden, & Hawkes, 2001; Zitova & Flusser, 2003). Em cada iteração uma medida de semelhança entre as imagens que se pretendem alinhar é calculada até que seja atingido um mínimo ou um máximo. No entanto, tais algoritmos de optimização podem convergir para uma solução incorrecta, i.e. para um extremo local, sendo muitas vezes pouco robustos face à existência destes extremos locais.

Alguns autores têm sugerido uma mudança de paradigma para o problema do alinhamento de imagens médicas, motivados principalmente pelos custos computacionais e pela robustez limitada face à presença de extremos locais (e.g. Chou & Pizer, 2013; Hoff, Komistek, Stefan, & Walker, 1998; Zhang et al., 2008). Nestes casos, o problema do alinhamento é encarado como um problema de regressão a partir de uma estratégia de aprendizagem, substituindo o cálculo da transformação por uma optimização iterativa. Em muitos destes trabalhos, as Redes Neuronais Artificiais, ou simplesmente Redes Neuronais, são o modelo de regressão escolhido para inferir os parâmetros da transformação geométrica.

No entanto, a literatura existente sobre métodos de alinhamento baseados em regressão deixa muitas questões em aberto, como quais as características das imagens (*image features*) a usar, qual o melhor modelo de regressão, bem como a relevância clínica dos mesmos. A investigação do alinhamento baseado na regressão é assim o principal tópico desta tese, focando-nos concretamente no alinhamento 3D/2D da angiografia por Tomografia Computadorizada com a angiografia por raios-X. Neste contexto é frequente a exigência da integração da informação obtida no período pre-operatório (e.g. imagens angiográficas 3D) com as imagens intra-operatórias. Por outro lado, o alinhamento de imagem por regressão nestas intervenções guiadas por imagem é exequível, uma vez que o processo de aprendizagem pode decorrer aquando do planeamento cirúrgico, o que ocorre antes do alinhamento propriamente dito.

Três questões de investigação foram assim formuladas:

- Como é influenciado o alinhamento pelas características das imagens (*image features*) usadas no processo de aprendizagem?
- Serão as redes neuronais um modelo de regressão adequado ao problema proposto?
- O alinhamento de imagem por regressão constitui uma solução clinicamente relevante para o alinhamento 3D/2D da angiografia por Tomografia Computorizada com a angiografia por raios-X angiografia das artérias coronárias, durante as intervenções guiadas por imagem?

Estas questões foram respondidas nos vários capítulos desta tese, sendo que o Capítulo 2 não incide directamente em nenhuma delas apesar do seu importante contributo. Este apresenta uma revisão de literatura das publicações sobre o alinhamento de imagem por regressão no contexto da imagem médica. A síntese das principais características dos métodos analisados também é realizada.

A primeira questão de investigação é abordada pelo Capítulo 3 e pelo Capítulo 5. O Capítulo 3 permite-nos compreender que a informação acerca da intensidade da imagem, associada a certas características, leva a uma muito boa aprendizagem. Mas o Capítulo 5 mostra-nos que esta dependência na intensidade não é compatível com o alinhamento de imagens 2D reais de angiografia. Por outro lado, algumas características não são afectadas pelas diferenças entre as imagens simuladas (usadas durante o treino) e as imagens reais, mas não contêm informação suficiente, pelo que não permitem obter resultados suficientemente precisos.

A segunda questão acerca das Redes Neuronais é estudada pelo Capítulo 4, onde diferentes modelos de regressão são comparados para este problema de alinhamento concreto.

O Capítulo 5 aborda a terceira e última questão de investigação, ou seja, a relevância clínica do alinhamento de imagem por regressão na integração automática da angiografia por Tomografia Computorizada com a angiografia por raios-X das artérias coronárias. Para este problema em particular, o método proposto com a sua actual configuração não é suficientemente preciso para ser usado na prática clínica. A principal dificuldade reside nas diferenças substanciais entre as imagens de treino (imagens de raios-X simuladas a partir de modelos 3D) e as imagens obtidas durante as intervenções (angiografias por raio-X). Tal significa que outros tipos de características têm que ser encontradas, mais robustas às diferenças entre as imagens, tais como a presença de cateteres nas imagens de raios-X ou a inexistência de vasos não visíveis durante oclusões totais crónicas (das artérias) coronárias, ou modelos 3D incompletos devido a uma visibilidade insuficiente nas angiografias por Tomografia Computadorizada (e.g. causada por ruído ou artefactos de movimento).

Os trabalhos recentes de Chou, Frederick, Mageras, & Chang (2013) e Chou & Pizer (2013) também sugerem algumas estratégias interessantes para melhorar a precisão, como as regressões em múltipla escala por um "treino hierárquico" (i.e. indo de um intervalo mais amplo para um intervalo mais estreito no que respeita às transformações usadas no treino) e correcção das diferenças de intensidade entre imagens simuladas e reais pela correspondência dos seus histogramas.

Por último, é recomendado testar o método de alinhamento por regressão desenvolvido noutras aplicações em que os dados de treino disponíveis sejam representativos dos dados de teste. Dois exemplos potencialmente interessantes são o alinhamento 3D/2D realizado em biopsias com agulha, guiadas por imagem, na coluna (van de Kraats, 2005; van de Kraats et al., 2006) e a correcção de movimento em tempo real de séries temporais de imagens (Luca, Tanner, & Székely, 2012).

Palavras-chave

Alinhamento de imagem 3D/2D, intervenções guiadas por imagem, artérias coronárias, regressão, Redes Neuronais, Perceptrão Multicamada, *gold standard*.

Abstract

This thesis focuses on the development and evaluation of a registration-by-regression approach for the 3D/2D registration of coronary Computed Tomography Angiography (CTA) and X-ray angiography. This regression-based method relates image features of 2D projection images to the transformation parameters of the 3D image by a nonlinear regression. It treats registration as a regression problem, as an alternative for the traditional iterative approach that often comes with high computational costs and limited capture range.

First we presented a survey of the methods with a regression-based registration approach for medical applications, as well as a summary of their main characteristics (Chapter 2). Second, we studied the registration methodology, addressing the input features and the choice of regression model (Chapter 3 and Chapter 4). For that purpose, we evaluated different options using simulated X-ray images generated from coronary artery tree models derived from 3D CTA scans. We also compared the registration-by-regression results with a method based on iterative optimization. Different image features of 2D projections and seven regression techniques were considered. The regression approach for simulated X-rays was shown to be slightly less accurate, but much more robust than the method based on an iterative optimization approach. Neural Networks obtained accurate results and showed to be robust to large initial misalignment.

Third, we evaluated the registration-by-regression method using clinical data, integrating the 3D preoperative CTA of the coronary arteries with intraoperative 2D X-ray angiography images (Chapter 5). For the evaluation of the image registration, a gold standard registration was established using an exhaustive search followed by a multi-observer visual scoring procedure. The influence of preprocessing options for the simulated images and the real X-rays was studied. Several image features were also compared. The coronary registration–by-regression results were not satisfactory, resembling manual initialization accuracy.

Therefore, the proposed method for this concrete problem and in its current configuration is not sufficiently accurate to be used in the clinical practice. The framework developed enables us to better understand the dependency of the proposed method on the differences between simulated and real images. The main difficulty lies in the substantial differences in appearance between the images used for training (simulated X-rays from 3D coronary models) and the actual images obtained during the intervention (real X-ray angiography). We suggest alternative solutions and recommend to evaluate the registration-by-regression approach in other applications where training data is available that has similar appearance to the eventual test data.

Keywords

3D/2D image registration, image guided interventions, coronary arteries, regression, Neural Networks, Multiple Layer Perceptron, gold standard.

Contents

Acronyms

AOC	Area Over the Curve
CLARET	Correction via Limited-Angle Residues in External Beam Therapy
CT	Computed Tomography
CTA	Computed Tomography Angiography
DRR	Digitally Reconstructed Radiographs
ECG	Electrocardiogram
fMRI	Functional Magnetic Resonance Image
ICP	Iterative Closest Point
IGRT	Image-Guided Radiation Therapy
IGT	Image-Guided Therapy
k-NN	k Nearest Neighbours
LIBVSM	Library for Support Vector Machines
LR	Linear Regression
LSM	Least Square Mean
MLP	Multi-Layer Perceptron
MLP-CG	Multi-Layer Perceptron with Conjugate Gradient Optimization
MLP-LM	Multi-Layer Perceptron with Levenberg-Marquardt Optimization
MRI	Magnetic Resonance Image
mTRE	Mean Target Registration Error
NN	Neural Network
PA	Principal Axis
PCA	Principal Component Analysis
PCI	Percutaneous Coronary Intervention
PET	Positron Emission Tomography
PR	Polynomial Regression
RBF	Radial Basis Function
REC	Regression Error Characteristic
SVR	Support Vector Regression
TKR	Total Knee Replacement

In Praise Of Learning

Learn the elementary things!
For those whose time has come
It is never too late!
Learn the ABC. It won't be enough,
But learn it! Don't be dismayed by it!
Begin! You must know everything.
You must lake over the leadership.

Learn, man in the asylum!
Learn, man in the prison!
Learn, woman in the kitchen!
Learn sixty year olds!
You must take over the leadership.
Seek out the school, you who are homeless!
Acquire knowledge, you who shiver!
You who are hungry, reach for the book:
it is a weapon.
You must take over the leadership.

Don't be afraid to ask, comrade!
Don't be talked into anything.
Check for yourself!
What you do not know yourself
you don't know.
Scrutinize the bill,
it is you who must pay it.
Put your finger on each item,
ask: how did this get there ?
You must take over the leadership.

Bertolt Brecht

Chapter 1 Introduction

Image registration is the process of establishing a spatial correspondence between the same element of tissue represented in two or more images, at different times, from different viewpoints, and/or by different modalities. The computational process of registration yields a geometric transformation between coordinate systems.

This geometric transformation is computed by optimizing a registration criterion, usually by an iterative search in which an initial estimate of the transformation is gradually refined by a numerical optimization procedure (Hill et al., 2001; Zitova & Flusser, 2003). In each iteration, a similarity measure between the images to be registered is computed until a minimum is reached. However, these optimization algorithms frequently converge to an incorrect solution, i.e. to a local optimum, leading to a small capture range. Different strategies have been implemented when the initial position is not close to correct alignment and to avoid possible local extremes, bringing the registration parameters within the capture range. Additionally to this problem, the use of image interpolation, in each iteration for the widely used voxel-based methods, can lead to high computational costs. An established approach to avoid false local optima improving the registration speed is the use of hierarchical multi-resolution and/or multi-scale search strategies. But such solutions are not suitable for all image data, because some image features may be suppressed and other non-corresponding image features may become more similar (Markelj, Tomaževič, Likar, & Pernuš, 2012).

Mainly motivated by the computational costs and limited capture range, some image registration methods have been presenting a different solution (e.g. Chou & Pizer, 2013; Hoff, Komistek, Stefan, & Walker, 1998; Zhang et al., 2008). They propose a paradigm change of the registration problem for medical applications. Instead of computing the transformation of the registration by an iterative optimization, they treat the registration as a regression problem by a learning strategy. In many cases, Artificial Neural Networks or simply Neural Networks (NNs) are chosen to infer the regression model. They have an excellent ability to learn the relationship between input and output from a data set without any prior knowledge, and without any assumptions about the statistical distribution of the data (Haykin, 1999). However, the existing literature on registration-by-regression methods leaves many open questions, such as the choice of input features, the choice of regression model, and the clinical applicability.

The investigation of the regression-based registration is the main topic of this Thesis. It focused on the 3D/2D registration of coronary CTA and X-ray angiography in the image guided context. The image guided interventions often require integration of preoperative image with

intraoperative images, which makes suitable a learning process during the intervention planning, before the registration is required. Three research questions are thus formulated attending different aspects:

- How is the registration influenced by the image features used in the learning process?
- Are Neural Networks an adequate regression model for this problem?
- Is the registration-by-regression method a clinically relevant solution for 3D/2D registration of coronary CTA and X-ray angiography during image-guided interventions?

The outline of this Thesis is as follows. In Chapter 2 we present a survey of the publications with a regression-based registration approach for medical applications, as well as a summary of their main characteristics. Chapter 3 presents in detail the framework of the 3D/2D image registration-by-regression method proposed in this thesis. The method is evaluated using simulated X-ray images and compared with a conventional registration method based on iterative optimization, developed for the specific application. In Chapter 4, a comparative evaluation of different regression techniques is performed for the application considered. In Chapter 5 we treat the extension of Chapter 3 to real X-ray images. Our aim is to integrate of 3D preoperative CTA of the coronary arteries with intraoperative 2D X-ray angiography images with the registration-by-regression method proposed. For the evaluation of the image registration, a gold standard registration was established using an exhaustive search followed by a multi-observer visual scoring procedure. Several image features were compared as well as different image preprocessing options. Finally the summary of the Thesis is presented in Chapter 6.

Perseverance

Say not that work is e'er ill-spent,
Say not that effort fails or seems;
Say not that he o'er labour bent
Is one in the world's many dreams.

Alexander Search

Chapter 2
Medical Image Registration based on Regression: a survey

This chapter is based on the manuscript:

Medical Image Registration based on Regression: a survey. Gouveia, A. R., Freire, L., Almeida, P., & Klein, S., *submitted*.

Abstract

In this chapter, we present a survey on regression-based image registration approaches for medical applications, as well as a summary of their main characteristics inspired by the classification proposed by Maintz & Viergever (1998). We adapted the criteria used by these authors, and we considered a) three aspects regarding the methodology of the registration-by-regression framework, namely the type of *features*, the regression *model*, and the *role* of the regression in the entire process, and b) five more general aspects, namely the *application* of the work, the *object imaged,* the *transformation nature*, the *evaluation* of the registration and the *motivation* for the approach.

This survey aims to provide insight into the different formulations of registration-by-regression methods, to give an overview of the applications for which this has been proposed, and to compare to what extent the different approaches have been evaluated. The main conclusion is that, while results on simulated data show promising accuracies, few registration-by-regression methods have been evaluated using real clinical data.

2.1 Introduction

"Image registration is the process of overlaying two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors" (Zitova & Flusser, 2003). It is often required in different areas, such as remote sensing, computer vision, and medical imaging among others, being a crucial step for the image analysis in multi-view, multitemporal and multimodal circumstances (Zitova & Flusser, 2003). In the specific context of medical image analysis, one may find many applications of image registration, such as to combine the information from multiple imaging modalities (e.g., anatomical body structure from magnetic resonance image (MRI) with functional and metabolic body activities from positron emission tomography (PET), or relating preoperative images and surgical plans to the physical space of the operating room during image-guided surgery, or even relating an individual's anatomy to a standardized atlas (Hajnal, Hawkes, & Hill, 2001; Zitova & Flusser, 2003).

Most medical image registration algorithms use an iterative approach, in which an initial estimate of the geometric transformation required to bring the images into correspondence is gradually refined by a numerical optimization procedure (Hill et al., 2001). In each iteration, a similarity measure between the images to be registered is computed until a minimum or a maximum value is reached. However, these optimization algorithms can convergence to an incorrect solution, *i.e.*, to a local optimum of the similarity measure's energetic landscape. As a result, such algorithms may have a small capture range (*i.e.*, large transformations cannot be recovered). Additionally, for the extensively used intensity-based registration methods, the use of image interpolation algorithms in each iteration implies high computational costs. Motivated by these disadvantages, image registration methods replacing iterative optimization by other strategies have been proposed in the literature. Among these works, several authors have explored the use of regression models for this purpose.

In this chapter, we present a survey on regression-based registration approaches proposed for medical applications. We provide a summary of their main characteristics and we classify the methods according to the three main methodological components: the choice of features, the regression model, and the way the regression takes part of the registration process. We also define five criteria to classify the presented methods, using a methodology inspired by the classification proposed by Maintz & Viergever (1998). To limit the scope of this survey, approaches for motion tracking are not included.

In the literature we find many reviews on image registration methods, either for generic applications (Glasbey & Mardia, 1998; Salvi, Matabosch, Fofi, & Forest, 2007; Zitova & Flusser, 2003), or for medical imaging context (Brown, 1992; Elsen, Pol, & Viergever, 1993; Hill et al., 2001; Maintz & Viergever, 1998; Maurer & Fitzpatrick, 1993; Oliveira & Tavares,

2014; Wyawahare, Patil, & Abhyankar, 2009) and for specific transformation nature (Sotiras, Davatzikos, & Paragios, 2013), or even for specific applications (Markelj et al., 2012), or focused on methods for medical images inspired by computational intelligence (Ramirez, Durdle, & Raso, 2003). However, to the best of our knowledge, no survey specifically dedicated to registration-by-regression methods can be found in the literature.

2.2 The image registration process

The term *registration* means to spatially align two or more images of the same or different subjects, acquired with the same or different imaging modalities, and also the registration of images with the coordinate system of a treatment device or tracked localizer (Hill et al., 2001). The spatial alignment implies the computation of a geometric transformation that relates corresponding points in the involved images. This transformation can denote two different mappings (Hill et al., 2001). One is a *spatial mapping*, which relates the position of features, in one image or coordinate space, with the position of the corresponding features in another image or coordinate space. The other relates both the position of corresponding features and the intensity information of corresponding positions. The first maps the coordinate systems of the images and the second maps an image to another image. This more complete mapping needs resampling and interpolation.

The image registration survey presented by Zitova & Flusser (2003) classifies the existing methods into *feature-based methods* and *area-based* . The former are recommended if the images contain enough distinctive and easily detectable -features, whereas the latter are best suited for the images without clearly identified features (Zitova & Flusser, 2003). The feature-based registration requires selection of a number of well-defined landmarks, which may be difficult to achieve automatically in the case of noisy medical images. However, in some cases this is solved by the interactive selection by an expert user (e.g. based on anatomical criteria) or by introducing extrinsic features, rigidly positioned with respect to the patient. Area-based methods depend on the intensity information of pixels (or voxels) of the images to be registered, considering the entire image or just a predefined region of interest. In the medical context, area-based image registration methods are usually named as *voxel-based* or *intensity-based*. In this work, we adopt this last term.

Different steps can be identified during the image registration process. Registration based on features relies on an initial identification of features (e.g. points or curves) or on the extraction of those features (e.g. by segmentation); then, the correspondence (or matching) between the features of the images to be registered is established by feature descriptors, similarity measures or spatial relationships among the features (Zitova & Flusser, 2003). For intensity-based methods, this correspondence estimation is performed by computing a

correlation between images or a template matching. Based on the feature or intensity information, a spatial coordinate transformation is estimated. Then, one image is mapped onto the coordinates of the other by image interpolation. Most medical image registration algorithms use an iterative approach, in which an initial estimate of the transformation is gradually refined by a numerical optimization procedure (Hill et al., 2001). In each iteration, the current estimate of the transformation is used to calculate a similarity measure. The optimization algorithm then makes another estimate of the transformation and evaluates the similarity measure again. This process continues on a trial-error basis until the algorithm converges, which means that no further transformation can be found that results in a better value of the similarity measure, to within a preset tolerance.

The registration methods focused on this survey propose a change of paradigm from an iterative optimization to a regression approach. The geometric transformation is estimated by a regression model, using a learning-based strategy. These methods are described in detail in the next section.

2.3 Regression approaches in image registration methods

Generally, image registration methods based on regression relate images or image features to the corresponding spatial transformations. In a library-based or template matching style, a training set is generated and the regression model is estimated by a learning process. Then the registration takes place by the computation of the transformation parameters using the function previously obtained. Mathematically, the regression approaches predict an output Y (e.g. transformation parameters) with a regression model f given an input vector $X = (X_1, \dots, X_P)$ (e.g., the image features), such as $Y = f(X)$. To estimate the parameters of the prediction model f, a set of N measurements (x_i, y_i), i.e. the training set, is used.

We find in the literature some authors adopting this regression approach, however their methods can differ in many aspects. They can use different *inputs*, such as points coordinates (e.g., Ryan, Heneghan, & de Chazal, 2004) or Fourier coefficients (e.g., Abche, Yaacoub, Maalouf, & Karam, 2006); different *outputs*, such as transformation parameters (e.g., Gouveia, Metz, Freire, & Klein, 2012a) or deformations coefficients (e.g., Kim, Wu, Yap, & Shen, 2012); and different *regression models* like MLP (e.g., Zhang et al., 2008) or k-NN (e.g., Hoff, Komistek, Dennis, Gabriel, & Walker, 1998). The way the regression takes part in the registration process, referred in this chapter as the *role* in the overall registration process, also differs. The registration-by-regression method can be used on its own (e.g., Liu, Yan, & Zhang, 2006), or as part of a more comprehensive multi-stage registration process (e.g., as initialization method or as refinement stage) (e.g., Li, Gao, Wu, & Li, 2010). In

Table 2.1 Regression approaches in medical image registration methods: the registration methodology. Papers from the same research group describing largely the same method were grouped per row.

	Features	Model	Role
(a) Yan et al. (2004)	Points	$(x,y,z) \rightarrow$ **MLP** $\rightarrow D_{surfaces}$	Entire process
(b) Zhang et al. (2008)			PA $\rightarrow$ Fine reg.
(c) Liu et al. (2006)			Entire process
(d) Li et al. (2010)	Points	$(x,y,z) \rightarrow$ **MLP** $\rightarrow (x',y',z')$	Initialization $\rightarrow$ ICP reg.
(e) Wang et al. (2003)			Entire process
(f) Wachowiak et al. (2002)	Points	$(x,y,z) \rightarrow$ **NNs** $\rightarrow (x',y',z')$	Entire process
Ryan et al. (2004) (g) Heneghan et Maguire (2002) Heneghan et al. (2002)	Points	$(x,y,z) \rightarrow$ **MODELS** $\rightarrow (x',y',z')$	Entire process
(h) Qi, Gu & Xu (2008) Qi, Gu & Zhao (2008)	Image	$T \rightarrow$ **SVR** $\rightarrow$ *Similarity Measure*	Similarity Measure computation
(i) Kim et al. (2012,2010)	Signature Vectors	PCA $\rightarrow$ **SVR** $\rightarrow$ *Deformations*	Initialization
(j) Chou et al (2013,2011,2010) Chou et Pizer (2013)	Image Difference	$I - I_0 \rightarrow$ **MODELS** $\rightarrow T$	Entire process
(k) Gouveia et al. (2012a,2012b)	2D moments, eigenvalues, eigenvectors	*2D M eigenvalues eigenvectors* $\rightarrow$ **MODELS** $\rightarrow T$	Entire process
(l) Freire et al. (2010a,2010b)	Fourier Coefficients	*Fourier Coef.* $\rightarrow$ **MLP** $\rightarrow T$	Entire process
(m) Abche et al. (2006)			
(n) Banks et al. (1996)	Fourier Coefficients	*Fourier Coef.* $\rightarrow$ **K-NN** $\rightarrow T_\theta$	Initialization
(o) Hoff et al. (1998)	Image difference $\downarrow$ number of pixels	$N \rightarrow$ **K-NN** $\rightarrow T_\theta$	Initialization

For the models: MLP means Multi-Layer Perceptron, NNs means Neural Networks, SVR means Support Vector Regression and k-NN means k Nearest Neighbours. For inputs/outputs: $\boldsymbol{T}$ means transformation parameters and $\boldsymbol{T_\theta}$ means the rotation parameters. For the role: PA means Principal Axis alignment, reg. means registration and ICP means Iterative Closest Point.

Table 2.2 Regression approaches in medical image registration methods: application, object imaged, transformation nature, evaluation and motivation.

	Application	Object imaged	Transf. nature	Evaluation	Motivation
(a) Yan et al. (2004)	3D Surface	Spine model	Rigid	Simulated	↑ Speed
(b) Zhang et al. (2008)	3D Surface	Spine model	Rigid	Simulated	↑ Speed
(c) Liu et al. (2006)	3D Surface	Ears model	Rigid + Scaling	Simulated	↑ Speed
(d) Li et al. (2010)	3D Surface	Head model	Rigid	Simulated	↑ Speed
(e) Wang et al. (2003)	3D/3D PET to MRI	Brain	Rigid	Simulated Real	-
(f) Wachowiak et al. (2002)	2D/2D MRI and CT	Brain Spine	Elastic	Simulated	↓ Complexity
Ryan et al. (2004) (g) Heneghan et Maguire (2002) Heneghan et al. (2002)	2D/2D Retinal images	Eye	Rigid Affine Non-rigid	Real	↑ Versatility
(h) Qi, Gu & Xu (2008) Qi, Gu & Zhao (2008)	3D/2D CT/MRI to X-ray	Brain Thorax	Rigid Non-rigid	Simulated	↑ Speed
(i) Kim et al. (2012,2010)	3D/3D MRI	Brain	Elastic	Simulated Real	↑ Speed ↑ Accuracy
(j) Chou et al (2013,2011,2010) Chou et Pizer (2013)	3D/2D IGRT CT to X-ray	Head-and-neck Lung	Rigid Non-rigid	Simulated Real	↑ Speed
(k) Gouveia et al. (2012a,2012b)	3D/2D IGT CT to X-ray	Heart (Coronary arteries)	Rigid	Simulated	↑ Range
(l) Freire et al. (2010a,2010b)	3D/3D fMRI	Brain	Rigid Affine	Simulated Real	↑ Speed
(m) Abche et al. (2006)	2D/2D MRI	Brain	Rigid + Scaling	Simulated	↑ Accuracy
(n) Banks et al. (1996)	3D/2D Model to X.-ray	Knee implants	Rigid	Simulated Real (in vitro)	↑ Range
(o) Hoff et al. (1998)	3D/2D Model to X-ray	Knee implants	Rigid	Simulated Real (in vitro)	↑ Range

For the applications: PET means positron emission tomography, MRI means magnetic resonance image, CT means computed tomography and fMRI means functional magnetic resonance image.

Section 2.3.1, we classify the methods according to these characteristics (all related to the registration methodology) and, when it is possible, we cluster the methods by their similarities on methodology aspects, describing what is common and what is not. Table 2.1 gives a schematic overview of all methods that are discussed, according to this categorization.

Next to this categorization on methodological properties, we adopted the criteria used by Maintz & Viergever (1998) to classify the literature on registration-by-regression with respect to the medical context of the work. Five items were considered: the medical *application* of the work, such as Image-Guided Treatments (IGT) (e.g., Yan et al., 2004); the *object imaged,* such as the brain (e.g., Wachowiak, Smolíková, Zurada, & Elmaghraby, 2002), the *transformation nature*, such as rigid (e.g., Banks & Hodge, 1996); some details about the *evaluation of the registration,* like if real clinical images are used (e.g., Freire, Gouveia, & Godinho, 2010b) and if the quantification of the error is performed (e.g., Chou, Frederick, Mageras, Chang, & Pizer, 2013); and the *motivation* of the registration-by-regression approach, such as registration speed (e.g., Qi, Gu, & Xu, 2008). In Sections 2.3.2 and 2.3.3, we briefly discuss and classify the methods using these criteria, and a schematic summary is presented in Table 2.2.

2.3.1 Registration methodology

Table 2.1 indicates schematically the *inputs*/*outputs* used, the *regression model*, and the *role* in the overall registration process.

A common registration-by-regression methodology was presented for the surface registration of 3D models (H. Liu et al., 2006; Yan et al., 2004; Zhang et al., 2008) (Table 2.1 – a, b, c). Their goal was not necessarily finding the transformation between surfaces explicitly from the used regression model but to find the closest surface to a reference surface by computing the distances between them, as for C. Yan et al. (2004) and Zhang et al. (2008). These authors used a Neural Network (NN) to create a distance function that maps 3D corresponding points, on the surface to be registered and the reference surface, to the distance between surfaces.

Yan et al. (2004) (Table 2.1.a) treated the 3D surface registration of spine models for its application in an image-guided scenario. Firstly, a surface extraction of two vertebrae bodies of a human spine is performed from a semi-automatic segmentation of the vertebrae in computed tomography (CT) images. Then, a Multi-layer Perceptron (MLP) with two hidden layers is used to model a distance function $d(x, y, z)$ which measures the distance of any spatial point (x, y, z) of a surface from the reference surface.

Zhang et al. (2008) (Table 2.1.b) proposed a 3D surface-based registration very similar to the previous approach. They only differed by a coarse registration based on principal axis alignment, before the fine registration and after the surface extraction. According to the authors, it is a computationally inexpensive procedure but capable of giving good starting points. Still for 3D surface registration, Liu et al. (2006) (Table 2.1.c) proposed a method where the rotation and translation matrices can be computed directly from the NN. In this case, and before the 3D surface registration, a mesh Principal Component Analysis (PCA) was used to analyze vertex data on the models and to extract the principal directions of the three largest spreads of the distributions of these vertices. Then, a MLP with only one hidden layer was used. The vertices' coordinates of the reference surface are the inputs and the distance between surfaces is the output. The connections between the layers (i.e. input, hidden and output layers) are defined in such a way that the weight matrix between the input and the hidden layer is the unknown registration matrix when the NN converges

Another 3D surface registration was proposed by Li, Gao, Wu, & Li (2010) (Table 2.1.d) but in this case using a NN relating the coordinates of corresponding points. This method is used as initialization procedure for a fine registration based on the Iterative Closest Point (ICP) method, commonly used for the registration of 3D shapes (Zhang et al., 2008). The corresponding pairs between surfaces to be matched were found according to the target points' curvature and color information. The NN differs from the previously used (H. Liu et al., 2006) by eliminating the distance computation. Therefore it is a MLP without any hidden layers, where inputs and outputs are the corresponding points' coordinates of the surface to be aligned and the reference surface, respectively. In this way, the initial estimate for the transformation is found by the computation of the weight matrix.

Other authors (W. F. Wang & et al, 2003) (Table 2.1.e) used a MLP to also relate the coordinates of corresponding points but for a 3D/3D multi-modal registration, namely for registration of 3D PET images to 3D MRI. They tested various and more complex MLP topologies than Li, Gao, Wu, & Li (2010) (Table 2.1.d), previously mentioned. The corresponding points were chosen taking into consideration anatomic criteria.

The same registration methodology of the last two works was used by Wachowiak, Smolíková, Zurada, & Elmaghraby (2002) (Table 2.1.f) but including elastic deformations. 2D Elastic registration based on landmarks was performed by a set of corresponding points manually selected from the image to be registered and from the reference image. They compared different NN, such as radial basis function networks, multi-layer perceptron networks with backpropagation and backpropagation with Bayesian regularization, using different levels of complexity. The input and output units are the corresponding points' coordinates.

Following previous approaches, we find in the literature some authors relating corresponding points by a regression model for the registration of images of the retina with different

ophthalmological purposes (Heneghan, Maguire, Ryan, & de Chazal, 2002; Heneghan & Maguire, 2002; Ryan et al., 2004) (Table 2.1.g). In the case of Ryan et al. (2004) the method comprises (1) the marking of control points (i.e., landmarks) in a reference image and a distorted image, (2) the matching of control points across the two images and then (3) the registration based on (2). Given a set of *k* matching control points in both images, the transformation parameters are estimated from the coordinates of the k matched control points in the reference image and in the distorted image. An affine transformation is used as the basis for the automatic matching of pairs of control points across two images, providing a good first order model of the transformation required for registration. But two other transformations are considered for more general geometric distortion correction (i.e., the bilinear and second order polynomial transformation), other than rotation, translation and scale. These authors applied the method in three different situations: the inter-modality image registration of an optical image and a fluorescein angiogram, temporal registration of two color fundus photographs images of an infant eye, and intra-modality registration of a set of seven standard field optical photographs. The preceding works (Heneghan et al., 2002; Heneghan & Maguire, 2002) are very similar, but fewer applications were considered for the evaluation of the method.

As well as the 3D surface registration methods proposed by C. Yan et al. (2004) and Zhang et al. (2008) referred before, Qi, Gu, & Xu (2008) and Qi, Gu, & Zhao (2008) (Table 2.1.h) used a regression model to compute the similarity measure values, rather than directly computing the transformation. With their approach, they avoided the time consuming generation of the 2D projections, a necessary procedure for 3D/2D image registration based on intensity. In a pre-operative stage, a 3D model is obtained from brain MRI images and simulated X-ray projection images (i.e., digitally reconstructed radiographs - DRRs) were rendered according to their real coordinates and position of the focal point of the X-ray scanner. In the intra-operative stage, the similarity metric between the pre-operative DRRs and the X-ray is computed; the Support Vector Regression (SVR) is trained by relating the transformation parameters previously applied to the rendered DRRs and the corresponding similarity metric values. Therefore, the time-consuming source is not the DRR rendering but the similarity measure computation and SVR learning, which are both performed during the intervention.

Other authors also used SVR but in a different context (Kim, Wu, Yap, & Shen, 2010; Kim et al., 2012) (Table 2.1.i). They presented a learning-based initialization for image registration of a population of subjects with respect to a template. The method starts with a training stage using the SVR to learn the correlation between image appearances from different subjects and their respective deformations. When an individual image needs to be registered, this learned correlation model automatically predicts an initial deformation field, which is then used to generate a corresponding intermediate template for registration refinement. Which means the registration will be performed onto a template much more similar to the

subject than without this initialization step. Thus, the registration algorithm only needs to estimate residual deformations, which allows to reduce the computational time and to avoid the local minima as the authors enhanced. To compute brain image appearances, they use PCA to capture the principal modes of brain deformations (i.e. brain outlines and boundaries along the different brain interfaces) using a finite set of parameters, which leads to low-dimensional features (referred as *signature vectors)*.

A regression strategy was also studied for image registration in Image-Guided Radiation Therapy (IGRT). Chou, Frederick, Chang, & Pizer (2010) (Table 2.1.j) proposed a regression-based matching methodology for patient re-positioning, which was inspired by a machine learning strategy (Cootes, Edwards, & Taylor, 2001), and later adapted to 3D/2D registration (Chou et al., 2011, 2013; Chou & Pizer, 2013). First, in the training stage, a range of transformations for patient movements is applied to a CT image in the planning position. 2D projections (DRRs) of transformed and non-transformed images are obtained and the intensity differences of DRRs with relation to the non-transformed DRR are computed. Then, a regression function is computed relating these intensity differences with the transformation values. This procedure was performed as a "hierarchical training", i.e. it was performed from large to small scales of training considering different transformation ranges, and leading to multi-scale regressions. In the treatment stage, the learned regression model is applied iteratively to the successive residuals between the DRR of the transformed CT and the 2D real-time projection, leading to the registration transformation. This method was firstly tested on simulated images of head-and-neck CTs, and considering a rigid transformation. In further works, this matching methodology, referred as CLARET, i.e. *Correction via Limited-Angle Residues in External Beam Therapy*, was adapted and its application extended to deformable registration and to real images (Chou et al., 2011, 2013; Chou & Pizer, 2013).

The registration-by-regression presented by (Gouveia et al., 2012a) (Table 2.1.k) for an image-guided application bears some similarities with the works just described. They treat the registration of 3D preoperative coronary CTA images to 2D intraoperative X-ray images like a nonlinear regression problem. The regression function is determined in a supervised learning stage using MLP; it relates image features (2D geometric moments, eigenvalues and eigenvector from PCA) of 2D projection images to the transformation parameters of the 3D image. For evaluation, simulated X-ray images (DRRs) were generated from coronary artery tree models derived from 3D CTA scans. This method has shown to be slightly less accurate than a conventional registration method based on iterative optimization, but much more robust. The same authors also investigated other regression methods for this problem, performing a comparative evaluation (Gouveia, Metz, Freire, & Klein, 2012b).

Another methodology treating the registration as a regression problem proposed in the literature is a MLP using Fourier coefficients as inputs and transformation parameters as outputs (Abche et al., 2006; Freire, Gouveia, & Godinho, 2010a; Freire et al., 2010b), for

fMRI and MRI images. The works of Freire, Gouveia, & Godinho, (2010a, 2010b) (Table 2.1.l) focused on 3D fMRI time series registration for motion correction. For the learning of the MLPs, an image training set was obtained by applying geometric transformations of typical motion amplitudes encountered in fMRI studies. The Fourier transform was applied on the images and the coefficients extracted. They did not use the overall image information, instead they confined to small subsets extracted from the images' 3D Fourier spaces. Different subsets' sizes were studied, which were comprised to the first octant of Fourier space (including the DC component). In these works, one MLP per each transformation parameter was used. Freire et al. (2010a) trained 6 different MLP for a 3D rigid geometric transformation, and in Freire et al. (2010b) 12 different MLP were used for a 3D affine geometric transformation (with 3 parameters for scaling and 3 parameters for shearing).

A similar approach was proposed by Abche et al. (2006) (Table 2.1.m) for the registration of 2D MRI images, but computing the transformation by using only one MLP. Considering translation, rotation and scaling parameters, the MLP had 5 output units. They studied the influence of the number of coefficients used (the window size considered on the Fourier space) and the influence of noise. Elhanany et al. (2000) proposed an identical method but not for medical images, paying a special attention to the relation between the method robustness and image noise. According to the authors, the noise introduction in the training set improves generalization and the NN becomes less sensitive to distorted inputs.

Two 3D/2D image registration methods based on template matching shall also be referred (Banks & Hodge, 1996; Hoff et al., 1998). In these methods, a library is first constructed by the features representing the 2D projections of the 3D object, for a predefined set of transformations, and the corresponding out-of-plane transformation values. Then, the out-of-plane transformation parameters of a given 2D projection are determined by a shape matching using the library 2D templates. Finally, the remaining and not yet computed parameters are obtained by a perspective imaging model.

Banks & Hodge (1996) (Table 2.1.n) studied the 3-D total knee replacement (TKR) kinematics during dynamic activities. This was achieved by imaging the knee joint as it moves, using X-ray fluoroscopy to obtain a sequence of images in which the prosthesis is projected. For this purpose, the silhouettes contours of the 2D projections were represented by Fourier coefficients, and together with the correspondent out-of-plane rotations, built a library. An initial estimation of the out-of-plane rotation of an unknown projection was yielded by computing a weighted average of the library coefficients interpolating between the three closest coefficients.

A similar method but with several new contributions in order to improve its accuracy is presented by Hoff, Komistek, Stefan, & Walker (1998) (Table 2.1.o). Their goal was to determine the relative pose (position and orientation) of two knee implant components with

respect to each other, under in vivo conditions, from X-ray fluoroscopy images. For the library, the 2D projections of the 3D models were represented as binary images after being centered, scaled to a constant area, and rotated so that their principal axis was aligned with the horizontal axis. To find the best match of a given 2D image, after the same preprocessing described, the image was systematically subtracted from each library entry and a "score" (i.e. the number of unmatched pixels) was generated for each. The best match corresponds to the smallest number of unmatched pixels.

Differently, Roche, Malandain, Pennec, & Ayache (1998) did not propose a new registration concept, but instead they used the regression idea to propose a new similarity measure to be optimized in an iterative scheme for multimodal registration. Addressing the cases when a functional correlation can be assumed, the degree of functional dependence between the images is considered as matching criterion. The evaluation of this functional dependence is reduced to a regression problem and the correlation ratio is used as similarity measure. Minimal assumptions regarding the nature of the function itself are made. The same authors expand this method performing the registration of intraoperative 3D US images with preoperative MR images by incorporating multivariate information from the MR data (Roche, Pennec, Malandain, & Ayache, 2001). They correlate the US intensity with both the MR intensity and the MR gradient magnitude, once the US images enhance the interface between anatomical structures; they also incorporate a robust intensity-based distance measure in order to handle with a variety of US artifacts.

2.3.2 Application, object imaged and transformation nature

Table 2.2 indicates schematically and in more detail the applications, object imaged and transformation natures.

The methods described in this survey use image registration for different applications. Most of them focus on IGT (Table 2.2 - a, b, c, h, j, k, n, and o), either for 3D surface matching (Table 2.2 – a, b, c, d) or for 3D/2D registration (Table 2.2 – h, j, k, n, o).

In the majority of the cases, the imaged objects were bone structures (Table 2.2 – a, b, c, d, f, h, j, m, n, o), like the knee (Table 2.2 – n, o) or the spine (Table 2.2 – a, b, f). Most of the works described used rigid (Table 2.2 – a, b, d, e, h, k, l, m, n, o), rigid with scaling (Table 2.2 – c, j), or affine transformations (Table 2.2 – g, l). The rigid transformation was used specially for the bone structures (Table 2.2 – a, b, d, h, l, n, o). However, some authors (5 out of 15) implemented their methods also considering non-rigid transformations (Table 2.2 – g, h, j) and some developed their methods originally for elastic deformations (Table 2.2 - f, i).

2.3.3 Motivation and evaluation

Table 2.2 indicates schematically the motivation and evaluation of the methods described.

The two main motivations of the cited works, and corresponding applications, are the high computational effort of iterative methods, which compromises the registration speed, and their small capture range.

When a learning strategy is used by inferring a regression model, the computation of the registration transformation is potentially very fast. The training or learning process, which usually occurs before the registration and when time is not a constraint, usually takes longer. All the works showed good or at least promising results, considering the range, the speed and/or the accuracy. Moreover, most works compared the registration results of their method with a popular alternative method for that application (Table 2.2 – b, c, d, g, h, i, j, k, l, n). However, only some of them (7 out of 15) registered clinical data (Table 2.2 – e, g, i, j, l, n, o), which is crucial for applicability in clinically relevant situations.

Five works (5 out of 15) presented a quantitative evaluation using real data and the experiments presented by (Heneghan et al., 2002; Heneghan & Maguire, 2002; Ryan et al., 2004) report a limited number of results. Below, the evaluation of the registration is described in more detail for Freire, Gouveia, & Godinho (2010a, 2010b) and for Chou et al. (2011, 2013, 2010) and Chou & Pizer (2013), in which regression assumes the principal role in image registration (Table 2.2 - j and l).

The method proposed in Freire, Gouveia, & Godinho (2010a, 2010b) focused on 3D fMRI time series registration for motion correction and their motivation was to develop a fast registration method that accomplishes a prospective registration relying on a limited number of Fourier coefficients of the images to be aligned. For simulated images, the error was computed and compared to an intensity-based registration method and yielded a similar accuracy. For real images, the error analysis was performed by the inspection of the evolution of the estimates of the parameters for all frames and for both methods, since they did not have a ground truth. The NN motion estimates presented a more irregular general behavior then the iterative method. And when affine registration was considered, there was a bias due to the presence of functional activation. The authors suggested that activation must be simulated and learned by the NN.

Chou et al. (2011, 2013, 2010) and Chou & Pizer (2013) studied 3D/2D registration methods in Image-guided Radiation Therapy (IGRT). Their methods showed to be fast and to have good results in localizing a tumor under rigid motion in the head and/or neck, and under respiratory deformation in the lung. They evaluated their methods with simulated and real images. The evaluation of real images performed in Chou & Pizer (2013) included the

quantification of the registration error computed on the 3D tumor centroid. According to the authors, the accuracy obtained in real-time, together with the low dose requirements thanks to the small number of 2D projection images required, indicate the high potential of the method.

2.4 Conclusion

This chapter presents a survey of registration-by-regression methods that use a learning strategy to infer a regression model for registration. The methodologies used in the methods studied inferred the regression model mainly by relating points to points, especially on the 3D surface registrations, and by relating the images or image features to the transformation parameters. In these last cases, strategies for data reduction were also implemented. The Neural Networks, and specially the MLPs, were the most used models. The regression approach was indicated as an initialization for some cases, but the majority of the authors used it for the entire registration.

Several applications and objects were studied by these methods. The 3D surface and the 3D/2D registrations in an image-guided context outstand, as well as the bone structures as imaged objects. Regarding the nature of transformation, most authors used rigid transformation, while 5 out of 15 used non-rigid transformations.

Some approaches are promising because of their good results with simulated data, but more studies with real data are needed. It should also be noted that some methods are rather specific for the applications and transformations used, and thus may not be straightforward to generalize to other settings (e.g., going from rigid to non-rigid transformations). Considering the methodological aspects, the ability to model non-rigid transformations, and the experiments performed with clinical validation, works that stand out are Kim et al. (2010, 2012), Chou et al. (2013) and Chou & Pizer (2013).

The two main motivations for the regression approach were a fast registration and a larger capture range. But although many authors compared the results of the proposed methods with a well-established iterative method for the application in study, they did it mainly for simulated data. Only in few works, the evaluation of the registration-by-regression methods was performed with clinical data. Future work in this research field should therefore aim at further understanding the strengths and weaknesses of the registration-by-regression paradigm when applied to clinical data, and thus moving forward towards clinical applicability of these promising methods.

Work

Thou wast not put on earth to ask
If there be God, or life or death.
Seize then thy tools and to thy task
And give to toil each panting breath.

Thy tools thou hast, nor needst to seek
Thy health or faith or useful art,
The strength to toil, the power to speak,
A mighty mind or kindly heart.

Alexander Search

Chapter 3

3D/2D Image Registration by Nonlinear Regression

This chapter is based on the manuscript:

3D/2D image registration by nonlinear regression. Gouveia, A. R., Metz, C., Freire, L., & Klein, S. (2012). *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)* (pp. 1343–1346).

Abstract

We propose a 3D/2D registration method that treats image registration as a nonlinear regression problem instead of using an iterative traditional approach. The nonlinear regression function is determined in a supervised learning stage and relates image features of 2D projection images to the transformation parameters of the 3D image by nonlinear regression. The method is compared with a conventional registration method based on iterative optimization. For evaluation, simulated X-ray images (DRRs) were generated from coronary artery tree models derived from 3D CTA scans. Registration of nine vessel trees was performed, and the alignment quality was measured by the mean target registration error (mTRE). The regression approach has shown to be slightly less accurate, but much more robust than the method based on an iterative optimization approach.

3.1 Introduction

Image guided interventions often require integration of preoperative image with intraoperative images. A well-known example is the registration of preoperative 3D CT images and intraoperative 2D X-ray images (Figure 3.2). Most methods proposed in the literature are based on simulated X-ray projection images – digitally reconstructed radiographs (DRRs) – computed from the preoperative CT scan. In these methods, the image registration parameters are estimated by iteratively optimizing a similarity measure, measuring the difference between the DRR and the X-ray image (Markelj et al., 2012). Due to local maxima of the similarity measure, such iterative optimization procedures usually have a small capture range and therefore require initialization close to the searched pose (Markelj et al., 2012; Van De Kraats, Penney, Tomaževič, & Van Walsum, 2005).

In this chapter, we present a robust 3D/2D registration method that addresses the capture range drawback. Our approach treats image registration as a nonlinear regression problem instead of using an iterative traditional approach. The nonlinear regression function is determined in a supervised learning stage and relates a DRR image to the 3D transformation parameters of the 3D object.

The proposed method is based on image features but not in the common meaning of being a point-to-point, curve-to-curve or surface-to-curve registrations. Instead, the registration parameters are obtained from a more similar to library-based approach where a library of 2D template is generated for a predefined set of transformation parameters. Each library template represents the expected 2D appearance of a 3D geometric feature for a particular transformation. The template that is most similar to the 2D geometric feature is then aligned with the 2D data to define the remaining unknown transformation parameters. In 3D/2D registration, templates like 2D silhouettes (Hermans, Claes, Bellemans, Vandermeulen, & Suetens, 2007), Fourier descriptors of a 2D silhouette (Banks & Hodge, 1996) and shock graph representation of the projected 3D shape (Cyr, Kamal, Sebastian, & Kimia, 2000) can be found in the literature, with different searching and/or matching schemes.

To the best of our knowledge, no such approach to the 3D/2D registration problem has been previously presented in the literature. The method proposed is quantitatively evaluated, and compared with a conventional iterative optimization-based registration method (Metz et al., 2011).

3.2 Methods

3.2.1 Registration by regression

Our proposed 3D/2D registration method consists of a nonlinear regression model (Figure 3.1) that relates image features of the 2D projection image to the (translation and rotation) transformation parameters of the 3D image required to render both images "matched". A training phase takes place before the intervention (e.g. surgery), when data are available and time constraints are not an issue. After the learning process, the nonlinear function is known and we are able to compute the transformation parameters for the 2D images acquired during the intervention.

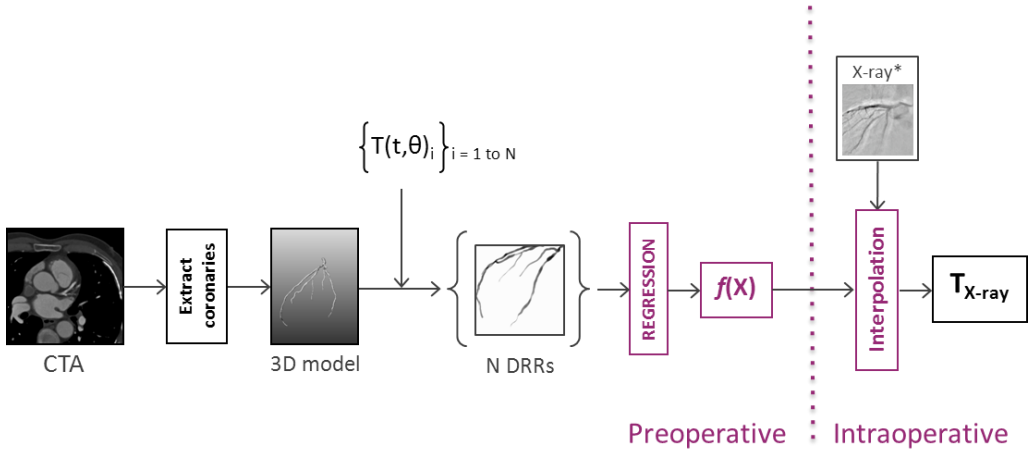


Figure 3.1 Registration-by-regression model.

The training set required for the learning process of our approach is a set of simulated 2D images (DRRs) obtained by manual transformations of the pre-interventional 3D image followed by projection. The features extracted from the DRR and its corresponding transformation parameters form an input-output pair in the training set. During the intervention, the image features of the 2D projection image (Figure 3.2) are computed and fed as input to the regression function, which returns the estimated 3D translation and rotation parameters of the 3D image.

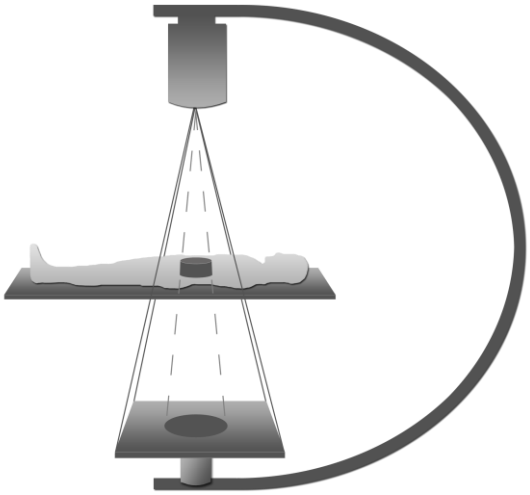


Figure 3.2 Geometry of a C-arm device, which makes 2D projection images of a 3D object.

The formulation we propose for the 3D/2D registration is similar to the viewpoint determination problem described in Liang & Chung (2008), where the solution is determined by an interpolation function considering input-output examples from 2D perspective projection of a 3D scene.

3.2.2 Input features

In this chapter we consider two different sets of image features. The first set consists of the three 2D (geometric) moments of the images (Liao, 1993). The moment of order zero, $\{M_{00}\}$, represents the total mass of a given image (or the area in the case of a binary image). The first order moments, $\{M_{01},M_{10}\}$, represent the centre of mass of the image and yield information about the object position; whereas the second order moments, $\{M_{02},M_{11},M_{20}\}$, also known as *moments of inertia*, contribute for its orientation.

The second set of features consists of the eigenvalues and eigenvectors computed from a Principal Component Analysis (PCA), on the pixels of the object of interest after a coarse segmentation of the image into object and background objects. The PCA was performed in two ways (which will be compared in the Section 3.3.5): a) on the 2D pixel position vectors (x,y) and b) on a combination of the pixel locations and their corresponding intensity values, i.e., a 3D vector with x, y and $I(x,y)$ as variables, where $I(x,y)$ is the intensity value of the point at position (x,y). In both cases, PCA was preceded by computing the z-score of the features, where we used an identical mean and standard deviation for x and y, to not loose pose information in this normalization procedure. The normalization is necessary since the pixel intensities $I(x,y)$ have a different range than the pixel positions x and y.

3.2.3 Nonlinear Regression Model

As a regression model we use Neural Networks (NN), which have a great ability of modelling complex (nonlinear) functions and are considered universal approximators (Haykin, 1999). The central idea of this nonlinear model is to extract linear combinations of the inputs as derived features, and then model the output as a nonlinear function of these features (Hastie, Tibshirani, & Friedman, 2009).

By applying a set of labelled training input-output examples (i.e., a set of inputs and corresponding desired responses), a neural network learns by the modification of its processing units or synaptic weights. The synaptic weights are optimized in such a way that the difference between the desired response and the actual response of the network is minimized.

The NN used is a two-stage regression model represented in a simplified and schematic way by the diagram below (Figure 3.3).

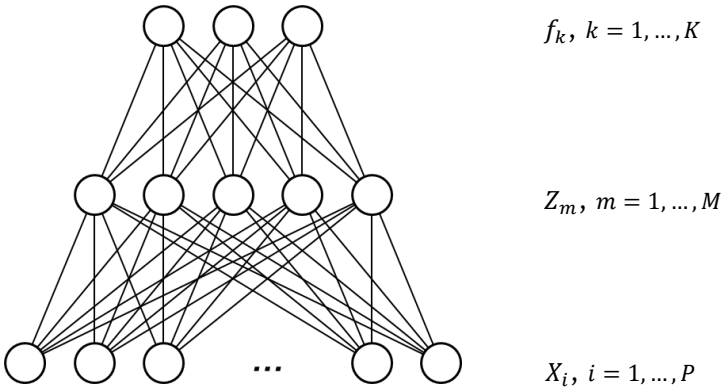


Figure 3.3 The network has an input layer, an intermediate layer (the hidden layer) and an output layer with$P$, M and K units X_i, Z_m and f_k, respectively.

First, the outputs of the hidden layer, working as derived features (Hastie et al., 2009), Z_m, are computed as a function of linear combinations of the inputs, X_i, and afterwards the outputs f_k are modeled as a function of linear combinations of Z_m, as shown by

$$Z_m = \varphi_m^{(1)}\left(\omega_{m0}^{(1)} + \sum_{i=1}^{P} \omega_{mi}^{(1)} X_i\right) \qquad \text{Equation 3.1}$$

and

$$f_k(X) = \varphi_k^{(2)}\left(\omega_{k0}^{(2)} + \sum_{m=1}^{M} \omega_{km}^{(2)} Z_m\right) \qquad \text{Equation 3.2}$$

or simply

$$f_k(X) = \varphi_k^{(2)}\left(\omega_{k0}^{(2)} + \sum_{m=1}^{M} \omega_{km}^{(2)}\left[\varphi_m^{(1)}\left(\omega_{m0}^{(1)} + \sum_{i=1}^{P} \omega_{mi}^{(1)} X_i\right)\right]\right) \qquad \text{Equation 3.3}$$

for $i = 1, \ldots, P$ and $X = (X_1, \ldots, X_P)$, for $m = 1, \ldots, M$ and $Z = (Z_1, \ldots, Z_M)$, and $k = 1, \ldots, K$. $\varphi_m^{(1)}$ and $\varphi_k^{(2)}$ are activation functions, $\omega_{mi}^{(1)}$ denotes the synaptic weights connecting the output of neuron i to the input of neuron m, $\omega_{km}^{(2)}$ denotes the synaptic weights connecting the output of neuron m to the input of neuron k, and the indexes (1) and (2) correspond to the connections between input-hidden layers and hidden-output layers, respectively.

To determine the function we must find the unknown parameters, i.e. the synaptic weights, which are achieved by the minimisation of the difference between the output f_k and the desired response, considering all neurons k and all examples of the training set.

3.3 Experiments and Results

3.3.1 Imaging data

In this work we focused on the registration of 3D preoperative coronary CTA and 2D intraoperative X-ray angiography. To evaluate the new registration method in a controlled setting, we do not use real X-ray images in our experiments, but simulated projection images of the coronary vessel tree, with known ground truth transformation. To this end, we made coronary segmentations at end-diastole of ten patients (Metz et al., 2011) to obtain binary vessel tree models (Figure 3.4. c). From these 3D models, DRRs were generated (Figure 3.4.d) using the computation procedure described in Metz, Schaap, Klein, et al. (2009). The projection geometry for the computation of the DRRs as well as the initial orientation of the preoperative data were derived from an interventional X-ray image, thereby mimicking a clinically relevant view. The field of view and voxel size of CTA were 256x256 voxels and 0.7x0.7x0.9 mm^3, respectively, and the field of view and voxel size of the DRR images were 512x512 voxels and 0.22x0.22 mm^2, respectively.

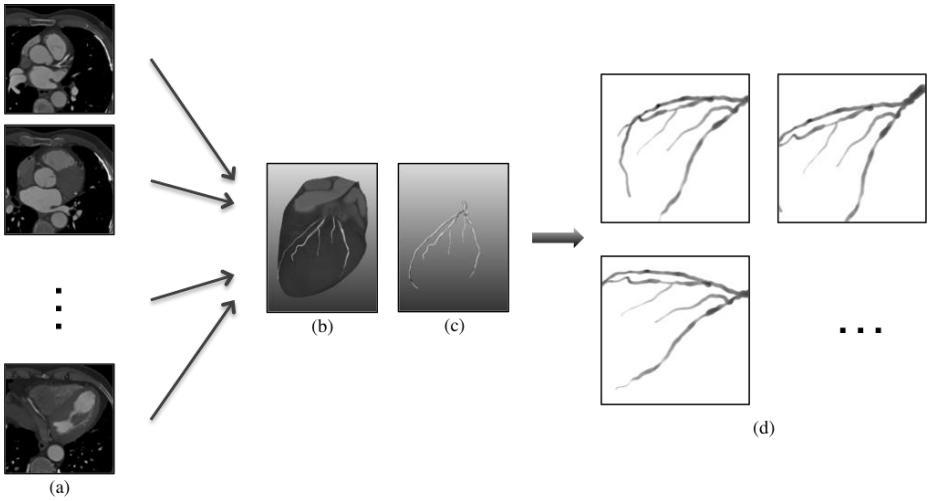


Figure 3.4 Coronary CTA slice (a), coronary 3D CTA with segmented coronary (b), 3D model (c), and DRR obtained from the model (d).

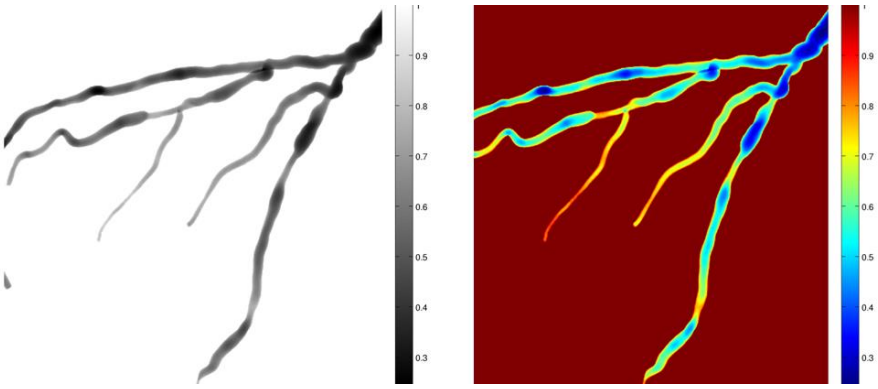


Figure 3.5 DRR obtained as represented in the previous figure but in more detail and shown in grey and colour scale. The DRR intensity depends on the ray length crossing the vessels: as much ray crosses the vessel, higher the intensity attenuation and lower the final intensity of the ray. The greater attenuation corresponds to the darker and blue regions, in the grey and coloured image respectively, and it is due to a thicker vessel region or to a bend of the vessel in 3D space.

For each patient, 11000 DRRs were generated, 10000 to train the regression model and 1000 to test the performance of our method. The transformations were drawn from a uniform distribution, with a wide yet relevant range, i.e. between -10 and 10 degrees for rotations and between -10 and 10 mm for translations.

3.3.2 Implementation details of features extraction and regression model

In order to select the points in the object for the computation of the PCA features, we segmented the DRR using a threshold value manually chosen. It should be noted that this step was trivial, since we used simulated projections of the vessel tree, which do not suffer from vessel-like structures in the background, as it can be seen in Figure 3.5. With real X-ray images, this is a major challenge, not yet addressed in this chapter.

The NN was designed following the general recommendations in the literature (Haykin, 1999; Sarle, 2005). Topology used is feed-forward, specifically a Multi-Layer Perceptron (MLP) with one hidden layer. The training algorithm is backpropagation (the generalization of the Widrow-Hoff learning rule) with conjugate gradient (MLP-CG) as the optimization technique (Molller, 1993). The weights are randomly initialized within small ranges ([-1,1]). The resampling technique used is the holdout method; the training set of 10000 images was split in two sets of 70% and 30% for training and validation[1], respectively (Cortez, 2008; Flexer, 1996). Additionally, we computed the average results obtained from 3 runs of the experiments in order to avoid random influences as recommended.

The activation functions for the hidden and output layers are the sigmoid (hyperbolic tangent) and linear functions, respectively. For the hidden layer, we performed an initial exploratory experiment to determine the number of units that yielded the best results (see Section 3.3.5), based on a trial-and-error approach. For this purpose, we used the data from the first patient. The output layer has one output unit per registration parameter, following general recommendations, which means our problem has one MLP per transformation parameter, i.e. our registration tool comprises six independent MLPs. All units are fully connected with each unit of the next layer.

In order to get a good generalization, the early stopping was considered. For a proper use by the NN, the input and output vector elements are rescaled (Haykin, 1999), to a range between -1 and 1.

The Neural Network was implemented in MatLab, version 7.11.0.584 (R2010b) 64 bits (MatLab, 2010).

---

[1] We considered three data sets: a *training set* to train the regression model, a *validation set* for testing purposes in the tuning experiments, and a *testing* set to compute the registration performance.

3.3.3 Conventional 3D/2D registration method

The registration method described in this chapter was compared to the method based on iterative optimization proposed in Metz et al. (2011), which was specifically designed for the application considered. This method uses a nonlinear conjugate gradient optimizer and a similarity metric based on a distance transform of a projection of the 3D coronary segmentation onto the X-ray image, and a fuzzy segmentation of vessel structures in the 2D image. As with our method, this fuzzy segmentation step became trivial, because we only used DRRs of the binary vessel tree, instead of real X-rays.

3.3.4 Evaluation methodology

The evaluation of the registration approaches was performed by the computation of the mean target registration error ($mTRE$) before and after registration. This enables us to assess registration success or failure and the consequent capture ranges for both registration approaches. The mean target registration error ($mTRE$) is computed as the mean distance to the ground truth (pose of the centerline tree at the known transformation parameters) of all center points of the 3D vessel centerline trees:

$$mTRE(T) = \frac{1}{N}\sum_{n=1}^{N}\left\|T(p_n) - T_{gold}(p_n)\right\| \qquad \text{Equation 3.4}$$

where T is the resulting transformation of one of the registration methods to assess, T_{gold} is the known transformation and p_n are points on the centerline of the 3D vessel tree. All reported $mTRE$ values in the following sections were computed on the test set of 1000 images, which were not used to train the regression model.

3.3.5 Optimization and parameter settings

Prior to the evaluation of the proposed registration method, some experiments were performed in order to optimize the method. To this end, the image set of one of the patients was used, whereas the sets of the nine remaining patients were used for the evaluation in Section 3.3.6.

Based on these tuning experiments, we set the number of hidden units to twice the number of input units (Appendix A). The stopping epoch (i.e., the epoch when the validation error started to grow) was tuned separately for each MLP. The effect of using different features (Section 3.2.2), individually and combined, was investigated and the results are summarised in Table 3.1. Individually, Set 1 was the best contributor to the registration, being especially

valuable for finding the translation parameters. Though Set 2 gave the worst results individually, when combined with Set 1 it helped to decrease the registration errors. The configuration Set1+Set3 was found to be the optimal combination of feature sets. In this configuration, the input layer of the neural network has thus 18 units: 6 moments, 3 eigenvalues and 3x3 components of the eigenvectors. This combined feature set (Set1+Set3) was used in all following experiments.

From Table 3.1 it can also be observed that a problematic transformation in the 3D/2D registration is the translation along the source-detector axis. Since the projection is in the y-axis direction, the corresponding translation parameter has less influence on the appearance of the projected model than the other parameters (Edwards, Hawkes, Penney, Clarkson, & London, 2001). Reduced performance in this direction has also been reported for the conventional registration approach described in Section 3.3.3 (Metz et al., 2011) and it is shown in Figure 3.6. In this figure the results per transformation parameter for the optimal configuration mentioned before (Set1+Set3) and for this conventional registration approach are shown.

Additional tuning experiments were performed to understand how large should be the training set. Four training set dimensions were considered and the results presented in Figure 3.7. Registration quality improved slightly with increasing training size. Since the training is a preoperative task, with no time constraints, our selection criteria were the accuracy and robustness of the registration and, therefore, we used the largest training set.

Table 3.1 Registration results for the feature sets considered: moments set (1), the eigenvalues and eigenvectors of the points in the object of interest (2) and the eigenvalues and eigenvectors set for the same points and corresponding intensity values (3).

Parameters	Set 1 Moments		Set 2 Eigen [*X Y*]		Set 3 Eigen [*X Y I*(*x*,*y*)]		Set 1 + Set 2		Set 1 + Set 3	
	mean	std	mean	std	mean	std	mean	std	mean	std
R_X (°)	0,29	0,29	3,01	2,31	0,23	0,36	0,13	0,15	0,05	0,05
R_Y (°)	0,30	0,26	2,84	2,14	0,36	0,43	0,12	0,12	0,07	0,08
R_Z (°)	0,41	0,39	1,79	1,47	0,25	0,22	0,14	0,18	0,07	0,09
T_X (mm)	0,15	0,12	4,86	2,94	2,68	2,45	0,04	0,06	0,02	0,03
T_Y (mm)	1,36	1,54	5,04	2,87	5,02	2,88	0,65	1,08	0,33	0,52
T_z (mm)	0,12	0,11	4,70	3,00	0,73	0,82	0,07	0,08	0,03	0,04
mTRE (mm)	1,45	1,55	9,76	3,02	6,31	2,85	0,69	1,09	0,35	0,55
mTRE (mm) (mTRE<2mm)	0,87	0,50	1,70	0,37	1,34	0,46	0,48	0,37	0,30	0,29

Parameters errors are given for rotations and translations (R_X, R_Y and R_Z – rotations in degrees around x-, y- and z-axis, respectively, and Tx, Ty and Tz – translations in x, y and z directions in mm, respectively) and computed from absolute values of errors; mean TRE errors (mTRE) for all registrations performed are presented, as well as just considering the successful registrations (i.e. mTRE<2mm). All reported results are based on the test set of 1000 images.

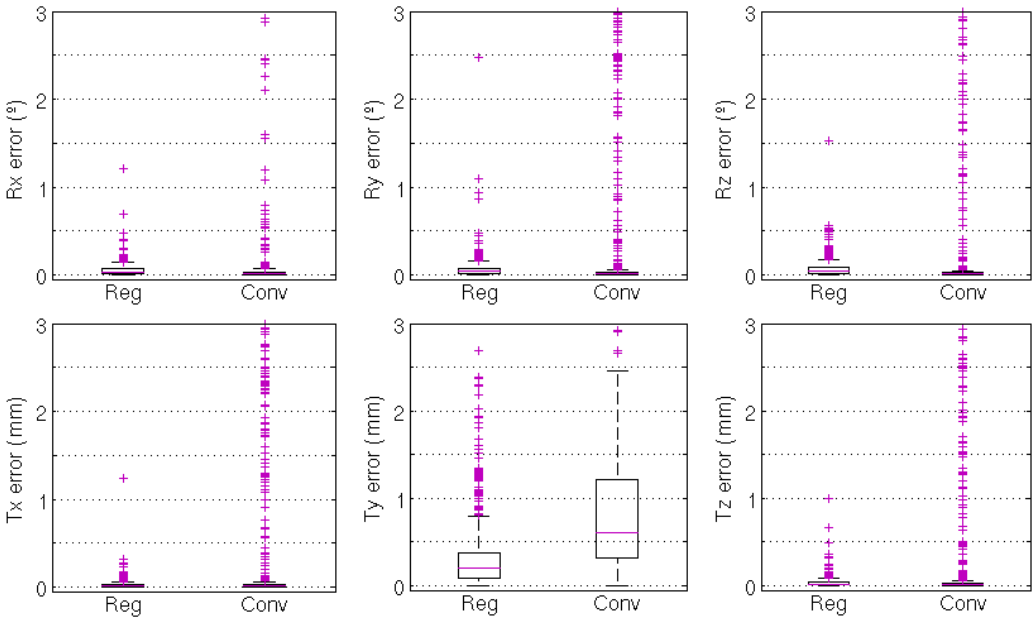


Figure 3.6 Registration results per transformation parameter for the feature set with best performance. Parameters errors are given for rotations and translations: R_X, R_Y and R_Z – rotations in degrees around x-, y- and z-axis, respectively, and Tx, Ty and Tz – translations in x, y and z directions in mm, respectively). The reported results are based on the test set of 1000 images.

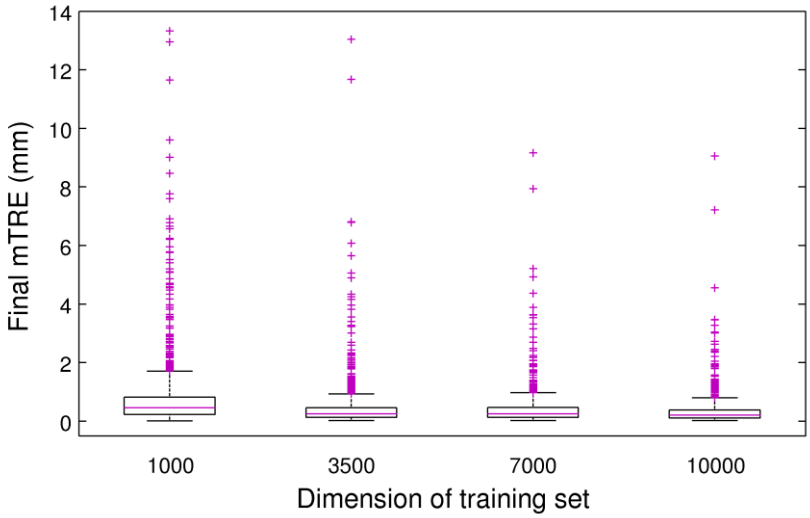


Figure 3.7 Registration results for different dimensions of the NN training set for Patient 0. Mean TRE errors (mTRE) for registrations performed are presented.

3.3.6 Test results

The results of the 3D/2D experiments for all patients and for both methods are summarized in Figure 3.8, Figure 3.9 and Figure 3.10.

Considering these registration results, the regression method was less accurate for the majority of datasets compared to the iterative optimization based registration method, with a median error larger for 7 of the 9 datasets (Figure 3.8). However, the number of outliers for both approaches clearly showed a higher robustness relatively to outliers for the regression approach (Figure 3.8). That robustness of the regression method is also stressed in Figure 3.10 and Figure 3.9 where the $mTRE$ values before registration against the $mTRE$ values after registration are shown, for each patient and for all patients, respectively.

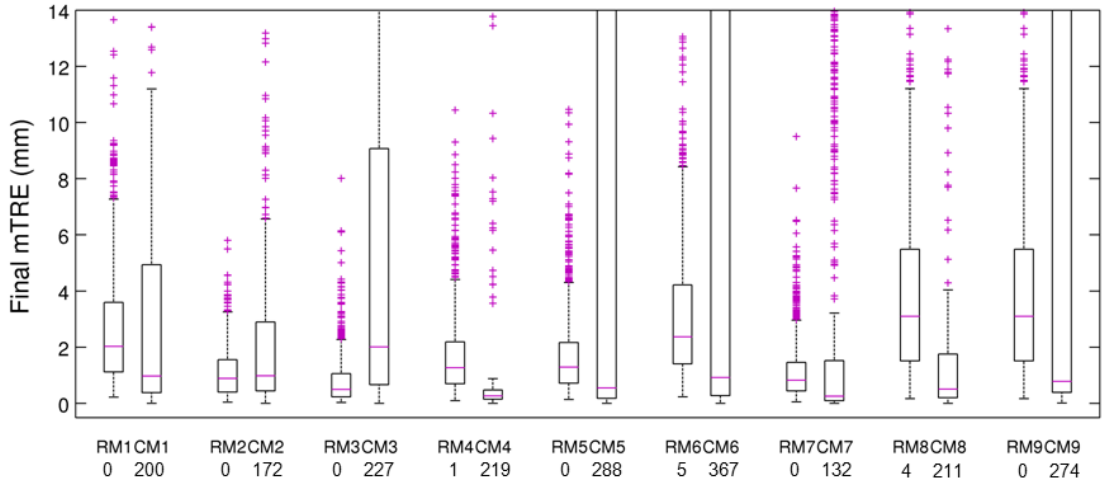


Figure 3.8 Comparison of the registration method by regression (RM) to the conventional registration method (CM), for each patient (labeled from 1 to 9). The number of points with $mTRE$ > 14mm that fall outside the plot is indicated below the correspondent labels.

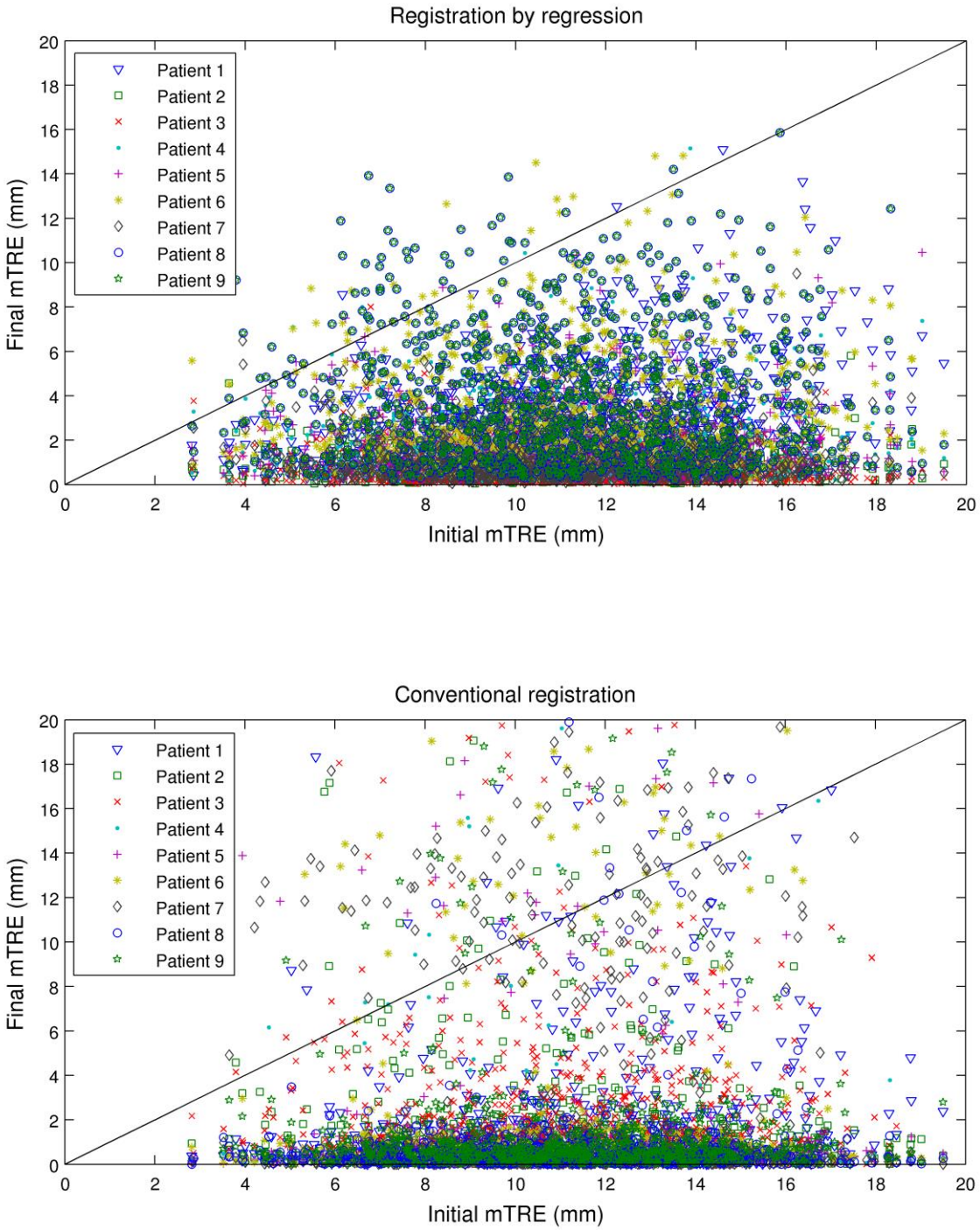


Figure 3.9 Comparison of the registration method by regression to the conventional registration method, considering the results for each patient (labeled from 1 to 9). The plot shows the $mTRE$ values before registration against the $mTRE$ values after.

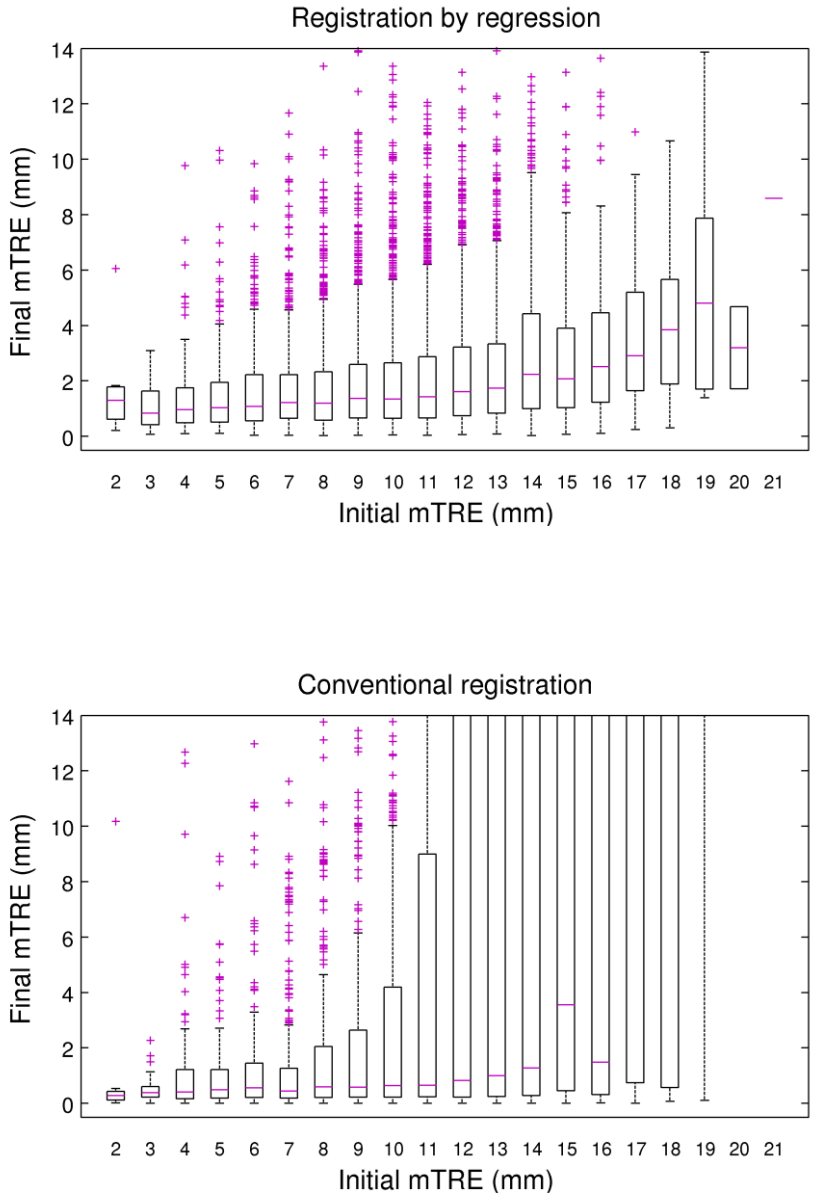


Figure 3.10 Comparison of the registration method by regression to the conventional registration method, considering the results for all patients. The plot shows the $mTRE$ values before registration against the $mTRE$ values after.

3.4 Discussion

The results presented in this chapter show that the proposed regression-based 3D/2D registration approach remains robust in case of large misalignments. It could be used either stand-alone, or as an initialization step for a conventional 3D/2D registration to improve the final accuracy.

In the literature some other work can be found with registration approaches based on neural network (Abche et al., 2006; Elhanany et al., 2000; Freire et al., 2010a; Yan et al., 2004; Zhang et al., 2008). However, the problem formulation is different (i.e., the function to be determined does not relate the same input-outputs that we do), the applications are different and most of them do not compare the results with a conventional method. The most similar approach (Freire et al., 2010a) presents a method designed for 3D rigid-body registration of fMRI time series, which relies on a limited number of Fourier coefficients of the images to be aligned.

3.5 Conclusion

We have proposed a novel approach for 3D/2D image registration, based on nonlinear regression. Promising results were demonstrated for the registration of coronary vessel trees.

Compared with a conventional approach, the number of misregistrations was reduced substantially. The following work includes the implementation and evaluation of different regression (other than NN), in Chapter 4 and experiments using real X-ray images in Chapter 4.

Seeker Of Truth

seeker of truth

follow no path
all paths lead where

truth is here

E. E. Cummings

Chapter 4
Comparative evaluation of regression methods for 3D/2D image registration

This chapter is based on the manuscript:

Comparative Evaluation of Regression Methods for 3D/2D Image Registration. Gouveia, A., Metz, C., Freire, L., & Klein, S. (2012). In A. E. Villa, W. Duch, P. Érdi, F. Masulli, & G. Palm (Eds.), *Artificial Neural Networks and Machine Learning – ICANN 2012* (Vol. 7553, pp. 238–245). Springer Berlin Heidelberg.

Abstract

We perform a comparative evaluation of different regression techniques for 3D/2D registration-by-regression. In registration-by-regression, image registration is treated as a nonlinear regression problem that relates image features of 2D projection images to the transformation parameters of the 3D image. In this chapter, we evaluate seven regression methods: Multiple Linear and Polynomial Regression (LR and PR), k-Nearest Neighbour (k-NN), Multi-Layer Perceptron with conjugate gradient optimization (MLP-CG) and Levenberg-Marquardt optimization (MLP-LM), Radial Basis Function networks (RBF) and Support Vector Regression (SVR). The experiments are performed with simulated X-ray images (DRRs) of nine vessel trees, so that we have a ground truth and can compute the mean target registration error (mTRE). MLP-LM and RBF obtained the highest accuracy but in general all methods were robust to large initial misalignment.

4.1 Introduction

Medical interventions can often benefit from the integration of preoperative (diagnostic) and intraoperative imaging data. Accurate and fast image registration is required to find the relation between the preoperative and the intraoperative images of the patient in the intervention room. A good example is the registration of preoperative 3D computed tomography angiography (CTA) images and intraoperative 2D X-ray images, for percutaneous coronary interventions. Most methods presented in the literature are based on simulated X-ray projection images – digitally reconstructed radiographs (DRRs) – computed from the preoperative CTA scan. In these cases, the difference between the DRR and the X-ray image is iteratively measured until a similarity metric is optimized (Markelj et al., 2012). This image registration approach, with such an iterative optimization procedure, usually has a small capture range because of the local maxima of the similarity measure (Markelj et al., 2012; Van De Kraats et al., 2005).

In our recently proposed registration-by-regression framework (Gouveia, Metz, Freire, & Klein, 2012a; Chapter 3), image registration is treated as a nonlinear regression problem, as an alternative for the iterative traditional approach. A Multi-Layer Perceptron (MLP) was used as the regression model relating the DRR image to the 3D transformation parameters of the 3D object.

In literature, few authors treat medical image registration as a regression problem. From those works, we conclude that the most common regression method used is MLP (Freire et al., 2010a; Zhang et al., 2008). Wachowiak, Smolíková, Zurada, & Elmaghraby (2002) studied different Neural Networks, including RBF networks. Image registration using SVR (Qi, Gu, & Zhao, 2008) and k-NN (Banks & Hodge, 1996) are also found. Although these works use regression for the registration problem, the application, the features used, and the function determined by the regression are different from our approach. Both Qi et al. (2008) and Banks & Hodge (1996) addressed 3D/2D registration: the first had a very different registration formulation and the second used regression to obtain an initial estimate of the rotation parameters.

In this chapter we perform a comparative evaluation of seven different regression techniques for the 3D/2D registration-by-regression problem, particularly for the registration of DRRs generated from coronary artery tree models derived from 3D CTA scans. The performances of registration-by-regression using Multiple Linear Regression, k–Nearest Neighbour, Multi-Layer Perceptron, Radial basis Function Network, and Support Vector Regression are computed. This quantitative evaluation identifies the most accurate method for this registration problem. For comparison purposes, the results of a conventional registration method (i.e. based on iterative optimization) obtained in Gouveia et al. (2012a) and Chapter 3 are also reported.

4.2 Method

4.2.1 Registration by regression

The 3D/2D registration method used in this chapter is based on regression as presented in Gouveia et al. (2012a) and Chapter 3. The used regression model relates image features of the 2D projection image to the (translation and rotation) transformation parameters of the 3D image required to bring both images (2D and 3D) into spatial correspondence. Before the intervention (e.g. surgery), a set of simulated 2D images (DRRs) is generated by applying random transformations of the pre-interventional 3D image followed by projection of its coronary artery segmentation. A set of features extracted from the DRR and their corresponding transformation parameters form an input-output pair in the training set for the learning process. During the intervention, the image features of the 2D projection image are computed and fed as input to the regression function, which returns the estimated 3D translation and rotation parameters of the 3D image.

Mathematically, given an input vector $X = (X_1, \ldots, X_P)$, we want to predict an output Y with a model f:

$$Y = f(X) \qquad \text{Equation 4.1}$$

To estimate the parameters of the prediction model f we use a set of measurements (x_i, y_i) for $i = \{1, \ldots, N\}$ (the training data). For each transformation parameter (three rotation angles, three translations) an independent regression model is trained.

4.2.2 Input features

In Gouveia et al. (2012a) and Chapter 3 different sets of image features were compared. In this chapter, we use the best performing set, which consists of two types of features: 1) 2D geometric moments and 2) the eigenvalues and eigenvectors computed from a Principal Component Analysis (PCA).

The 2D geometric moments included are the moment of order zero (representing the total mass of a given image), the two first order moments (representing the centre of mass of the image, yielding information about the object position) and the three second order moments (or moments of inertia, sensitive to changes in orientation). The PCA features are computed on the pixels of the object of interest after a coarse segmentation of the image into object and background objects, as described in Gouveia et al. (2012a) and Chapter 3. The PCA is performed on a combination of the pixel locations and their corresponding intensity values, i.e., a 3D vector with x, y and $I(x,y)$ as variables, where $I(x,y)$ is the intensity value of the point at position (x,y). PCA was preceded by computing the z-score of the features, where we

used an identical mean and standard deviation for *x* and *y*, to prevent loosing pose information in this normalization procedure. The normalization is necessary since the intensities *I*(*x*,*y*) have a different unit than the pixel positions *x* and *y*.

In total, the input vectors have 18 units (i.e. *P* = 18 input features) which were standardized by z-score for a proper use by the regression techniques (Hair, Black, Babin, Anderson, & Tatham, 2005; Haykin, 1999).

4.2.3 Regression Models

The following regression methods are compared in this chapter.

4.2.3.1 Multiple Regression: Linear and Polynomial

We considered two Multiple Regression cases. The first is Multiple Linear Regression (LR) where the model f in Equation 4.1 has the form:

$$f(X) = \beta_0 + \sum_{i=1}^{P} X_i \beta_i \qquad \text{Equation 4.2}$$

with β_i the model parameters to be trained (Hastie et al., 2009).

The second model is the Polynomial Regression (PR) model, which uses basis expansions X_i^2 and interactions between variables $X_i X_j$:

$$f(X) = \beta_0 + \sum_{i=1}^{P} X_i \beta_i + \sum_{i<j}^{P} X_i X_j \beta_{ij} + \sum_{i=1}^{P} X_i^2 \beta_{ii} \qquad \text{Equation 4.3}$$

Both regression models are linear in the parameters and we used the Least Squares Method (LSM) to compute the parameters β (Hastie et al., 2009).

4.2.3.2 k-Nearest Neighbour

The k-nearest neighbour (k-NN) method makes predictions by averaging the responses Y_i of the *k* nearest points X_i in the training set. The model can be written as:

$$f(X) = \frac{1}{k} \sum_{X_i \in N_k(X)} Y_i \qquad \text{Equation 4.4}$$

where $N_k(X)$ is the neighbourhood of X defined by the k nearest points X_i in the training set (Hastie et al., 2009). The metric used is the Euclidean distance.

4.2.3.3 Multi-Layer Perceptron

This Neural Network learns by the modification of its synaptic weights using the error back-propagation algorithm (Haykin, 1999). A MLP with a hidden layer is considered with the hyperbolic tangent function as activation function. For the output layer a linear function is used. This MLP model for *P* input features and *H* hidden units is given by:

$$f(X) = \omega_0^{(2)} + \sum_{j=1}^{H} \omega_j^{(2)} \left[\varphi^{(1)} \left(\omega_{j0}^{(1)} + \sum_{i=1}^{P} \omega_{ji}^{(1)} X_i \right) \right] \qquad \text{Equation 4.5}$$

with $\varphi^{(1)}$ being the hyperbolic tangent function (Hastie et al., 2009; Haykin, 1999). The model parameters ω's represent the synaptic weights connecting the output of neurons of one layer to the input of neurons of the following layer, and the indexes (1) and (2) correspond to the connections between input-hidden layers and hidden-output layers, respectively.

The MLP was designed following the general recommendations in the literature (Hastie et al., 2009; Sarle, 2005). Two optimization techniques for training are evaluated: conjugate gradient (MLP-CG) (Molller, 1993), as in Gouveia et al. (2012a) and Chapter 3, and Levenberg-Marquardt (MLP-LM) (Hagan, 1994). All units are fully connected with each unit of the next layer. The weights are randomly initialized within a range [-1,1].The regularization parameter in the Levenberg-Marquardt method is initialized at 10^{-3} and increases by a factor of 10 until the change introduced by it results in a reduced performance value (Beale, Hagan, & Demuth, 2012). In this case it decreases by a factor of 10^{-1}. The algorithm is stopped if the parameter becomes larger than 10^{10} (maximum value allowed).

4.2.3.4 Radial Basis Function Network

The RBF network has a similar architecture as the MLP. It has one hidden layer; the output layer has one unit and the activation function uses the linear case. However, the activation function of the hidden layer is different (i.e. a Gaussian function) as well as the way the inputs are connected to the units in this layer (i.e. the way the argument of the activation function is constructed). The argument of the activation function in RBF networks is the Euclidean distance between the input vector X and the synaptic weight vector (the RBF centre) of that neuron (Haykin, 1999; Sarle, 2005). The RBF network model can be written as

$$f(X) = \omega_0^{(2)} + \sum_{j=1}^{H} \omega_j^{(2)} \cdot e^{-\left(\omega_0^{(1)} \left\| X - \omega_j^{(1)} \right\|\right)^2} \text{ where } \omega_0^{(1)} = \frac{\sqrt{-\log(0.5)}}{spread}, \qquad \text{Equation 4.6}$$

where the $spread$ (or radius) parameter controls the amplitude of the Gaussian function (W. Wang, Xu, Lu, & Zhang, 2003), represented in Figure 4.1. The bias for the hidden layer (i.e. $\omega_0^{(1)}$) adjustes the sensitivity of the neuron j.

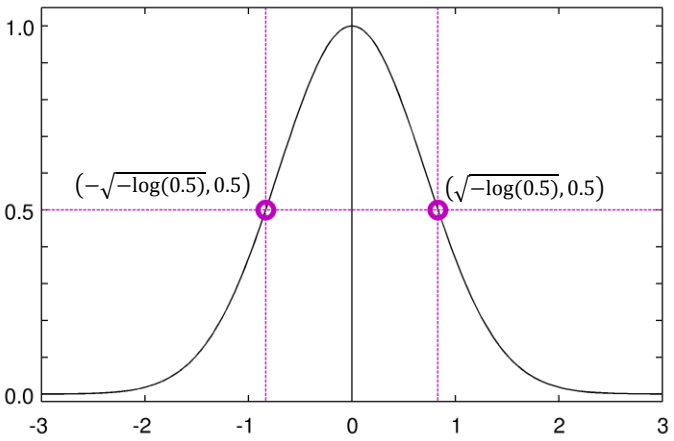


Figure 4.1 Gaussian function $G(x) = e^{-x^2}$ used in a RBF network.

The RBF uses the Orthogonal Least Squares as training algorithm where RBF centres are chosen one by one from the input data. Each selected centre maximizes the increment to the explained variance of the desired output (Chen, Cowan, & Grant, 1991).

4.2.3.5 Support Vector Regression

In Support Vector Regression (SVR), a linear regression function f in a high N-dimensional feature space is computed as

$$f(X) = \omega_0 + \sum_{i=1}^{N} \omega_i \phi(X_i) \qquad \text{Equation 4.7}$$

where the input data are mapped via a nonlinear transformation ϕ according to the kernel function $K(X_i, X_j) = \phi(X_i)^{\mathrm{T}} \phi(X_j)$ (Basak, Pal, & Patranabis, 2007). The ϵ-SVR model is used (Vapnik, 1995), which finds ω's such that $f(X)$ has at most ϵ deviation from the actually obtained targets for all the training data, being as flat as possible (i.e. small ω's) (Smola & Schölkopf, 2004). The loss function considered is ϵ-insensitive (Figure 4.2), a commonly used function (Smola & Schölkopf, 2004), and the kernel function is a RBF:

$$K(X_i, X_j) = e^{-\gamma \|X_i - X_j\|^2}, \gamma > 0. \qquad \text{Equation 4.8}$$

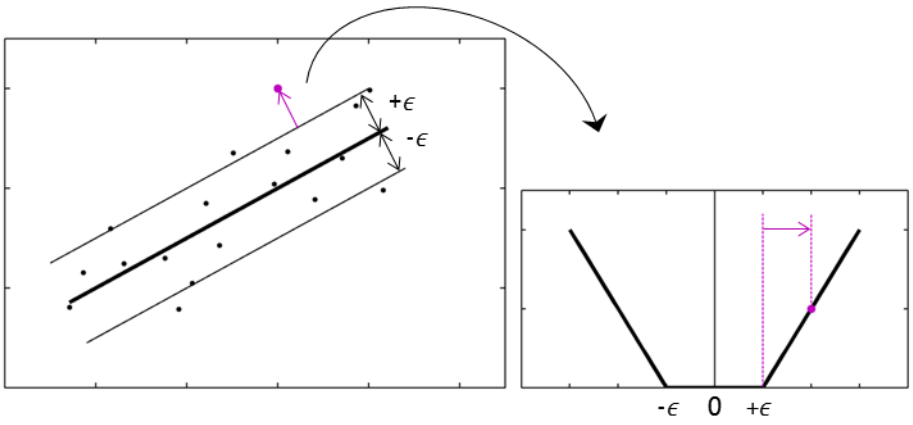


Figure 4.2 Linear SVM regression (left) and ϵ-insensitive function (right).

4.3 Experiments and Results

4.3.1 Imaging data

We used 3D preoperative coronary CTA data with 2D intraoperative X-ray angiography of ten patients. To evaluate the regression methods in a controlled setting, we used simulated projection images of the coronary vessel tree (Figure 4.3), with known ground truth transformation, instead of real X-ray images. To this end, we made binary vessel tree models by segmenting coronary arteries at the end-diastole phase (Metz et al., 2011). From these 3D models, DRRs were generated using the computation procedure described in Metz, Schaap, Klein, et al. (2009). The projection geometry for the computation of the DRRs and the initial orientation of the preoperative data were derived from an interventional X-ray image, and so simulating a clinically relevant view.

The size and voxel spacing of the CTA images were 256x256x[99-184] voxels and 0.7x0.7x[0.8-1.0] mm^3, respectively, and the size and pixel spacing of the DRR images were 512x512 pixels and 0.22x0.22 mm^2, respectively.

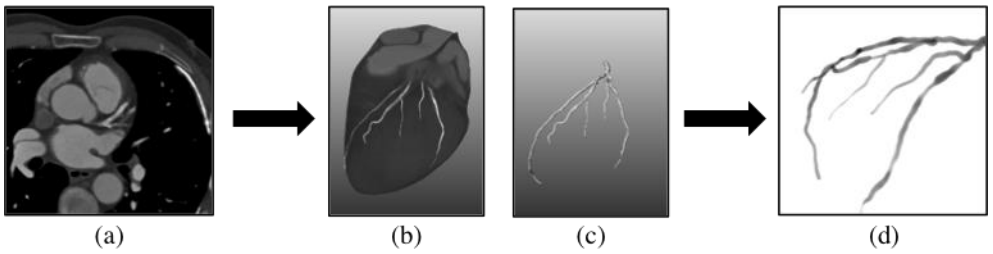


Figure 4.3 Coronary CTA slice (a), coronary 3D CTA with segmented coronary (b), 3D model (c), and DRR obtained from the model (d).

For each patient, 11000 DRRs were generated, 10000 to obtain the regression model and 1000 to test the performance of our method. The transformations were drawn from a uniform distribution, with a wide yet relevant range, i.e. between -10 and 10 degrees for rotations and between -10 and 10 mm for translations.

4.3.2 Conventional 3D/2D registration method

The different registration-by-regression models were compared to a method based on iterative optimization. This method was proposed by Metz et al. (2011) and it was designed for the application considered. It uses a nonlinear conjugate gradient optimizer and a similarity metric based on the distance transform of the projection of the 3D coronary

segmentation onto the X-ray image, and a fuzzy segmentation of vessel structures in the 2D image.

4.3.3 Evaluation methodology

The evaluation of the different registration approaches was performed by the computation of the mean target registration error ($mTRE$) before and after registration, which is a well-known evaluation criterion for registration methods (Fitzpatrick & West, 2001). It computes the distance between corresponding points, to assess the accuracy of the registration. The $mTRE$ is computed as the mean 3D distance of the centreline tree at the ground truth position and orientation to the centreline tree at the position and orientation determined by the registration-by-regression method, given by

$$mTRE(T) = \frac{1}{N}\sum_{n=1}^{N}\left\|T(p_n) - T_{gold}(p_n)\right\| \qquad \text{Equation 4.9}$$

where T is the transformation resulting from one of the registration methods, T_{gold} is the known ground-truth transformation and p_n are points on the centrelines of the 3D vessel tree. All reported $mTRE$ values in the following sections were computed on the test set of 1000 images, which was not used for the construction of the regression models.

4.3.4 Optimization and parameter settings

Prior to the evaluation of the regression models, some experiments were performed to optimize k-NN, MLP, RBF and SVR. To this end, the image set of one of the patients (named as patient 0) was used, whereas the sets of the nine remaining patients were used for the evaluation in Section 4.3.5. For each patient, the set of 10000 images, for the construction of the regression model, was split in two sets of 70% and 30%. The set of 7000 images was used to train the regression models; the remaining 3000 images were considered for validation purposes, to select tuning parameters. The performance of the models was computed by the mean absolute difference between the model prediction and the known output for each regression. Table 4.1 presents the ranges used for the tuning parameters for each regression method, as well as the final values chosen for the parameters, representing a scheme of the optimization process.

For the k-NN model, the search range for the optimal value of k was limited to $[1,50]$, based on experiments performed in patient 0 considering a coarse grid-search in a larger range. For

every patient and for each transformation parameter, the optimal value of $k \in [1,50]$ was chosen as the point where the validation error started to grow with increasing k.

For the MLPs, we have tried several numbers of units for the hidden layer, $\{9, 18, 36, 54\}$, using the image set of patient 0 (Appendix A). The number of hidden units was set to 36, which leads to a topology of P=18 input units, 36 hidden units and 1 output unit. The number of epochs was defined separately for each MLP by a stopping epoch (i.e., the epoch when the validation error started to grow) with a maximum of 1000.

For RBF networks, a two-level grid-search for the spread of radial basis functions was performed for patient 0. First, a rough spread estimate sp was computed as

$$sp = \frac{d_{max}}{\sqrt{P}} \qquad \text{Equation 4.10}$$

where d_{max} represents the maximum distance between the inputs (Haykin, 1999). The first-level search set was defined as $sp \times \{0.5, 1, 1.5, 2, 3, 4, 5\}$. Based on the result, the second-level grid search was then defined as $sp \times \{0.5, 0.6, \ldots, 1.4, 1.5\}$. From this search, the value *sp* turned out to be optimal. For the remaining patients the spread was therefore set to sp, computed on their respective input sets. The optimum number of RBF centres was determined for each patient (and each transformation parameter) as the point where the validation error started to grow, with a maximum of 1000 neurons.

In the SVR case, the problem of optimal parameter selection is further complicated by the fact that SVM model complexity (and hence its generalization performance) depends on three parameters: ϵ, C and γ (Cherkassky & Ma, 2004). Parameter ϵ is the width of the insensitive zone, parameter C a regularization parameter, determines the trade-off between the model complexity (flatness) and the degree to which deviations larger than ϵ are tolerated in the optimization formulation, and γ is the parameter of the kernel, in this case the inverse of the RBF spread. All three parameters are tuned simultaneously. We use a coarse-to-fine grid-search as recommended by Chang & Lin (2011) and Hsu, Chang, & Lin (2010), considering exponentially growing sequences of the parameter values (Table 4.1). To assure the complete space is covered and considering the computational costs, we performed a wide range three-level grid-search in patient 0 and the values obtained were used for the other patients. In the first level, we used ranges $\{2^{-13}, 2^{-11}, \ldots, 2^{5}\}$ for ϵ, $\{2^{-1}, 2^{1}, \ldots, 2^{15}\}$ for C and $\{2^{-9}, 2^{-7}, \ldots, 2^{9}\}$ for γ. For the next levels, the limits of the search ranges were selected from the previous level as the nearest values of the parameters combination with best validation error in that level. A finer search was then performed increasing the parameters by exponential factors of 1 and -0.25 for the second and the third levels, respectively.

All regressions experiments were performed using MatLab, version 7.11.0.584 (R2010b) 64 bits (MatLab, 2010). For SVR we used version 3.1 of the LIBVSM tool (Chang & Lin, 2011).

Table 4.1 Parameters for the optimization of the regression models k-NN, MLP, RBF and SVR.

<table>
<tr><th></th><th>Patient 0</th><th>Patients 1 to 9</th><th>For each regression (or parameter)</th></tr>
<tr><td>k-NN</td><td>1st level search: range search
Range search
$[0,7000] \rightarrow [1,50]$</td><td>2nd level search: k
$k \in [1,50]$</td><td>k</td></tr>
<tr><td>MLP</td><td>Number of hidden units h
$(18{:}\,h{:}\,1)$:
$h \in \{9, 18, 36, 54\}$</td><td>h=36</td><td>Maximum Epoch</td></tr>
<tr><td>RBF</td><td>Spread:
$sp \times \{0.5, \mathbf{1}, 1.5, 2, 3, 4, 5\}$
$\downarrow$
$sp \times \{0.5, 0.6, \ldots, 1.5\}$</td><td>$sp \times 1$
(specific for each patient by Equation 4.10)</td><td>Maximum number of neurons</td></tr>
<tr><td>SVR</td><td>Simultaneous tuning:
ϵ, C and γ
$\epsilon \in \{2^{-13}, 2^{-11}, \ldots, 2^{5}\}$
$C \in \{2^{-1}, 2^{1}, \ldots, 2^{15}\}$
$\gamma \in \{2^{-9}, 2^{-7}, \ldots, 2^{9}\}$
$\downarrow$
$\epsilon \in \{2^{s_{1\epsilon}}, 2^{s_{1\epsilon}+1}, \ldots\}$
$C \in \{2^{s_{1c}}, 2^{s_{1c}+1}, \ldots\}$
$\gamma \in \{2^{s_{1\gamma}}, 2^{s_{1\gamma}+1}, \ldots\}$
$\downarrow$
$\epsilon \in \{2^{s_{2\epsilon}}, 2^{s_{2\epsilon}+0.25}, \ldots\}$
$C \in \{2^{s_{2c}}, 2^{s_{2c}+0.25}, \ldots\}$
$\gamma \in \{2^{s_{2\gamma}}, 2^{s_{2\gamma}+0.25}, \ldots\}$</td><td>
<table>
<tr><th></th><th>ϵ</th><th>C</th><th>γ</th></tr>
<tr><td>R_X</td><td>$2^{-7.50}$ (0.0055)</td><td>2^{13} (8192)</td><td>$2^{-7.50}$ (0.0055)</td></tr>
<tr><td>R_Y</td><td>$2^{-6.25}$ (0.0131)</td><td>2^{13} (8192)</td><td>$2^{-7.25}$ (0.0066)</td></tr>
<tr><td>R_Z</td><td>$2^{-6.50}$ (0.0110)</td><td>2^{13} (8192)</td><td>$2^{-7.50}$ (0.0055)</td></tr>
<tr><td>T_X</td><td>$2^{-7.50}$ (0.0055)</td><td>2^{12} (4096)</td><td>$2^{-8.00}$ (0.0039)</td></tr>
<tr><td>T_Y</td><td>$2^{-7.00}$ (0.0078)</td><td>2^{13} (8192)</td><td>$2^{-6.00}$ (0.0156)</td></tr>
<tr><td>T_Z</td><td>$2^{-6.75}$ (0.0093)</td><td>2^{13} (8192)</td><td>$2^{-8.25}$ (0.0033)</td></tr>
</table>
</td><td></td></tr>
</table>

4.3.5 Results

The $mTRE$ values of the different regression methods and of the conventional registration method, considering all patients except the one used for parameter optimization, are shown in (Figure 4.4). We also present, for all patients except the one used for parameter optimization (Figure 4.5) and for each of these nine patients (Figure 4.6 to Figure 4.8), the Regression Error Characteristic (REC) curves (Bi & Bennett, 2003) for all methods[2]. These REC curves are a customization of receiver operating characteristic (ROC) curves to regression and were previously used to compare regression models (Cortez, Portelinha, Rodrigues, Cadavez, & Teixeira, 2006; Pina & Zaverucha, 2008). They plot the error tolerance on the x-axis (expressed as mTRE in our case) and the accuracy of a regression function on the y-axis (i.e.

[2] For these patients and for all methods, boxplot figures like Figure 4.4 are shown in Appendix A.

the percentage of points that lie within the tolerance). The REC curve estimates the cumulative distribution function of the error and the area over the curve (AOC) is a measure of the expected error for a regression model.

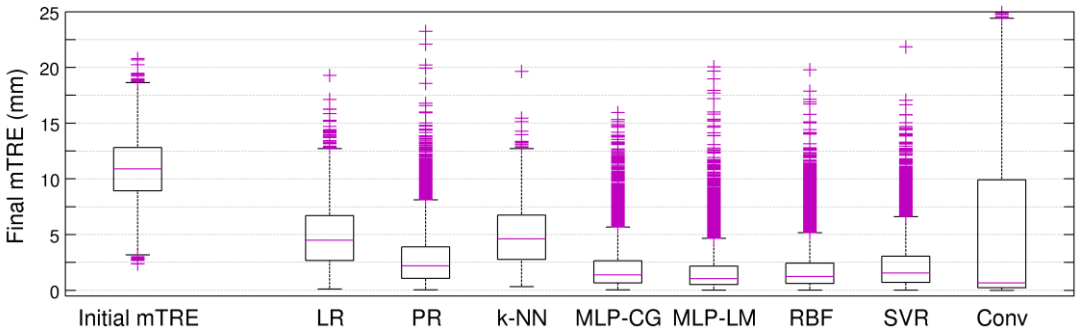


Figure 4.4 Comparison of registration-by-regression results for all methods and considering all patients, except patient 0 (which was used for parameter optimization). The graphic also shows the initial mTRE before registration.

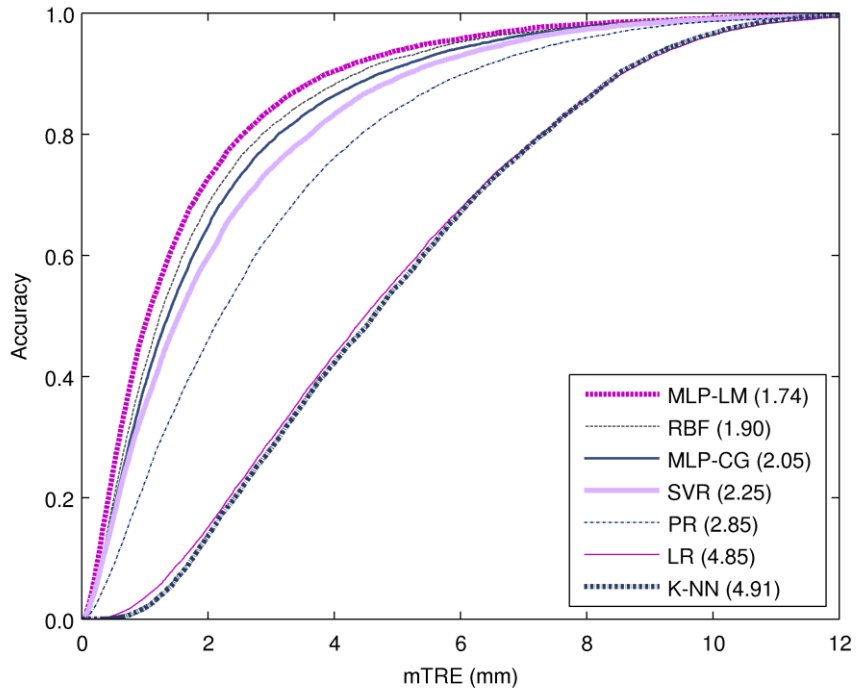


Figure 4.5 REC curve for all methods and considering all patients, except patient 0 (which was used for parameter optimization) and AOC values.

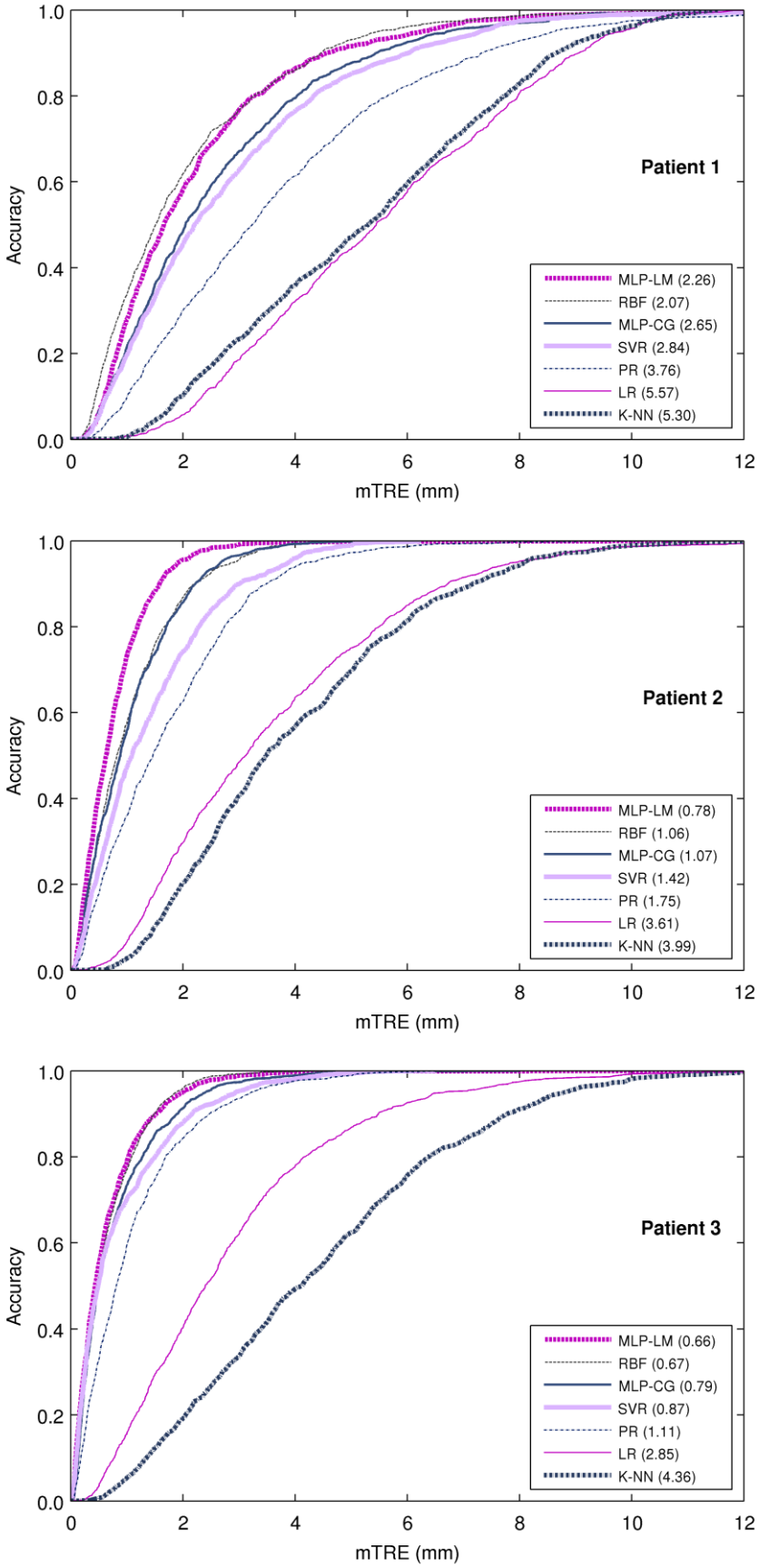


Figure 4.6 REC curves for all methods and for each patient (1 to 3) and AOC values.

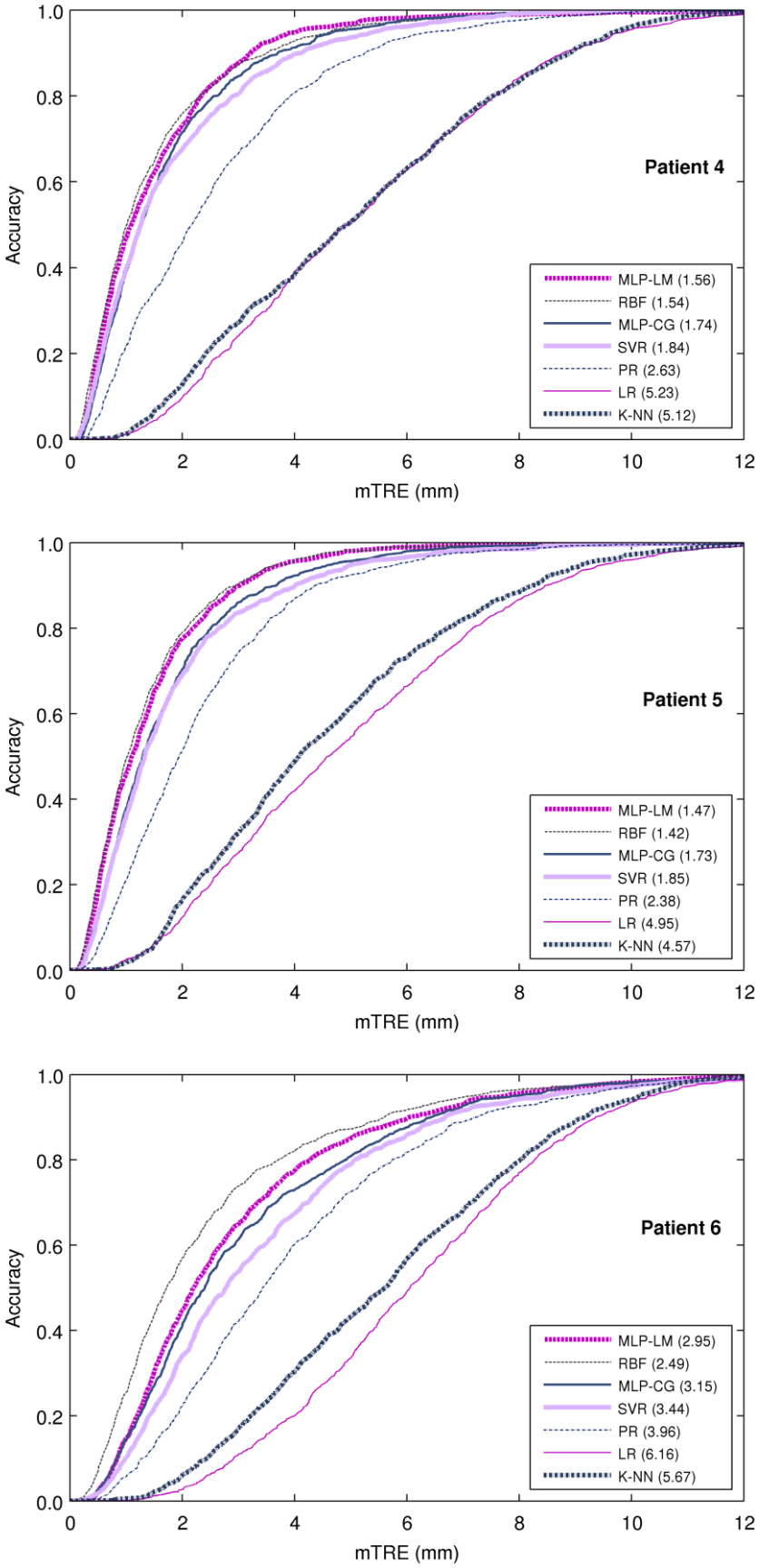


Figure 4.7 REC curves for all methods and for each patient (4 to 6) and AOC values.

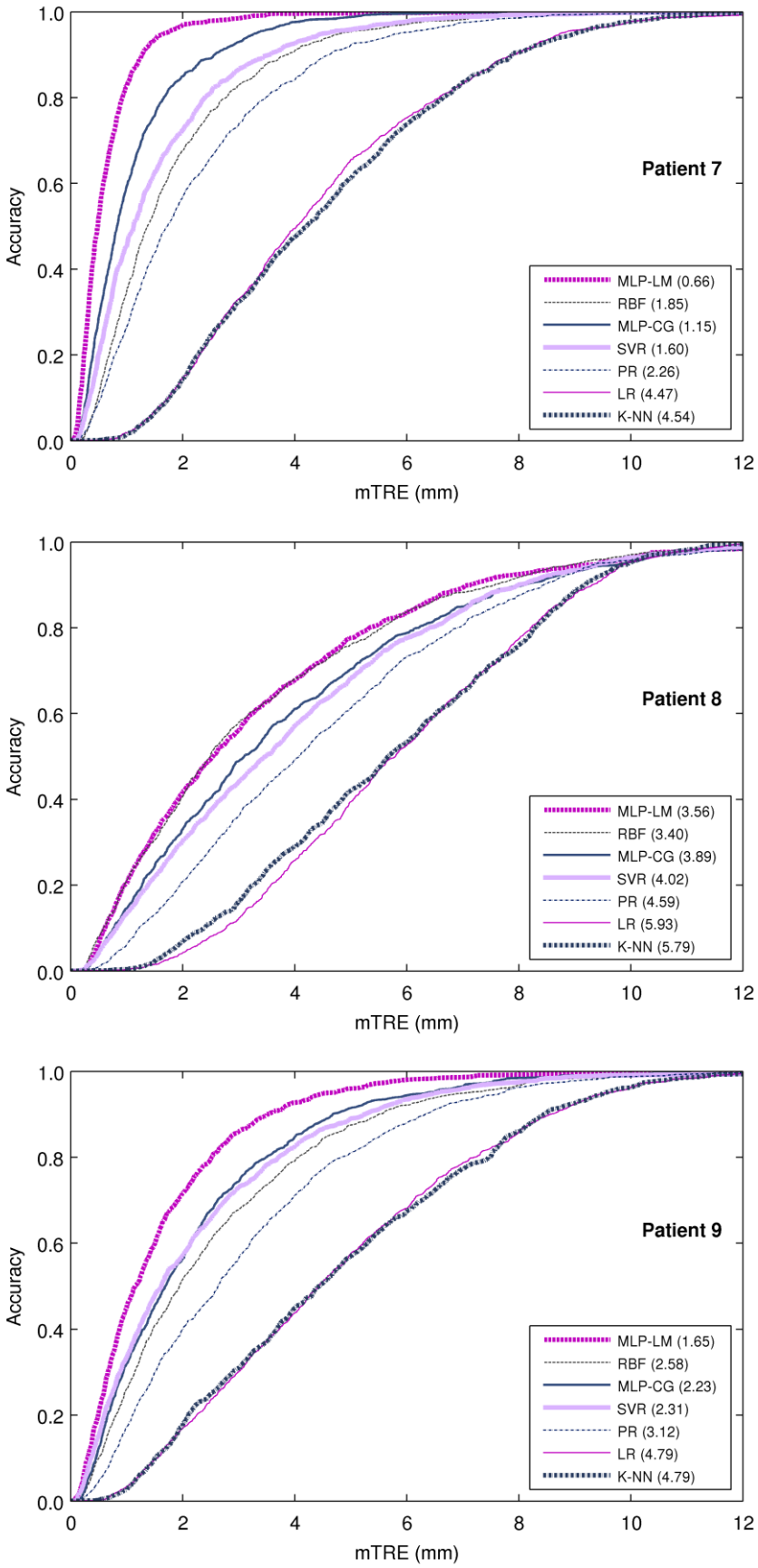


Figure 4.8 REC curves for all methods and for each patient (7 to 9) and AOC values.

4.4 Discussion and Conclusion

We compared different regression methods for the registration-by regression approach for 3D/2D image registration of coronary vessel trees. Figure 4.4 shows that Neural Networks (MLP and RBF) and SVR behave similarly and perform better than Multiple Regressions and k-NN. The registration-by-regression approach is less accurate but has a smaller variance when compared to the iterative registration approach, confirming the results obtained in Gouveia et al. (2012a) and Chapter 3 using the MLP-CG strategy

The evaluation per patient (Figure 4.6 to Figure 4.8) shows that MLP-LM and RBF have the best performances for all patients except for patient 7 and 9 (Figure 4.8), followed by MLP-CG and SVR. For patients 7 and 9, MLP-LM still gives the best result, followed by MLP-CG, SVR and finally RBF. For all patients, the worst results were obtained with k-NN and LR. The REC curve computed for all patients (Figure 4.5) also indicates the best and worst performance is MLP-LM followed by RBF and for k-NN/LR, respectively.

In general, the MLP-LM gives the best results using an optimization strategy which performs better than the MLP-CG approach in Gouveia et al. (2012a) and Chapter 3. The relatively low performance of LR and PR suggests that a highly nonlinear regression model is required for the registration-by-regression method.

Registration-by-regression with real X-ray images will be the focus of the next chapter (Chapter 5).

No meio do caminho

No meio do caminho tinha uma pedra
tinha uma pedra no meio do caminho
tinha uma pedra
no meio do caminho tinha uma pedra.

Nunca me esquecerei desse acontecimento
na vida de minhas retinas tão fatigadas.

Nunca me esquecerei que no meio do caminho
tinha uma pedra

Tinha uma pedra no meio do caminho
no meio do caminho tinha uma pedra.

Carlos Drummond de Andrade

Chapter 5
Registration-by-Regression of coronary CTA and X-ray angiography

This chapter is adapted into the manuscript in Appendix C:

Registration-by-regression of coronary CTA and X-ray angiography. Gouveia, A.R., Metz, C., Freire, L., Almeida, P., & Klein, S., *submitted*.

Abstract

We evaluate the integration of 3D preoperative CTA of the coronary arteries with intraoperative 2D X-ray angiography images using a registration-by-regression method. In this approach, image registration is estimated by a nonlinear regression model that is trained with DRRs. Image features of 2D projection images are related to the transformation parameters of the 3D image. The simulated 2D images were generated from coronary artery tree models derived from 3D CTA scans.

In this chapter we compared several image features: the 2D image moments and the geometric centers, the eigenvectors and eigenvalues from a PCA of the object imaged, and the projected images in the space of principal components. Moreover the influence of preprocessing options on simulated images used for the training set, as well as on real data, was studied. For the registration evaluation, a gold standard was developed from 9 X-ray angiography sequences from 7 different patients. The alignment quality was measured by the mean target registration error (mTRE).

The coronary registration–by-regression results were not satisfactory, resembling manual initialization accuracy. The framework performed showed the inadequacy of the features used to solve the registration problem. It enabled us to better understand the dependency of the proposed method on the differences between simulated and real images. Alternative solutions are suggested.

5.1 Introduction

Angioplasty, also called percutaneous coronary intervention (PCI), is a minimally invasive coronary intervention frequently used for the treatment of the coronary artery disease (CAD) (Kasper et al., 2005). The CAD results from the build-up of fatty deposits (or plaques) in the coronary arteries walls and the vessels become narrow and rigid, restricting the blood flow to the heart (Kasper et al., 2005). During PCI, the severally narrowed vessel is widen by inflation of a balloon at the location of the lesion. This is often combined with the placement of a stent (i.e. an expandable wire mesh tube) to prevent re-occlusion of the treated vessel segment (Meier, Bachmann, & Lüscher, 2003). X-ray imaging with contrast injection is the modality of choice for the guidance of PCI (Rivest-Hénault, Sundar, & Cheriet, 2012); it is used to visualize the vessels and guidewires, since the intervention is percutaneous. However, the navigation through the vascular system is particularly challenging in some situations since it is visualized by a projective imaging technique. This is especially the case for chronic total occlusions, in which to cross the lesion with a guidewire[3] can be a difficult task. A better visualization of the pathological region and some information about the lesion density is desirable (Metz, 2011)

Therefore, the integration of preoperative computed tomographic angiography (CTA) data, containing the occluded vessel segment, with the intraoperative X-rays can be very useful during the guidance and treatment of these pathologies. Any PCI can also benefit by increasing the accuracy of the stent placement, using pre-annotated CTA data, e.g., indicating plaque locations (Ruijters, ter Haar Romeny, & Suetens, 2009). Furthermore, the CTA can eventually give information about the density of the lesion contributing to the chronic total occlusions cases.

Our motivation is to achieve the integration of 3D preoperative CTA of the coronary arteries with intraoperative 2D X-ray angiography images with a registration-by-regression method (Gouveia et al., 2012a; Chapter 3). Whereas image registration is usually estimated by iteratively optimizing a similarity measure, in this approach it is treated as a nonlinear regression problem. In this chapter we extend the work developed by (Gouveia, Metz, Freire, & Klein, 2012a; Chapter 3), where simulated images were used, to real X-ray images.

Although with different formulations, we can find in the literature some authors adopting the regression solution for the 3D/2D registration problem (e.g. Chou & Pizer, 2013; Hoff, Komistek, Stefan, & Walker, 1998). However, to the best of our knowledge, no work with a

[3] In a PCI, a guidecatheter is inserted into a blood vessel in the upper thigh via an incision in the groin and moved through the vascular system towards the beginning of the coronary artery. Then, the guidewire is introduced in the body using the guidecatheter as a transport canal, and it is moved through the coronary artery to the site of the lesion.

registration-by-regression approach studied the coronary registration and, even for other applications, very few works with such a regression approach are found in the literature using real images in a clinically relevant context.

In general, for 3D/2D registration, a similarity measure is iteratively optimized between simulated X-ray projection images – digitally reconstructed radiographs (DRRs) – computed from the preoperative CT scan and an X-ray image (Markelj et al., 2012). Even for this more conventional iteratively-based approach, we cannot find in the literature many studies of 3D/2D coronary registration. Turgeon et al. (2005) presented a quantitative evaluation for 3D/2D coronary registration with simulated data, and some more recent works also included clinical data, as Metz, Schaap, Klein, et al. (2009), Metz et al. (2011) and Rivest-Hénault, Sundar, & Cheriet (2012), Metz et al. (2013), Baka et al. (2013), Baka et al. (2014), Aksoy et al. (2013). Although Ruijters, ter Haar Romeny, & Suetens (2009) considered clinical data, they only performed a qualitatively evaluation of the registrations (by an expert observer).

For the experiments with real X-ray data, a ground truth registration is not available. Therefore we constructed a gold standard by means of an exhaustive search routine, matching the projected centerlines as good as possible to manually annotated vessels on the X-ray. Similarly to some of the studies mentioned (Metz et al., 2009; Ruijters et al., 2009; Turgeon et al., 2005), we registered coronary CTA to X-ray angiography considering a rigid transformation. However, we find deformations sources like cardiac and breathing motion leading to non-rigid misalignments. Addressing the possible time differences with respect to cardiac cycle between the CTA and the X-ray scans, we constructed a gold standard for each X-ray of the time-sequences collected and performed a systematic visual inspection with five observers to select the best matching X-ray phase.

The main contributions of this work are:

- the evaluation of the registration-by-regression approach for a clinical relevant situation, and particularly a 3D/2D problem;
- the construction of a gold standard for 3D/2D coronary registration;
- we provide insight in the influence of the choice of features on the accuracy of the registration-by-regression approach.

The remainder of this chapter is structured as follows. In *Methods*, the next section, we explain the registration-by-regression methodology, the imaging data used, a detailed explanation of the gold standard proposed and the evaluation metrics. Next, three related experiments are described in *Experiment I*, *Experiment II* and *Experiment III*, where different features for the registration and/or different image preprocessing operations are evaluated. In *Experiment I* we used the best set of input features previously found and in *Experiment II* we extended the search to other related features, in combination with a variety of

preprocessing operations. In *Experiment III*, we tried a different kind of features and we also modified the imaged objects on the training dataset; an overall comparison of the features studied in the chapter was performed confined to one of the preprocessing options. The conclusion chapter summarizes the main developments of the work.

5.2 Methods

5.2.1 Registration by nonlinear regression

In this chapter we use the 3D/2D registration by nonlinear regression method presented in Gouveia et al. (2012a) and Chapter 3, schematically represented in Figure 3.1. This regression model relates image features of the 2D projection images to the transformation parameters of the 3D image. Once the nonlinear regression function is computed, which takes place before the intervention, the estimate of the 3D transformation parameters of the 3D image can be obtained by applying the regression function to the features computed on the intraoperative X-ray. During the preoperative phase, the training set for the learning process is built comprising a set of input-output pairs: the features extracted from the simulated 2D images (DRRs) as inputs and the corresponding 3D transformation parameters as outputs. The simulated 2D images (DRRs) are generated by applying random transformations to the pre-interventional 3D image (coronary CTA) followed by projection of its coronary artery segmentation. During the intervention, the image features of the intraoperative X-ray are computed and the 3D translation and rotation parameters of the 3D image are estimated.

The regression model used is a feed-forward Neural Network, as described in detail in Sections 3.2.3 and 3.3.2. An additional study (Gouveia, Metz, Freire, & Klein, 2012b; Chapter 4) compared different regression models for this registration-by-regression problem. It was concluded that the Multi-Layer Perceptron is a good choice, having the highest performances of all methods studied. In Experiment I and Experiment II (Sections 5.3 and 5.4), the MLP with Levenberg-Marquardt optimization was used, whereas in Experiment III (Section 5.5) the conjugate-gradient optimization was used because of the higher number of input units in this case.

5.2.2 Imaging data

For this chapter, we performed the registration of 3D preoperative coronary CTA data with 2D intraoperative X-ray angiography using 9 time-sequences. These 9 X-ray angiography sequences were collected from 7 different patients, where the different sequences of the same patients correspond to different intervention dates. The regression function for the registration-by-regression method was obtained by using a training set of simulated projection images of the coronary vessel tree with known ground truth transformation. From the retrospectively ECG-gated 4D CTA images available, we segmented coronary arteries (at the end-diastole phase) and made binary vessel tree models (Metz et al., 2011), as it is shown in Figure 4.3. DRRs were generated using these 3D models with a computation procedure described in Metz, Schaap, Klein, et al. (2009). To represent a clinically relevant view, the projection geometry for the computation of the DRRs and the initial orientation of the preoperative data were derived from an interventional X-ray image.

The CTA images were obtained in the end-diastole phase, more precisely at 70% of the cardiac cycle considering a R-wave peak-to-peak cycle. The X-rays images used for registration were the ones from the second-half of the cardiac cycle, i.e. from 50% to 100% of the cardiac cycle, which contained on average 6 time frames.

CTA images were acquired using a Siemens Definition or Siemens Definition Flash and X-ray images using a Siemens Axiom Artis biplane system. As calibration data was not available for the X-ray acquisition, only monoplane experiments using the sequences of the primary C-arm were considered. The field of view of the CTA images used for the reconstruction was 256x256x[99-184] voxels and the resulting voxel size was 0.7x0.7x [0.8-1.0] mm^3; for DRR images and X-rays images the same quantities were 512x512 pixels and 0.22x0.22 mm^2, respectively.

For each patient, 11000 DRRs were generated, 10000 to obtain the regression model and 1000 to test the performance of our method before the registration of real data. The transformations were drawn from a uniform distribution, with values between -13 and 13 degrees for rotations and between -16 and 16 mm for translations. The transformations ranges used in this chapter differ from Chapter 3 (and Chapter 4) where we used ranges from -10 to 10 degrees and millimetres for rotations and translations, respectively. In these chapters, the regression methods were evaluated in a controlled setting, using simulated projection images of the coronary vessel tree; in the current chapter, we are using real X-ray images. The gold standard transformation parameters were computed according to the procedure described in Section 5.2.3.1. The transformation ranges were conditioned to those values to ensure the training set contains the transformation parameters of the X-ray images to be registered.

5.2.3 Evaluation of registration

5.2.3.1 Gold standard

A gold standard for evaluation of the registration-by-regression method is needed. A problem thereby is the occurrence of non-rigid deformation due to cardiac and respiratory motion. Particularly, time differences with respect to cardiac cycle could exist since we matched an end-diastolic 3D model to all frames of a X-ray time-sequence of the second half of the cardiac cycle. Therefore we constructed a gold standard for each X-ray frame, asking five observers to rank the quality of alignment.

For each X-ray frame of the 9 X-ray time-sequences collected, we first manually annotated the centerlines of the coronary arteries. Then we implemented an exhaustive search for each X-ray frame based on the 2D distances between vessel centerlines, matching the projected centerlines as good as possible to manually annotated vessels on the X-ray. Then, 5 different observers analyzed the result of this full-search through a systematic visual inspection of the X-ray frame with both manually annotated and projected centerlines, in order to select the best matching X-ray phase. A score was attributed to each case, and we only used the most coincident X-ray frame with respect to the phase cycle of each time-sequence.

For each X-ray frame the initial exhaustive search was done as following:

(1) We started by manually annotating the centerlines in each X-ray angiography of all images collected.

(2) We projected the vessel centerlines from the correspondent coronary artery 3D model (as explained in Metz, Schaap, Weustink, et al., 2009), after a given transformation has been applied.

(3) We measured the shortest distance between annotated and projected centerlines. See Section 5.2.3.1.1 for details.

(4) Second and third steps were then exhaustively repeated over a large range of rotations/translations of the 3D model to find which transformation minimizes the distance. See Section 5.2.3.1.2 for details.

After having performed this procedure for all time frames of the X-ray sequence, a manual scoring was done, in which the X-ray time frames were ranked based on the quality of the alignment with the projected CTA centerlines. The observers were asked (a) to rank the images of the time-sequence on alignment quality, i.e. to order all images from best to worst match; and (b) to identify unsatisfactory alignments. The best case was accepted as a gold standard, but only if at least half of the observers scored the alignment as satisfactory. In further registration experiments, we only use this best case time frame, as only for that time

frame we can be confident about the gold standard, and we are sure that the registration is not hampered by non-rigid deformations introduced by cardiac and respiratory motion. Section 5.2.3.1.3 reports the results of the scoring procedure.

5.2.3.1.1 Distance between annotated and projected centerlines

The gold standard of 3D/2D coronary registration requires the determination of the shortest distance between annotated and projected centerlines. For this purpose, we first matched the segments of manually annotated centerlines to segments of projected centerlines[4] and then distances between the centerlines were computed. Figure 5.1 shows an example of segments manually annotated on a X-ray image.

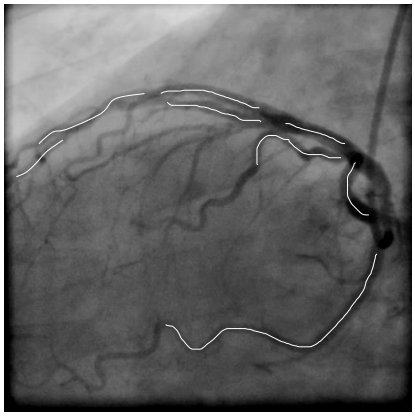


Figure 5.1 Example of manual annotated segments (in white) in a X-ray image.

Given a point on a manually annotated segment (Figure 5.2.a), we compute the distances to every point on a projected segment and take the minimum distance (Figure 5.2.b). This minimum distance calculation is repeated for every point of the annotated segment and the average distance over all points is calculated (Figure 5.2.b). In the same way, the average distance for all annotated segments to all projected segments is computed and the projected segment with minimum average distance is taken as the corresponding segment (Figure 5.2.c).

[4] The projected segments are determined by averaging all centerline points from different centerlines that are closer than 0.5 mm apart.

After segment matching, we can determine the distance for every point of every segment on the annotated centerline to the projected centerline, which is the minimum distance to the matched segment (Figure 5.2.d). The average and standard deviation over all the points of the annotated centerline is used as a measure of the misalignment between the X-ray and the DRR.

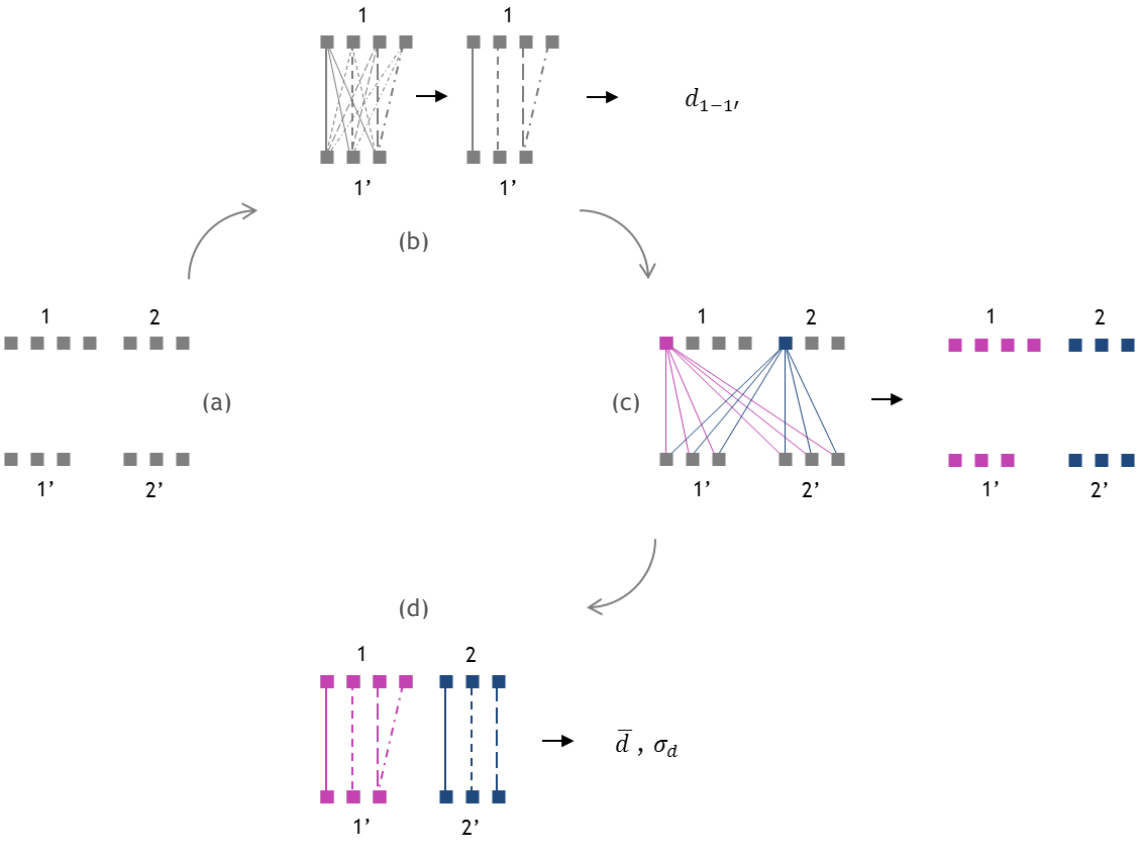


Figure 5.2 Shortest distance determination between (a) a manually annotated centerline with two segments, named 1 and 2, to a projected centerline with two projected segments, named 1' and 2'. (b) The distance of each point of 1 to each point of 1´is determined, the shortest distance of each point of 1 is taken and the average distance for the shortest distances of the 4 points of segment 1 is computed ($d_{1-1'}$). (c) The average distances of all annotated segments to all projected segments are computed ($d_{1-1'}$, $d_{1-2'}$, $d_{2-1'}$, $d_{2-2'}$) and the segments are matched by closeness. (d) Considering the minimum distance of each point to the matched segment, the average and standard deviation is computed over the 7 points of the annotated centerline.

5.2.3.1.2 Search for the closest projected centerline

For the exhaustive search of the closest projected centerline to the annotated centerline, we used a multi-resolution scheme for practical reasons, keeping the computation time within reasonable limits. A coarse-to-fine search was performed, with higher deviations and steps in the first 3 levels followed by smaller values in the final 3 levels (Table 3.1). Moreover, we manually initialized the parameters to limit the required search range.

Table 5.1 Multi-resolution scheme for the closest projected centerline search.

Levels	1		2		3		4		5		6	
	dev	step	*dev*	step	*dev*	step	*dev*	step	*dev*	step	*dev*	step
T_X,T_Z (mm)	10	1	-	-	-	-	2	0.5	-	-	0.3	0.1
T_Y (mm)	-	-	10	1	-	-	2	0.5	-	-	0.3	0.1
R_X,R_Y,R_Z (º)	-	-	-	-	0.5	0.05	0.1	0.05	0.05	0.02	0.02	0.005

The search was performed within the range $[X_i - dev_i\,,\, X_i + dev_i\,]$ by an indicated step, where X_i is the parameter value found for a manual initialization and dev_i is the deviation for each of the transformation parameters.

5.2.3.1.3 Gold standard results

For each of the 9 time-sequences analyzed, all frames were ranked from best to worst match. When a frame was measured as the best frame for each observer, it got a score of 0; the second best got a score of 1, and so on. The final score was the sum of scores of a frame over all observers, and the ranking was established ordering the final scores from minimum to maximum, where a score of 0 means all observers have considered it the best alignment of all frames of the time-sequence. The values of the final scores and the ranking order are shown in Table 5.2. Simultaneously, for each frame, the alignment was classified as satisfactory or unsatisfactory; the number of observers which considered the alignment as satisfactory is presented (Table 5.2) over a total of 5. This analysis should be in a rigid alignment sense, defining as unsatisfactory alignments those which position and/or orientation of the 3D model is wrong (e.g. when a vessel bifurcation is not well rotated). Thus, local misalignments of the centerlines must be taken into account carefully because they could be due to non-rigid deformation. The quality of this local alignment will be captured by the ranking of the images. If an observer is not able to judge if the orientation is correct, then the alignment should be marked as unsatisfactory.

Table 5.2 Final scores and ranking of each frame of each time-sequence analysed

	Time-sequences of X-rays								
	Quality of each alignment over the time-sequence: frames scores								
Ranking	1	2	3.1	3.2	4.1	4.2	5	6	7
1	**0**	**1**	**2**	**2**	**0**	**1**	**2**	**5**	**5**
2	7	7	5	5	8	5	5	6	6
3	11	18	13	8	8	10	13	10	10
4	16	19	15	15	19	14	15	13	12
5	20	23	17	20	20	20	17	16	17
6	21	23	24	25	20	29	23		
7		24	29			29			
8		32				32			
9		40							
10		41							
	Quality of the best scored alignment of each time-sequence								
#/5	5	5	5	5	4	5	5	4	2

The lower the score the better the alignment, and number of observers in a total of 5 that judged as satisfactory the best scored alignment per time-sequence.

For one of the time-sequences collected, none of the frames attained the gold standard criteria defined before. Even for its best case the number of satisfactory judgments was too low, i.e. only 2 in a total of 5 observers set it as an satisfactory alignment. Figure 5.3 shows the best and worst alignment for time-sequence 3.2, where although the best reached the gold standard criteria the worst did not. The same cases for the remaining time-sequences are shown in Appendix C.

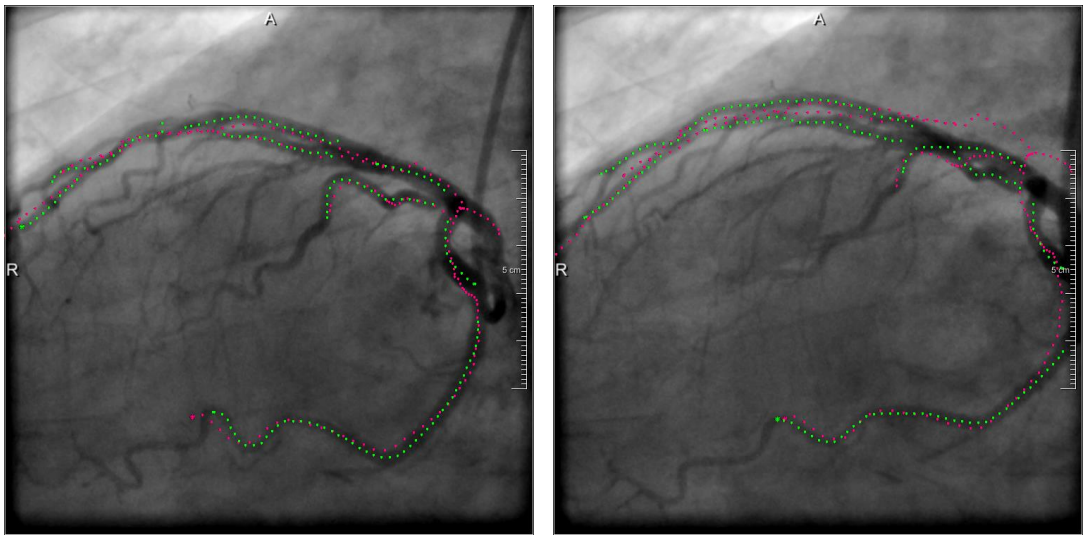


Figure 5.3 Examples of two images used for the evaluation of the alignment quality. These are X-ray frames of the time-sequence 3.2 where the manual annotated centerline and the closest projected centreline (obtained by the full-search process) are highlighted in green and in pink, respectively. They represent the best (left) and the worst (right) alignment according to the ranking established.

Since one of the time-sequences did not attain the gold standard criteria, in the remaining chapter just 8 time-sequences of 6 patients will be considered.

5.2.3.2 Evaluation of registration

The evaluation of the registration approaches was performed by the computation of the mean target registration error ($mTRE$) before and after registration (Fitzpatrick & West, 2001). The $mTRE$ is computed as the mean 3D distance to the gold standard position and orientation (computed as described in Section 5.2.3.1) of all points of the vessel centerline trees:

$$mTRE(T) = \frac{1}{N}\sum_{n=1}^{N}\left\|T(p_n) - T_{gold}(p_n)\right\| \qquad \text{Equation 5.1}$$

where T is the resulting transformation from the registration method to assess, T_{gold} is the known gold standard transformation and p_n are points on the centrelines of the 3D vessel tree.

$mTRE$ for simulated images were computed on the test set of 1000 images, which was not used for the construction of the regression models. For real data, $mTRE$ was determined for one X-ray of each time-sequence as defined before, in Section 5.2.3.1. The gold standard of simulated images is the known set of transformation parameters of the test images, and for real images it is the set obtained from the process of Section 5.2.3.1.

5.3 Experiment I

In the registration-by-regression method a regression function is determined using a training set of simulated images. In previous chapters, the method was tested on simulated images, taking an unseen set for this purpose. However, in the present chapter, real images are also used and the problem becomes much more challenging. Preprocessing tasks like background elimination and vessel segmentation may be required to establish a correct relationship between images and allowing a successful registration.

In this first experiment we tested the registration-by-regression method with real images, studying different preprocessing options. We used the best set of input features previously found for simulated images (Gouveia et al., 2012a; Chapter 3).

5.3.1 Input features I

The features extracted from the 2D projection images, either for DRRs in the training phase or for X-rays during the intervention, were 1) the 2D geometric moments of order 0, 1 and 2 and 2) the eigenvalues and eigenvectors computed from a Principal Component Analysis (PCA) on the pixels of the object of interest, after image segmentation as described in Section 5.3.2. This set was the best performing option obtained from a comparison study of different image features in Gouveia et al. (2012a) and Chapter 3 for this problem.

The PCA was performed on 3D vectors with x, y and $I(x,y)$ as variables, where $I(x,y)$ is the intensity value of the point at position (x,y). Before the PCA computation and due to the unit difference between the intensities $I(x,y)$ and the pixel positions x and y, a normalization was required. Therefore, a z-score was applied to the set of 3D vectors with an identical mean and standard deviation for x and y, to prevent loosing pose information.

5.3.2 Image preprocessing I

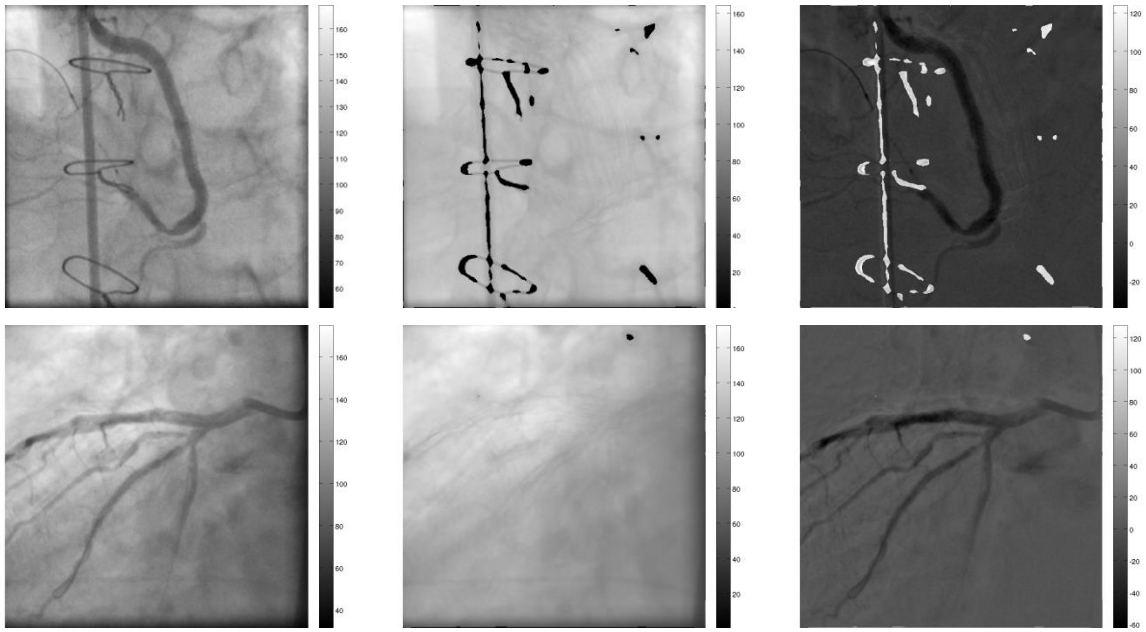
Two main preprocessing steps were implemented: the background subtraction and the vessel segmentation of coronary arteries. Vessel segmentation is mandatory since only the information in X-rays about the coronary arteries matters. However, the presence in the image background of some structures (i.e., like ribs, wires and other vessels) could influence the automatic vessel segmentation. Therefore we applied a background subtraction.

We used the automatic segmentation method proposed by Frangi et al. (1998) which treats vessel enhancement as a filtering process searching for tubular geometrical structures. A

range of scales at which relevant structures are expected to be found were defined, since vessels appear in different sizes. For each of those scale values, named as s, the authors proposed to measure the contrast between the regions inside and outside the range $(-s, s)$ in the direction of the second derivative. This 2D vesselness filter was used for the background segmentation step and also for the vessel segmentation.

5.3.2.1 Background subtraction

For the background subtraction (Figure 5.4), we used the information of the X-ray sequence over time. Given a time-sequence of X-rays, we defined a vessel mask for each X-ray image using the 2D vesselness filter proposed by Frangi et al. (1998). We determined the intensity average over time of each pixel outside the mask, ignoring the pixels that lie inside the vessel mask. Then, the average was subtracted from the X-ray image chosen by the gold standard process (Section 5.2.3.1), and the resulting image was used as input for the vessel segmentation.



(a) (b) (c)

Figure 5.4 Background subtraction for X-rays of time-sequences 1 and 2 (in 1st and 2nd rows respectively): (a) original X-ray, (b) average over time, (c) resulting subtracted image. The white pixels in the final images, as well as black pixels in the average images, correspond to fixed structures for the automatic segmentation process, i.e. these pixels have never been considered background by the automatic function over the all sequence. The existence of these pixels leads to an enlargement of the intensity windows, giving a darker and brighter appearance, respectively.

The average was determined by dividing the sum of all time-sequence images backgrounds by the sum of the masks of all time-sequence images, represented by Figure 5.5.e and Figure 5.5.c respectively.

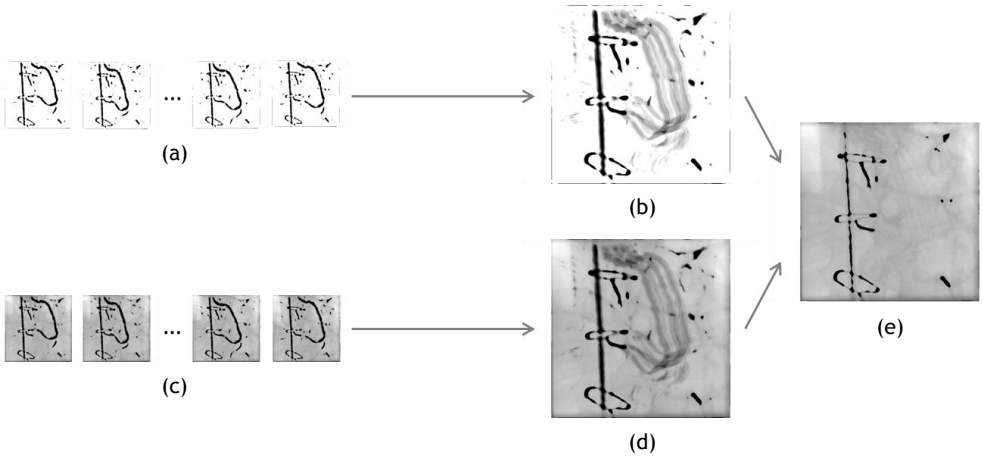


Figure 5.5 Determination of the image average over time.(a) Computation of vessel masks for each X-ray of the time-sequence, (b) sum of all masks. (c) Background determination for each X-ray, by setting to 0 the intensity of the pixels inside the correspondent mask, and (d) sum of the backgrounds. (e) Image average resulting by the division of (d) by (b) (assuming as 1 the values smaller than 1).

An alternative to this subtraction process would be the subtraction of the average over time using the entire images of the X-ray time-sequences, without the previous vessel masks (Figure 5.6.S2). Although this subtraction would also eliminate the static structures, it would interfere with the intensity inside the vessels since it does not use masks.

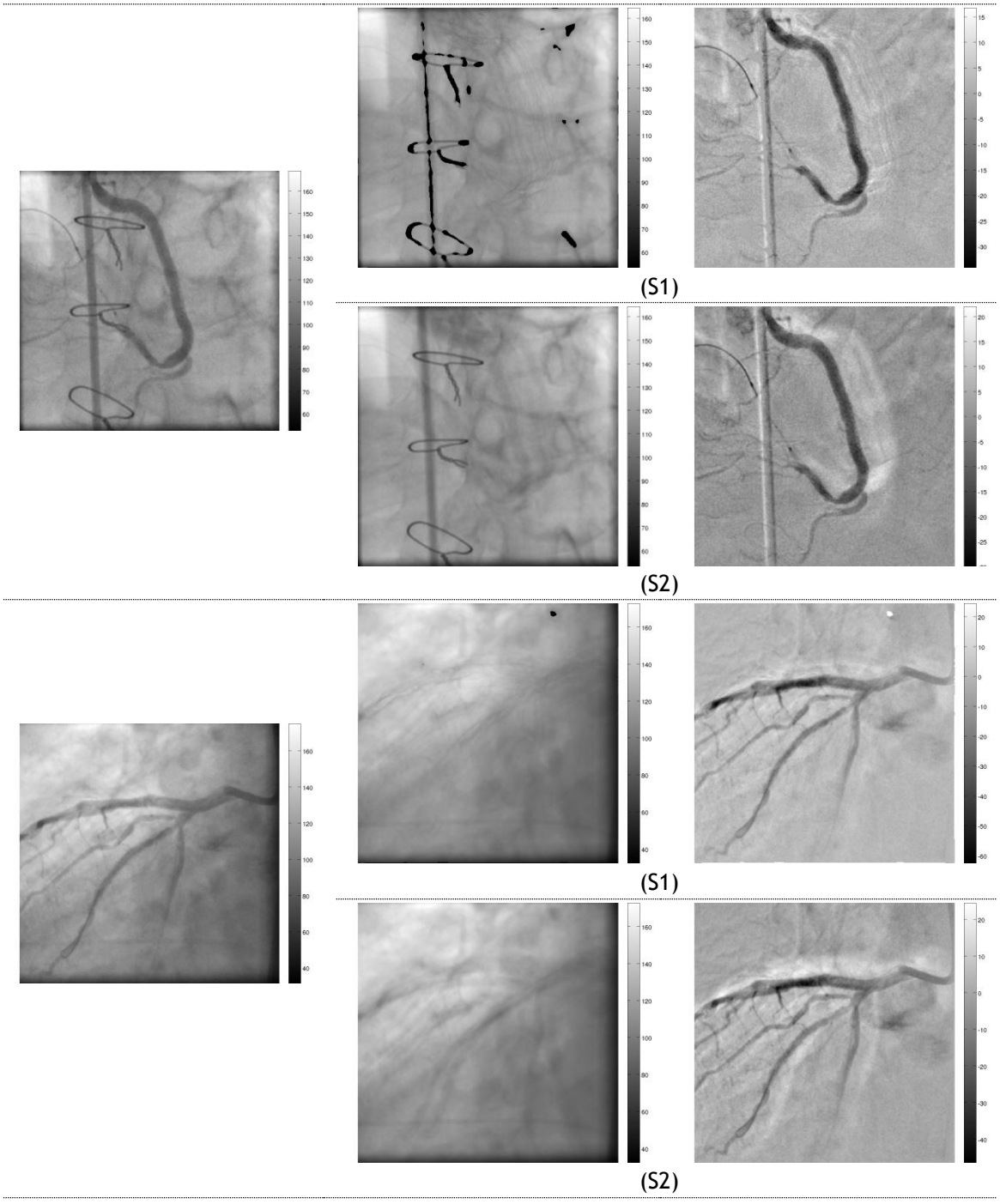


Figure 5.6 Background subtraction with the average of the image backgrounds (S1) and with the average of entire images (S2), for the X-rays 1 and 2 presented in Figure 5.4 (i.e., of time-sequences 1 and 2 respectively): (a) original X-ray, (b) average over time, (c) resulting subtracted image. The intensity windows of images (b) and (c) for S1 were modified in this figure, to a better comparison between S1 and S2. For the remaining X-rays the same figures are shown in Appendix D.

5.3.2.2 Vessel segmentation

The vessel segmentation comprises a mask computation process, with 3 options (named as 1, 2, 3), followed by the preprocessing of intensity values inside the mask with 4 options (named as A, B, C, D).

Three masks were determined (Figure 5.7) by:

1) Using a manual segmentation;

2) Using an automatic segmentation by a thresholding operation with a value t1;

3) Using an automatic segmentation by a thresholding operation with a value t2.

(1) (2) (3)

Figure 5.7 Manual mask (1), automatic mask corresponding to threshold t1 (2) and automatic mask corresponding to threshold t2 (3) for X-ray 1. For the remaining X-rays the same figures are shown in Appendix E.

Option (1) was obtained by the analysis of two observers, and for options (2) and (3) we used the 2D vesselness filter (Frangi et al., 1998) with two different thresholds t1 and t2. Figure 5.8 shows the manual masks for two X-ray images.

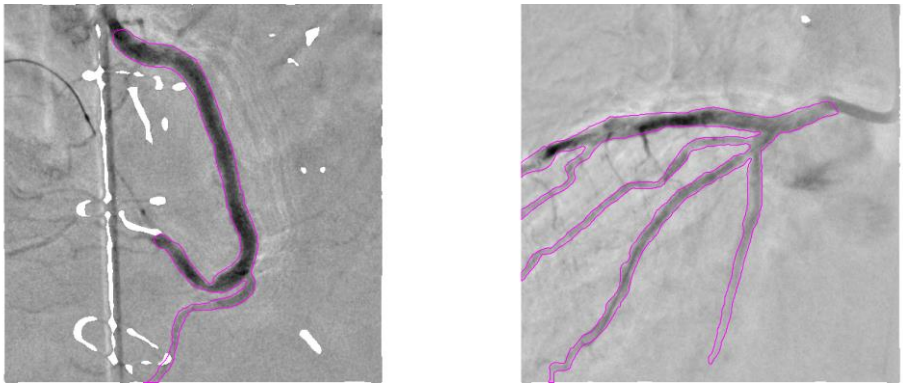


Figure 5.8 Manual masks (colour lines) on images after background subtraction for X-ray 1 (left) and X-ray 2 (right). The manual masks for the remaining X-rays are presented in Appendix F.

We computed the automatic masks (2) and (3) in 4 steps: (i) outlier rejection where we applied a Z standardization (z-score) on the image, changed the mean and standard deviation from (0, 1) to (0.5, 0.25) respectively and eliminated the tails above 0.75 and below 0.25; (ii) vessel segmentation by applying the 2D vesselness filter on the output of the previous step; (iii) z-score normalization of the vessel segmentation result; (iv) and intensity thresholding by only considering the values above a specific value. The mask is a binary image whose pixels found in step (iv) have an intensity value 1 and the remaining an intensity value 0.The first step addresses the existence of intensity outliers after background subtraction noticed in the preceding section, and shown in Figure 5.4. The masks options (2) and (3), differ in their threshold values used in step iv.

Since the input features of Experiment I are dependent of the coronary arteries intensity, differences between intensity values must be carefully prevented. Therefore we considered different options for the intensity values of pixels that lie inside the mask:

A) Using the X-ray image values;

B) Using the values of the 2D vesselness filtered image, where the filter was applied on the background corrected X-ray image;

C) Using the values of the 2D vesselness filtered image, where the filter was applied on a masked X-ray image, using the mask described above;

D) Using the segmentations directly (totally ignoring intensity information)

The option C is considered because the values of the filtered image could be corrupted because of the information outside the vessels. For options A, B and C some preprocessing operations were applied: a Z standardization (z-score) on the image pixels inside the mask, a mean and standard deviation change from (0, 1) to (0.5, 0.25) respectively and the eliminations of the tails above 1 and below 0. The final image assumes the intensity values determined by this process in the pixels within the mask and set to zero the remaining. In the case of option D the intensity within the mask is 1.

Figure 5.9 and Figure 5.10 shows the 4x3 images obtained from the vessel segmentation process for one X-ray image in a grey and coloured scale, respectively.

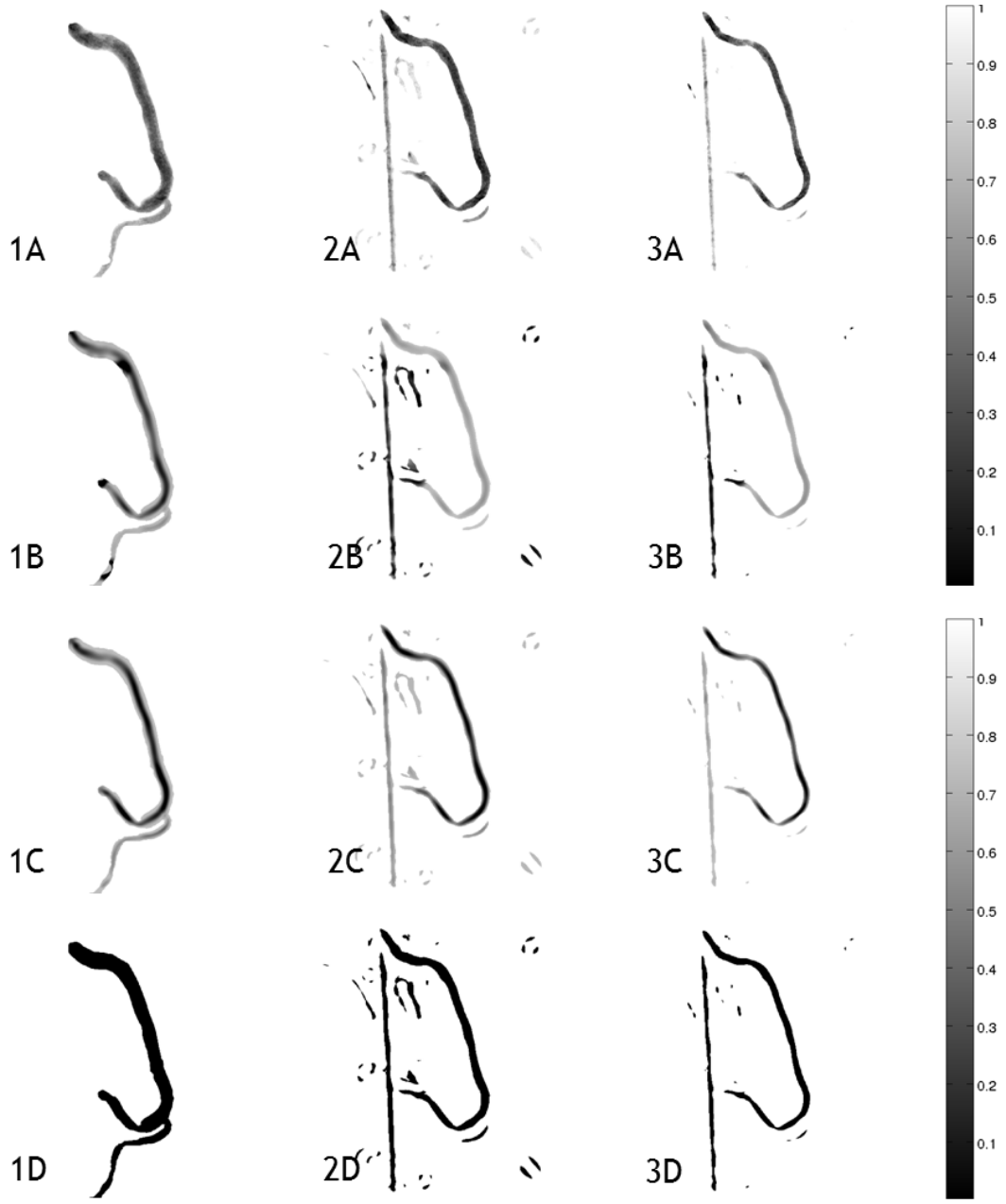


Figure 5.9 Vessel segmentation results for X-ray 1 in grey scale considering the different combinations used: columns (1,2,3) correspond to the mask options and rows (A,B,C,D) to the intensity values of pixels within the masks.

Although the image preprocessing was designed for X-ray angiographies, the same preprocessing operations for the intensity values (A, B, C, D) must be applied on the DRR training set to obtain the regression model for the CTA/X-rays registration. For the mask

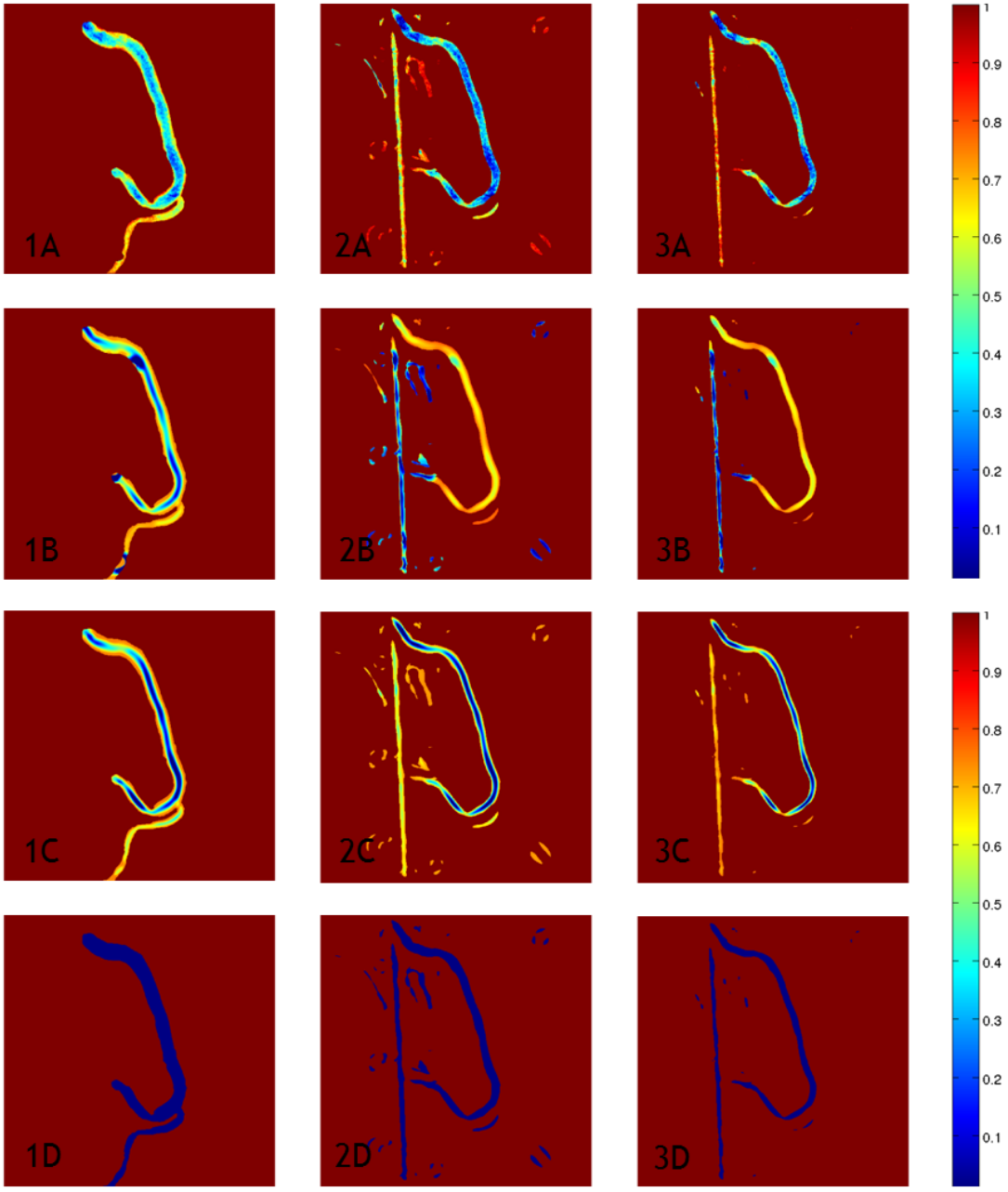


Figure 5.10 Vessel segmentation results for X-ray 1, now in coloured scale, considering the different combinations used: columns (1,2,3) correspond to the mask options and rows (A,B,C,D) to the intensity values of pixels within the masks. For the remaining X-rays, images are shown in Appendix G.

determination, and once these images are simulated projections of the vessel tree without any structures in the background (as it is noticed in Figure 3.5), we segmented the images simply using a threshold value manually chosen.

5.3.3 Implementation details I

For image preprocessing, the range of scales defined in the 2D vesselness filter (Frangi et al., 1998) was {5,5.5,6,6.5,7,7.5,8}. These values were empirically chosen so that they will cover the range of the vessel widths. The remaining parameters required by the vesselness filter were chosen following the authors recommendations. In vessel segmentation, the thresholds values for the masks (determined by automatic segmentation) were empirically chosen: 0.1 and 0.2 as t1 and t2, respectively.

For each registration, 6 regression models were used, one per transformation parameter, using identical Neural Networks. Each one has one hidden layer with fully connected units and a topology [18:36:1]: 18 input units (i.e., 18 input features), 36 hidden units (from optimization studies in Chapter 3 and Chapter 4) and 1 output unit (the correspondent transformation parameter). The activation functions are the hyperbolic tangent and the linear functions for the hidden and output layers, respectively. The input units were standardized by a z-score (Haykin, 1999) and the weights were randomly initialized within a range [-1,1]. The number of epochs was defined separately for each MLP by a stopping epoch (i.e., the epoch when the validation error started to grow) with a maximum of 1000.

For each patient, the set of 10000 images, for the construction of the regression model, was split in two sets: 7000 images to train the regression models and the remaining 3000 to select the stopping epoch. The performance of the models was computed by the mean absolute difference between the model prediction and the known output for each regression.

The features computation, the image preprocessing and the implementation of NN were performed using MatLab, version 7.11.0.584 (R2010b) 64 bits (MatLab, 2010). The function for the 2D vesselness filter used was obtained from the MatLab Central - File Exchange[5].

5.3.4 Results I

The results of the 3D/2D coronary registration-by-regression using as feature set the first three 2D geometric moments and the eigenvalues and eigenvectors (from the PCA on the pixels of the object of interest) are presented in this section. The $mTRE$ before and after registration is computed as defined in 5.2.3.2, for simulated images on the test set of 1000 unseen images and for real data on one X-ray per time-sequence selected from gold standard procedure (Section 5.2.3.1).

[5] This function was written by Marc Schrijver in 2001 and re-written by D.Kroon in 2009, updated in Mars of 2010. It can be found in http://www.mathworks.com/matlabcentral/fileexchange/24409-hessian-based-frangi-vesselness-filter.

5.3.4.1 Simulated data

In this chapter, we enlarged the range of the transformation parameters (Table 5.3) for training purposes with respect to the previous work, as explained in Sections 5.2.2 . The mTRE became higher, as it is shown in Table 5.3 and in Figure 5.11 for one patient's data, where values from the previous chapter and current values are compared. The mTRE values were computed on the test set considering the DRR without any of the preprocessing steps described in *Image preprocessing I* (Section 5.3.2).

Table 5.3 Ranges of the transformation parameters used for the training set currently and in the previous chapters, and respective means and standard deviations of mTRE values before and after the registration.

	Previously	Presently
Rotations range (°)	[-10,10]	[-13,13]
Translations range (mm)	[-10,10]	[-16,16]
Initial mTRE (mm)	10 (3)	17 (4)
Final mTRE (mm)	0.35 (0.55)	2.17 (2.80)

The values presented were computed for Patient 1.

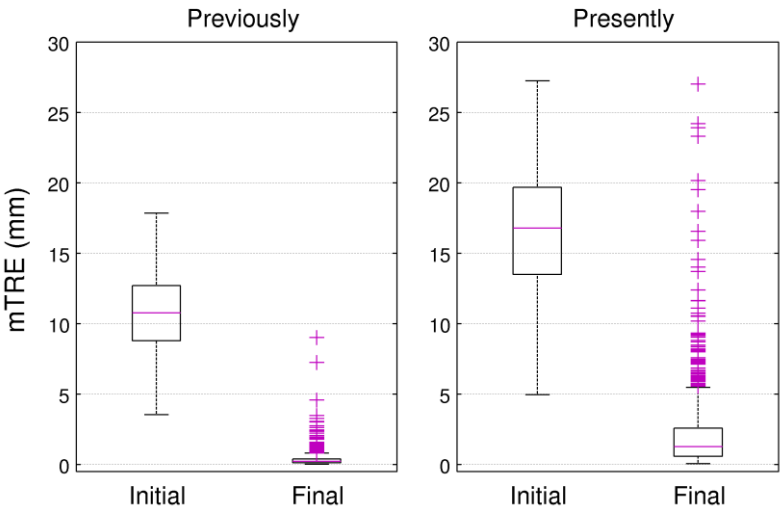


Figure 5.11 Comparison of the transformation parameters ranges used currently and in the previous chapters for the training set. The values presented were computed before and after the registration for Patient 1.

The registration results using simulated images and considering the different preprocessing options for the intensity values, defined for this *Experiment I* (Section 5.3.2), are presented in Figure 5.12. The option which ignored the intensity information (D) had the worst results, whereas option A was in general the most accurate, which included the intensity values of the image without any vesselness filter (as B and C included).

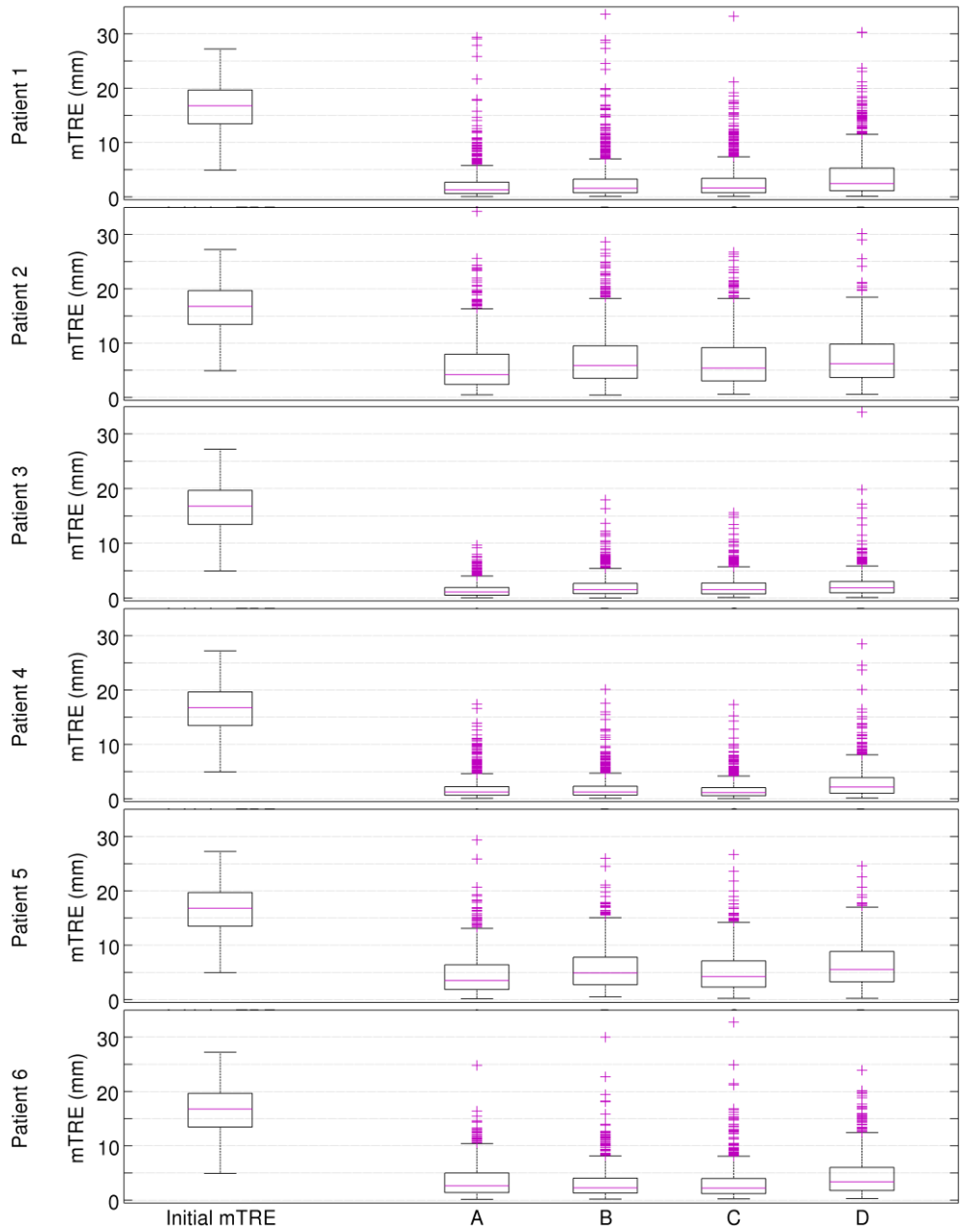


Figure 5.12 Comparison of the preprocessing options applied to the DRRs of the training set for each patient, labelled as A, B, C and D. The mTRE values before the registration (*initial mTRE*) are also shown.

5.3.4.2 Real data

Figure 5.13 and Figure 5.14 summarize the results of 3D/2D coronary registration-by-regression with X-ray angiographies after the image preprocessing defined for Experiment I has been applied (Section 5.3.2), with different masks (1, 2, 3) and intensity values (A, B, C, D). The grey map presents in an overall way the magnitude of 3D mTRE for real data up to 50 mm. Additionally, mTRE values from a manual registration were also displayed in the plots (Figure 5.14), which contextualizes the registration results. These *manual* values were computed before the registration but after a manual initialization and would be the initialization position for an iteratively approach. The manual initialization was performed on the projection plan (XY) while observing a X-ray image, after moving the centre of gravity of the model to the origin of the intraoperative coordinate system (Metz, 2011).

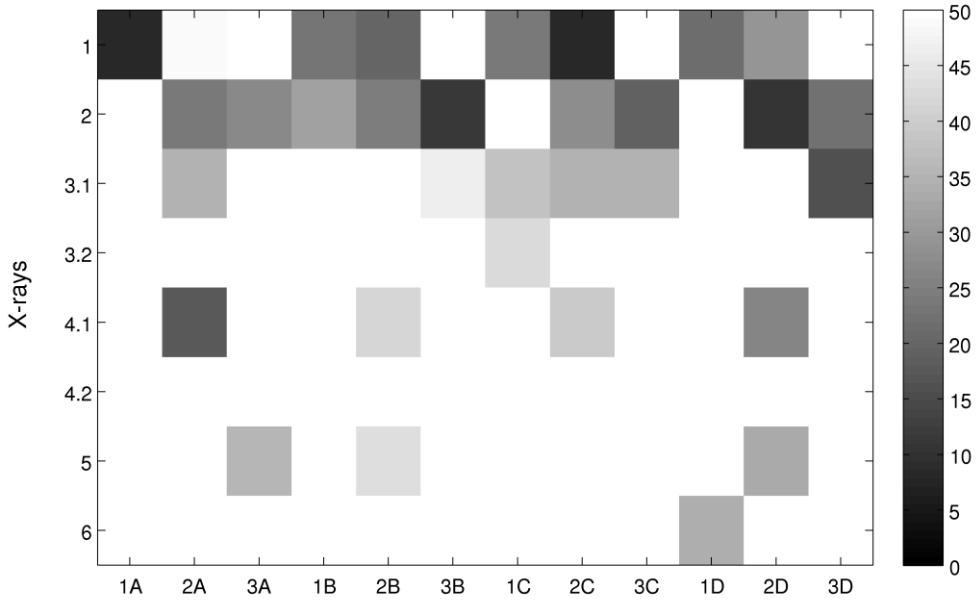


Figure 5.13 Grey map with coronary 3D/2D registration results for Experiment I. 8 X-rays of 6 patients were used for all preprocessing options, labelled with a number from {1, 2, 3} for the masks and a letter from {A, B, C, D} for the intensity values (as defined in Section 5.3.2.2). Only mTRE below 50 mm are shown.

The mTRE values obtained were very high and just half of the X-rays presented at least one value at the range of the manual mTRE, i.e. just for half of the X-rays the registration-by-regression got the accuracy of a manual initialization (Figure 5.14). Furthermore, for 5 of the 8 X-rays, the majority of preprocessing options resulted in mTRE values above 50 mm, as it is noticed by white spots in the grey map (Figure 5.13).

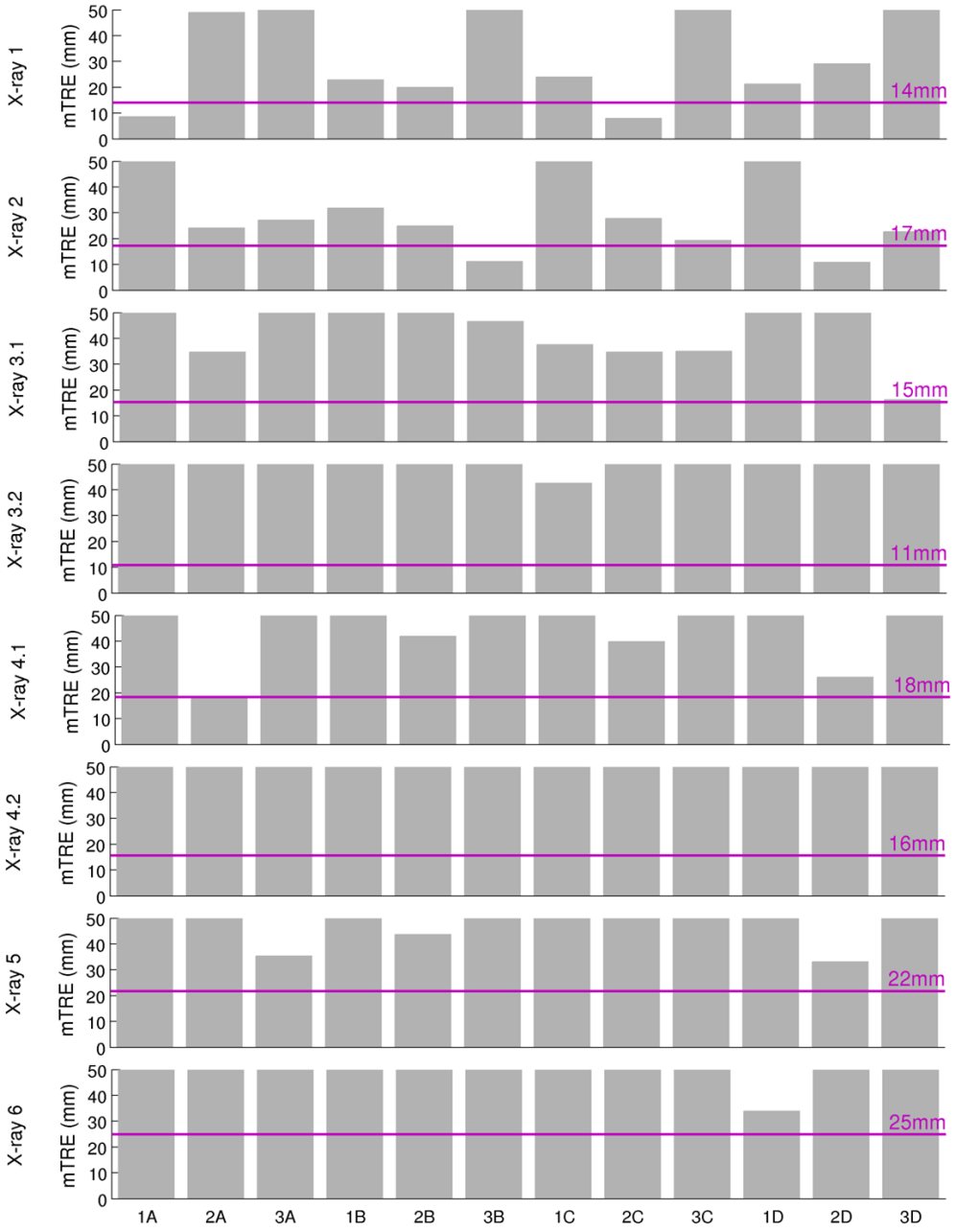


Figure 5.14 Coronary 3D/2D registration results for Experiment I with mTRE values from a manual registration. 8 X-rays of 6 patients were used for all preprocessing options, labelled with a number from {1, 2, 3} for the masks and a letter from {A, B, C, D} for the intensity values. Only mTRE below 50 mm are shown.

5.3.5 Discussion I

When we performed the registration-by-regression of 3D CTA with X-rays angiographies, the definition of the transformation parameters range came as an inherent limitation of the registration method proposed. The larger is the range, the higher will be the probability of the transformation parameters belonging to the range and, consequently, more general the learning will be. However, if the training set dimension and the model complexity is maintained, the training performance will be lower and the regression model less accurate, as it happened in our case (Figure 5.11). But improving the dimension and the complexity of the model was not feasible, because of computational constraints. Some authors overcame this limitation using a coarse-to-fine range strategy (Cyr et al., 2000) or using this approach as an initialization (Hermans et al., 2007).

The registration-by-regression of 3D CTA with X-rays angiographies using this features set was not successful, not even being as accurate as a manual initialization (Figure 5.14). This is mainly justified by the differences between the angiographies and DRRs computed from CTA, and also by the incapacity of the features set in use to deal with them. These differences are a problem reported in the literature for the iteratively approaches of 3D/2D registration using DRRs (Penney et al., 1998), where a good similarity measure should be able to achieve accurate registrations despite of these differences. Some differences are intrinsic to image formation, i.e. are caused by the inherent dissimilarity between CTA and X-ray images, and other are related to the imaged object, when for example interventional instruments are present.

As inherent differences, we can usually find overlaid structures and different contrast acquisition processes described in the literature (Imamura, Ida, & Sugimoto, 2002). In our case, the DRRs were rendered from a 3D CTA-derived coronary model. In the DRRs rendered from CTAs (and not from models), the coronary arteries are difficult to visualize because of the high intensity structures being over-projected on them (Metz, Schaap, Klein, et al., 2009). This procedure however just addresses one side of the problem, the DRR construction, and not the formation of the X-ray images.

Because of the differences between DRRs and X-ray images, image segmentation with different masks and different options for intensity values was considered. The automatic masks for some X-rays had serious problems. By their observation, it is obvious the confusion between the vessels and the catheters (which look like vessels) in some images by the 2D vesselness filter used, as well as a deficient background elimination. A more dedicated solution for the segmentation of catheters could be studied. Additionally, the 2D segmentation algorithm proposed by Schneider & Sundar (2010) could also be implemented, which showed to be more robust against the background artifacts than Frangi's vesselness. A

parameter optimization for the vesselness filter used could also be tried for each X-ray image, since we chose similar parameters for all X-ray cases.

Although the differences just mentioned between DRRs and X-rays are likely an important error source for the registration-by-regression method and the proposed solutions didn't solve them (Figure 5.13), the way the features rely on the intensities and on the vessels size are a crucial factor. If we restrict the analysis to the binary images (option D) and in particularly to the manual mask option (1D), we realize the results with real images were still not good. And for this specific case, the existing differences are the size of the masks and/or the presence/absence of some tiny vessels. Of course that in this case the learning itself is not so accurate because the information about the intensities is not contributing for the regression model, as we can see by the analysis of the results computed on the test set of simulated images (Figure 5.12).

In fact, this high sensitivity of the features in use on minor details between the DRRs and X-ray images is probably one of the main problems that justifies the bad results. Therefore, although these features allowed reaching an accurate regression model for simulated images, as it was studied in Gouveia et al. (2012a) and Chapter 3, their information is not adequate when the regression model is used for the registration of real images and a new set of features must be found.

5.4 Experiment II

In Experiment I, the registration-by-regression method with real images didn't perform well for the set of features used. In the current experiment, new sets of features were studied to find a robust solution to deal with the differences between the simulated images (used for the learning process to compute a regression model) and the real X-rays angiographies. We tested the registration-by-regression method with simulated and real images for the features sets studied, applying the same preprocessing options as in the previous experiment.

5.4.1 Input features II

Additionally to the features of the previous experiment, we have also considered in this experiment the geometric centre, or centre of mass, of the 2D image, computed as (Liao, 1993):

$$X_m = \frac{M_{10}}{M_{00}} \quad \text{and} \quad Y_m = \frac{M_{01}}{M_{00}} \quad . \qquad \text{Equation 5.2}$$

This is the point where all mass of the image could be concentrated without changing its first order moment, and can be used as a reference point to describe the position of the object imaged.

Thus, four features were combined as shown in Table 5.4. The first combination (F1) corresponds to the feature set used in *Experiment I*.

Table 5.4 Features and respective combinations for Experiment II.

	Moments	Eigenvectors	Eigenvalues	Geometric Centers	#features
F1	✓	✓	✓		18
F2		✓	✓		12
F3		✓			9
F4		✓		✓	11
F5				✓	2

5.4.2 Image preprocessing II

The image preprocessing applied in this experiment was the same as in the preceding one, described in Section 5.3.2. After a background subtraction, the vessels were segmented on the images. For that purpose, masks were determined by manual and automatic segmentation, labeled as options 1, 2 and 3, respectively; and then four possibilities for the intensity values of pixels that lie inside the mask were considered, named as A, B, C and D (Section 5.3.2.2).

5.4.3 Implementation details II

All the implementation details of the current experiment were similar to Experiment I but a remark must be done about the MLP topology (Table 5.5), particularly about the number of hidden units used. When this parameter was studied (Chapter 3), the optimum number of hidden units was set to the double of the number of input units. In this Experiment II, we studied different but yet similar features sets. Even if the number of features in each set is different and there are new features included, the features number and nature do not differ that much. For that reason, the same criterion was used to choose the number of hidden units in the current experiment except for the feature set labelled as F5 where we considered 10 hidden units since we expected that 4 hidden units would be too few.

Table 5.5 Topology of MLPs used in *Experiment II*.

	Units of input:hidden:output layers
F1	18:36:1
F2	12:24:1
F3	9:18:1
F4	11:22:1
F5	2:10:1

5.4.4 Results II

The results of the 3D/2D coronary registration-by-regression for the features sets F1, F2, F3, F4, and F5 are presented in this section. The $mTRE$ before and after registration is computed as defined in 5.2.3.2, for simulated images on the test set of 1000 unseen images and for real data on one X-ray per time-sequence selected from gold standard procedure (Section 5.2.3.1).

5.4.4.1 Simulated data

The registration results for the features sets studied (from F1 to F5) and for the different preprocessing options for the intensity values (A, B, C, D) are shown below, where each boxplot is based on all patients' results. Figure 5.15 presents the comparison between the preprocessing options for each feature set and Figure 5.16 presents the comparison between the features sets for each preprocessing option.

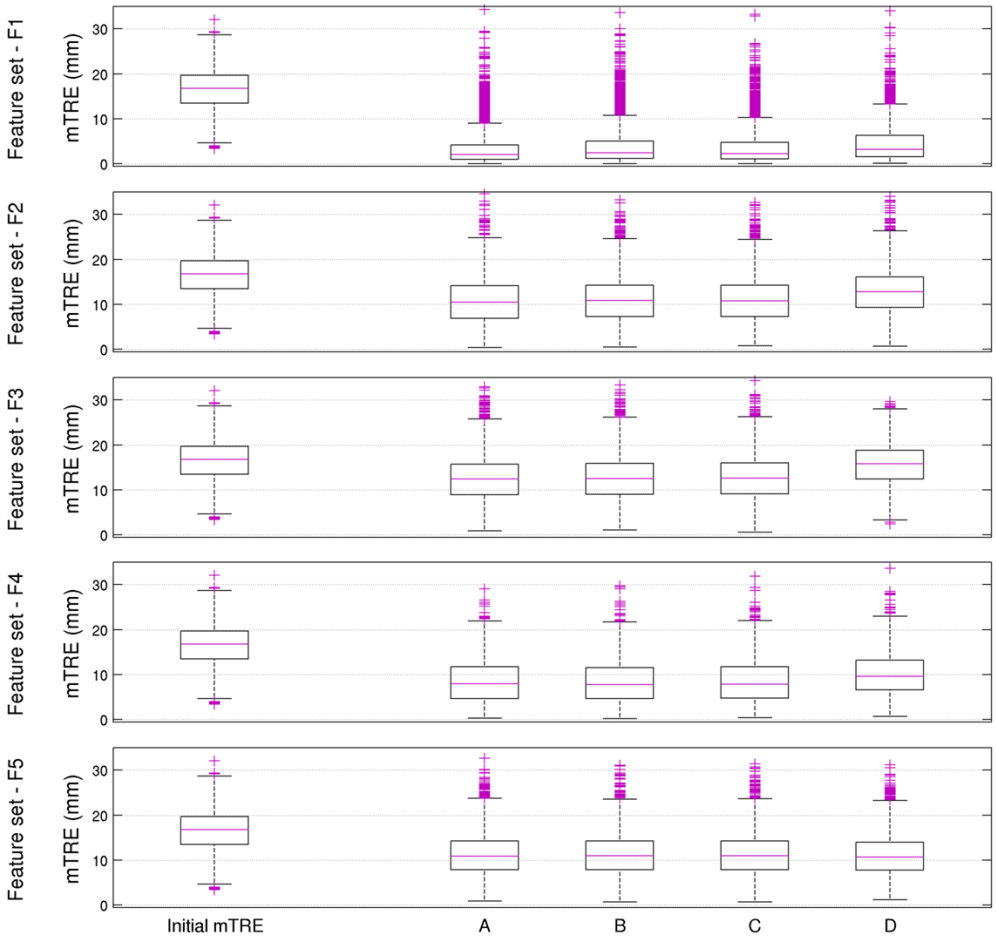


Figure 5.15 Comparison of the preprocessing options (A, B, C, D) applied to DRRs for each feature set (F1 to F5). All patients were considered and the mTRE values before the registration (*initial mTRE*) are also shown.

Unlike the feature set F1, analysed in detail in Experiment I, the other sets do not present a pronounced difference between the preprocessing option D and the remaining ones (Figure 5.15) due to the lack of the first three moments and their strong intensity dependence. For F5, which corresponds to the geometric center of the image, the differences are almost inexistent. The non-contribution of moments also explains the high mTRE values for options F2 to F5 for all preprocessing options, since the information hold by the moments revealed to be determinant for a high accuracy of the registration of simulated images as discussed in Section 3.3.5.

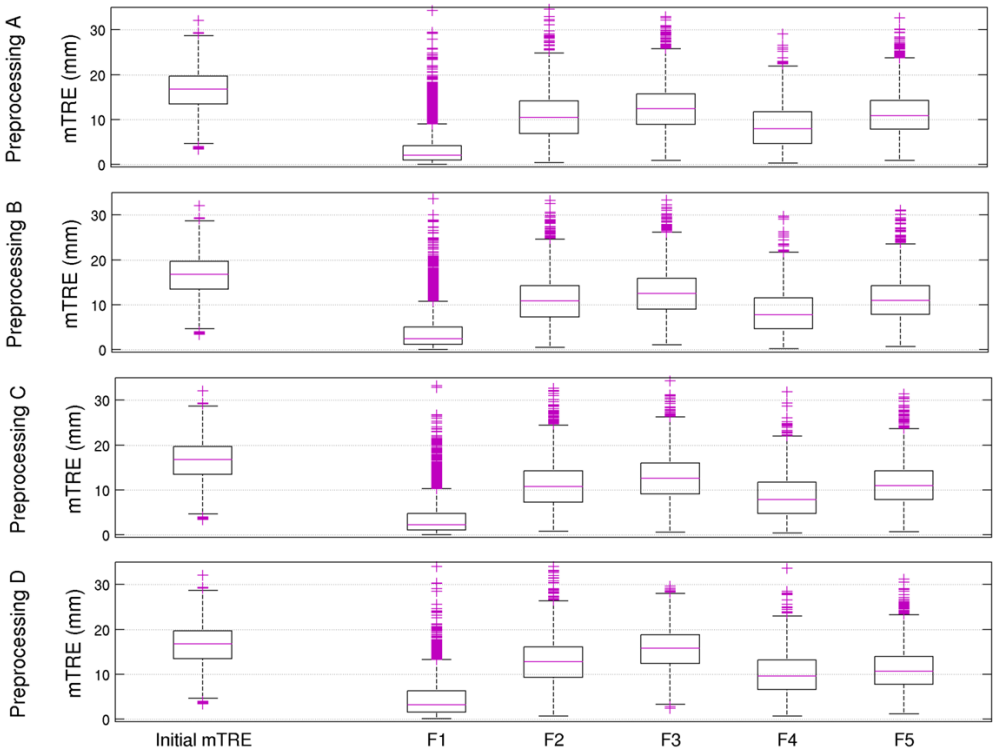


Figure 5.16 Comparison of the features sets (F1 to F5) for each preprocessing options (A, B, C, D) applied to DRRs. All patients were considered and the mTRE values before the registration (*initial mTRE*) are also shown.

Besides feature set F1, the set with best results for simulated images was F4 (Figure 5.16), i.e., geometric center and eigenvectors from the PCA on the pixels of the object of interest. When binary images were used, the results of F4 and F5 become very close.

5.4.4.2 Real data

The grey maps of Figure 5.17 gives an overall view of the 3D/2D coronary registration-by-regression with real data over the features sets in study. The 3D mTRE values up to 20 mm are presented for all preprocessing operations applied in X-ray angiographies. These were defined in Section 5.3.2 and regards to different masks (1, 2, 3) and intensity values (A, B, C, D).

From the observation of darker and lighter spots, we conclude the geometric center (feature set F5) obtained the highest number of values under 20 mm whereas the feature set of Experiment I, and the most accurate for simulated images, obtained the lowest frequency, just 7 in 96 values.

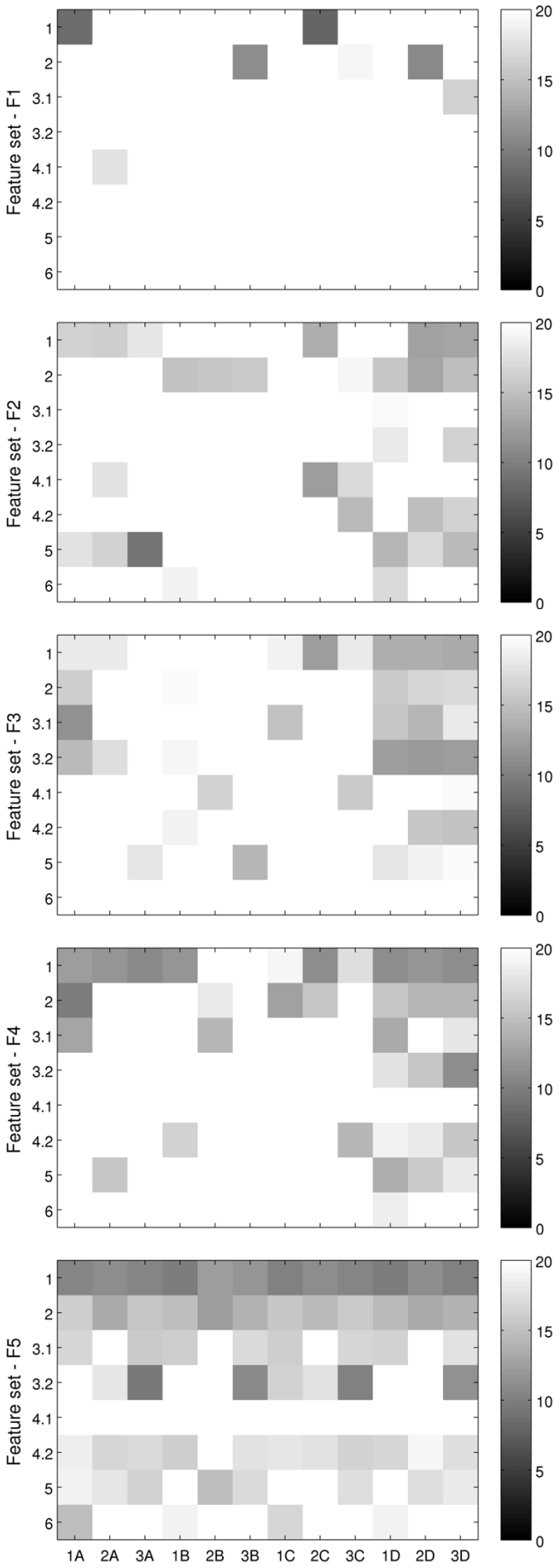


Figure 5.17 Grey maps with coronary 3D/2D registration results for each features sets (F1 to F5) compared in *Experiment II*. 8 X-rays of 6 patients were used for all preprocessing options, labelled with a number from {1, 2, 3} for the masks and a letter from {A, B, C, D} for the intensity values. Only mTRE below 20 mm are shown.

This global view allows us to qualify the robustness of features sets over the preprocessing options, more than evaluate the accuracy of the methods for real images. For feature sets F2, F3 and F4 there is clearly a higher frequency of results under 20 mm for option D, i.e. for binary images, not being evident if there is any difference between the masks. However, that difference does not exist for F5, where the results of option D seems to be similar to option A, B and C, and the main variety is observed across masks 1, 2 and 3. This robustness of geometric center with relation to intensity was to be expected because the intensity options for pixels values do not change the geometric position of the object in the overall image.

5.4.5 Discussion II

The performance of the registration-by-regression of real images improved with the new features sets proposed. However results are still not satisfactory.

We confirmed that the best solution for simulated images is not necessarily the best solution for real images. For the features sets studied, the set composed by moments, eigenvectors and eigenvalues was clearly the most accurate when only DRRs were involved, but this same set was the worst when the X-ray images were registered by the regression model DRRs-built. On the contrary, the best features for real data were the geometric centers, very robust to the intensity values, but not providing the information needed for a good learning to the model.

The balance between the quality of learning and the robustness to the difference between simulated and real images was not achieved. A different feature nature should be investigated to better understand the limitations of the problem in study.

5.5 Experiment III

In this experiment, for the CTA/X-rays registration-by-regression problem we used a different feature approach and we modified the imaged objects on the training dataset to provide different information to the regression model than before. The purpose was the registration of real images but the results for a test set of simulated images are also reported. We used binary images, either simulated or real, and performed a manual segmentation on X-rays. Restricting the problem to these intensity values and this vessel segmentation, we intended to control some of the differences between DRRs and X-rays.

So far, the features studied in this work introduced valuable information to the regression model, i.e. quantities with a geometric nature about translation and rotation. However, the registration results were not as accurate as we expected and a different direction in the feature search was followed in this experiment.

Moreover, modifications in the dataset were implemented attending the presence/absence of vessels in simulated and real images, i.e. some vessels clearly identified in the X-rays of a patient did not appear in the correspondent DRRs and vessels identified in DRRs did not appear in the correspondent X-rays. This difference is clearly identified by the visualization of the images, for example in Figure 5.18, and varies case by case.

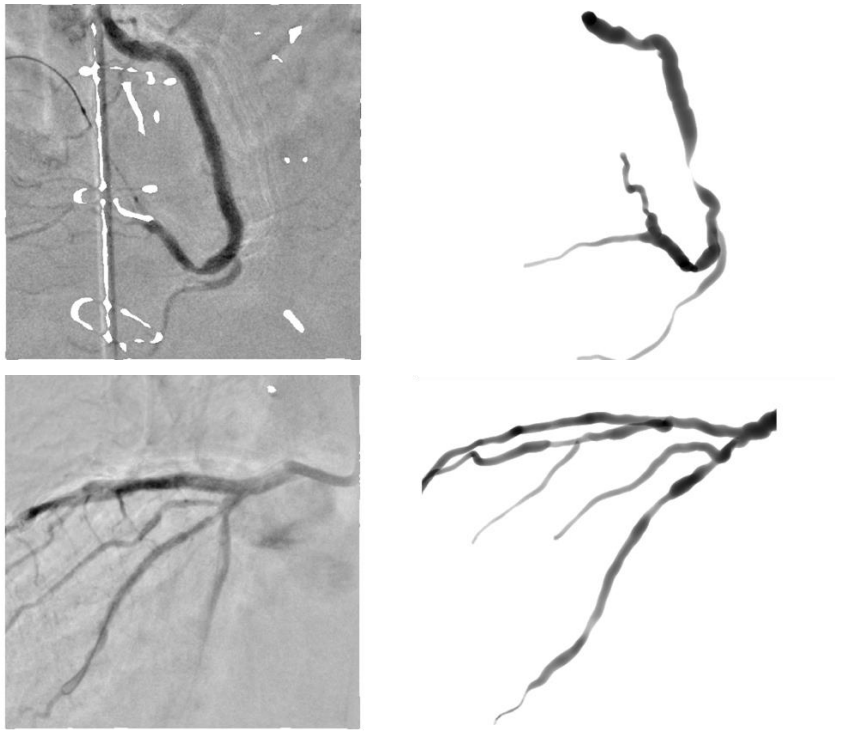


Figure 5.18 X-rays and correspondent DRR (of training set) with closer transformation parameters. For patient 1 (1st row), there are two tiny vessels on the left of the DRR image not present in X-ray image; and for patient 2 (2nd row), for example one vessel on the inferior right part of the X-ray image does not appear in DRR.

5.5.1 Input features III

In this experiment the features used to feed the NN were the projected images in the space of principal components (Shlens, 2005). Considering each image one observation and each image pixel one variable, we defined a MxN dataset with N observations and M variables each, where each image in a matrix format is firstly transformed into an 1xM array. A principal component analysis (PCA) was applied to this dataset and, after the determination of the rank-ordered variance, the meaningful principal components were selected ensuring a variance representation above 50% of the analysed dataset.

5.5.2 Image preprocessing III

In this experiment we used binary images, both from simulated and real images, i.e. images with intensity values 0 and 1. As we are only interested in the information contained in the coronary arteries for image registration, a procedure for vessel segmentation was necessary. For the X-ray angiographies, the masks were determined by manual segmentation from the analysis of two observers, and for the DRRs by a segmentation with manual threshold, a trivial process giving the images concerned. These preprocessing procedures correspond to the "1D" option for X-rays and "D" option for DRR of image preprocessing options in the previous experiment (Section 5.3.2).

5.5.3 Training set preprocessing

Two modifications were exclusively implemented on simulated images in order to introduce information into the regression model for the eventual absence of parts of the coronary arteries in real images, when compared to the simulated images. The first is the application of a binary erosion to the image to remove the tiniest vessels present, followed by a binary dilation operation to restore the vessel size (Gonzalez & Woods, 2008). The second is the elimination of a part of the image, with a random size and position. Additionally, a Gaussian filter was applied to the binary images creating a blurring effect and differentiating the borders from the inner regions.

Therefore, for each of the 6 patients in study, 9 training sets were considered with different combinations of preprocessing operations, schematized in Table 5.6. For an easier comparison, these different training sets will be labelled as *Fp* and numbered from 1 to 9, indicating the preprocessing operation (Table 5.6). Two different standard deviations for the Gaussian filter were set; the cases of unmodified and unfiltered images were also studied.

Table 5.6 Combinations used for preprocessing operations applied, resulting in 9 different training sets per patient.

	Gaussian filter		
Operations	-	$\sigma 1$	$\sigma 2$
-	Fp1	Fp2	Fp3
Erosion/dilation	Fp4	Fp5	Fp6
Elimination of vessel parts	Fp7	Fp8	Fp9

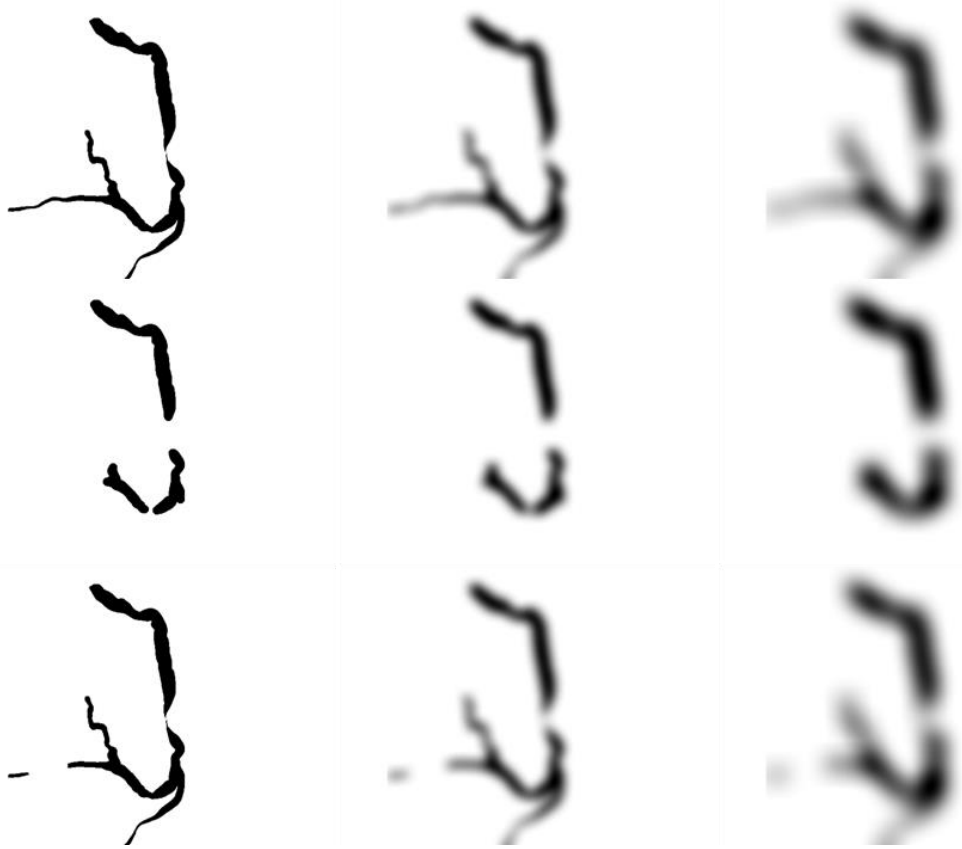


Figure 5.19 Resulting images from the training set preprocessing options defined for Experiment III for a DRR of patient 1. The order of the images presented corresponds to the order of options in Table 5.6.

5.5.4 Implementation details and parameters optimization III

Considering the preprocessing operations applied to the images referred in the previous section, some parameters were defined. The erosion and dilation operations were performed by using a disk-shape structuring element with a specific radius for each patient. These values were empirically determined and set as $\{8,6,0,5,5,8\}$ pixels for patients 1 to 6, respectively. On the other hand, the elimination of parts of the vessels was implemented by removing a circle with a random diameter up to 1/3 of the image size, randomly centred within the coronary arteries. About the filters 1 and 2 (Table 5.6), we used two low-pass Gaussian filters with different standard deviations, where σ was empirically defined as 10 and 20 pixels.

The images preprocessing and the implementation of NN were performed using MatLab, version 7.11.0.584 (R2010b) 64 bits (MatLab, 2010).

5.5.4.1 Input features computation

For each of the 9 training sets per patient defined in Section 5.5.3, the PCA-based features were derived:

(1) We decreased the number of pixels per image for practical computational reasons, by downsampling the images from 512x512 to 128x128 which resulted in images with 16384 pixels. We set the 11000 images on the MxN format, i.e. we transformed the 128x128 images in 1x16384 vectors and then we constructed a 16384x11000 dataset.

(2) Then, the dataset was split in three sets: 16384x7000, 16384x3000 and 16384x1000. The first set was used to train the regression model, the second set was used for validation purposes and the last one for testing the registration performance on simulated unseen images.

(3) Before the PCA computation, a z-score transformation is applied to each variable over all the observations of the training set. The PCA is computed on the z-score output and a set of principal components and the respective cumulative percentages obtained. The remaining data was centred and scaled using the z-score parameters found before, and then it was projected using the principal components.

5.5.4.2 Regression model implementation

Globally, we proceed similarly to the previous experiments regarding the implementation options of the Neural Networks. We considered 6 MLPs per registration, one per transformation parameter. Each one has one hidden layer with fully connected units and 1 output unit (the corresponding transformation parameter).

After the PCA computation and dataset projection, we had 16384 features per observation. Each one of these features corresponds to a principal component which represents a percentage of the variance of the original data. To decide how many principal components will be necessary for the regression model, and therefore what percentage of the total variance of original data will be considered in it, we used the cumulative percentage variances associated with the principal components. We searched for the best cumulative percentage in the range $\{50\%, 60\%, 70\%, 80\%, 90\%\}$, also leading to the number of input units. Simultaneously, we looked for the optimum number of hidden units considering the range $\{8,16,32\}$ as possibilities. This process was necessary for each of the 9 training sets studied per patient and for each of the 6 patients, leading to an exhaustive search. To limit computation time, the tuning was only performed for one of the transformation parameters, the translation parameter in x, and its results were extended to the other transformation parameters. The search ranges were defined after some preliminary tests.

For the tuning process we analysed the mean absolute difference between the model prediction and the known output, for translations in xx and computed on the validation set. This set was also used to select the stopping epoch for each MLP run occurred during and after the tuning with a maximum of 1000 and 3000 epochs, respectively.

5.5.5 Results III

5.5.5.1 Tuning parameter results

The cumulative percentages and the number of input and hidden units obtained by the search process are presented in Table 5.7 for every dataset of every patient.

Table 5.7 Cumulative percentages, number of input units and number of hidden units for all datasets from 1 to 9 and for all patients named from 1 to 6.

Cumulative percentage (%)									
	Datasets								
Patients	1	2	3	4	5	6	7	8	9
1	60	90	90	70	90	80	50	70	90
2	50	70	90	50	60	90	50	70	90
3	50	70	80	50	70	80	50	50	80
4	60	50	90	50	70	90	50	60	90
5	50	90	90	50	60	90	50	80	90
6	50	70	70	50	70	50	50	60	70
Input units									
	Datasets								
Patients	1	2	3	4	5	6	7	8	9
1	155	160	58	148	138	36	101	65	65
2	107	48	45	105	33	52	115	54	53
3	147	45	27	147	45	27	161	21	33
4	206	22	50	88	52	53	135	39	59
5	121	150	37	108	23	38	133	82	46
6	114	54	20	108	77	13	127	41	24
Hidden units									
	Datasets								
Patients	1	2	3	4	5	6	7	8	9
1	16	16	32	8	8	32	32	16	32
2	16	32	16	32	32	16	32	32	32
3	32	8	32	32	8	32	32	8	16
4	16	16	16	32	32	32	32	16	8
5	32	32	32	32	32	32	32	32	32
6	32	16	16	32	32	32	32	32	8

5.5.5.2 Comparison of training set preprocessing operations

The registration results, for simulated images and real images, with the different preprocessing options defined in this experiment are presented in Figure 5.20 and Figure 5.21 respectively.

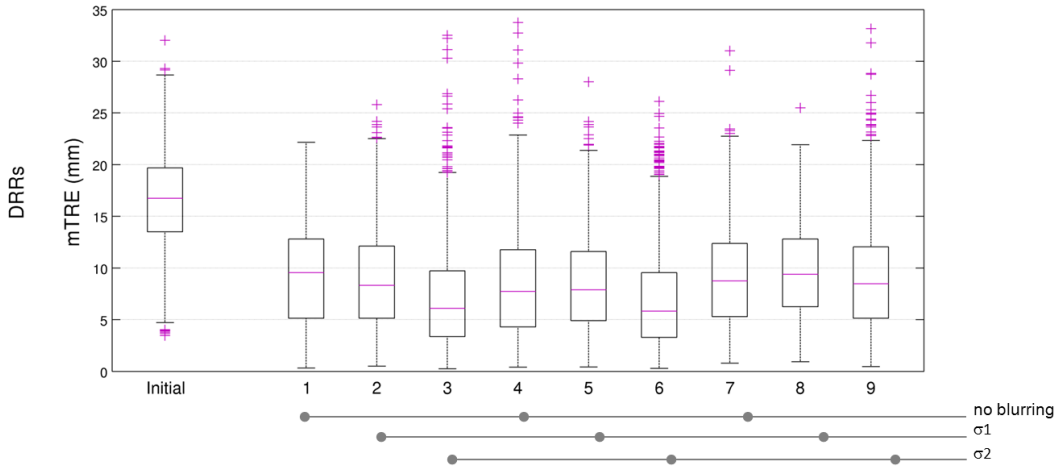


Figure 5.20 Registration results for simulated images for all patients comparing the different preprocessing operations labelled as 1 to 9, according to Table 5.6. The options (4,5,6) correspond to erosion/dilation operations and (7,8,9) to the elimination of parts of the vessels. The cases with similar blurring options are indicated in the image below by points on the lines.

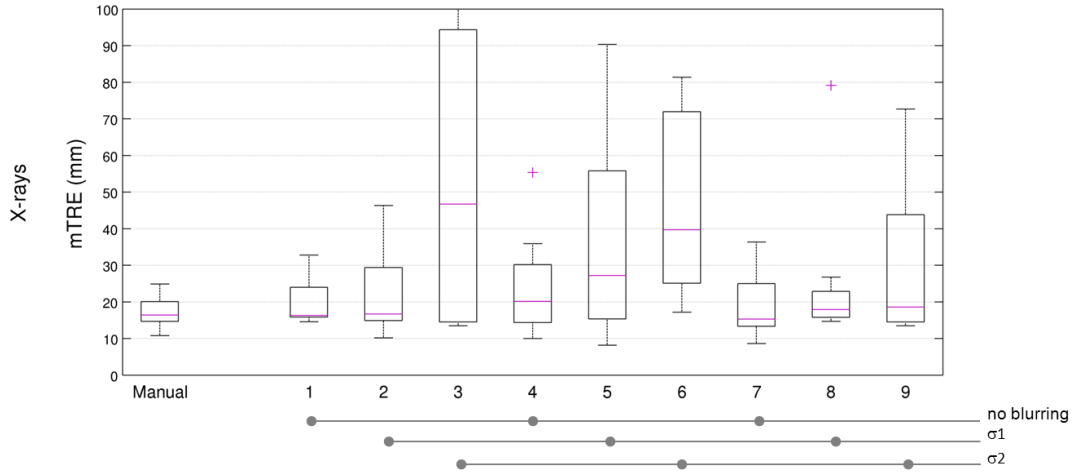


Figure 5.21 Registration results for all real images comparing the different preprocessing operations labelled as 1 to 9, according to Table 5.6. The options (4,5,6) correspond to erosion/dilation operations and (7,8,9) to the elimination of parts of the vessels. The cases with similar blurring options are indicated in the image below by points on the lines.

From the mTRE distributions for X-ray angiographies (Figure 5.21), it is evident the blurring is not helping the registration of real images, although for simulated images it does (Figure 5.20), i.e. as much blurred the DRRs are the better the registration performance.

5.5.5.3 Comparison of all experiments' feature sets for real data

Figure 5.22 summarizes the more relevant results of 3D/2D coronary registration-by-regression obtained in Chapter 5, comparing for the image preprocessing option 1D the results of all features considered in this chapter. The features introduced in Experiment III are

labelled in the table mentioned as *Fp*, followed by the number of the processing option (Table 5.6) defined in Section 5.5.2. The number of results at least as accurate as the manual registration improved considerably, and features like geometric center and eigenvectors from the PCA on the pixels of the object of interest (F4), projected images with elimination of parts of the vessels in DRRs (Fp7) and projected images with no modification of DRRs (Fp1) must be highlighted.

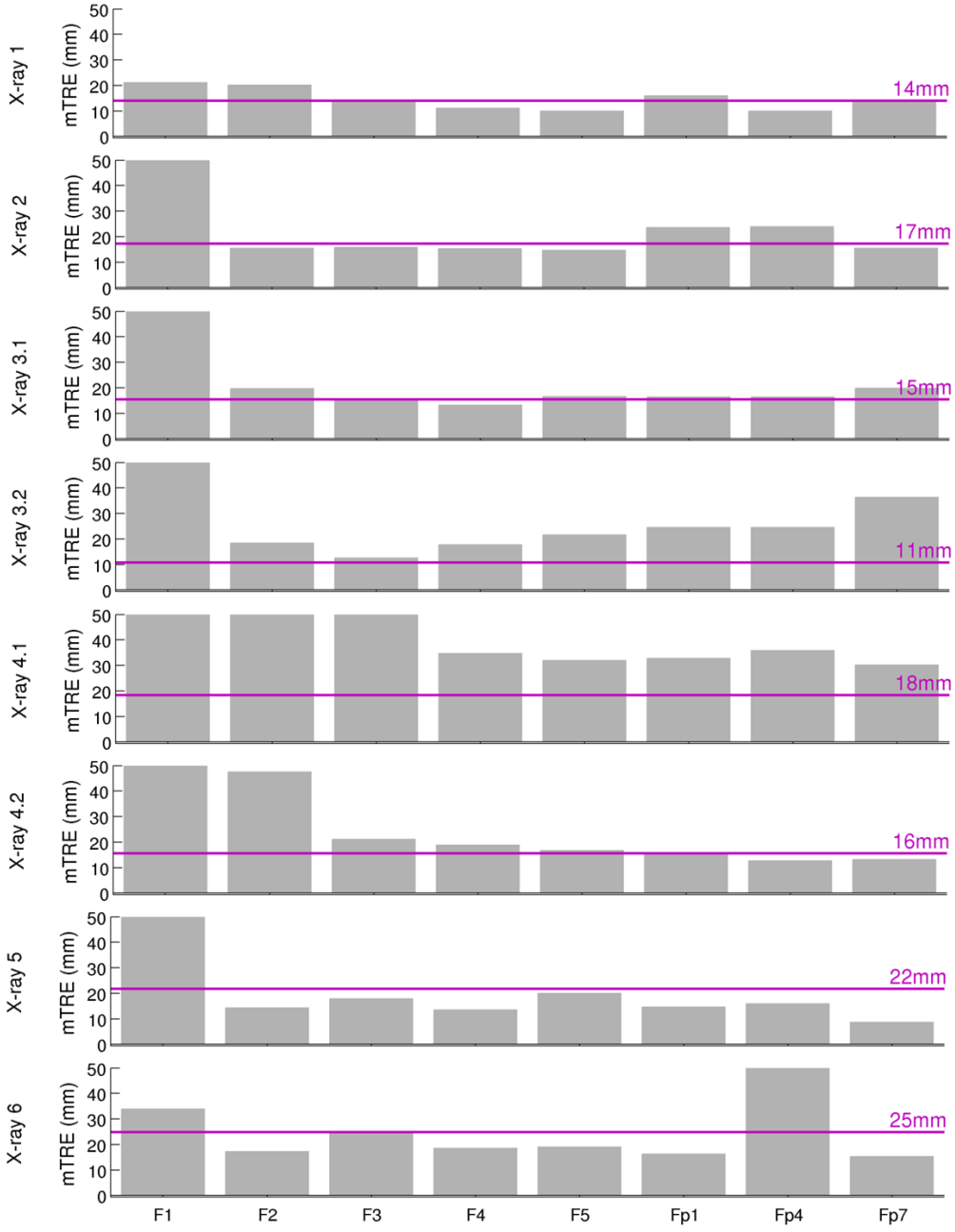


Figure 5.22 Summary of the main coronary 3D/2D registration results for the 8 X-rays of 6 patients with mTRE values from a manual registration. The features sets of Chapter 5 are compared for the preprocessing option 1D defined in *Image preprocessing I* (Section 5.3.2); F1 corresponds to Experiment I, (F2, F3, F4, F5) to Experiment II, and (Fp1, Fp4, Fp7) to the options 1, 4 and 7 of Experiment III, respectively.

5.5.6 Discussion III

In this experiment we studied a very different feature than before. Instead of using a geometric quantity, we used the intensity variability of a pixel in the image for different transformations. Of course, when we restricted the preprocessing options to the binary images, we are not using exactly the intensity variability but the variability of a pixel corresponding to the object imaged or to the background (0/1 intensity) for different transformations. And it would be of interest to see how this feature performs for the other intensity values options (A, B, and C). But to work just with binary images (with a manual mask) was the way to confine the problem to only a few differences between simulated and real images regarding to all differences discussed before.

We also tested the influence of a training set artificially modified aiming to simulate one of the differences between real and simulated images, i.e. the eventual absence of parts of the coronary arteries. The results indicated the learning performance with this extra information was as good as before, and the results for the registration of real images did not differ either.

However, even without some of the crucial differences between the X-rays and the DRRs, as the one related to the intensity values of the object imaged, the results were not promising. This regression model with a better feature set and considering the modification of the training set does not constitute a global solution for the 3D/2D registration of CTA with X-rays angiographies. But it could be an initialization solution to replace the need of manual initialization.

In fact, the 3D/2D iteratively registration approaches based on DRRs usually requires initialization close to the searched pose because of the numerous local maxima of intensity or gradient-based similarity measures (Markelj et al., 2012). The initialization influence is investigated in the literature for 3D/2D registration applications (e.g. Aouadi & Sarry, 2008), and in particularly for the case of CTA/X-ray angiographies (e.g. Metz et al., 2011). The initialization can be accomplished by methods of patient positioning, less accurate, used in radiation therapy, like Khamene, Bloch, Wein, Svatos, & Sauer (2006) and Mu, Fu, & Kuduvalli (2008). Moreover, registration methods based on templates (Banks & Hodge, 1996; Cyr et al., 2000; Hermans et al., 2007; Hoff et al., 1998), with some similarities with the registration-by-regression approach in study, present some accuracy limitations which make them more suitable for initialization purposes (Markelj et al., 2012).

The registration approach in study only used 6 transformation parameters – three rotations and three translations – which means we considered a rigid transformation. The presence of non-rigid deformations due to cardiac and respiratory motion is thus neglected. We took this into account when creating the gold standard, by selecting the X-ray time frame that most accurately matched the CTA-derived coronary artery model for further registration

experiments. This manual selection would have to be replaced by an automated procedure when implementing this method in clinical practice. Note that in principle the registration-by-regression method could be extended to a non-rigid registration. For instance, instead of computing the features on the entire images, we could divide the images and compute the features on small patches, i.e. on small sub-blocks of the images. This could be an interesting direction for future research, and could be of interest also for other image registration applications than the currently studied 3D/2D registration problem.

A final limitation of this experiment is related to the tuning of parameters. The topology of MLPs for all transformation parameters was obtained by a tuning process performed only for one of the transformation parameters. This procedure was motivated by time constraints since we tried for each different cumulative percentage variances (associated with the principal components) different number of hidden units. A better solution might be obtained by doing an exhaustive search for all parameters.

5.6 Conclusion

Performing the interpolation of real X-ray images with the regression model obtained by a set of DRRs was a real challenge. Many problems were faced, solutions investigated and evaluated: (1) a real image segmentation with a not trivial mask computation and a problematic determination of vessel intensities, especially regarding the differences between the values for a DRR and a real-image, (2) a search for a features set which lead to a good learning (attaining an accurate model) and simultaneously is robust enough to deal with the inherent differences between real and simulated images, (3) the representativeness of a simulated dataset used for the construction of regression model.

The validation of a registration method is essential, especially to its clinical relevance. This is the process which shows the application needs are met, regarding the accuracy and other performance criteria such as error tolerance, time and rate of failure (Hajnal et al., 2001). The accuracy assessment requires the knowledge of a gold standard or ground truth registration, which is difficult to achieve with clinical images (Hajnal et al., 2001). The construction of a gold standard to this nontrivial application was an important contribution of this work.

The work presented in this chapter enabled us to know better the problem of registration of real images. The intensity information introduced into the regression model by certain features allowed a very good learning and an accurate training performance, however the nature of this intensity-dependence is not adequate for the registration of the X-rays. On the contrary, some features had a robust behavior to the differences between simulated and real images, but they did not contain sufficient information for a good learning.

The registration-by-regression method for the automatic integration of coronary CTA with X-ray angiographies was not accurate with the tested feature sets. The results showed the inadequacy of the features used to solve this particular registration problem. The current work, rather than showing the registration approach is not able to achieve an accurate registration of images clinically relevant, pushes forward to other kind of feature solutions.

Considering the framework developed with this investigation, a different direction could be considered for the regression-by-registration of CTA with X-rays of the coronary arteries. Analysing the iterative approaches, especially the cases where real images are used and the reported results are good, a different kind of features can be tried where the intensity information is carefully included. An example could be a distance transform inspired by the work of Ruijters, ter Haar Romeny, & Suetens (2009), weighted by the image intensity. The inclusion of the intensity must be done carefully and a better vesselness filter must be attempted, as for example from Schneider & Sundar (2010). Another idea to explore could be the use of vessels' centerlines and vessels' skeletons (Elizabeth Bullitt et al., 1999).

Sísifo

Recomeça...
Se puderes,
Sem angústia e sem pressa.
E os passos que deres,
Nesse caminho duro
Do futuro,
Dá-os em liberdade.
Enquanto não alcances
Não descanses.
De nenhum fruto queiras só metade.

E, nunca saciado,
Vai colhendo
Ilusões sucessivas no pomar
E vendo
Acordado,
O logro da aventura.
És homem, não te esqueças!
Só é tua a loucura
Onde, com lucidez, te reconheças.

Miguel Torga

Chapter 6 Summary and Future Perspectives

This Thesis focused on the development and evaluation of a registration-by-regression approach for 3D/2D image registration of coronary vessel trees, where image registration is treated as a nonlinear regression problem, as an alternative for the traditional iterative approach.

Below we present a summary of the Thesis:

Chapter 2: A survey of the publications with a regression-based registration approach for medical applications is presented, as well as a summary of their main characteristics inspired by the classification proposed by Maintz & Viergever (1998). We adapted the criteria used by these authors, and we considered a) three aspects regarding the methodology of the registration-by-regression framework, being the type of *features*, the regression *model*, and the *role* of the regression in the entire process, and b) five more general aspects, begging the *application* of the work, the *object imaged,* the *transformation nature*, the *evaluation* of the registration and the motivation for this approach.

This survey aims to provide insight in the different formulations of registration-by-regression methods, to given an overview of the applications for which this has been proposed, and to compare to what extent the different approaches have been evaluated. A major conclusion here is that there are few registration-by-regression methods evaluated on real clinical data.

Chapter 3: We propose a 3D/2D registration method that treats image registration as a nonlinear regression problem instead of using an iterative traditional approach. The nonlinear regression function is determined in a supervised learning stage and relates image features of 2D projection images to the transformation parameters of the 3D image by nonlinear regression. The method is compared with a conventional registration method based on iterative optimization. For evaluation, simulated X-ray images (DRRs) were generated from coronary artery tree models derived from 3D CTA scans. Registration of nine vessel trees was performed, and the alignment quality was measured by the mean target registration error (mTRE). The regression approach has shown to be slightly less accurate, but much more robust than the method based on an iterative optimization approach.

Chapter 4: We perform a comparative evaluation of different regression techniques for 3D/2D registration-by-regression. In registration-by-regression, image registration is treated as a nonlinear regression problem that relates image features of 2D projection images to the transformation parameters of the 3D image. In this chapter, we evaluate seven regression methods: Multiple Linear and Polynomial Regression (LR and PR), k-Nearest Neighbour (k-NN), Multi-Layer Perceptron with conjugate gradient optimization (MLP-CG) and Levenberg-

Marquardt optimization (MLP-LM), Radial Basis Function networks (RBF) and Support Sector Regression (SVR). The experiments are performed with simulated X-ray images (DRRs) of nine vessel trees, so that we have the ground truth and can compute the mean target registration error (mTRE). MLP-LM and RBF obtained the highest accuracy but in general all methods were robust to large initial misalignment.

Chapter 5: We evaluate the integration of 3D preoperative CTA of the coronary arteries with intraoperative 2D X-ray angiography images using a registration-by-regression method. In this approach, image registration is estimated by a nonlinear regression model that is trained with DRRs. Image features of 2D projection images are related to the transformation parameters of the 3D image. The simulated 2D images were generated from coronary artery tree models derived from 3D CTA scans.

In this chapter we compared several image features: the 2D image moments and the geometric centers, the eigenvectors and eigenvalues from a PCA of the object imaged, and the projected images in the space of principal components. Moreover the influence of preprocessing options on simulated images used for the training set, as well as on real data, was studied. For the registration evaluation, a gold standard was developed from 9 X-ray angiography sequences from 7 different patients. The alignment quality was measured by the mean target registration error (mTRE).

The coronary registration–by-regression results were not satisfactory, resembling manual initialization accuracy. The framework performed showed the inadequacy of the features used to solve the registration problem. It enabled us to better understand the dependency of the proposed method on the differences between simulated and real images. Alternative solutions are suggested.

In the introduction of this work, we formulated three research questions:

- How is the registration influenced by the image features used in the learning process?
- Are Neural Networks an adequate regression model for this problem?
- Is the registration-by-regression method a clinically relevant solution for 3D/2D registration of coronary CTA and X-ray angiography during image-guided interventions?

The question about the Neural Networks is answered by Chapter 4 with the comparison of regression models for this particular registration problem.

The first research question was addressed in Chapter 3 and Chapter 5. In Chapter 3 we realized that the intensity information associated with certain features allowed a very good learning and an accurate training performance. However chapter 5 showed that this intensity-dependence was not adequate for the registration of real X-rays. On the other hand, some features were not affected by the differences between simulated images (used during training) and real images, but they did not contain sufficient information for a good learning, and therefore did not achieve a very high accuracy.

Chapter 5 addressed the final question about the clinical application of the registration-by-regression method. For this concrete problem, i.e. for the automatic integration of coronary CTA with X-ray angiographies, the proposed method in its current configuration is not sufficiently accurate to be used in practice. The main difficulty lies in the substantial differences in appearance between the images used for training (simulated X-rays from3D coronary models) and the actual images obtained during the intervention (real X-ray angiography). This indicates that other types of image features need to be developed, more robust to the typical differences in appearance, such as the presence of catheters in X-ray, missing data in X-ray (i.e., vessels not visible on X-ray due to coronary chronic total occlusions), or incomplete 3D coronary models caused by insufficient visibility in (noisy, motion-blurred) CTA data. The recent works of Chou, Frederick, Mageras, & Chang (2013) and Chou & Pizer (2013) also suggest some interesting strategies to improve accuracy, like multi-scale regressions by a "hierarchical training" (i.e. large to small scales of transformation ranges for training), and correction of intensity differences between simulated and real images by histogram matching.

Finally, it is recommended to test the registration-by-regression framework in other applications, where training data is available that is representative for the test data. Potentially interesting examples of such applications are 3D/2D registration for image-guided needle biopsy interventions in the spine (van de Kraats, 2005; van de Kraats et al., 2006) and motion estimation from real-time image streams (Luca et al., 2012)

References

Abche, A. B., Yaacoub, F., Maalouf, A., & Karam, E. (2006). Image registration based on neural network and Fourier transform. *Conference Proceedings: 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, *1*, 4803–6. doi:10.1109/IEMBS.2006.260342

Aouadi, S., & Sarry, L. (2008). Accurate and precise 2D–3D registration based on X-ray intensity. *Computer Vision and Image Understanding*, *110*(1), 134–151. doi:10.1016/j.cviu.2007.05.006

Banks, S. a, & Hodge, W. a. (1996). Accurate measurement of three-dimensional knee replacement kinematics using single-plane fluoroscopy. *IEEE Transactions on Bio-Medical Engineering*, *43*(6), 638–49. doi:10.1109/10.495283

Basak, D., Pal, S., & Patranabis, D. C. (2007). Support Vector Regression. *Neural Information Processing – Letters and Reviews*, *11*(10), 207–224.

Beale, M. H., Hagan, M. T., & Demuth, H. B. (2012). Neural Network Toolbox: User's Guide R2012a. *The Mathworks, Inc.* Retrieved April 02, 2012, from http://www.mathworks.com/help/pdf_doc/nnet/nnet_ug.pdf

Bi, J., & Bennett, K. P. (2003). Regression error characteristic curves. In *Twentieth International Conference on Machine Learning (ICML-2003)*. Washington, DC.

Brown, L. (1992). A survey of image registration techniques. *ACM Computing Surveys (CSUR)*, 1–60.

Bullitt, E., Liu, A., Aylward, S. R., Coffey, C., Stone, J., Mukherji, S. K., Muller, Keith E., Pizer, S. M. (1999). Registration of 3D Cerebral Vessels with 2D Digital Angiograms. *Academic Radiology*, *6*(9), 539–546.

Chang, C., & Lin, C. (2011). LIBSVM: A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology*, *2*(3), 1–39. doi:10.1145/1961189.1961199

Chen, S., Cowan, C. F. N., & Grant, P. M. (1991). Orthogonal least squares learning algorithm for radial basis function networks. *Neural Networks, IEEE Transactions on*, *2*(2), 302–309.

Cherkassky, V., & Ma, Y. (2004). Practical selection of SVM parameters and noise estimation for SVM regression. *Neural Networks: The Official Journal of the International Neural Network Society*, *17*(1), 113–26. doi:10.1016/S0893-6080(03)00169-2

Chou, C.-R., Frederick, B., & Liu, X. (2011). Claret: A fast deformable registration method applied to lung radiation therapy. In *Fourth International (MICCAI) Workshop on Pulmonary Image Analysis* (pp. 113–124).

Chou C-R, Frederick B, Mageras G, Chang S, Pizer S. (2013). 2D/3D image registration using regression learning. Comput Vis Image Underst; 117:1095–1106. doi:10.1016/j.cviu.2013.02.009

Chou, C.-R., Frederick, C., Chang, S., & Pizer, S. (2010). A learning-based patient repositioning method from limited-angle projections. In J. Angeles & B. Boulet (Eds.), *International Symposium on the Occasion of the 25th Anniversary of McGill University Centre for Intelligent Machines*.

Chou, C.-R., & Pizer, S. (2013). Real-time 2D/3D deformable registration using metric learning. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 7766 LNCS, pp. 1–10). doi:10.1007/978-3-642-36620-8_1

Cootes, T., Edwards, G. J., & Taylor, C. J. (2001). Active appearance models. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *23*(6), 681–685.

Cortez, P. (2008). Data Mining with MLPs. In P. S. of E. of P. (ISEP) & P. U. (FEUP) Faculty of Engineering (Eds.), *Summer School NN2008 - Neutral Networks in Classification, Regression and Data Mining*. Porto.

Cortez, P., Portelinha, M., Rodrigues, S., Cadavez, V., & Teixeira, A. (2006). Lamb Meat Quality Assessment by Support Vector Machines. *Neural Processing Letters*, *24*(1), 41–51. doi:10.1007/s11063-006-9009-6

Cyr, C. M., Kamal, A. F., Sebastian, T. B., & Kimia, B. B. (2000). 2D-3D registration based on shape matching. In *Proceedings IEEE Workshop on Mathematical Methods in Biomedical Image Analysis. MMBIA-2000 (Cat. No.PR00737)* (pp. 198–203). IEEE Comput. Soc. doi:10.1109/MMBIA.2000.852378

Edwards, P. J., Hawkes, D. J., Penney, G. P., Clarkson, M. J., & London, C. (2001). Guiding Therapeutic Procedures. In J. V. Hajnal, D. J. Hawkes, & D. L. G. Hill (Eds.), *Medical image registration* (pp. 253–270). CRC Press.

Elhanany, I., Sheinfeld, M., Beck, A., Kadmon, Y., Tal, N., & Tirosh, D. (2000). Robust image registration based on feedforward neural networks. In *SMC 2000 Conference Proceedings. 2000 IEEE International Conference on Systems, Man and Cybernetics. "Cybernetics Evolving to Systems, Humans, Organizations, and their Complex Interactions" (Cat. No.00CH37166)* (Vol. 2, pp. 1507–1511). IEEE. doi:10.1109/ICSMC.2000.886068

Elsen, P. Van den, Pol, E.-J. D., & Viergever, M. A. (1993). Medical image matching-a review with classification. *Engineering in Medicine and Biology*, (March), 26–39.

Fitzpatrick, J. M., & West, J. B. (2001). The distribution of target registration error in rigid-body point-based registration. *IEEE Transactions on Medical Imaging*, *20*(9), 917–27. doi:10.1109/42.952729

Flexer, A. (1996). Statistical evaluation of neural networks experiments: Minimum requirements and current practice. In *Proceedings of the 13th European Meeting on Cybernetics and Systems Research* (Vol. 2, pp. 1005–1008).

Frangi, A. F., Niessen, W. J., Vincken, K. L., & Viergever, M. A. (1998). Multiscale vessel enhancement filtering. In W. M. Wells, A. Colchester, & S. L. Delp (Eds.), *Medical Image Computing and Computer-Assisted Intervention - MICCAI'98* (Vol. 1496, pp. 130–137). Lecture Notes in Computer Science.

Freire, L. C., Gouveia, A. R., & Godinho, F. M. (2010a). A Neural Network-Based Method for Affine 3D Registration of FMRI Time Series Using Fourier Space Subsets. In K. Diamantaras, W. Duch, & L. S. Iliadis (Eds.), *Artificial Neural Networks – ICANN 2010* (Vol. 6352, pp. 22–31). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-15819-3

Freire, L. C., Gouveia, A. R., & Godinho, F. M. (2010b). FMRI 3D registration based on Fourier space subsets using neural networks. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society.* (pp. 5624–7). doi:10.1109/IEMBS.2010.5628038

Glasbey, C. A., & Mardia, K. V. (1998). A review of image warping methods. *Journal of Applied Statistics*, *25*(2), 155–171.

Gonzalez, R. C., & Woods, R. E. (2008). *Digital Image Processing* (3rd ed.). Upper Saddle River, NJ, USA: Prentice Hall.

Gouveia, A. R., Metz, C., Freire, L., & Klein, S. (2012a). 3D-2D image registration by nonlinear regression. In *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)* (pp. 1343–1346). IEEE. doi:10.1109/ISBI.2012.6235814

Gouveia, A. R., Metz, C., Freire, L., & Klein, S. (2012b). Comparative Evaluation of Regression Methods for 3D-2D Image Registration. In A. E. Villa, W. Duch, P. Érdi, F. Masulli, & G. Palm (Eds.), *Artificial Neural Networks and Machine Learning – ICANN 2012* (Vol. 7553, pp. 238–245). Springer Berlin Heidelberg. doi:10.1007/978-3-642-33266-1_30

Hagan, M. (1994). Training feedforward networks with the Marquardt algorithm. *Networks, IEEE Transactions on*, *5*(6), 2–6.

Hair, J. F., Black, B., Babin, B., Anderson, R. E., & Tatham, R. L. (2005). *Multivariate Data Analysis* (6th ed., p. 928). Upper Saddle River, NJ, USA: Prentice Hall.

Hajnal, J. V., Hawkes, D. J., & Hill, D. L. G. (2001). *Medical image registration, Biomedical engineering series*. (J. V. Hajnal, D. J. Hawkes, & D. L. G. Hill, Eds.). CRC Press LLC.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: data mining, inference and prediction* (2nd. ed.). Springer.

Haykin, S. (1999). *Neural networks: a comprehensive foundation* (2nd. ed.). Delhi: Prentice hall.

Heneghan, C., & Maguire, P. (2002). Registration of ophthalmic images using control points. In *OPTP-IRELAND 2002: OPTICAL METROLOGY, IMAGING, AND MACHINE VISION; Proceedings of SPIE - The International Society for Optical Engineering* (Vol. 4877, pp. 238–247). Department of Electronic Engineering, University College Dublin, Belfield, Dublin 4, Ireland.

Heneghan, C., Maguire, P., Ryan, N., & de Chazal, P. (2002). Retinal image registration using control points. In *Proceedings IEEE International Symposium on Biomedical Imaging* (pp. 349–352). 345 E 47TH ST, NEW YORK, NY 10017 USA: IEEE. doi:10.1109/ISBI.2002.1029265

Hermans, J., Claes, P., Bellemans, J., Vandermeulen, D., & Suetens, P. (2007). Robust initialization for 2D/3D registration of knee implant models to single-plane fluoroscopy. In *Proceedings of SPIE* (Vol. 6512, pp. 651208–651208–12). SPIE. doi:10.1117/12.708163

Hill, D. L. G., Batchelor, P. G., Holden, M., & Hawkes, D. J. (2001). Medical image registration. *Physics in Medicine and Biology*, *46*(3), R1–R45. doi:10.1088/0031-9155/46/3/201

Hoff, W. A., Komistek, R. D., Dennis, D. A., Gabriel, S. M., & Walker, S. A. (1998). Three-dimensional determination of femoral-tibial contact positions under in vivo conditions using fluoroscopy. *Clinical Biomechanics*, *13*, 455–472.

Hsu, C., Chang, C., & Lin, C. (2010). A Practical Guide to Support Vector Classification. *Bioinformatics*, *1*(1), 1–16.

Imamura, H., Ida, N., & Sugimoto, N. (2002). Registration of preoperative CTA and intraoperative fluoroscopic images for assisting aortic stent grafting. In T. Dohi & R. Kikinis (Eds.), *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2002* (Vol. 2489, pp. 477–484). Lecture Notes in Computer Science. doi:10.1007/3-540-45787-9_60

Kasper, D. L., Braunwald, E., Fauci, A. S., Hauser, S. L., Longo, D. L., & Jameson, J. L. (2005). *Harrison's Principles of Internal Medicine* (2nd. ed.). New York: McGraw-Hill.

Khamene, A., Bloch, P., Wein, W., Svatos, M., & Sauer, F. (2006). Automatic registration of portal images and volumetric CT for patient positioning in radiation therapy. *Medical Image Analysis*, *10*(1), 96–112. doi:10.1016/j.media.2005.06.002

Kim, M., Wu, G., Yap, P.-T., & Shen, D. (2010). A generalized learning based framework for fast brain image registration. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 6362 LNCS, pp. 306–314). Department of Radiology, BRIC, University of North Carolina, Chapel Hill, NC, United States. doi:10.1007/978-3-642-15745-5_38

Kim, M., Wu, G., Yap, P.-T., & Shen, D. (2012). A general fast registration framework by learning deformation-appearance correlation. *IEEE Transactions on Image Processing*, *21*(4), 1823–1833.

Li, Q., Gao, C., Wu, X., & Li, D. (2010). A method for range image registration via neural network and ICP algorithm. *Wuhan University Journal of Natural Sciences*, *15*(5), 398–402. doi:10.1007/s11859-010-0673-z

Liang, B., & Chung, R. (2008, July 2). Viewpoint determination of image by interpolation over sparse samples. *Image and Vision Computing*. doi:10.1016/j.imavis.2007.10.008

Liao, S. X. (1993, March). *On image analysis by moments*. University of Manitoba, Canada.

Liu, A., Bullitt, E., & Pizer, S. M. (1998). 3D / 2D Registration Via Skeletal Near Projective Invariance in Tubular Objects. In *First International Conference on Medical Image*

Computing and Computer-Assisted Intervention—MICCAI'98 (pp. 952–963). Cambridge MA, USA: Springer.

Liu, H., Yan, J., & Zhang, D. (2006). Three-dimensional surface registration: A neural network strategy. *Neurocomputing*, *70*(1-3), 597–602. doi:10.1016/j.neucom.2006.04.004

Luca, V. De, Tanner, C., & Székely, G. (2012). Speeding-up image registration for repetitive motion scenarios. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on* (pp. 1355–1358).

Maintz, J. B. A., & Viergever, M. A. (1998). A survey of medical image registration. *Medical Image Analysis*, *2*(1), 1–36. doi:10.1016/S1361-8415(01)80026-8

Markelj, P., Tomaževič, D., Likar, B., & Pernuš, F. (2012). A review of 3D/2D registration methods for image-guided interventions. *Medical Image Analysis*, *16*(3), 642–61. doi:10.1016/j.media.2010.03.005

MatLab. (2010). Natick, MA, USA: The Mathworks Inc.

Maurer, C. R., & Fitzpatrick, J. M. (1993). A review of medical image registration. In R. J. Maciunas (Ed.), *Interactive image-guided neurosurgery* (pp. 17–44). Park Ridge, IL, US: American Association of Neurological Surgeons.

Meier, B., Bachmann, D., & Lüscher, T. (2003). 25 Years of Coronary Angioplasty: Almost a Fairy Tale. *Lancet*, *361*(9356), 527.

Metz, C. T. (2011). *Coronary Motion Modeling for CTA to X-Ray Angiography Registration*.

Metz, C. T., Schaap, M., Klein, S., Rijnbeek, P., Neefjes, L., Mollet, N., … Walsum, T. Van. (2011). Alignment of 4D Coronary CTA with Monoplane X-Ray Angiography. *6th International Workshop, AE-CAI (Augmented Environments for Computer-Assisted Interventions), Held in Conjunction with MICCAI 2011*, 106–116.

Metz, C. T., Schaap, M., Klein, S., Weustink, A. C., Mollet, N. R., Schultz, C., … Niessen, W. J. (2009). GPU accelerated alignment of 3-D CTA with 2-D X-ray data for improved guidance in coronary interventions. In *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (pp. 959–962). IEEE. doi:10.1109/ISBI.2009.5193213

Metz, C. T., Schaap, M., Weustink, a. C., Mollet, N. R., van Walsum, T., & Niessen, W. J. (2009). Coronary centerline extraction from CT coronary angiography images using a minimum cost path approach. *Medical Physics*, *36*(12), 5568. doi:10.1118/1.3254077

Molller, M. F. (1993). A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning. *Moller, Neural Networks*, *6*, 525–533.

Mu, Z., Fu, D., & Kuduvalli, G. (2008). A probabilistic framework based on hidden markov model for fiducial identification in image-guided radiation treatments. *IEEE Transactions on Medical Imaging*, *27*(9), 1288–300. doi:10.1109/TMI.2008.922693

Oliveira, F. P. M., & Tavares, J. M. R. S. (2014). Medical image registration: a review. *Computer Methods in Biomechanics and Biomedical Engineering*, *17*(2), 73–93. doi:10.1080/10255842.2012.670855

Penney, G. P., Weese, J., Little, J. a, Desmedt, P., Hill, D. L., & Hawkes, D. J. (1998). A comparison of similarity measures for use in 2-D-3-D medical image registration. *IEEE Transactions on Medical Imaging*, *17*(4), 586–95. doi:10.1109/42.730403

Pina, A. C., & Zaverucha, G. (2008). Applying REC analysis to ensembles of particle filters. *Neural Computing and Applications*, *18*(1), 25–35. doi:10.1007/s00521-008-0199-x

Qi, W., Gu, L., & Xu, J. (2008). Non-rigid 2D-3D Registration Based on Support Vector Regression Estimated Similarity Metric. In T. Dohi, I. Sakuma, & H. Liao (Eds.), *Medical Imaging and Augmented Reality* (pp. 339–348). Tokyo: Springer.

Qi, W., Gu, L., & Zhao, Q. (2008). Effective 2D-3D medical image registration using Support Vector Machine. In *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 5386–9). Vancouver, British Columbia, Canada. doi:10.1109/IEMBS.2008.4650432

Ramirez, L., Durdle, N., & Raso, V. (2003). Medical image registration in computational intelligence framework: a review. In *Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on* (Vol. 2, pp. 1021–1024). IEEE. doi:10.1109/CCECE.2003.1226069

Rivest-Hénault, D., Sundar, H., & Cheriet, M. (2012). Nonrigid 2D/3D registration of coronary artery models with live fluoroscopy for guidance of cardiac interventions. *IEEE Transactions on Medical Imaging*, *31*(8), 1557–72. doi:10.1109/TMI.2012.2195009

Roche, A., Malandain, G., Pennec, X., & Ayache, N. (1998). The correlation ratio as a new similarity measure for multimodal image registration. *Medical Image Computing …*, *1496*.

Roche, A., Pennec, X., Malandain, G., & Ayache, N. (2001). Rigid registration of 3-D ultrasound with MR images: a new approach combining intensity and gradient

information. *IEEE Transactions on Medical Imaging*, *20*(10), 1038–49. doi:10.1109/42.959301

Ruijters, D., ter Haar Romeny, B. M., & Suetens, P. (2009). Vesselness-based 2D-3D registration of the coronary arteries. *International Journal of Computer Assisted Radiology and Surgery*, *4*(4), 391–7. doi:10.1007/s11548-009-0316-z

Ryan, N., Heneghan, C., & de Chazal, P. (2004). Registration of digital retinal images using landmark correspondence by expectation maximization. *Image and Vision Computing*, *22*(11), 883–898. doi:10.1016/j.imavis.2004.04.004

Salvi, J., Matabosch, C., Fofi, D., & Forest, J. (2007). A review of recent range image registration methods with accuracy evaluation. *Image and Vision Computing*, *25*(5), 578–596. doi:10.1016/j.imavis.2006.05.012

Sarle, W. (2005). Neural network frequently asked questions ftp://ftp.sas.com/pub/neural/FAQ.html. 2005.

Schneider, M., & Sundar, H. (2010). Automatic global vessel segmentation and catheter removal using local geometry information and vector field integration. In *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (pp. 45–48). IEEE. doi:10.1109/ISBI.2010.5490419

Shi, Y., Wu, G., Song, Z., & Shen, D. (2012). Dense deformation reconstruction via sparse coding. *Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. Digital Medical Research Center, Shanghai Key Lab. of MICCAI, Fudan University, Shanghai, 200032, China.

Shlens, J. (2005). A Tutorial on Principal Component Analysis. *Measurement*, *51*(10003), 52. doi:10.1.1.115.3503

Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, *14*(3), 199–222. doi:10.1023/B:STCO.0000035301.49549.88

Sotiras, A., Davatzikos, C., & Paragios, N. (2013). Deformable medical image registration: a survey. *IEEE Transactions on Medical Imaging*, *32*(7), 1153–90. doi:10.1109/TMI.2013.2265603

Turgeon, G.-A., Lehmann, G., Guiraudon, G., Drangova, M., Holdsworth, D., & Peters, T. (2005). 2D-3D registration of coronary angiograms for cardiac procedure planning and guidance. *Medical Physics*, *32*(12), 3737. doi:10.1118/1.2123350

Van de Kraats, E. B. (2005). *3D Rotational X-Ray Guidance for Surgical Interventions*.

Van De Kraats, E. B., Penney, G. P., Tomaževič, D., & Van Walsum, T. (2005). Standardized Evaluation Methodology for 2-D – 3-D Registration. *IEEE Transactions on Medical Imaging*, *24*(9), 1177–1189.

Van de Kraats, E. B., van Walsum, T., Verlaan, J.-J., Voormolen, M. H. J., Mali, W. P. T. M., & Niessen, W. J. (2006). Three-Dimensional Rotational X-Ray Navigation for Needle Guidance in Percutaneous Vertebroplasty: An Accuracy Study. *Spine*, *31*(12), 1359–1364. doi:10.1097/01.brs.0000218580.54036.1b

Vapnik, V. N. (1995). *The nature of statistical learning theory*. New York, NY, USA: Springer-Verlag New York, Inc.

Wachowiak, M. P., Smolíková, R., Zurada, J. M., & Elmaghraby, A. S. (2002). A supervised learning approach to landmark-based elastic biomedical image registration and interpolation. In *Neural Networks, 2002. IJCNN'02. Proceedings of the 2002 International Joint Conference on* (Vol. 2, pp. 1625–1630). IEEE.

Wang, W., Xu, Z., Lu, W., & Zhang, X. (2003). Determination of the spread parameter in the Gaussian kernel for classification and regression. *Neurocomputing*, *55*, 643–663.

Wang, W.-F., & et al. (2003). PET-MRI image registration and fusion using artificial neural networks. *Biomedical Engineering: Applications, Basis and Communications*, *15*(3), 95–99.

Wyawahare, M. V, Patil, P. M., & Abhyankar, H. K. (2009). Image Registration Techniques : An overview. *Pattern Recognition*, *2*(3), 11–28.

Yan, C., Ong, S. H., Ge, Y., Zhang, J., Teoh, S. H., & Okker, B. H. (2004). A neural network approach for 3d surface modeling and registration. In *IEEE International Workshop on Biomedical Circuits and Systems*. (pp. 119–122). IEEE. doi:10.1109/BIOCAS.2004.1454111

Zhang, J., Ge, Y., Ong, S., Chui, C., Teoh, S., & Yan, C. (2008). Rapid surface registration of 3D volumes using a neural network approach. *Image and Vision Computing*, *26*(2), 201–210. doi:10.1016/j.imavis.2007.04.003

Zitova, B., Flusser J. (2003). Image registration methods: a survey. *Image and Vision Computing*, *21*(11), 977–1000. doi:10.1016/S0262-8856(03)00137-9

Publications

Gouveia, A. R., Metz, C., Freire, L., & Klein, S. (2012). Comparative Evaluation of Regression Methods for 3D-2D Image Registration. In A. E. Villa, W. Duch, P. Érdi, F. Masulli, & G. Palm (Eds.), *Artificial Neural Networks and Machine Learning – ICANN 2012* (Vol. 7553, pp. 238–245). Springer Berlin Heidelberg. doi:10.1007/978-3-642-33266-1_30

Gouveia, A. R., Metz, C., Freire, L., & Klein, S. (2012). 3D-2D image registration by nonlinear regression. *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)* (pp. 1343–1346). IEEE. doi:10.1109/ISBI.2012.6235814

Freire, L. C., **Gouveia, A. R.**, & Godinho, F. M. (2010b). A Neural Network-Based Method for Affine 3D Registration of FMRI Time Series Using Fourier Space Subsets. In K. Diamantaras, W. Duch, & L. S. Iliadis (Eds.), *Artificial Neural Networks – ICANN 2010* (Vol. 6352, pp. 22–31). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-15819-3

Freire, L. C., **Gouveia, A. R.**, & Godinho, F. M. (2010a). FMRI 3D registration based on Fourier space subsets using neural networks. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society.* (pp. 5624–7). doi:10.1109/IEMBS.2010.5628038

List of Figures

Figure 5.22 Summary of the main coronary 3D/2D registration results for the 8 X-rays of 6 patients with mTRE values from a manual registration. The features sets of Chapter 5are compared for the preprocessing option 1D defined in *Image preprocessing I* (Section 5.3.2); F1 corresponds to Experiment I, (F2, F3, F4, F5) to Experiment II, and (Fp1, Fp4, Fp7) to the options 1, 4 and 7 of Experiment III, respectively. 98

List of Tables

Appendices

Appendix A

Table A.1 Mean and variance results for the selection of the number of hidden units for the feature sets considered in Chapter 3.

Mean

Feature sets	Set 1 Moments		Set 2 Eigen [*X Y*]		Set 3 Eigen [*X Y I*(*x*,*y*)]		Set 1 + Set 2		Set 1 + Set 3	
Units of input:hidden:output layers	6:6:1	6:12:1	6:6:1	6:12:1	12:12:1	12:24:1	12:12:1	12:24:1	18:18:1	18:36:1
R_X (°)	0,380	0,289	3,032	3,012	0,254	0,234	0,156	0,131	0,069	0,053
R_Y (°)	0,355	0,295	2,843	2,841	0,380	0,361	0,147	0,118	0,085	0,070
R_Z (°)	0,802	0,414	1,804	1,786	0,265	0,255	0,173	0,141	0,083	0,071
T_X (mm)	0,183	0,146	4,878	4,863	2,790	2,681	0,058	0,043	0,034	0,024
T_Y (mm)	1,981	1,356	5,049	5,042	5,022	5,017	0,753	0,649	0,383	0,332
T_z (mm)	0,135	0,115	4,675	4,698	0,793	0,738	0,083	0,069	0,034	0,033

Variance

Feature sets	Set 1 Moments		Set 2 Eigen [*X Y*]		Set 3 Eigen [*X Y I*(*x*,*y*)]		Set 1 + Set 2		Set 1 + Set 3	
Units of input:hidden:output layers	6:6:1	6:12:1	6:6:1	6:12:1	12:12:1	12:24:1	12:12:1	12:24:1	18:18:1	18:36:1
R_X (°)	0,119	0,085	5,330	5,322	0,136	0,128	0,034	0,023	0,006	0,003
R_Y (°)	0,095	0,070	4,561	4,557	0,207	0,184	0,023	0,015	0,009	0,006
R_Z (°)	0,397	0,152	2,143	2,161	0,049	0,047	0,040	0,031	0,009	0,008
T_X (mm)	0,020	0,015	8,538	8,655	6,241	6,000	0,004	0,003	0,002	0,001
T_Y (mm)	3,384	2,380	8,322	8,270	8,342	8,296	1,141	1,166	0,326	0,279
T_z (mm)	0,016	0,013	8,963	9,001	0,713	0,670	0,007	0,006	0,002	0,002

Table A.2 Mean and variance results for the selection of the number of hidden units for the MLPs considered in Chapter 4.

	Mean			
Units of Input:hidden:output layers	18:9:1	18:18:1	18:36:1	18:54:1
R_X (°)	0,087	0,069	0,053	0,049
R_Y (°)	0,119	0,085	0,070	0,068
R_Z (°)	0,108	0,083	0,071	0,071
T_X (mm)	0,042	0,034	0,024	0,024
T_Y (mm)	0,650	0,383	0,332	0,332
T_z (mm)	0,047	0,034	0,033	0,032

	Variance			
Units of Input:hidden:output layers	18:9:1	18:18:1	18:36:1	18:54:1
R_X (°)	0,009	0,006	0,003	0,004
R_Y (°)	0,014	0,009	0,006	0,005
R_Z (°)	0,015	0,009	0,008	0,007
T_X (mm)	0,002	0,002	0,001	0,001
T_Y (mm)	0,805	0,326	0,279	0,311
T_z (mm)	0,003	0,002	0,002	0,002

Appendix B

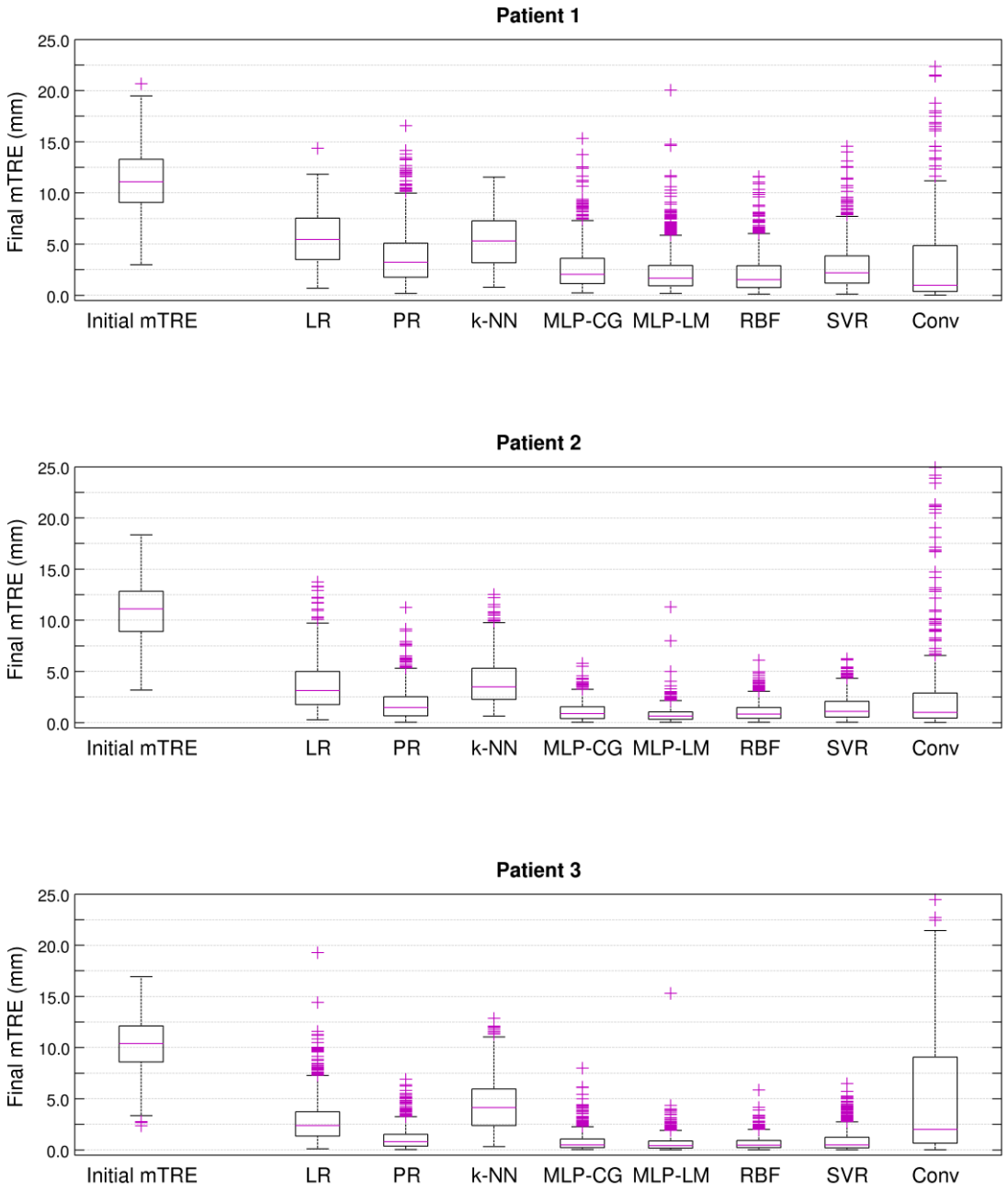


Figure B.1 Comparison of registration-by-regression results for all methods and for patients 1, 2 and 3. The graphic also shows the initial mTRE before registration.

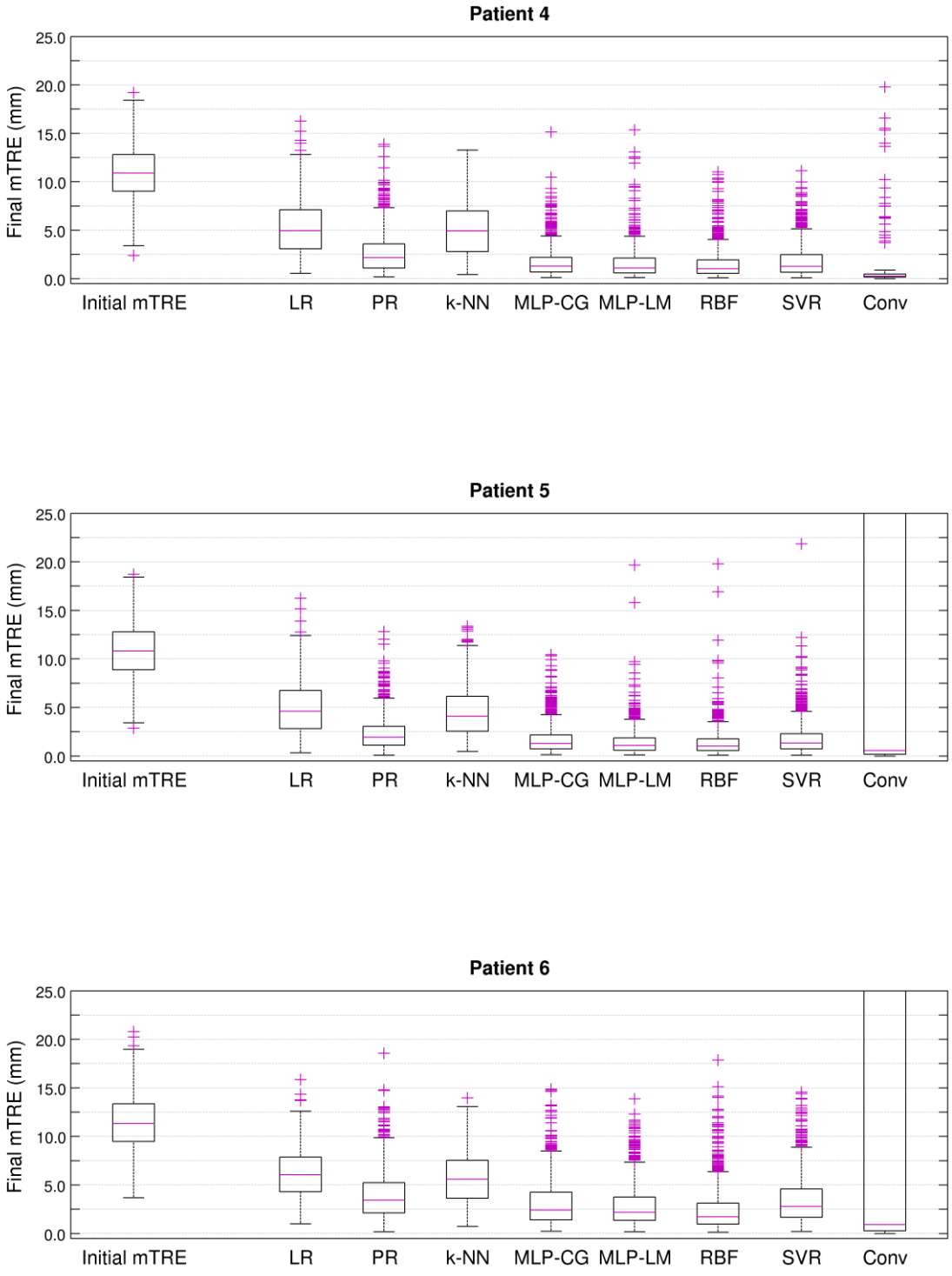


Figure B.2 Comparison of registration-by-regression results for all methods and for patients 4, 5 and 6. The graphic also shows the initial mTRE before registration.

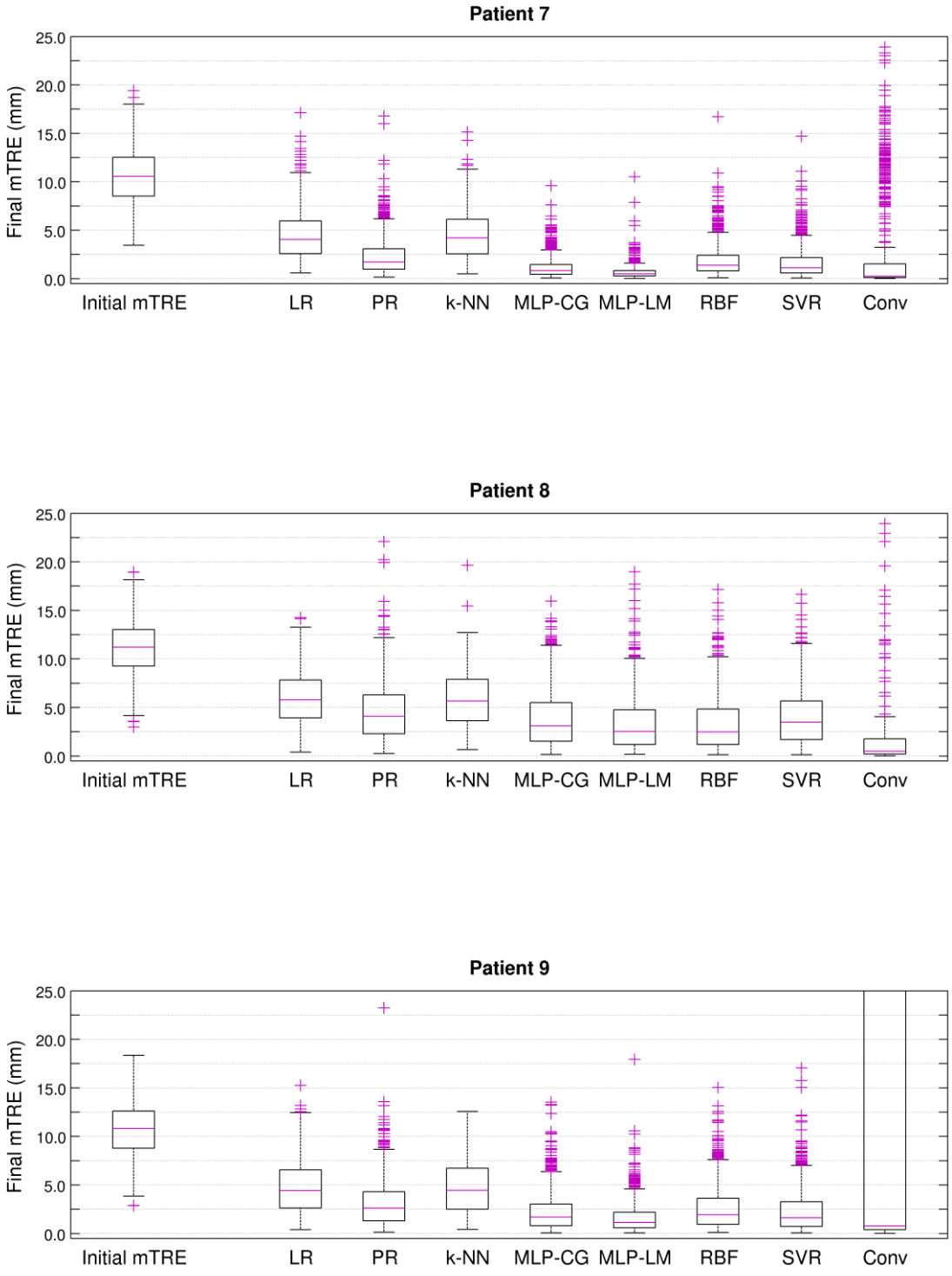


Figure B.3 Comparison of registration-by-regression results for all methods and for patients 7, 8 and 9. The graphic also shows the initial mTRE before registration.

Appendix C

Registration-by-regression of coronary CTA and X-ray angiography

Ana R. Gouveia[1,2], Coert Metz[3], Luís Freire[4], Pedro Almeida[2], Stefan Klein[3]

[1] CICS-UBI – Health Sciences Research Centre, University of Beira Interior, Covilhã, Portugal
[2] Institute of Biophysics and Biomedical Engineering, Faculty of Sciences, University of Lisbon, Lisbon, Portugal
[3] Biomedical Imaging Group Rotterdam, Depts. of Medical Informatics & Radiology, Erasmus MC, Rotterdam, the Netherlands
[4] Escola Superior de Tecnologia da Saúde de Lisboa, Instituto Politécnico de Lisboa, Lisbon, Portugal

Abstract

We evaluate the integration of 3D preoperative computed tomography angiography (CTA) of the coronary arteries with intraoperative 2D X-ray angiographies by a recently proposed novel registration-by-regression method. The method relates image features of 2D projection images to the transformation parameters of the 3D image. We compared different sets of features and studied the influence of preprocessing the training set. For the registration evaluation, a gold standard was developed from 8 X-ray angiography sequences from 6 different patients. The alignment quality was measured using the 3D mean target registration error (mTRE). The registration–by-regression method achieved moderate accuracy (median mTRE of 15mm) on real images. It does therefore not provide yet a complete solution to the 3D-2D registration problem but it could be used as an initialization method to eliminate the need for manual initialization.

Keywords

3D/2D image registration, image guided interventions, coronary arteries, regression, neural networks.

1 Introduction

Angioplasty, also called percutaneous coronary intervention (PCI), is a minimally invasive coronary intervention frequently used for the treatment of coronary artery disease (CAD) (Kasper et al., 2005). During a PCI, the clinician follows the path of the catheter in the patient's body with the help of an imaging technique, usually X-ray imaging with contrast injection (Rivest-Hénault, Sundar, & Cheriet, 2012). Although PCI is a not trivial procedure (Markelj, Tomaževič, Likar, & Pernuš, 2012), it has very high success rate (Soon et al., 2007). Nevertheless, there are certain particularly challenging lesions such as bifurcation lesions and chronic total occlusions (CTO) (Metz, 2011). Magnetic navigation could potentially improve the success rate of such complicated procedures (Ramcharitar, van der Giessen, van der Ent, Serruys, & van Geuns, 2011; Serruys, 2006). However, this technique requires the availability of a 3D roadmap, derived from pre-interventional images (e.g. 3D CTA). This roadmap then needs to be related to the patient by image registration to intra-operative imaging data (e.g. x-ray imaging). Several authors (e.g., Ramcharitar, Patterson, van Geuns, van Meighem, & Serruys, 2008) also highlight the importance of multi-slice CT in interventional cardiology because of the additional information it provides with respect to angiography. They point out its ability to accurately identify the vessel's border (even the occluded vessel segments that are not seen angiographically), and to provide information about coronary plaque composition which may be helpful to facilitate crossing the lesion in the chronic total occlusions cases. Furthermore, any PCI can benefit of using pre-annotated CT data by increasing the accuracy of the stent placement, e.g., indicating plaque locations (Ruijters, ter Haar Romeny, & Suetens, 2009). To conclude, the registration of preoperative 3D computed tomographic angiography (CTA) data, containing the occluded vessel segment, with intraoperative 2D X-ray can be very useful for guiding the treatment of these pathologies.

Generally, 3D/2D image registration is estimated by iteratively optimizing a similarity measure between simulated X-ray projection, computed from the preoperative CT scan, and an X-ray image (Markelj et al., 2012). However, such iterative optimization procedures usually have a small capture range due to local maxima of the similarity measure, and therefore require initialization close to the searched pose (Markelj et al., 2012; Van De Kraats, Penney, Tomaževič, & Van Walsum, 2005). We previously proposed a registration-by-regression approach for 3D/2D coronary registration (Gouveia et al., 2012a) addressing the capture range drawback, which showed promising results for simulated X-ray images (Digitally Reconstructed Radiographs - DRRs). It relates image features of 2D projection images to the transformation parameters of the 3D image by nonlinear regression in a supervised learning stage. Our motivation in the current work is the clinical extension of this registration-by-regression method. Thorough evaluation of a registration process is crucial for its clinical use (Jannin et al., 2002; Markelj et al., 2012). The most straightforward method for estimating the registration error is by a "ground truth" transformation, which may be obtained from a gold

standard registration system (Fitzpatrick, 2001). Although gold standards based on patients are the most difficult to establish according to the same authors, the other possibilities (e.g. computer simulations, phantoms or cadavers) cannot take into account all the true variability encountered in clinical situations (Jannin et al., 2002).

Therefore, we implement the registration-by-regression methodology on real clinical data and we construct a gold standard based on patient data. The main contributions of this paper are:

(1) The evaluation of registration-by-regression of 3D preoperative CTA of the coronary arteries with intraoperative 2D X-ray angiographies.

(2) Exploring novel feature definitions (based on principal component analysis of projected image data) and by investigating various combinations of features. We provide insight in the influence of features on the accuracy of the method.

(3) The construction of a gold standard for 3D/2D coronary registration by means of an exhaustive search routine, matching the projected centerlines to manually annotated vessels on the X-ray, followed by systematic visual inspection.

2 Methods and materials

2.1 Registration by nonlinear regression framework

In this work we consider the 3D/2D registration by nonlinear regression method presented in Gouveia et al. (2012a) (Figure C.4). This regression model relates image features of the 2D projection images to the transformation parameters of the 3D image. Once the nonlinear regression function is computed, which takes place before the intervention (e.g., PCI), the estimate of the 3D transformation parameters of the 3D image can be obtained by applying the regression function to the features computed on the intraoperative X-ray. During the preoperative phase (Figure C.4), a 3D vessel tree model is obtained by the segmentation of coronary arteries from 4D CTA images; a set of N simulated 2D images (DRRs) is generated by applying random transformations $T(t,\theta)$, followed by projection. Then a training set is built comprising a set of N input-output pairs: the features extracted from the N simulated 2D images (DRRs) as inputs and the corresponding 3D transformation parameters ($T(t,\theta)$) as outputs. This training data is used for the learning process to estimate the parameters of the regression function ($f(X)$ in Figure C.4). During the intraoperative phase (Figure C.4), the image features of the X-ray acquired are computed and fed as input to the regression function (i.e., the interpolation *step* in Figure C.4), which returns the estimated 3D translation and rotation parameters of the 3D image ($T_{X\text{-}ray}$).

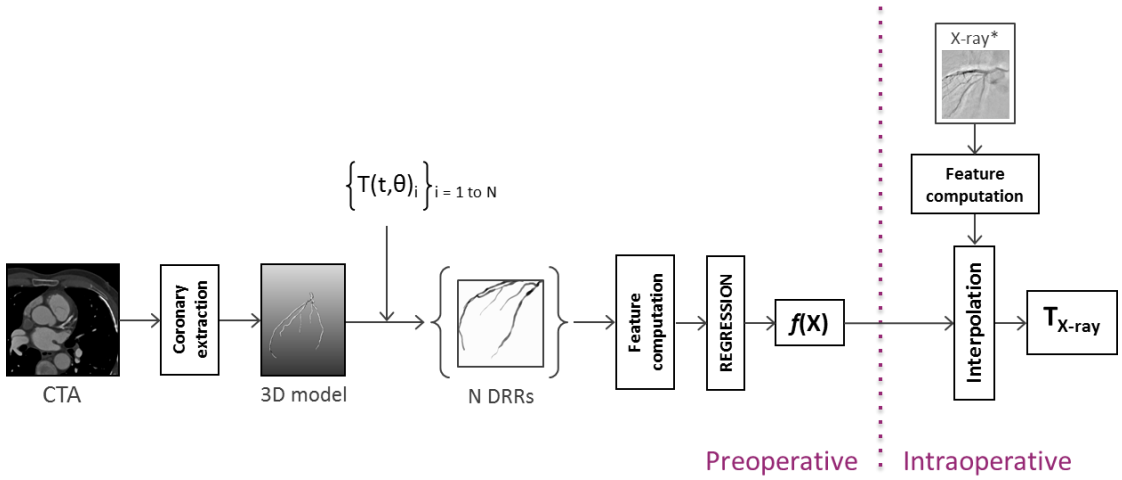


Figure C.4 Registration-by-regression schematic model.

The regression model used during the preoperative phase is the Multi-Layer Perceptron (MLP), a feed-forward Neural Network as described in detail in Gouveia et al. (2012a). For this regression problem with such images and input-output set, the MLP is an adequate choice with higher performance when compared with other suitable possibilities (Gouveia, Metz, Freire, & Klein, 2012b).

2.2 Imaging data

We performed the registration of 3D preoperative coronary CTA data with 2D intraoperative X-ray angiography, using 8 time-sequences collected from 6 different patients. For two of those patients, two different time-sequences were considered for the registration using the same CTA data, which correspond to different intervention dates.

CTA images were acquired using a Siemens Definition or Siemens Definition Flash scanner and X-ray images using a Siemens Axiom Artis biplane system. Calibration data was not available for the X-ray acquisition and thus only monoplane experiments using the sequences of the primary C-arm were considered. The field of view of the CTA images used for the reconstruction was 256×256×[99-184] voxels and the resulting voxel size was 0.7×0.7×[0.8-1.0] mm^3; for DRR images and X-rays images the same quantities were 512×512 pixels and 0.22×0.22 mm^2, respectively. The CTA images were obtained in the end-diastole phase, more precisely at 70% of the cardiac cycle considering a R-wave peak-to-peak cycle. The X-rays images used for registration were the ones from the second-half of the cardiac cycle, i.e. from 50% to 100% of the cardiac cycle, which contained on average 6 time frames.

2.2.1 Image preprocessing

We used binary images, either simulated or real. As we are only interested in the information contained in the coronary arteries for image registration, a procedure for vessel segmentation was necessary. For the X-ray angiographies, the masks were determined by manual segmentation from the analysis of two observers; Figure C.5 shows the manual masks for two X-ray images. For the DRRs, the masks were determined by threshold based segmentation using a manually defined threshold value. Restricting the problem to binary images and using manual segmentations, differences in appearance between DRRs and X-rays were reduced, resulting in an idealized setting for experimentation.

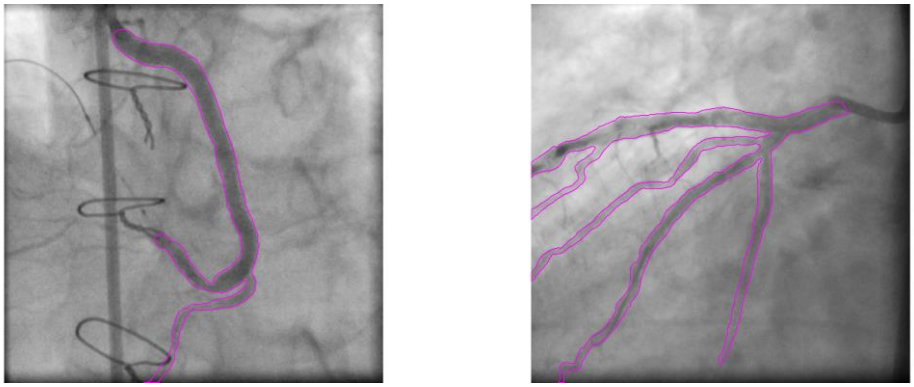


Figure C.5 Manual masks (colour lines) on X-ray 1 (left) and X-ray 2 (right).

2.3 Construction of training data

The regression function for the registration-by-regression method was calculated by using a training set of simulated projection images obtained from the application of known geometric transformations to the 3D coronary vessel tree. From the retrospectively ECG-gated 4D CTA images available, we segmented coronary arteries (at the end-diastole phase) and made binary vessel tree models (Metz et al., 2011). DRRs were generated using these 3D models with a computation procedure described in Metz, Schaap, Klein, et al. (2009). To represent a clinically relevant view, the projection geometry for the computation of the DRRs and the initial orientation of the preoperative data were derived from an interventional X-ray image. DRRs were generated using MeVisLab.

For each patient, 11000 DRRs were generated, 10000 to obtain the regression model and 1000 to test the performance of our method before the registration of real data. The 6 rigid transformation parameters were drawn from uniform distributions, with values between -13

and 13 degrees for rotations and between -16 and 16 mm for translations (to ensure the training set contains the transformation parameters of the X-ray images to be registered).

2.3.1 Training set preprocessing

Modifications in the dataset were implemented attending the presence/absence of vessels in simulated and real images, i.e. some vessels clearly identified in the X-rays of a patient did not appear in the correspondent DRRs and vessels identified in DRRs did not appear in the correspondent X-rays. This difference is clearly recognized by the visualization of the images, for example in Figure C.6, and varies case by case.

Two modifications were exclusively implemented on simulated images. The first is the application of a morphological opening operation to the image to remove the tiniest vessels present (Gonzalez & Woods, 2008). The second is the elimination of a part of the image, with a random size and position, simulating occlusion. Additionally, a Gaussian filter was applied to the binary images creating a blurring effect and differentiating the borders from the inner regions

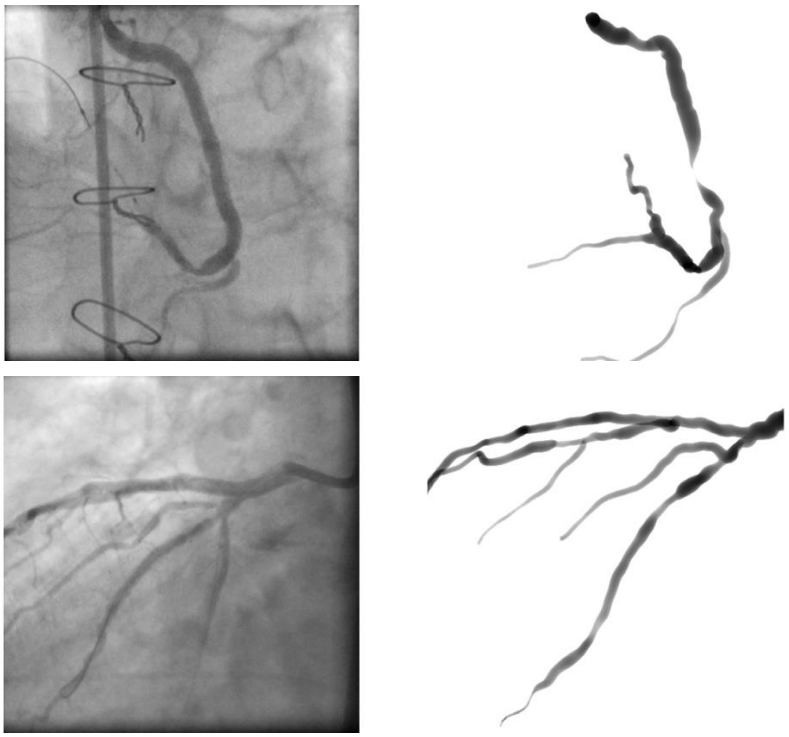


Figure C.6 X-rays and correspondent DRR (of training set) with closer transformation parameters. For patient 1 (1st row), there are two tiny vessels on the left of the DRR image not present in X-ray image; and for patient 2 (2nd row), for example one vessel on the inferior right part of the X-ray image does not appear in DRR.

For each of the 6 patients in the study, 9 training sets were considered with different combinations of preprocessing operations, schematized in Table C.2. For an easier comparison, these different training sets will be labelled as *P* and numbered from 1 to 9, indicating the preprocessing operation (Table C.2). Two different standard deviations for the Gaussian filter were used; the cases of unmodified images and unfiltered images were also studied. Figure C.7 shows the resulting images after the application of these operations on a DRR.

Table C.2 Combinations used for preprocessing operations applied, resulting in 9 different training sets per patient.

	Gaussian filter		
Operations	-	$\sigma 1$	$\sigma 2$
-	P1	P2	P3
Erosion/dilation	P4	P5	P6
Elimination of vessel parts	P7	P8	P9

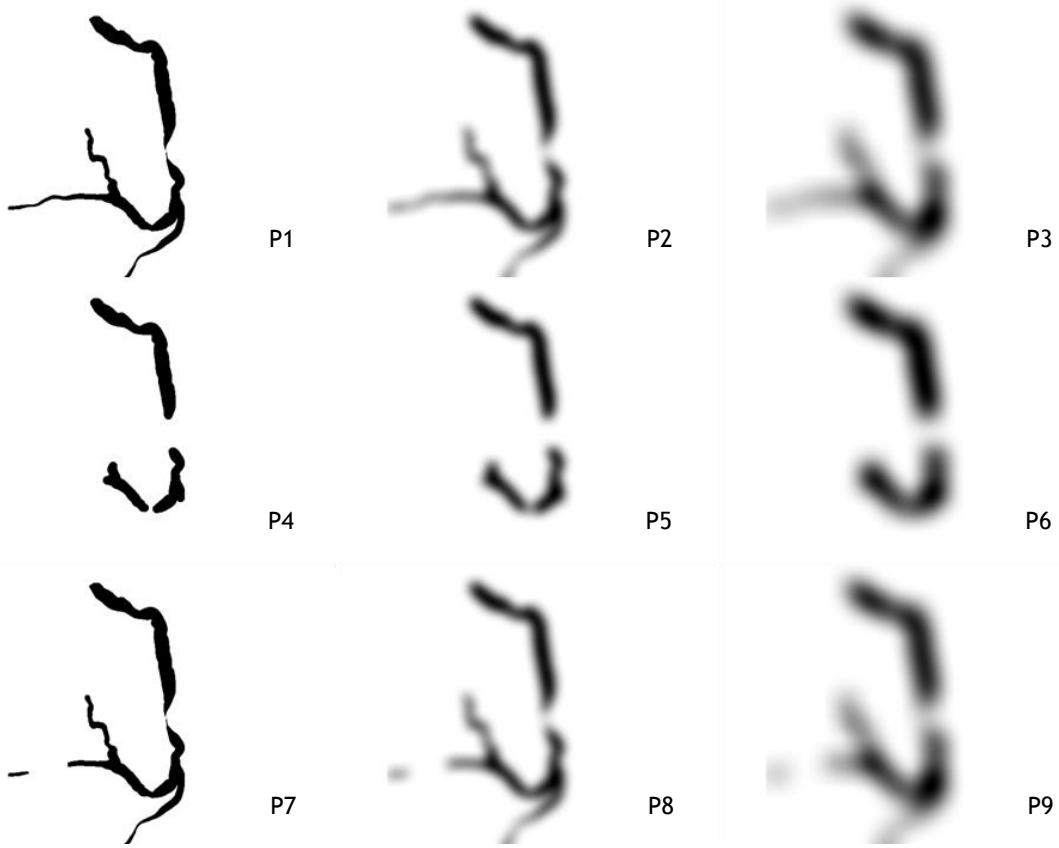


Figure C.7 Resulting images from the training set preprocessing options for a DRR of patient 1. The order of the images presented corresponds to the order of options in Table C.2.

For these preprocessing operations applied to the images, parameters were defined. The erosion and dilation operations were performed by using a disk-shape structuring element with a specific radius for each patient. These values were empirically determined and set as $\{8,6,0,5,5,8\}$ pixels for patients 1 to 6, respectively. On the other hand, the elimination of parts of the vessels was implemented by removing a circle with a random diameter up to 1/3 of the image size, randomly centred within the coronary arteries. About the filters 1 and 2 (Table C.2), we used two low-pass Gaussian filters with different standard deviations, where σ was empirically defined as 10 and 20 pixels.

2.4 Input features

Two different feature approaches are considered in this work. One is to compute a limited number of features with a geometric nature, and the other is to use the entire projected images in the space of principal components.

In the first case, we computed from the 2D projection images (a) the 2D geometric moments of order 0, 1 and 2; (b) the eigenvalues and eigenvectors computed from a Principal Component Analysis (PCA) on the pixels of the object of interest, after a coarse segmentation of the image into object and background objects (Gouveia et al. 2012a); (c) the geometric centre, or centre of mass, i.e. the point where all mass of the image could be concentrated without changing its first order moment, which can be used as a reference point to describe the position of the object imaged (Liao, 1993).

These four features were combined as shown in Table C.3, where the first combination (F1) corresponds to the best performing feature set used in Gouveia et al. (2012a) for this problem.

In the second case, the features were the projected images in the space of principal components (Shlens, 2005). These features will be referred in the text as *PCA features* and labelled as F6. Considering each image one observation and each image pixel one variable, we defined a M×N dataset with N observations and M variables each, where each image in a matrix format is firstly transformed into an 1×M array. A PCA was applied to this dataset and, after the determination of the rank-ordered variance, the highest *V*% principal components were selected, where the setting of *V* was optimized (Section 2.5).

Table C.3 Geometric features and respective combinations (F1 to F5); the mean and standard deviation of the number of PCA features (F6), computed over the different training preprocessing operations P1-P9 and over all patients, is presented (mean,standard deviation).

	Moments	Eigenvectors	Eigenvalues	Geometric Centers	#features
F1	✓	✓	✓		18
F2		✓	✓		12
F3		✓			9
F4		✓		✓	11
F5				✓	2
F6	(PCA-based features)				(70,47)

For practical computational reasons image were downsampled from 512×512 to 128×128 before applying the PCA.

2.5 Regression model

For each registration, 6 regression models were constructed, one per transformation parameter, using identical Neural Networks with one hidden layer with fully connected units. The activation functions are the hyperbolic tangent and the linear functions for the hidden and output layers, respectively. The input units were standardized by a z-score (Haykin, 1999) and the weights were randomly initialized within a range [-1,1]. The number of epochs was defined separately for each MLP by a stopping epoch (i.e., the epoch when the validation error started to grow) with a maximum of 1000.

For each patient, the set of 10000 images, for the construction of the regression model, was split in two sets: 7000 images to train the regression models and the remaining 3000 to select the stopping epoch. The performance of the models was computed by the mean absolute difference between the model prediction and the known output for each regression.

For combinations F1 to F5 of geometric features (Table C.3), the MLP topologies are described in Table C.4. All cases have one output unit, which is the correspondent transformation parameter. From optimization studies of Gouveia et al. (2012a, 2012b), the optimum number of hidden units for combination F1 was set to the double of the number of input units. The remaining combinations are different but yet similar features sets. Even if the number of features is different and there are new features included, the features number and nature do not differ that much. Then the same criterion was used to choose the number of hidden units

for combinations F2, F3 and F4; for the feature set labelled as F5, we considered 10 hidden units since we expected that 4 hidden units would be too few.

Table C.4 Topology of MLPs used for different combinations Table C.3 (i.e., number of input units : number of hidden units :number of output units).

F1	F2	F3	F4	F5
18:36:1	12:24:1	9:18:1	11:22:1	2:10:1

For the PCA features, the cumulative percentage *V* was tuned (for each subject) in the range $\{50\%, 60\%, 70\%, 80\%, 90\%\}$. Simultaneously, we looked for the optimum number of hidden units considering the range $\{8,16,32\}$ as possibilities. To limit computation time, the tuning was only performed for one of the transformation parameters, the translation parameter in *x*, and its results were extended to the other transformation parameters. The search ranges were defined after some preliminary tests.

Two optimization schemes for the MLP were used depending on the feature approach. For the combinations of geometric features, it was used the MLP with Levenberg-Marquardt optimization; whereas for the PCA features, we choose the conjugate-gradient optimization due to the much higher number of input units (Gouveia et al., 2012b).

For the tuning process we analysed the mean absolute difference between the model prediction and the known output, for translations over the *x* direction and computed on the validation set. This set was also used to select the stopping epoch for each MLP run, which occurred during and after the tuning with a maximum of 1000 and 3000 epochs, respectively.

The implementation of NN were performed using MatLab, version 7.11.0.584 (R2010b) 64 bits (MatLab, 2010).

2.6 Gold standard definition

A gold standard for evaluation of the registration-by-regression method is needed when registering real data. A problem thereby is the occurrence of non-rigid deformation due to cardiac and respiratory motion. Particularly, time differences with respect to cardiac cycle could exist since we matched an end-diastolic 3D model to all frames of a X-ray time-sequence of the second half of the cardiac cycle. Therefore, we constructed a gold standard for each X-ray frame of the 8 time-sequences by implementing an exhaustive search based on the 2D distances between vessel centerlines, and asked five observers to rank the quality of alignment.

For each X-ray frame the initial exhaustive search was done as follows: (1) we started by manually annotating the centerlines in each X-ray angiography of all images collected; (2) we projected the vessel centerlines from the correspondent coronary artery 3D model (as explained in Metz, Schaap, Weustink, et al., 2009), after a given transformation has been applied; (3) we measured the average of shortest distances between annotated and projected centerlines; (4) second and third steps were then exhaustively repeated over a large range of rotations/translations of the 3D model to find which transformation minimizes the distance.

After having performed this procedure for all time frames of each X-ray sequence, five different observers analyzed the result of this full-search through a systematic visual inspection of the X-ray frame with both manually annotated centerline and registered and projected centerline. A manual scoring was done, in which the X-ray time frames were ranked based on the quality of the alignment with the projected CTA centerlines. The observers were asked (a) to rank the images of the time-sequence on alignment quality, i.e. to order all images from best to worst match; and (b) to identify unsatisfactory alignments. The best case was accepted as a gold standard, but only if at least half of the observers scored the alignment as satisfactory. In further registration experiments, we only use this best case time frame, as only for that time frame we can be confident about the gold standard, and we are sure that the registration is not hampered by non-rigid deformations introduced by cardiac and respiratory motion. Figure C.8 shows the best and worst alignment for one of the time-sequences studied, where the best alignment was scored satisfactory and the worst not.

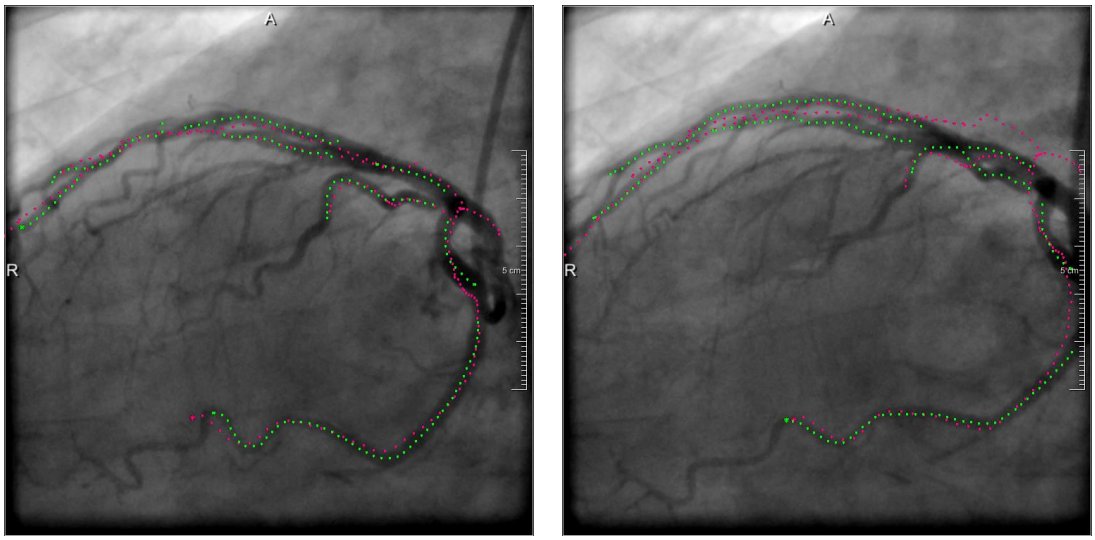


Figure C.8 Examples of two images used for the evaluation of the alignment quality. These are X-ray frames of the time-sequence 3.2 where the vessel centerlines are highlighted: the green is the manual annotated centerline in X-ray and the pink is the closest registered and projected centerline from the CTA model (obtained by the exhaustive-search process). They represent the best (left) and the worst (right) alignment according to the established ranking.

2.7 Evaluation of registration

The evaluation of the registration approaches was performed by the computation of the mean target registration error ($mTRE$) before and after registration (Fitzpatrick & West, 2001; Van De Kraats et al., 2005). The $mTRE$ is computed as the mean 3D distance to the gold standard position and orientation (computed as described in Section 2.6) of all points, p_k, of the vessel centerline trees:

$$mTRE(T) = \frac{1}{K}\sum_{k=1}^{K}\left\|T(p_k) - T_{gold}(p_k)\right\| \qquad \text{Equation C.1}$$

where T is the resulting transformation from the registration method to assess and T_{gold} is the known gold standard transformation.

$mTRE$ for simulated images were computed on the test set of 1000 images, which was not used for the construction of the regression models. For real data, $mTRE$ was determined for one X-ray of each time-sequence as defined before, in Section 2.6. The ground truth of each simulated image is the known set of transformation parameters applied in order to generate it and, for real images, it is the set obtained from the process described in Section 2.6.

3 Experiments and results

We compared the effect of using different features F1-F6. Subsequently, the influence of the preprocessing options P1-P9 was investigated using feature set F6 (the PCA-based features).

3.1 Comparison of input features

The registration results for simulated and real images for each input feature studied (from F1 to F6) are shown in Figure C.9, where each boxplot is based on all patients' results. Additionally, mTRE values from a manual registration initialization are also displayed in the plot of real images results (the first column). The manual initialization was performed on the projection plane (XY) while observing the X-ray image, after moving the centre of gravity of the model to the origin of the intraoperative coordinate system (Metz, 2011).

For simulated images, the F2 to F5 options yielded similar mTRE values but much higher than the F1 option, the only set with the contribution of the moments. Actually, the information held by the moments revealed to be determinant for a high accuracy of the registration of

simulated images as discussed in Gouveia et al. (2012a). However, this same feature set F1 gave the worst results for real images.

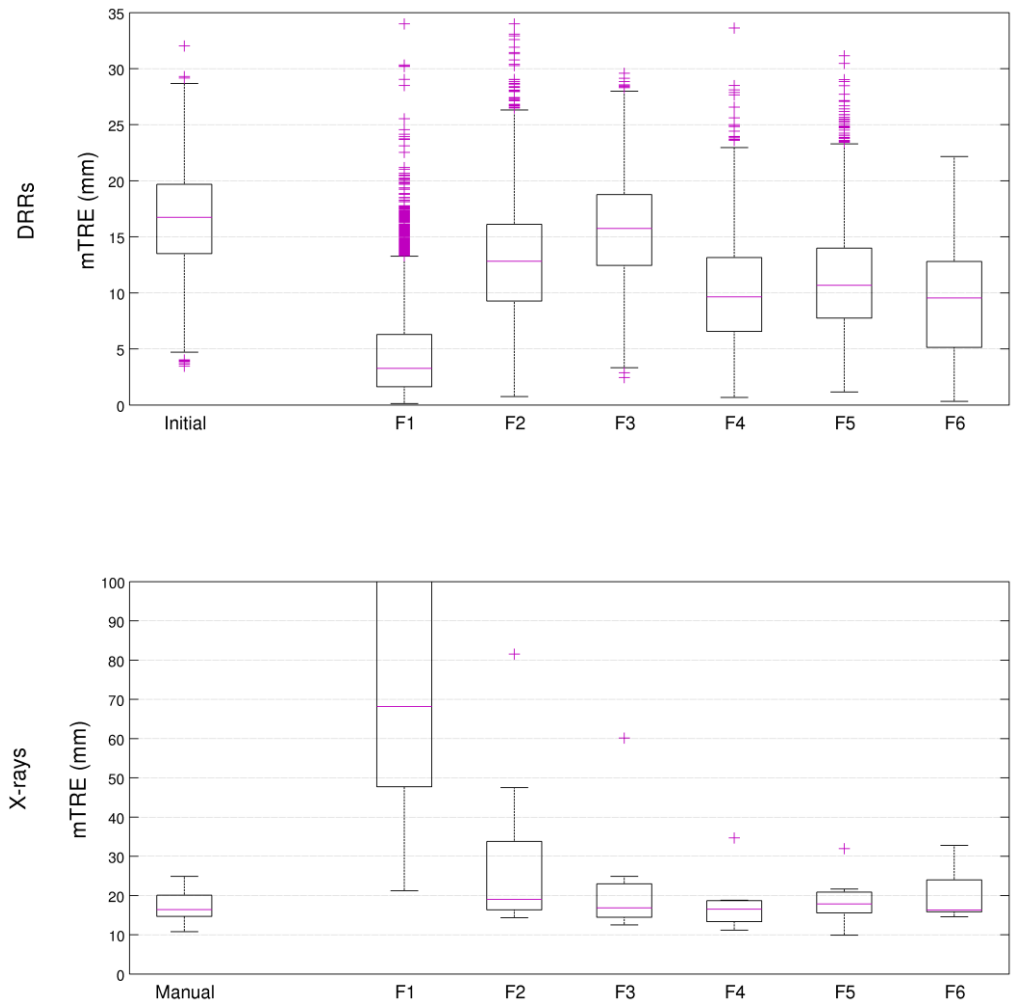


Figure C.9 Registration results for simulated (DRRs) and real (X-rays) images, for all patients, comparing the different input features studied (F1 to F6). The mTRE values before the registration (*initial*) and the mTRE values from a manual registration (*manual*) are also shown.

3.2 Effect of training set preprocessing

For the PCA features (F6) the registration results are presented in Figure C.10, for simulated and real images. Each boxplot is based on all patients' results and corresponds to the different preprocessing options as defined in Section 2.3.1 (P1 to P9). For the real images plot, mTRE values from manual registration are also displayed.

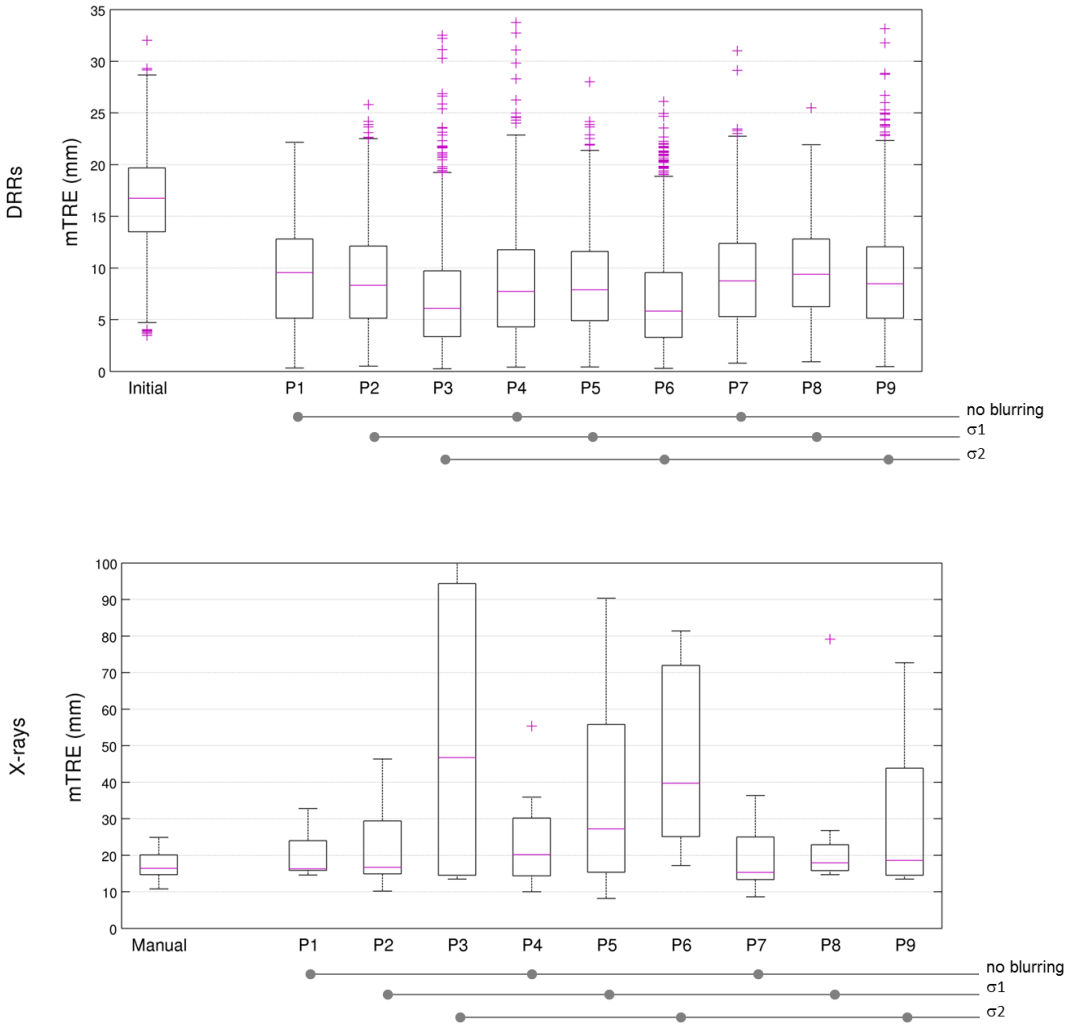


Figure C.10 Registration results for simulated and real images for all patients comparing the different preprocessing operations labelled as P1 to P9, according to Table C.2, for the PCA features (F6). The options (P4,P5,P6) correspond to erosion/dilation operations and (P7,P8,P9) to the elimination of parts of the vessels. The cases with similar blurring options are indicated in the image below by points on the lines. The mTRE values before the registration (*initial*) and the mTRE values from a manual registration (*manual*) are also shown.

From the mTRE distributions for X-ray angiographies (Figure C.10), it is evident the blurring is not helping the registration of real images, although for simulated images it does, i.e. the more blurred the DRRs are, the better the registration performance. On another hand, it is not clear if the erosion/dilation (F6-P4 to F6-P6) or the elimination of vessel parts (F6-P7 to F6-P9) introduce any improvement on the method performance. Therefore, for a better assessment, the registration results without considering the blurring pre-processing steps will be separately presented for each X-ray case in the next section.

3.3 Individual accuracies on real data

Figure C.11 summarizes the results of 3D/2D coronary registration-by-regression obtained for all combinations of features (F1 to F5) in study and considering the most relevant cases of PCA features with image preprocessing (F6-P1, F6-P4 and F6-P7). Additionally, manual mTRE values were displayed in the plots (by horizontal lines), which contextualizes the registration results.

The mTRE values obtained are high and for 2 of the 8 cases the X-ray images presented values above the range of the manual mTRE. Moreover, for the majority of the X-ray images, the registration-by-regression achieved the accuracy of a manual initialization. Comparing the different features options, the F1 set is the least accurate option although it presented the best results for simulated images as mentioned before. In fact, for all X-rays cases, its mTRE values were much higher than manual values. However, some features options like geometric center and eigenvectors from the PCA on the pixels of the object of interest (F4), PCA features with elimination of parts of the vessels in DRRs (F6-P7) and PCA features with no modification of DRRs (F6-P1) must be highlighted.

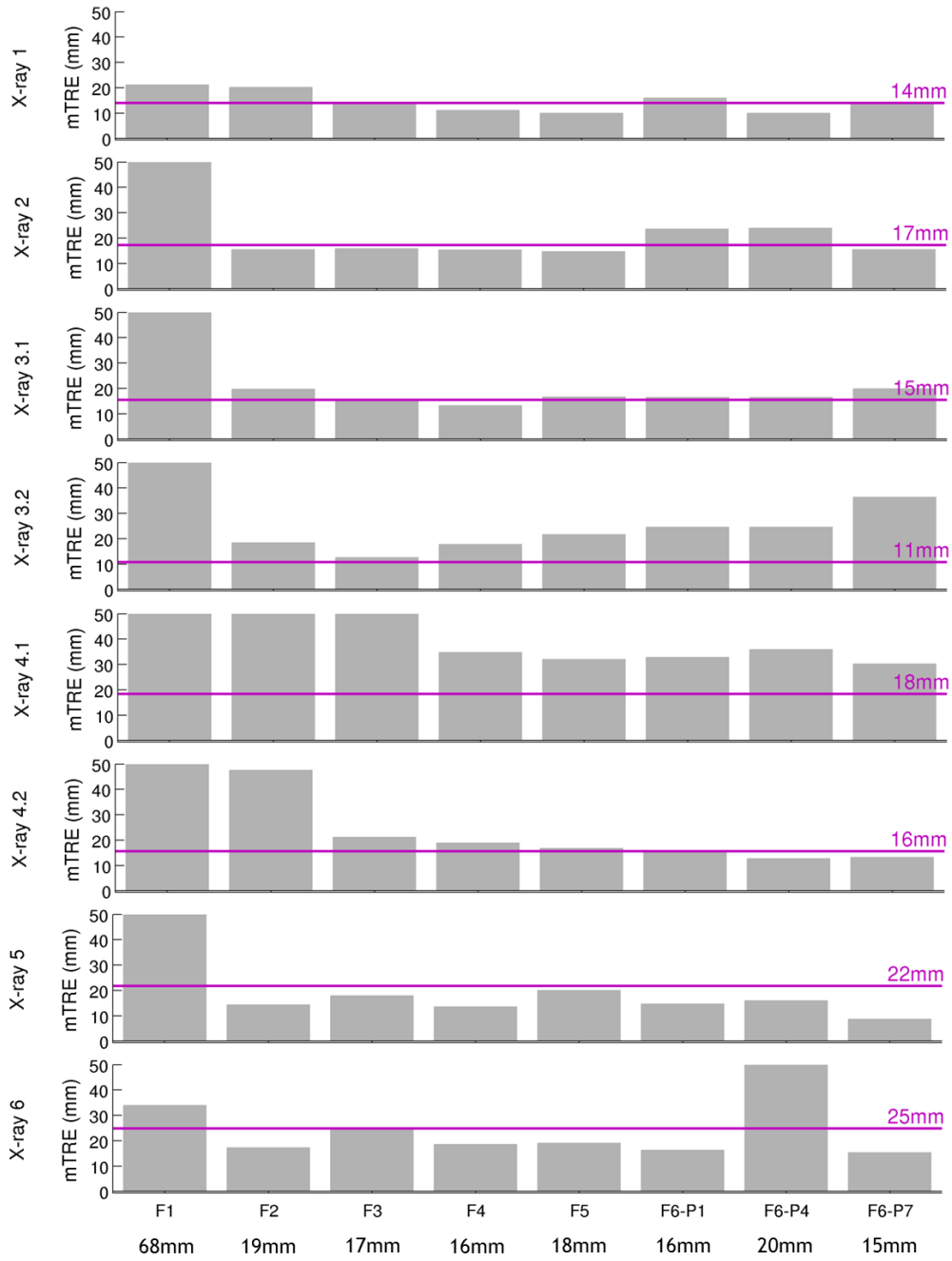


Figure C.11 Coronary 3D/2D registration results for the 8 X-rays of 6 patients with mTRE values from a manual registration (represented by the horizontal line). (F1, F2, F3, F4, F5) correspond to geometric features sets and (F6-P1, F6-P4, F6-P7) to the PCA features with the best preprocessing options (P1, P4 and P7). The median of mTRE values are presented below the features labels.

4 Discussion

The registration of 3D CTA with X-rays angiographies is a very challenging problem. The differences between the angiographies and DRRs computed from CTA are a problem reported in the literature for the iterative approaches of 3D/2D registration using DRRs (Penney et al., 1998), where a good similarity measure should be able to achieve accurate registrations despite these differences. For a registration-by-regression approach, these differences are also an important challenge. Some differences are intrinsic to image formation, i.e. are caused by the inherent dissimilarity between CTA and X-ray images, while others are related to the imaged object, when for example interventional instruments are present.

As inherent differences, we can usually find overlaid structures and different contrast acquisition processes described in the literature (Imamura, Ida, & Sugimoto, 2002). In our case, the DRRs were rendered from a 3D CTA-derived coronary model. In the DRRs rendered from CTAs (and not from models), the coronary arteries are difficult to visualize because of the high intensity structures being over-projected on them (Metz, Schaap, Klein, Weustink, et al., 2009). This procedure however just addresses one side of the problem, the DRR construction, and not the formation of the X-ray images.

The nature of the feature sets studied was different: on one side few geometric quantities and, on the other side, a high number of features representing the intensity variability of a pixel in the image for different transformations. Of course, when we restricted the preprocessing options to the binary images, we are not using exactly the intensity variability but the variability of a pixel corresponding to the object imaged or to the background (0/1 intensity) for different transformations. We performed some preliminary tests conducting to this work, where we used these geometric feature sets but considering image segmentation with automatic masks and different options for intensity values. Those tests showed that although the differences mentioned before between DRRs and X-rays are an important error source, the way the features rely on the intensities and on the vessels size are a crucial factor. Therefore, to work just with binary images (with a manual mask) was the way to confine the problem to only a few differences, allowing us to evaluate the capability of the features chosen, disregarding differences between DRR and X-rays. However, an automatic vessel segmentation method would be needed to make this approach fully automatic. Examples of such methods are the vesselness filter proposed by Frangi, Niessen, Vincken, & Viergever (1998) or the segmentation algorithm proposed by Schneider & Sundar (2010), which was designed to be more robust against the background artifacts.

In this work, we tested several feature sets and it was found that the best solution for simulated images is not necessarily the best solution for real images. Among the various studied features, the set composed by moments, eigenvectors and eigenvalues (F1 in Section 3) was

clearly the most accurate when only DRRs were involved, but this same set was the worst when the X-ray images were registered by the regression model built from DRRs. The high sensitivity of these features on minor details between the DRRs and X-ray images (e.g. the vessels size) especially when including the moments and/or the eigenvalues (F1 and F2 in Section 3), is probably one of the main problems that justifies the poor results for this feature.

For some features, we also tested the influence of artificially modifying the training set aiming to simulate the typical differences between real and simulated images, i.e. the eventual absence of parts of the coronary arteries. The results indicated the learning performance with this extra information was as good as before, and the results for the registration of real images did not differ either.

The 3D/2D iterative registration approaches based on DRRs usually require initialization close to the searched pose because of the numerous local maxima of intensity or gradient-based similarity measures (Markelj et al., 2012). For the 3D/2D registration of CTA with X-rays angiographies, the proposed regression models, considering the best feature sets and the modification of the training set, could constitute an initialization solution to replace the need of manual initialization. We find in the literature registration methods with some similarities with the registration-by-regression approach in study, which are suitable for initialization purposes (Markelj et al., 2012). This is the case of the methods based on templates (Banks & Hodge, 1996; Cyr, Kamal, Sebastian, & Kimia, 2000; Hermans, Claes, Bellemans, Vandermeulen, & Suetens, 2007; Hoff, Komistek, Dennis, Gabriel, & Walker, 1998).

Although with different formulations, we find in the literature some authors adopting the regression solution for the 3D/2D registration problem (e.g. Chou & Pizer, 2013; Hoff, Komistek, Stefan, & Walker, 1998). However very few works with such a regression approach use real images in a clinically relevant context. To the best of our knowledge, only one work used a resembling approach to partially solve (specifically) the 3D/2D coronary registration problem (Aksoy, Unal, Demirci, Navab, & Degertekin, 2013). This method decouples rotation and translation estimation into frequency and spatial domain, respectively. In a prior step, they built a library of DRRs obtained by generating different rotational poses of CTA vessels. These templates are compared with the segmented X-ray vessels in Fourier domain and the closest DRR found is used to compute the similarity measure to estimate the translation component. As for the majority of the 3D/2D coronary registration methods, the similarity measure (iteratively optimized) depends on the distance between 3D and 2D vessel centerline (Ruijters et al., 2009).

Regarding the evaluation of registration, most of 3D/2D coronary registration studies in the literature used clinical data (e.g., Baka et al., 2013; Metz et al., 2011, 2013; Metz, Schaap, Klein, Neefjes, et al., 2009; Metz, Schaap, Klein, Weustink, et al., 2009), and some used both clinical and computer simulated data (Baka et al., 2014; Rivest-Hénault et al., 2012; Ruijters et

al., 2009). Turgeon et al. (2005) considered computer simulated data generated from an excised human heart and Aksoy et al.(2013) considered images from a phantom vessel but also clinical data. Even though visual inspection was considered (Aksoy et al., 2013; Ruijters et al., 2009), a quantitative evaluation was performed by all authors mentioned. Some computed 3D registration errors, but generally the evaluation metric was based on the projection distance (Van De Kraats et al., 2005), i.e. a 2D distance between the projection of a 3D centerline (from CTA or CTA model) at registration position, and the projection of the same centerline at gold standard position. For clinical validation, vessel centerlines were manually annotated on X-rays and served as the ground truth, rather than building a gold standard system to compute the ground truth transformation as we did and present in this work.

Some authors considered non-rigid registration in temporal and spatial domain caused by cardiac and respiratory motion (Baka et al., 2013, 2014; Metz et al., 2011, 2013; Rivest-Hénault et al., 2012) or rigid registration at multiple points in the cardiac cycle (Metz, Schaap, Klein, Neefjes, et al., 2009). Other authors presented rigid solutions where coronary matching was performed in images acquired at one point in the cardiac cycle, i.e. at the same cardiac phase (Metz et al., 2009; Ruijters et al., 2009; Turgeon et al., 2005); a recent work developed a rigid method without having this source of temporal misalignment into account (Aksoy et al., 2013). Similarly to other authors (Metz et al., 2009; Ruijters et al., 2009; Turgeon et al., 2005), we registered coronary CTA to X-ray angiography considering a rigid transformation. The presence of non-rigid deformations due to cardiac and respiratory motion is thus neglected. However, we took this into account when creating the gold standard, by selecting the X-ray time frame that most accurately matched the CTA-derived coronary artery model for further registration experiments. This manual selection would have to be replaced by an automated procedure when implementing this method in clinical practice.

Note that in principle the registration-by-regression method could be extended to a non-rigid registration. For instance, instead of computing the features on the entire images, we could divide the images and compute the features on small patches, i.e. on small sub-blocks of the images. This could be an interesting direction for future research, and could be of interest also for several image registration applications other than the currently studied.

A final limitation of this experiment is related to the tuning of parameters. The topology of MLPs for all transformation parameters was obtained by a tuning process performed only for one of the transformation parameters. This procedure was motivated by time constraints since, for features in the principal components space, we tried different cumulative percentage variances and for each of those we still tried different numbers of hidden units. A better solution might be obtained by doing an exhaustive search for all parameters.

5 Conclusion

We performed the interpolation of real X-rays with the regression model obtained by a set of DRRs in order to integrate 3D CTA with X-rays angiographies. For the registration-by-regression method we searched for an adequate feature set, which means a good learning (attaining an accurate model) and, simultaneously, to be robust enough to deal with the inherent differences between real and simulated. The information introduced into the regression model, by features like moments, allowed a very good learning and an accurate training performance. However, they are not adequate for the registration of the X-ray images. On the contrary, features like eigenvectors had a robust behavior to the differences between simulated and real images, but they did not contain sufficient information for a good learning.

Thorough quantitative validation of a registration method is essential for clinical usage. The accuracy assessment requires the knowledge of a gold standard or ground truth registration, which is difficult to achieve with clinical images (Hawkes, 2001). The construction of a gold standard for registration of 3D CTA with 2D X-ray of the coronary arteries was an important contribution of this work.

The registration-by-regression method for the automatic integration of coronary CTA with X-ray angiographies was not highly accurate with the tested feature sets (the best one with a median mTRE of 15mm). It does therefore not provide yet a complete solution to the problem but it could be an initialization solution to replace the need of manual initialization. More research into novel and robust image features is needed to make the framework suitable for clinical practice. Interesting directions to explore could be the use of vessels' centerlines and vessels' skeletons (Elizabeth Bullitt et al., 1999). The recent works of Chou, Frederick, Mageras, & Chang (2013) and Chou & Pizer (2013) also suggest some interesting strategies to improve accuracy, like multi-scale regressions by a "hierarchical training" (i.e. large to small scales of transformation ranges for training), and correction of intensity differences between simulated and real images by histogram matching.

Additionally, it would be interesting to test the registration-by-regression framework in other applications, where training data is available that is representative for the test data. Some examples are 3D/2D registration for image-guided needle biopsy interventions in the spine (van de Kraats, 2005; van de Kraats et al., 2006) and motion estimation from real-time image streams (Luca, Tanner, & Székely, 2012).

Acknowledgments

The authors are grateful to Dr. Theo van Walsum for the data and, together with Dr. Nóra Baka, for the help on the development of the gold standard of this work.

References

Aksoy, T., Unal, G., Demirci, S., Navab, N., & Degertekin, M. (2013). Template-based CTA to x-ray angio rigid registration of coronary arteries in frequency domain with automatic x-ray segmentation. *Medical Physics*, *40*(10), 101903. doi:10.1118/1.4819938

Baka, N., Metz, C. T., Schultz, C. J., van Geuns, R.-J., Niessen, W. J., & van Walsum, T. (2014). Oriented gaussian mixture models for nonrigid 2D/3D coronary artery registration. *IEEE Transactions on Medical Imaging*, *33*(5), 1023–34. doi:10.1109/TMI.2014.2300117

Baka, N., Metz, C. T., Schultz, C., Neefjes, L., van Geuns, R. J., Lelieveldt, B. P. F., Niessen, W.S., van Walsum, T., de Bruijne, M. (2013). Statistical coronary motion models for 2D+t/3D registration of X-ray coronary angiography and CTA. *Medical Image Analysis*, *17*(6), 698–709. doi:10.1016/j.media.2013.03.003

Banks, S. a, & Hodge, W. a. (1996). Accurate measurement of three-dimensional knee replacement kinematics using single-plane fluoroscopy. *IEEE Transactions on Bio-Medical Engineering*, *43*(6), 638–49. doi:10.1109/10.495283

Bullitt, E., Liu, A., Aylward, S. R., Coffey, C., Stone, J., Mukherji, S. K., Muller, Keith E., Pizer, S. M. (1999). Registration of 3D Cerebral Vessels with 2D Digital Angiograms. *Academic Radiology*, *6*(9), 539–546.

Chou, C.-R., Frederick, B., & Liu, X. (2011). Claret: A fast deformable registration method applied to lung radiation therapy. In *Fourth International (MICCAI) Workshop on Pulmonary Image Analysis* (pp. 113–124).

Chou C-R, Frederick B, Mageras G, Chang S, Pizer S. 2013. 2D/3D image registration using regression learning. Comput Vis Image Underst; 117:1095–1106. doi:10.1016/j.cviu.2013.02.009

Chou, C.-R., & Pizer, S. (2013). Real-time 2D/3D deformable registration using metric learning. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial*

Intelligence and Lecture Notes in Bioinformatics) (Vol. 7766 LNCS, pp. 1–10). doi:10.1007/978-3-642-36620-8_1

Cyr, C. M., Kamal, A. F., Sebastian, T. B., & Kimia, B. B. (2000). 2D-3D registration based on shape matching. In *Proceedings IEEE Workshop on Mathematical Methods in Biomedical Image Analysis. MMBIA-2000 (Cat. No.PR00737)* (pp. 198–203). IEEE Comput. Soc. doi:10.1109/MMBIA.2000.852378

Fitzpatrick, J. M. (2001). Detecting Failure, Assessing Success. In J. V. Hajnal, D. J. Hawkes, & D. L. G. Hill (Eds.), *Medical image registration, Biomedical engineering series*. CRC Press LLC.

Fitzpatrick, J. M., & West, J. B. (2001). The distribution of target registration error in rigid-body point-based registration. *IEEE Transactions on Medical Imaging*, *20*(9), 917–27. doi:10.1109/42.952729

Frangi, A. F., Niessen, W. J., Vincken, K. L., & Viergever, M. A. (1998). Multiscale vessel enhancement filtering. In W. M. Wells, A. Colchester, & S. L. Delp (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI'98* (Vol. 1496, pp. 130–137). Lecture Notes in Computer Science.

Gonzalez, R. C., & Woods, R. E. (2008). *Digital Image Processing* (3rd ed.). Upper Saddle River, NJ, USA: Prentice Hall.

Gouveia, A. R., Metz, C., Freire, L., & Klein, S. (2012a). 3D-2D image registration by nonlinear regression. In *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)* (pp. 1343–1346). IEEE. doi:10.1109/ISBI.2012.6235814

Gouveia, A. R., Metz, C., Freire, L., & Klein, S. (2012b). Comparative Evaluation of Regression Methods for 3D-2D Image Registration. In A. E. Villa, W. Duch, P. Érdi, F. Masulli, & G. Palm (Eds.), *Artificial Neural Networks and Machine Learning – ICANN 2012* (Vol. 7553, pp. 238–245). Springer Berlin Heidelberg. doi:10.1007/978-3-642-33266-1_30

Hair, J. F., Black, B., Babin, B., Anderson, R. E., & Tatham, R. L. (2005). *Multivariate Data Analysis* (6th ed., p. 928). Upper Saddle River, NJ, USA: Prentice Hall.

Hawkes, D. J. (2001). Registration Methodology: Introduction. In J. V. Hajnal, D. J. Hawkes, & D. L. G. Hill (Eds.), *Medical image registration, Biomedical engineering series*. CRC Press LLC.

Haykin, S. (1999). *Neural networks: a comprehensive foundation* (2nd. ed.). Delhi: Prentice hall.

Hermans, J., Claes, P., Bellemans, J., Vandermeulen, D., & Suetens, P. (2007). Robust initialization for 2D/3D registration of knee implant models to single-plane fluoroscopy. In *Proceedings of SPIE* (Vol. 6512, pp. 651208–651208–12). SPIE. doi:10.1117/12.708163

Hoff, W. A., Komistek, R. D., Dennis, D. A., Gabriel, S. M., & Walker, S. A. (1998). Three-dimensional determination of femoral-tibial contact positions under in vivo conditions using fluoroscopy. *Clinical Biomechanics*, *13*, 455–472.

Imamura, H., Ida, N., & Sugimoto, N. (2002). Registration of preoperative CTA and intraoperative fluoroscopic images for assisting aortic stent grafting. In T. Dohi & R. Kikinis (Eds.), *Medical Image Computing and Computer-Assisted Intervention — MICCAI 2002* (Vol. 2489, pp. 477–484). Lecture Notes in Computer Science. doi:10.1007/3-540-45787-9_60

Jannin, P., Fitzpatrick, J. M., Hawkes, D. J., Pennec, X., Shahidi, R., & Vannier, M. W. (2002). Validation of medical image processing in image-guided therapy. *IEEE Transactions on Medical Imaging*, *21*(12), 1445–9. doi:10.1109/TMI.2002.806568

Kasper, D. L., Braunwald, E., Fauci, A. S., Hauser, S. L., Longo, D. L., & Jameson, J. L. (2005). *Harrison's Principles of Internal Medicine* (2nd. ed.). New York: McGraw-Hill.

Liao, S. X. (1993, March). *On image analysis by moments*. University of Manitoba, Canada.

Liu, A., Bullitt, E., & Pizer, S. M. (1998). 3D / 2D Registration Via Skeletal Near Projective Invariance in Tubular Objects. In *First International Conference on Medical Image Computing and Computer-Assisted Intervention—MICCAI'98* (pp. 952–963). Cambridge MA, USA: Springer.

Luca, V. De, Tanner, C., & Székely, G. (2012). Speeding-up image registration for repetitive motion scenarios. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on* (pp. 1355–1358).

Markelj, P., Tomaževič, D., Likar, B., & Pernuš, F. (2012). A review of 3D/2D registration methods for image-guided interventions. *Medical Image Analysis*, *16*(3), 642–61. doi:10.1016/j.media.2010.03.005

MatLab. (2010). Natick, MA, USA: The Mathworks Inc.

Metz, C. T. (2011). *Coronary Motion Modeling for CTA to X-Ray Angiography Registration*.

Metz, C. T., Schaap, M., Klein, S., Baka, N., Neefjes, L. A., Schultz, C. J., … van Walsum, T. (2013). Registration of 3D+t coronary CTA and monoplane 2D+t X-ray angiography. *IEEE Transactions on Medical Imaging*, *32*(5), 919–31. doi:10.1109/TMI.2013.2245421

Metz, C. T., Schaap, M., Klein, S., Neefjes, L. a, Capuano, E., Schultz, C., … Niessen, W. J. (2009). Patient specific 4D coronary models from ECG-gated CTA data for intra-operative dynamic alignment of CTA with X-ray images. In *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention* (Vol. 12, pp. 369–76).

Metz, C. T., Schaap, M., Klein, S., Rijnbeek, P., Neefjes, L., Mollet, N., … Walsum, T. Van. (2011). Alignment of 4D Coronary CTA with Monoplane X-Ray Angiography. *6th International Workshop, AE-CAI (Augmented Environments for Computer-Assisted Interventions), Held in Conjunction with MICCAI 2011*, 106–116.

Metz, C. T., Schaap, M., Klein, S., Weustink, A. C., Mollet, N. R., Schultz, C., … Niessen, W. J. (2009). GPU accelerated alignment of 3-D CTA with 2-D X-ray data for improved guidance in coronary interventions. In *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (pp. 959–962). IEEE. doi:10.1109/ISBI.2009.5193213

Metz, C. T., Schaap, M., Weustink, a. C., Mollet, N. R., van Walsum, T., & Niessen, W. J. (2009). Coronary centerline extraction from CT coronary angiography images using a minimum cost path approach. *Medical Physics*, *36*(12), 5568. doi:10.1118/1.3254077

Penney, G. P., Weese, J., Little, J. a, Desmedt, P., Hill, D. L., & Hawkes, D. J. (1998). A comparison of similarity measures for use in 2-D-3-D medical image registration. *IEEE Transactions on Medical Imaging*, *17*(4), 586–95. doi:10.1109/42.730403

Ramcharitar, S., Patterson, M. S., van Geuns, R. J., van Meighem, C., & Serruys, P. W. (2008). Technology Insight: magnetic navigation in coronary interventions. *Nature Clinical Practice. Cardiovascular Medicine*, *5*(3), 148–56. doi:10.1038/ncpcardio1095

Ramcharitar, S., van der Giessen, W. J., van der Ent, M., Serruys, P. W., & van Geuns, R. J. (2011). Randomized comparison of the magnetic navigation system vs. standard wires in the treatment of bifurcations. *European Heart Journal*, *32*(12), 1479–83. doi:10.1093/eurheartj/ehq123

Rivest-Hénault, D., Sundar, H., & Cheriet, M. (2012). Nonrigid 2D/3D registration of coronary artery models with live fluoroscopy for guidance of cardiac interventions. *IEEE Transactions on Medical Imaging*, *31*(8), 1557–72. doi:10.1109/TMI.2012.2195009

Ruijters, D., ter Haar Romeny, B. M., & Suetens, P. (2009). Vesselness-based 2D-3D registration of the coronary arteries. *International Journal of Computer Assisted Radiology and Surgery*, *4*(4), 391–7. doi:10.1007/s11548-009-0316-z

Schneider, M., & Sundar, H. (2010). Automatic global vessel segmentation and catheter removal using local geometry information and vector field integration. In *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (pp. 45–48). IEEE. doi:10.1109/ISBI.2010.5490419

Serruys, P. W. (2006). Fourth annual American College of Cardiology international lecture: a journey in the interventional field. *Journal of the American College of Cardiology*, *47*(9), 1754–68. doi:10.1016/j.jacc.2005.12.051

Shlens, J. (2005). A Tutorial on Principal Component Analysis. *Measurement*, *51*(10003), 52. doi:10.1.1.115.3503

Soon, K. H., Selvanayagam, J. B., Cox, N., Kelly, A.-M., Bell, K. W., & Lim, Y. L. (2007). Percutaneous revascularization of chronic total occlusions: review of the role of invasive and non-invasive imaging modalities. *International Journal of Cardiology*, *116*(1), 1–6. doi:10.1016/j.ijcard.2006.03.023

Turgeon, G.-A., Lehmann, G., Guiraudon, G., Drangova, M., Holdsworth, D., & Peters, T. (2005). 2D-3D registration of coronary angiograms for cardiac procedure planning and guidance. *Medical Physics*, *32*(12), 3737. doi:10.1118/1.2123350

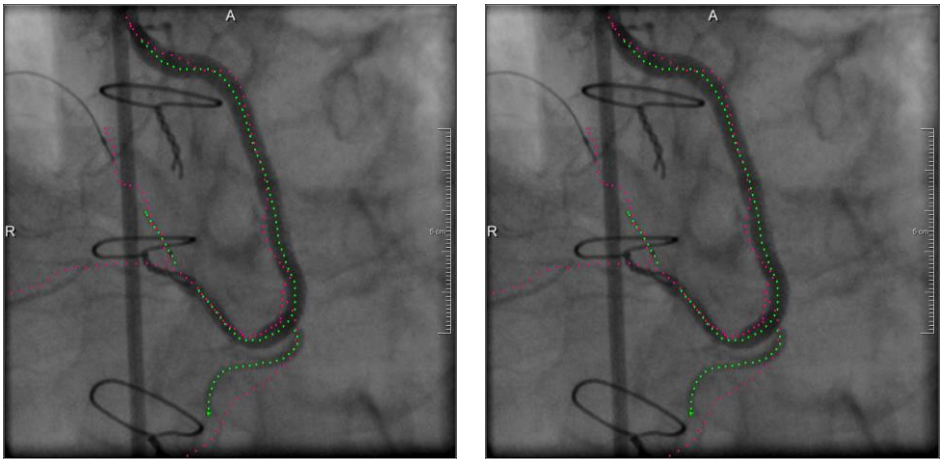
Van de Kraats, E. B. (2005). *3D Rotational X-Ray Guidance for Surgical Interventions*.

Van De Kraats, E. B., Penney, G. P., Tomaževič, D., & Van Walsum, T. (2005). Standardized Evaluation Methodology for 2-D – 3-D Registration. *IEEE Transactions on Medical Imaging*, *24*(9), 1177–1189.

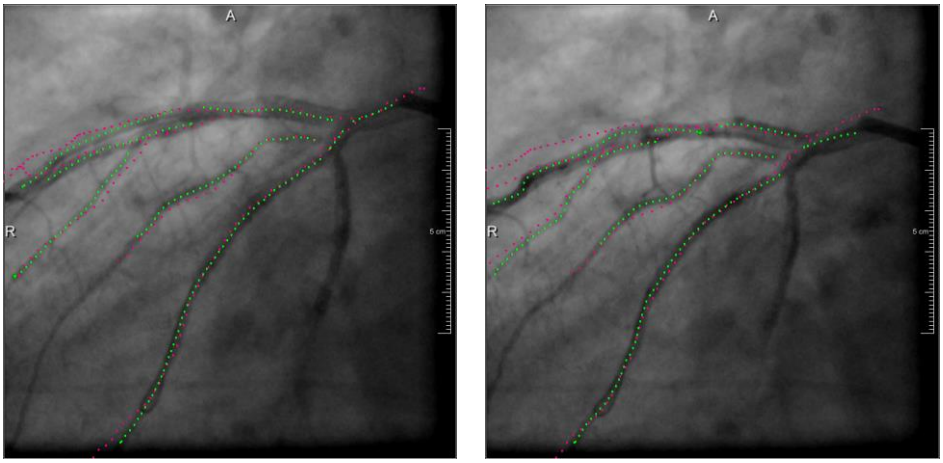
Van de Kraats, E. B., van Walsum, T., Verlaan, J.-J., Voormolen, M. H. J., Mali, W. P. T. M., & Niessen, W. J. (2006). Three-Dimensional Rotational X-Ray Navigation for Needle Guidance in Percutaneous Vertebroplasty: An Accuracy Study. *Spine*, *31*(12), 1359–1364. doi:10.1097/01.brs.0000218580.54036.1b

Zhang, J., Ge, Y., Ong, S., Chui, C., Teoh, S., & Yan, C. (2008). Rapid surface registration of 3D volumes using a neural network approach. *Image and Vision Computing*, *26*(2), 201–210. doi:10.1016/j.imavis.2007.04.003

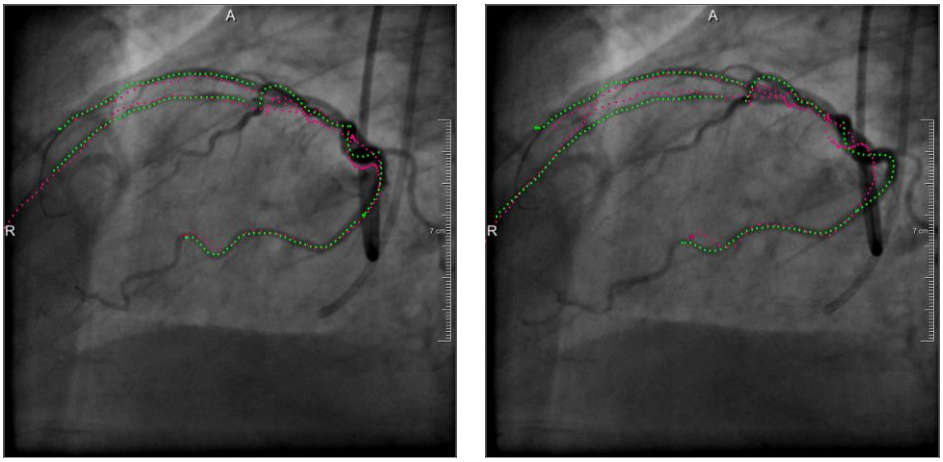
Appendix D



X-ray 1

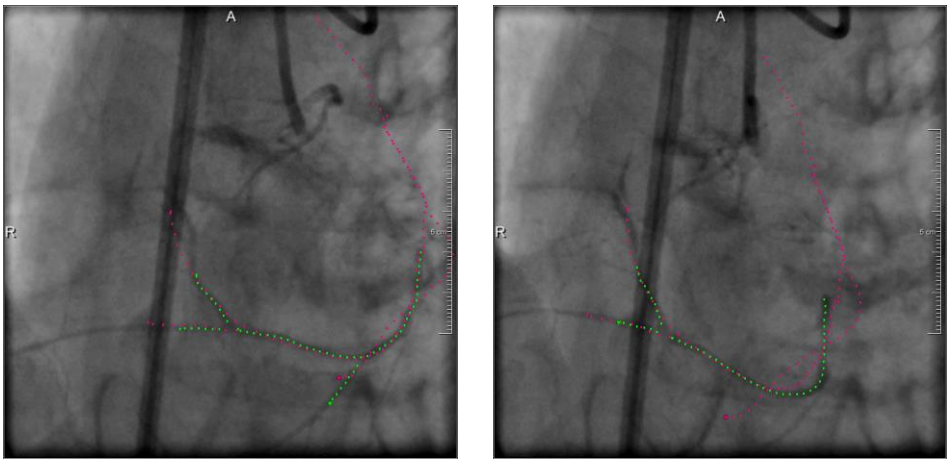


X-ray 2

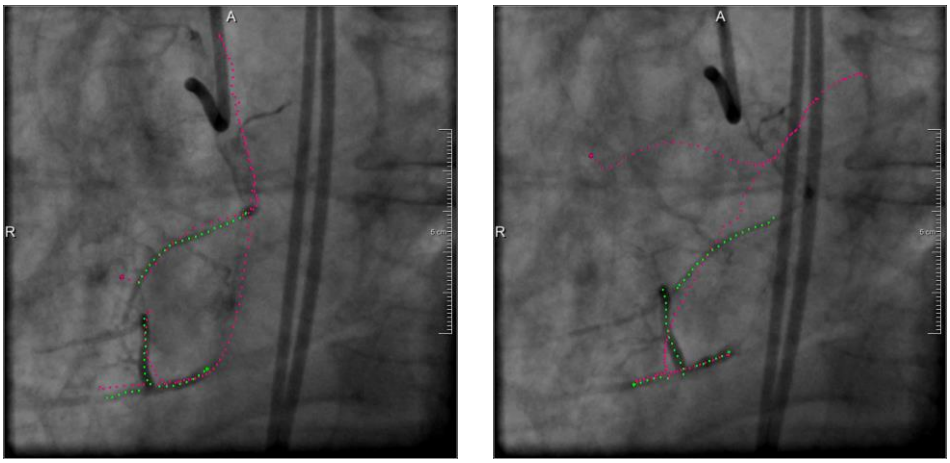


X-ray 3.1

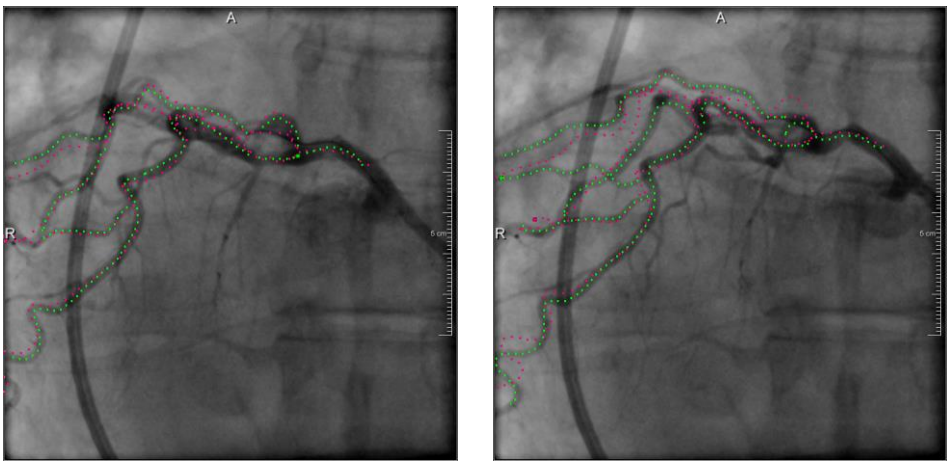
Figure D.12 Best (left) and worst (right) alignments according to the ranking established for time-sequences 1, 2 and 3.1. The images are X-ray frames with manual annotated centerline in green and the closest projected centreline in pink.



X-ray 4.1

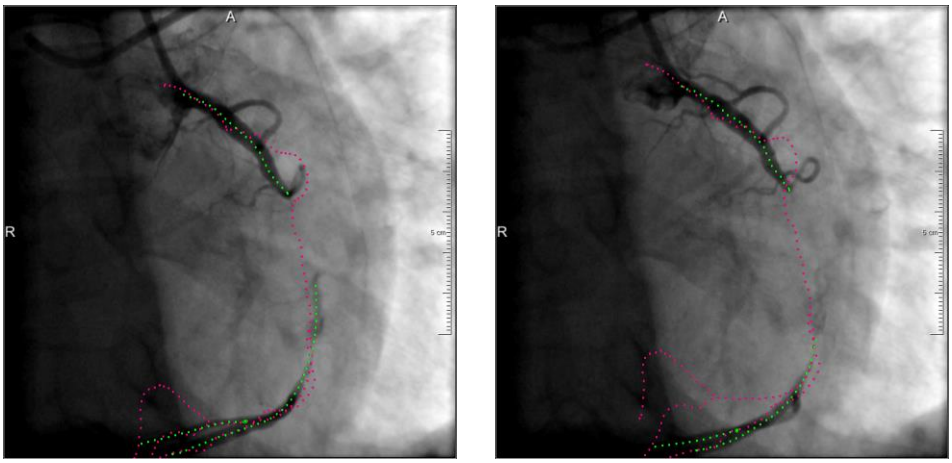


X-ray 4.2

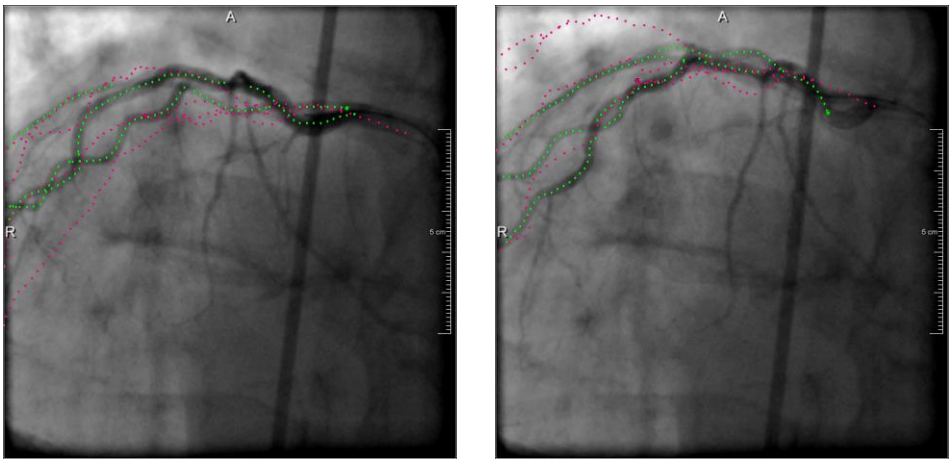


X-ray 5

Figure D.13 Best (left) and worst (right) alignments according to the ranking established for time-sequences 4.1, 4.2 and 5. The images are X-ray frames with manual annotated centerline in green and the closest projected centreline in pink.



X-ray 6



X-ray 7

Figure D.14 Best (left) and worst (right) alignments according to the ranking established for time-sequences 6 and 7. The images are X-ray frames with manual annotated centerline in green and the closest projected centreline in pink.

Appendix E

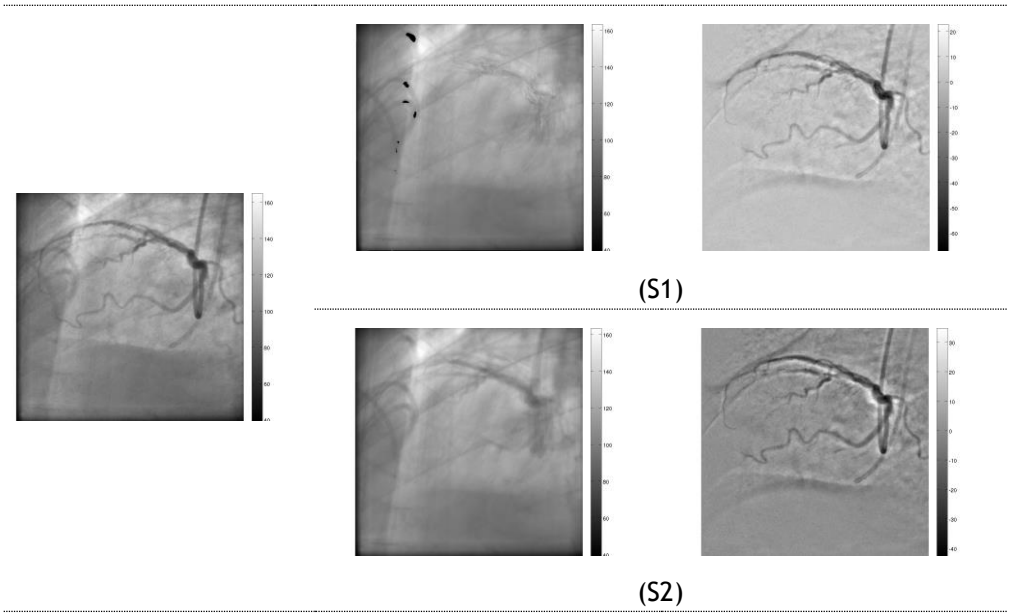


Figure E.15 Background subtraction with the average of the image backgrounds (S1) and with the average of entire images (S2), for the x-ray 3.1: (a) original x-ray, (b) average over time, (c) resulting subtracted image.

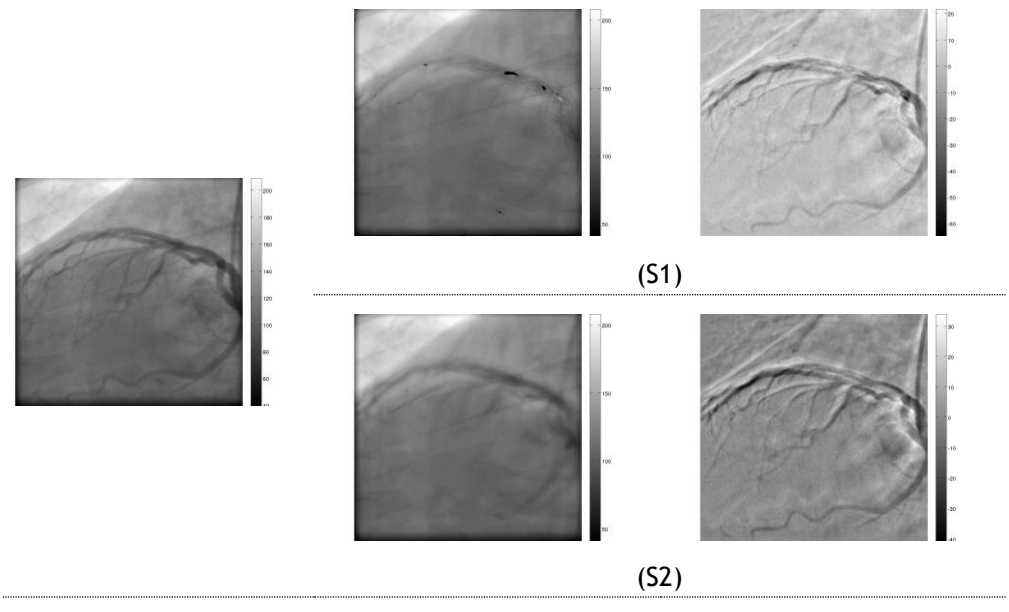


Figure E.16 Background subtraction with the average of the image backgrounds (S1) and with the average of entire images (S2), for the x-ray 3.2: (a) original x-ray, (b) average over time, (c) resulting subtracted image.

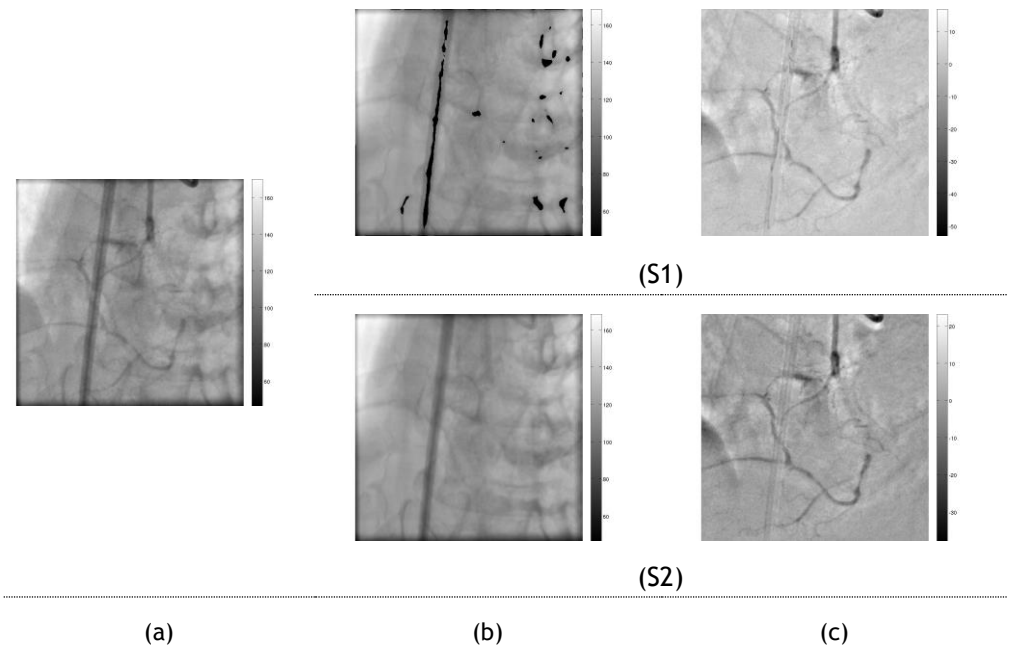


Figure E.17 Background subtraction with the average of the image backgrounds (S1) and with the average of entire images (S2), for the x-ray 4.1: (a) original x-ray, (b) average over time, (c) resulting subtracted image.

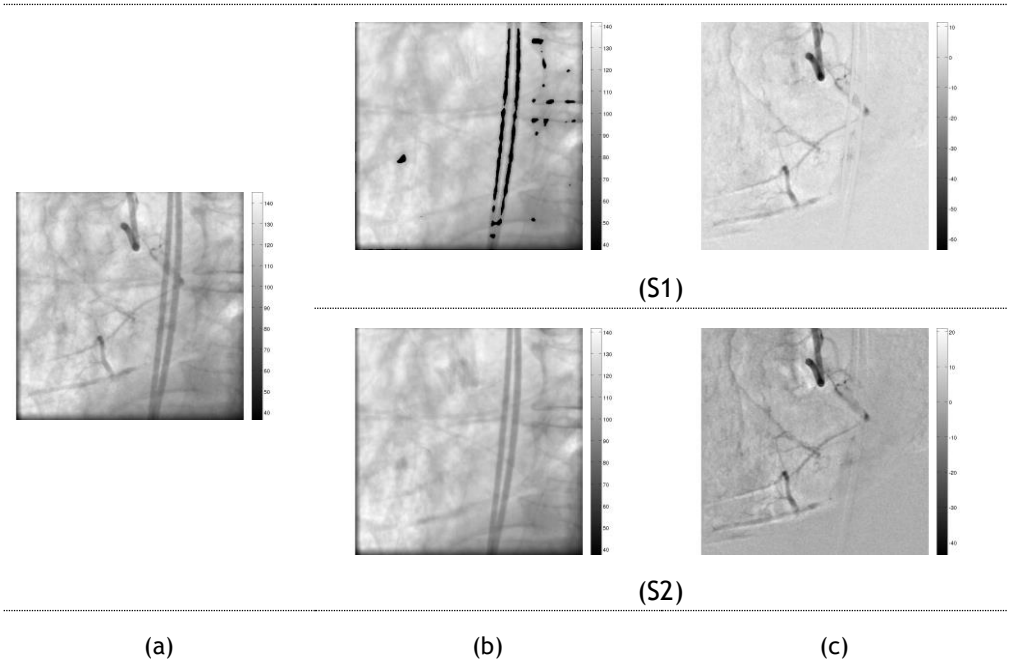


Figure E.18 Background subtraction with the average of the image backgrounds (S1) and with the average of entire images (S2), for the x-ray 4.2: (a) original x-ray, (b) average over time, (c) resulting subtracted image.

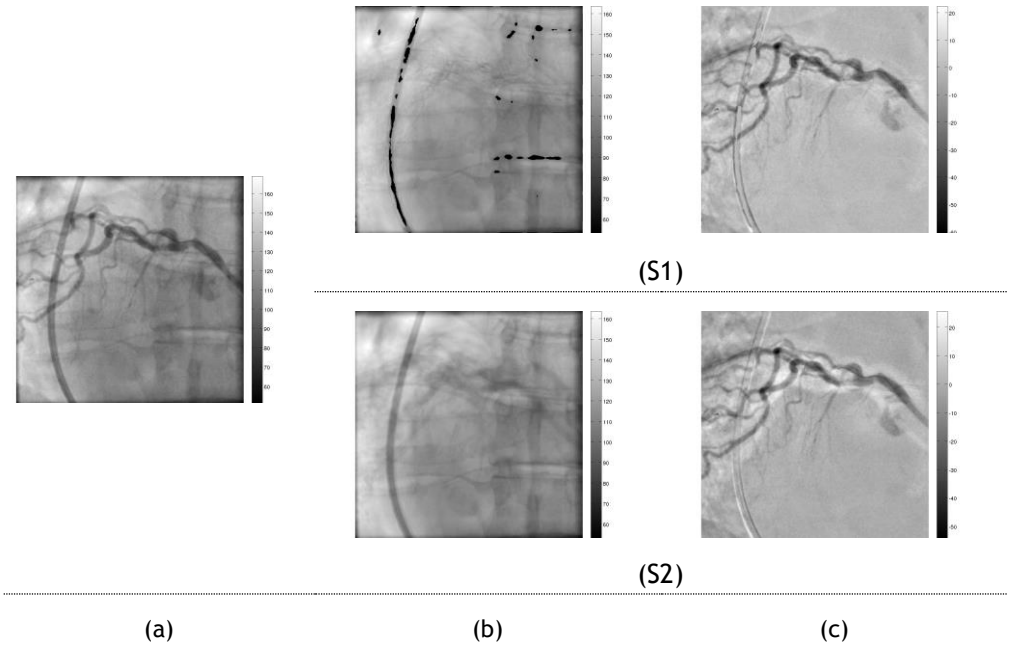


Figure E.19 Background subtraction with the average of the image backgrounds (S1) and with the average of entire images (S2), for the x-ray 5: (a) original x-ray, (b) average over time, (c) resulting subtracted image.

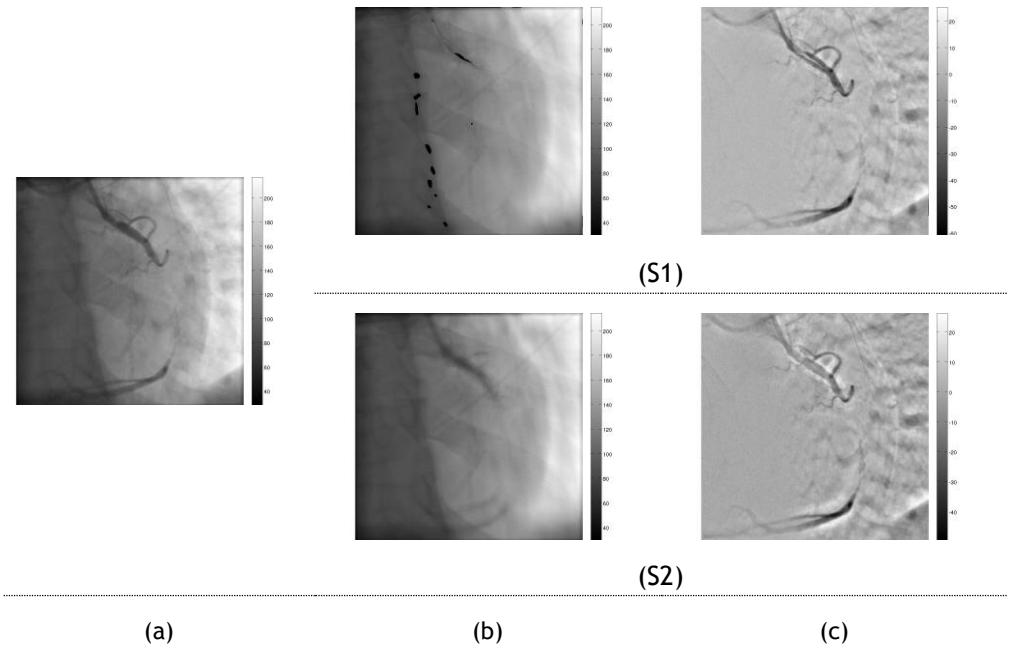


Figure E.20 Background subtraction with the average of the image backgrounds (S1) and with the average of entire images (S2), for the x-ray 6: (a) original x-ray, (b) average over time, (c) resulting subtracted image.

Appendix F

X-ray 2

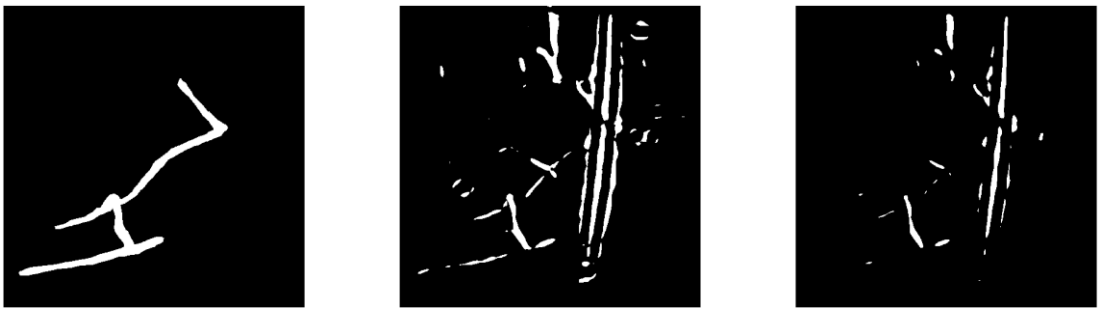
X-ray 3.1

X-ray 3.2

(1) (2) (3)

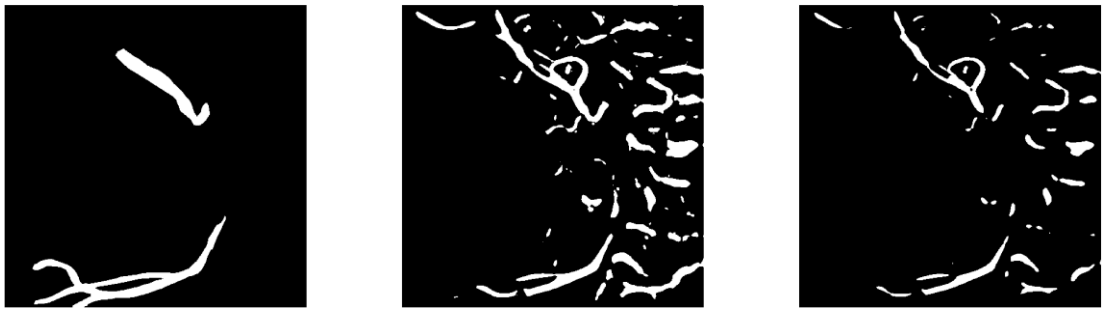
Figure F.21 Manual mask (1), automatic mask corresponding to threshold t1 (2) and automatic mask corresponding to threshold t2 (3) for x-rays 2, 3.1 and 3.2.

X-ray 4.1



X-ray 4.2

X-ray 5

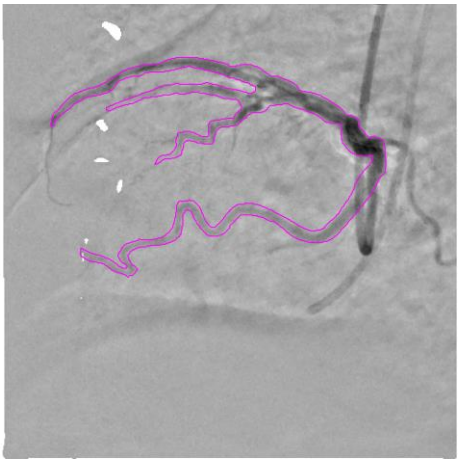


X-ray 6

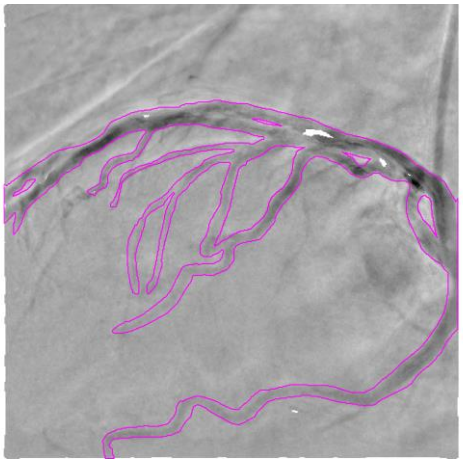
(1) (2) (3)

Figure F.22 Manual mask (1), automatic mask corresponding to threshold t1 (2) and automatic mask corresponding to threshold t2 (3) for x-rays 4.1, 4.2, 5 and 6.

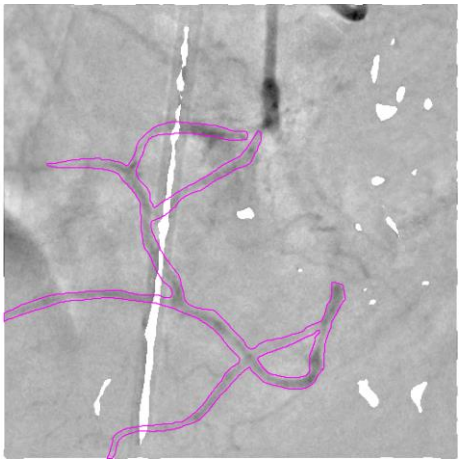
Appendix G



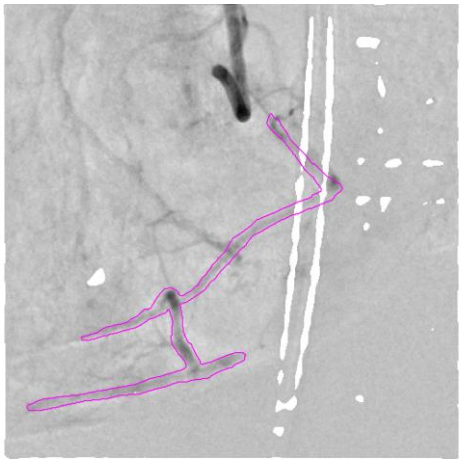
X-ray 3.1



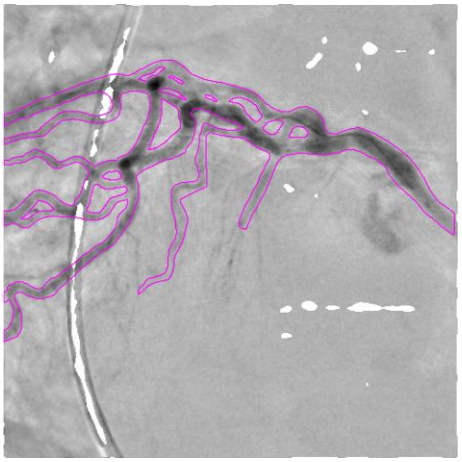
X-ray 3.2



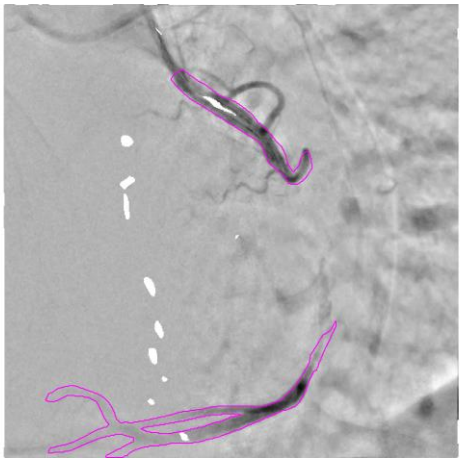
X-ray 4.1



X-ray 4.2



X-ray 5



X-ray 6

Figure F.23 Manual masks (colour lines) on images after background subtraction for the x-rays indicated.

Appendix H

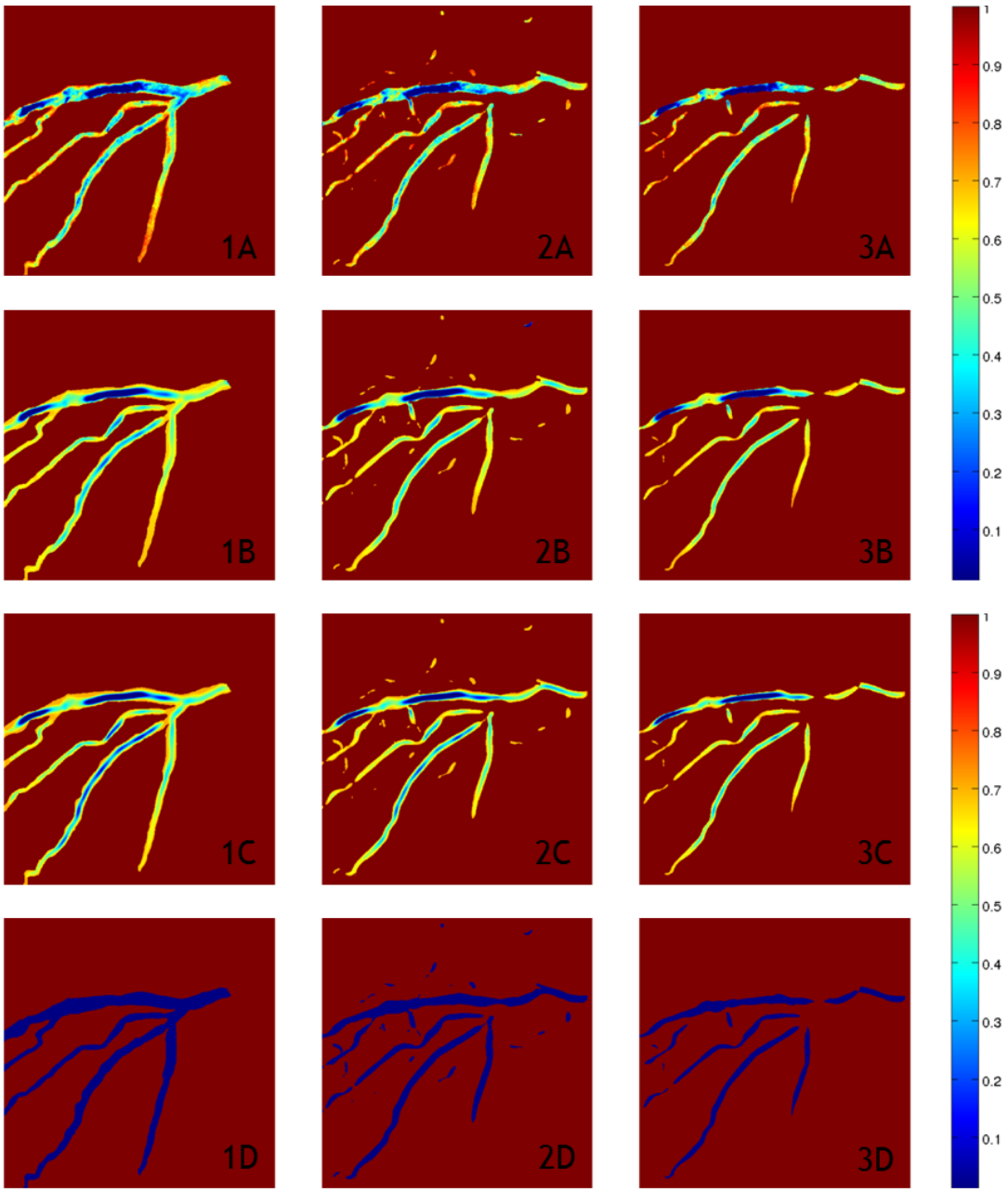


Figure H.16 Vessel segmentation results for X-ray 2 in coloured scale, considering the different combinations used: columns (1,2,3) correspond to the mask options and rows (A,B,C,D) to the intensity values of pixels within the masks.

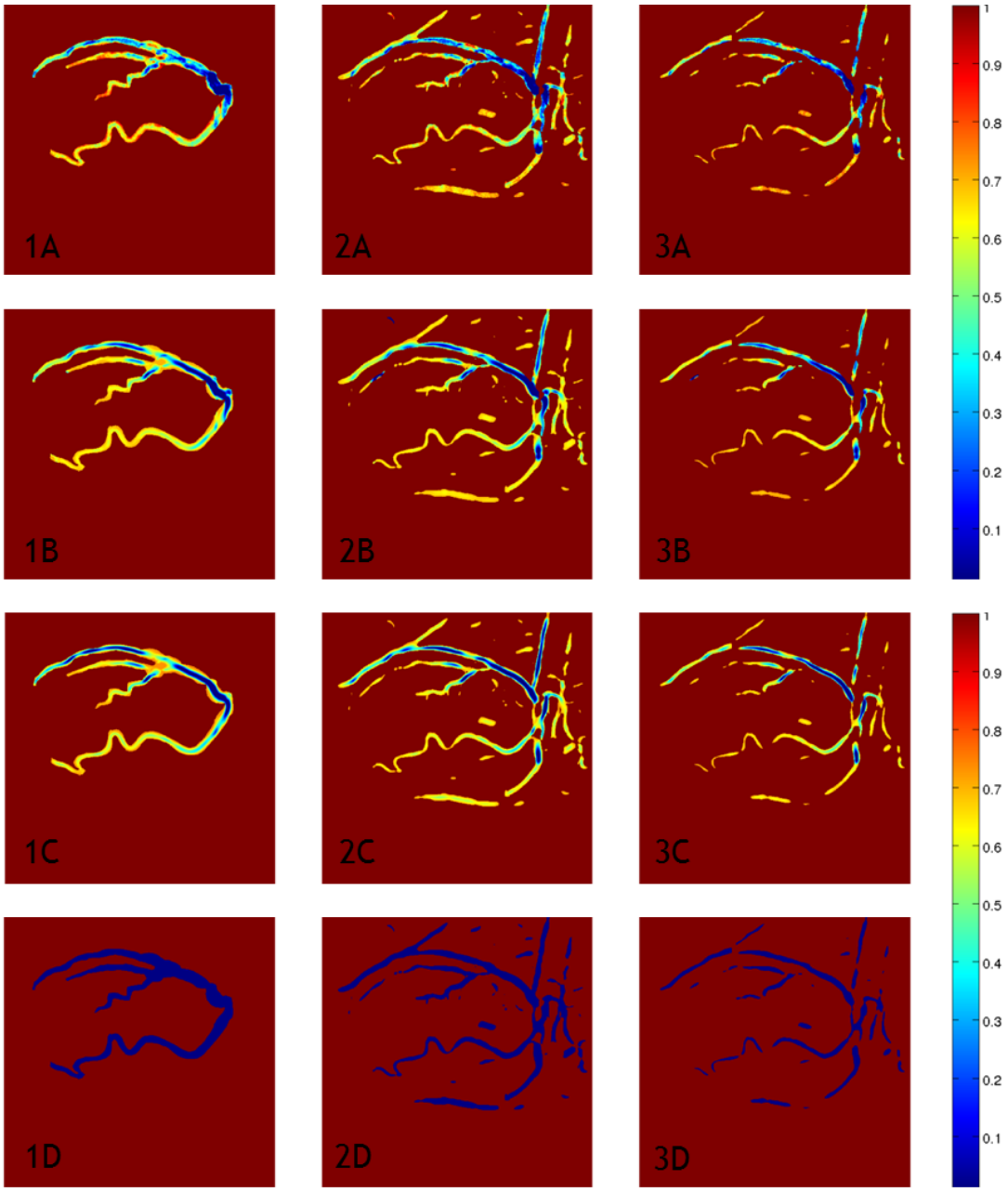


Figure H.17 Vessel segmentation results for X-ray 3.1 in coloured scale, considering the different combinations used: columns (1,2,3) correspond to the mask options and rows (A,B,C,D) to the intensity values of pixels within the masks.

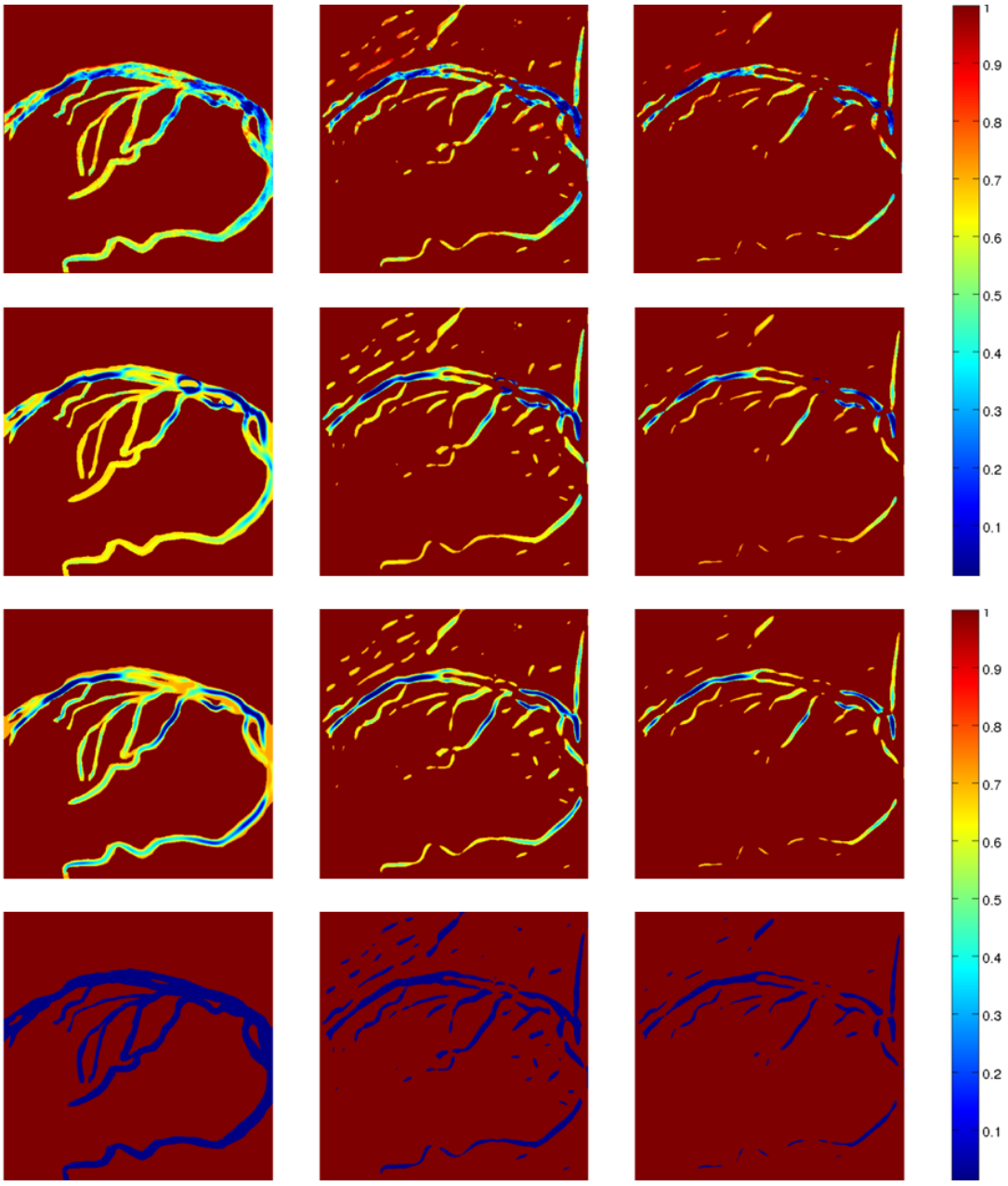


Figure H.18 Vessel segmentation results for X-ray 3.2 in coloured scale, considering the different combinations used: columns (1,2,3) correspond to the mask options and rows (A,B,C,D) to the intensity values of pixels within the masks.

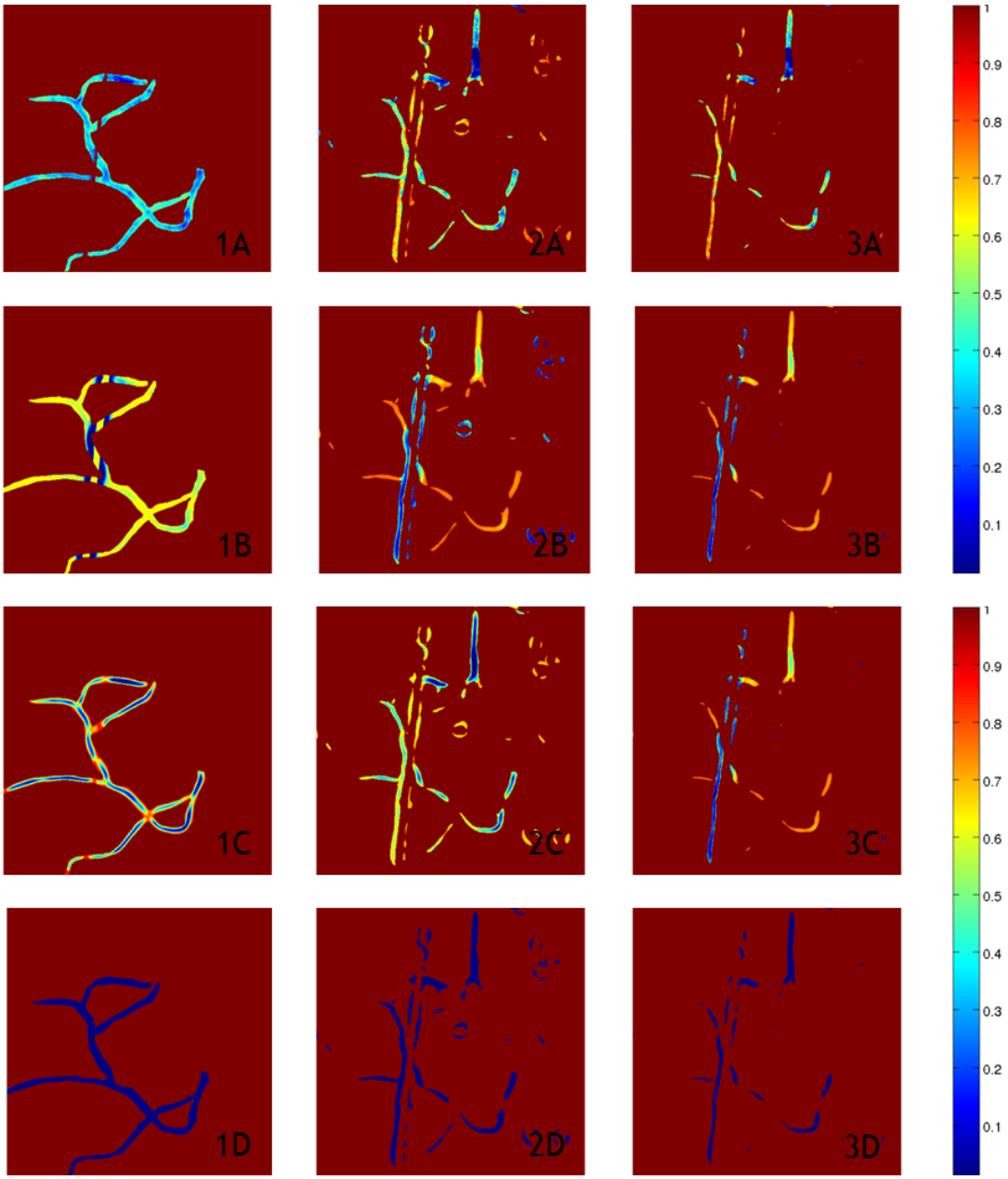


Figure H.19 Vessel segmentation results for X-ray 4.1 in coloured scale, considering the different combinations used: columns (1,2,3) correspond to the mask options and rows (A,B,C,D) to the intensity values of pixels within the masks.

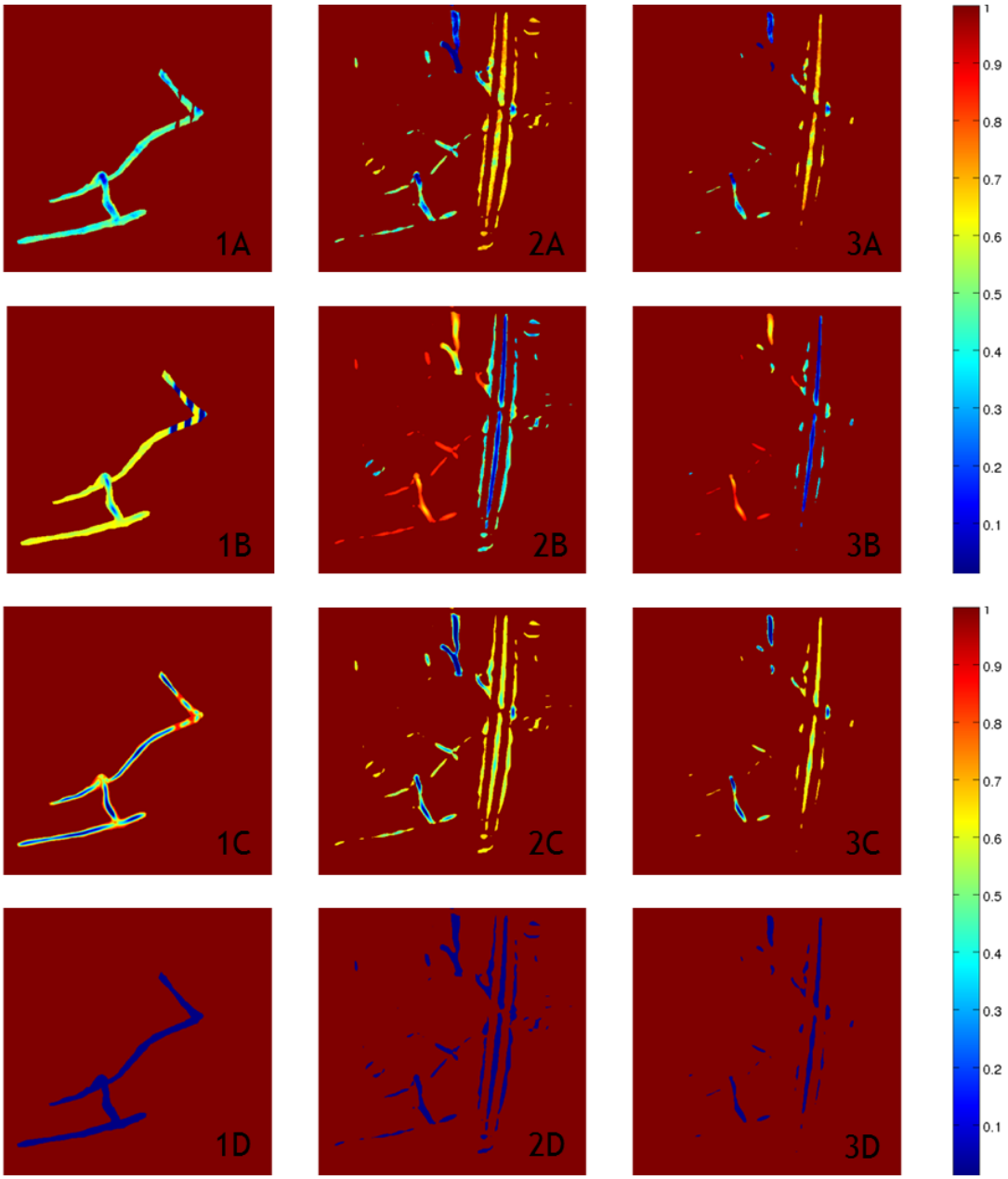


Figure H.20 Vessel segmentation results for X-ray 4.2 in coloured scale, considering the different combinations used: columns (1,2,3) correspond to the mask options and rows (A,B,C,D) to the intensity values of pixels within the masks.

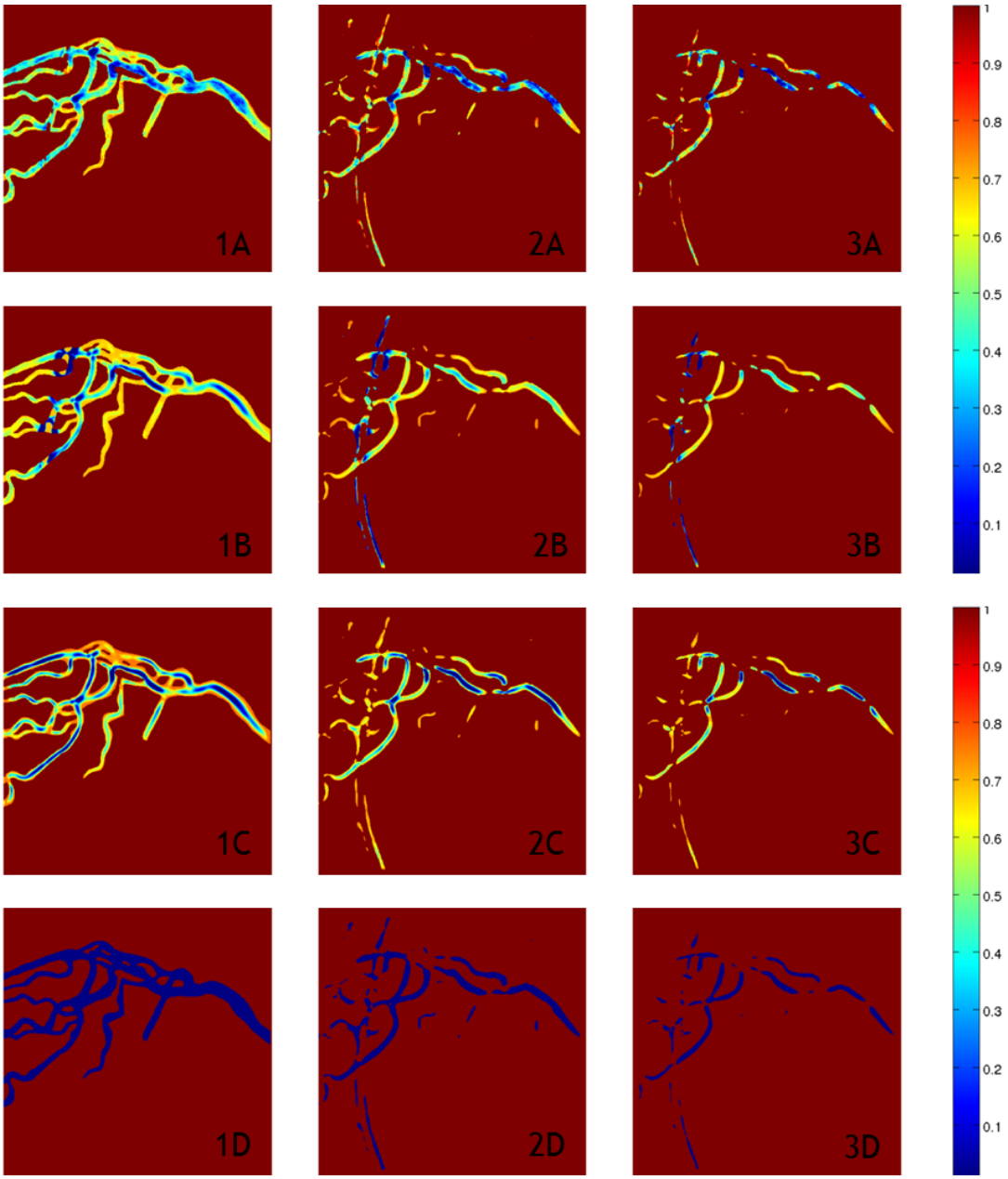


Figure H.21 Vessel segmentation results for X-ray 5 in coloured scale, considering the different combinations used: columns (1,2,3) correspond to the mask options and rows (A,B,C,D) to the intensity values of pixels within the masks.

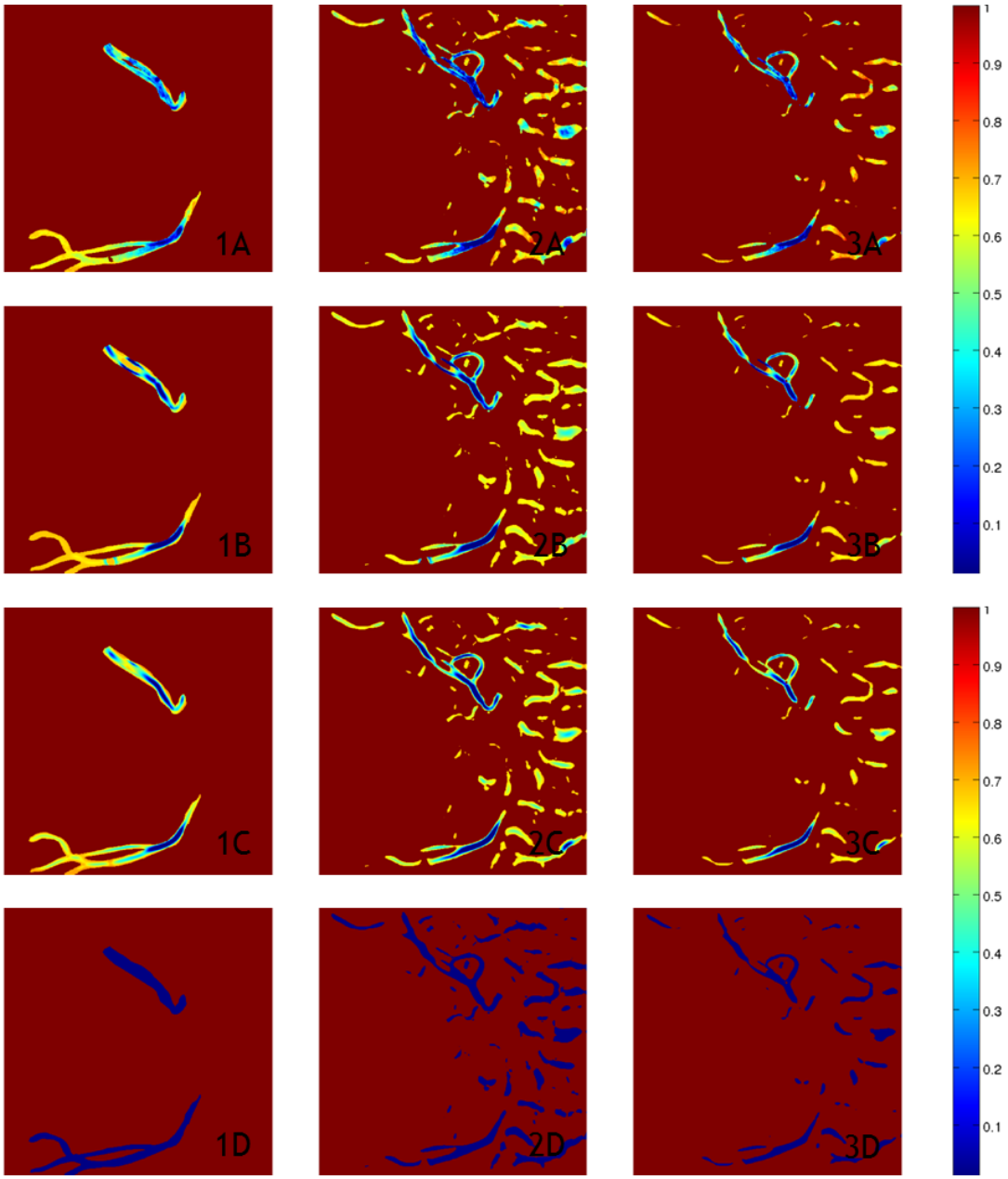


Figure H.22 Vessel segmentation results for X-ray 6 in coloured scale, considering the different combinations used: columns (1,2,3) correspond to the mask options and rows (A,B,C,D) to the intensity values of pixels within the masks.