



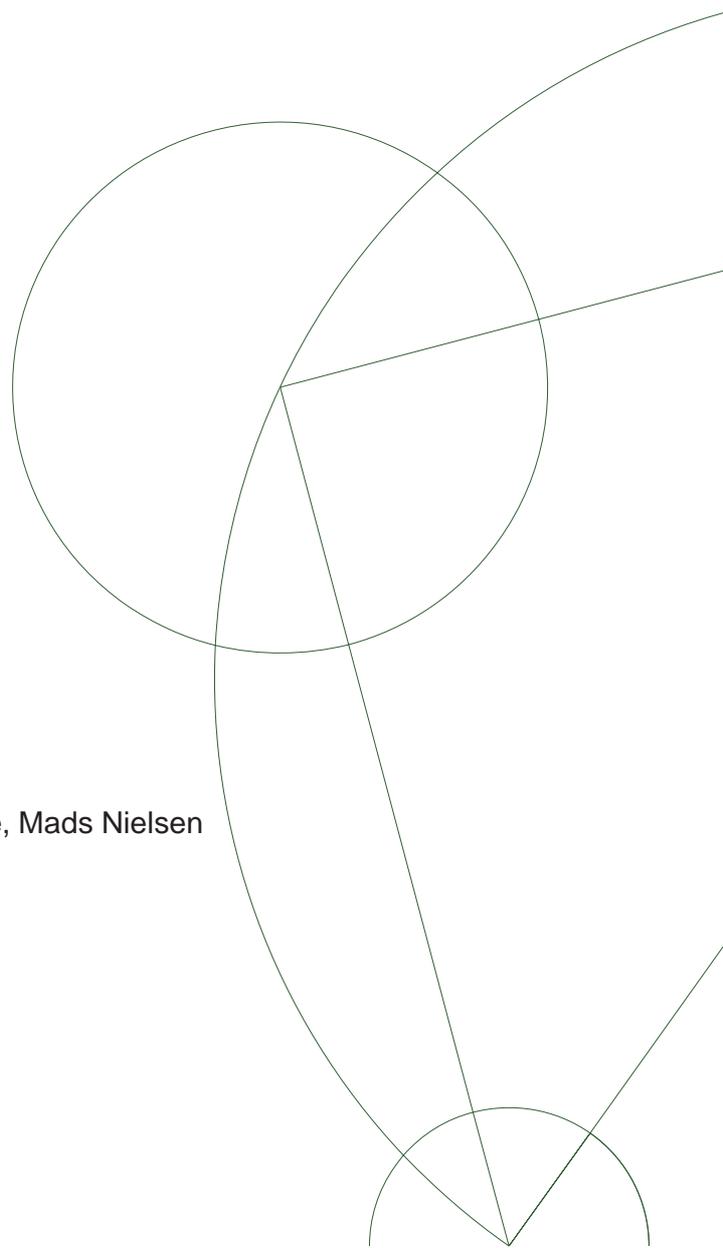
## PhD thesis

Melanie Ganz

# Automated Detection and Analysis of Calcified Deposits and their relation to Cardiovascular Diseases

Academic advisors: Sami Brandt, Marleen de Bruijne, Mads Nielsen

Submitted: 28/02/2011



©2011 by Melanie Ganz.  
All rights reserved.

Department of Computer Science  
*Image Group*  
University of Copenhagen  
Universitetsparken 1  
2100 København Ø  
Denmark  
Telephone: + 45 (0)3532 1400

# Contents

<b>List of Figures</b>	<b>vi</b>
<b>List of Tables</b>	<b>vii</b>
<b>Abstract</b>	<b>ix</b>
<b>Resume</b>	<b>xi</b>
<b>Preface</b>	<b>xiii</b>
<b>List of Publications</b>	<b>xv</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Purpose . . . . .	2
1.2 Outline . . . . .	2
<b>2 Risk Assessment of Cardiovascular Diseases</b>	<b>5</b>
2.1 Causes of Cardiovascular Diseases . . . . .	5
2.2 Introduction to Part I:	
Medical Imaging of Atherosclerosis . . . . .	6
2.2.1 Choice of Modality and Region of Interest in the Body . . . . .	9
2.2.2 Automating the Measurement of Imaging Biomarkers on Lumbar Aortic Radiographs . . . . .	11
2.2.3 Application of Renormalization Group Theory in Medical Imaging . . . . .	15
2.3 Introduction to Part II:	
Diagnostic Tools for Cardiovascular Disease Risk . . . . .	17
2.3.1 Current Metabolic Biomarkers for CVDs . . . . .	17
<b>I Medical Imaging of Lumbar Aortic Calcifications</b>	<b>19</b>
<b>3 Dense Iterative Contextual Pixel Classification Using Kriging</b>	<b>23</b>

3.1	Introduction . . . . .	23
3.2	Problem Description . . . . .	24
3.2.1	Statistical Interpretation . . . . .	25
3.2.2	Kriging . . . . .	26
3.2.3	Application of Kriging . . . . .	26
3.3	DICPC - Dense Iterative Contextual Pixel Classification . . . . .	27
3.3.1	Gaussian Distributions . . . . .	27
3.3.2	Other Distributions . . . . .	28
3.3.3	DICPC algorithm . . . . .	29
3.4	Methods . . . . .	30
3.4.1	Study population . . . . .	30
3.4.2	Evaluation Methods . . . . .	31
3.4.3	Classification and Kriging Settings . . . . .	31
3.4.4	Comparison to other methods . . . . .	31
3.5	Results . . . . .	32
3.6	Discussion and Conclusion . . . . .	32
<b>4</b>	<b>A Texton-based Generative Shape Model and MDL Model Selection for Statistical Analysis of Archipelago-Like Textures</b>	<b>37</b>
4.1	Introduction . . . . .	37
4.2	Statistical Objective . . . . .	39
4.3	Generative Shape Model . . . . .	40
4.3.1	Texton Dictionary . . . . .	40
4.3.2	Texton Grammar Models . . . . .	42
4.3.3	Visiting Order . . . . .	43
4.3.4	Model Selection . . . . .	44
4.4	Sampling from the Posterior with the Shape Prior . . . . .	46
4.5	Experiments . . . . .	46
4.5.1	Modeling different Shape Distributions . . . . .	46
4.5.2	Statistical Shape Segmentation . . . . .	48
4.6	Discussion . . . . .	50
4.7	Conclusion . . . . .	55
<b>5</b>	<b>Application of Renormalization Group Theory to Multi-Scale Image Segmentation Problems</b>	<b>57</b>
5.1	Introduction . . . . .	57
5.2	Notation . . . . .	58
5.3	Multi-scale Image Segmentation . . . . .	59
5.4	Using RGT for Multi-scale Image Segmentation . . . . .	60
5.4.1	Energy Function Model . . . . .	60
5.4.2	Kadanoff's Block Spin Method . . . . .	62
5.5	Experiments and Results . . . . .	64

5.5.1	Segmentation Quality . . . . .	64
5.5.2	Segmentation Time . . . . .	65
5.6	Discussion and Conclusion . . . . .	66
 <b>II Clinical Application of the Imaging of Lumbar Aortic Calcifications</b>		<b>69</b>
<b>6</b>	<b>Abdominal Aortic Calcified Deposits and their Relationship to Mortality in Post Menopausal Women</b>	<b>73</b>
6.1	Introduction . . . . .	74
6.2	Materials and Methods . . . . .	75
6.2.1	Study Population . . . . .	75
6.2.2	Metabolic and Physical Measurements . . . . .	75
6.2.3	Radiographic Analysis . . . . .	76
6.2.4	AAC Markers . . . . .	77
6.2.5	Statistical Analysis . . . . .	79
6.3	Results . . . . .	81
6.4	Discussion . . . . .	84
6.5	Conclusion . . . . .	87
<b>7</b>	<b>Distribution, Size, Shape, Growth Potential and Extent of Abdominal Aortic Calcified Deposits Predict Mortality in Post Menopausal Women</b>	<b>89</b>
7.1	Background . . . . .	90
7.2	Methods . . . . .	91
7.2.1	Subjects . . . . .	91
7.2.2	Markers . . . . .	91
7.2.3	Statistical Analysis . . . . .	94
7.3	Results . . . . .	94
7.4	Discussion . . . . .	97
7.5	Conclusion . . . . .	100
<b>8</b>	<b>Discussion and Conclusions</b>	<b>103</b>
8.1	Summary . . . . .	103
8.2	Discussion and Conclusions . . . . .	104
8.3	Future Work . . . . .	106
 <b>Acknowledgments</b>		<b>119</b>



# List of Figures

2.1	The anatomy of an artery . . . . .	6
2.2	Plaque Development . . . . .	7
2.3	A schematic description of AC24 . . . . .	8
2.4	Abdominal aortic calcifications (AAC) in a lumbar aortic X-ray	10
2.5	How we apply the CVD pipeline . . . . .	11
2.6	The CVD pipeline . . . . .	11
2.7	Example X-ray image . . . . .	14
2.8	A typical pipeline result . . . . .	15
2.9	X-ray of our study population . . . . .	16
3.1	A manual as well as an automatic segmentation of a calcification	30
3.2	Results for different morphological operations and DICPC . . .	33
3.3	Comparison of different methods . . . . .	35
4.1	A spectrum of texture . . . . .	38
4.2	Illustrations of archipelago-like structures . . . . .	38
4.3	Training images and an example of a dictionary . . . . .	41
4.4	The causal neighborhoods . . . . .	42
4.5	Different visiting orders . . . . .	44
4.6	Texture synthesis results I . . . . .	48
4.7	Texture synthesis results II . . . . .	49
4.8	Calcification segmentation result I . . . . .	52
4.9	Calcification segmentation result II . . . . .	53
4.10	Calcification segmentation result III . . . . .	54
5.1	Problem description . . . . .	59
5.2	The Kadanoff Triangulation . . . . .	62
5.3	Man segmentation results . . . . .	65
5.4	Bear segmentation results . . . . .	66
5.5	Plant segmentation results . . . . .	67
5.6	Change in accuracy of the partial solutions . . . . .	68
5.7	Comparison of computational time . . . . .	68

6.1	Manual annotation of an X-ray . . . . .	76
6.2	A schematic overview of the AC24 scoring . . . . .	77
6.3	The simulated area percentage . . . . .	78
6.4	Different AAC markers . . . . .	79
6.5	A schematic overview of the study population . . . . .	81
6.6	X-ray of a participant in the EPI follow-up population . . . . .	83
7.1	An example of a calcification annotation . . . . .	91
7.2	A comparison of AC24, NCD and MACD . . . . .	93
7.3	An overview of the study population . . . . .	95

# List of Tables

1.1	Causes of death in the U.S. in 2005 . . . . .	1
2.1	Imaging Techniques for Atherosclerosis . . . . .	9
2.2	The EU SCORE card and the Framingham Coronary Heart Disease Risk Score . . . . .	17
3.1	Area overlap results . . . . .	33
4.1	The code length per pixel in bits. . . . .	47
4.2	Evaluation of the segmentation results . . . . .	51
5.1	Evaluation of the segmentation results . . . . .	66
6.1	Metabolic and physical markers of the study population . . . . .	82
6.2	The inter- and intra-observer mean coefficients of variation . . . . .	83
6.3	The mean $\pm$ one standard deviation of all the imaging markers . . . . .	84
6.4	The hazard ratio per standard deviation increase in marker values . . . . .	85
6.5	The individual hazard ratios . . . . .	86
7.1	Population characteristics of the study population . . . . .	96
7.2	Hazard ratios of all-cause death of the metabolic/physical markers . . . . .	97
7.3	Stratification of abdominal aortic calcification marker values . . . . .	98
7.4	Hazard ratio per standard deviation increase in marker value . . . . .	99
7.5	Hazard ratio for high risk subjects . . . . .	100



# Abstract

According to the World Health Organization cardiovascular diseases (CVDs) are the number one cause of death globally. Two thirds of women and half of men who die suddenly resulting from CVDs have no previously recognized symptoms. This indicates that much more people are actually at risk than the current methods identify, and therefore it is relevant to look into new methods. Lumbar aortic calcified deposits have been shown to relate to CVD risk, but are not yet widely used. Hence, we investigate methods that can help to detect lumbar aortic calcified deposits and examine their influence on CVD risk to optimally target timely intervention and to better identify people at risk.

The goal is to create an automatic system for the detection of lumbar aortic calcifications, which can find the aorta and subsequently calcifications in the aorta on X-ray images, and convert the findings to a CVD risk. To improve segmentation of calcifications in the aorta two calcification shape priors are developed: The first is inspired from geostatistics and based on a method called Kriging, while the second is a texton-based generative shape model using a minimal description length model selection, which is applied to both lumbar aortic calcification data as well as other natural image data. Furthermore a new methodology based on the concept of renormalization group theory (RGT) is introduced and used in the context of multi-scale image segmentation. Another important aspect is the clinical relevance of lumbar aortic calcifications. Several biomarkers are developed and their clinical relevance tested. The biomarkers are also combined to form the morphological atherosclerotic calcification distribution (MACD) index, and the newly formed MACD index, and its relation to mortality in post menopausal women, is analyzed.

The proposed calcification shape priors produce promising results. The method based on Kriging improves the sensitivity and the Jaccard index of calcification test data, but the largest improvement comes with the texton-based generative shape model. It improves the Jaccard index of calcification segmentations by 50%, which indicates that it can be successfully used as a prior distribution in statistical segmentation of calcifications on X-ray image data. Additionally, RGT provides a new approach for multi-scale image segmentation that could be an alternative to parts of the fully automatic system. The examination of different biomarkers shows that simple statistical modeling can help to identify

potential imaging markers. Especially, the MACD index seems to be a more sensitive predictor of CVD mortality based on lumbar X-rays than the current gold standard, the AC24 radiographic scoring of atherosclerotic plaques.

We can conclude that there is still some way to a fully automatic system of detecting aortic calcified deposits, but that the assessment of the shape, size, number, distribution, and extent of lumbar aortic calcifications may aid in identifying people at risk of dying from CVDs and thus in the future help those in most need of treatment.

# Resume

Ifølge Verdenssundhedsorganisationen er hjertekarsygdomme (HKS) globalt set den største årsag til dødsfald. To tredjedele af kvinder og halvdelen af mænd som dør pludseligt pga. hjertekarsygdomme har ikke tidligere haft nogen symptomer. Det betyder at mange flere mennesker har risiko for få hertekarsygdomme end nuværende metoder identificerer, og derfor er det relevant at forske i nye metoder. Forkalkninger i lumbal aorta har vist sig at være relateret til risikoen for at få hjertekarsygdomme, men brugen heraf er endnu ikke særlig udbredt. Derfor arbejder vi på metoder som kan finde forkalkninger i lumbal aorta automatisk og undersøge deres betydning for risikoen for at få hjertekarsygdomme, så man kan rette tidlige indgreb imod de rigtige mennesker og identificere dem som har den største risiko at udvikle hjertekarsygdomme. Vores formål er at lave et fuldt automatisk system til at finde forkalkninger i røntgenbilleder. Det skal dels kunne finde lumbal aorta og efterfølgende forkalkninger deri og konvertere det til en risiko for at udvikle hjertekarsygdomme. For at forbedre segmenteringen af forkalkninger i aorta, udvikles to modeller for forkalkningernes form: den første er inspireret af geostatistik og baseret på en metode som man kalder Kriging, mens den anden er baseret på textons og bruger en minimal description length model selection og er anvendt på røntgenbilleder af forkalkninger i aorta og andre naturlige billeder. Desuden introducerer vi en ny metode baseret på et koncept som kaldes renormaliseringsgruppeteorie (RGT) og bruger metoden til multi-scale billede segmentering. Et andet aspekt er den kliniske relevans af forkalkninger i aorta. Flere biomarkører bliver udviklet og deres kliniske betydning undersøgt. Biomarkørerne er også kombineret til det morphological atherosclerotic calcification distribution (MACD) indeks og dets relation til dødeligheden af kvinder som har passeret overgangsalderen er analyseret.

Modellerne for forkalkningernes form, som vi forslår, giver lovende resultater. Den metode som er baseret på Kriging forbedrer sensitiviteten og Jaccard indekset på vores trænings data, men den største forbedring opnår modellen baseret på textons. Den forbedrer Jaccard indekset med 50%, som indikerer at vores model kan blive brugt til statistisk segmentering af forkalkninger i røntgenbilleder. Yderligere forsyner RGT os med en ny måde at lave multi-scale billede segmentering på og kunne være et alternativ til dele af vores

automatiske system. Undersøgelsen af forskellige biomarkører viser at simple statistiske modeller kan hjælpe med at identificere potentielle billede markører. Især MACD ser ud til at være en mere sensitiv måling til at forudsige dødeligheden fra hjertekarsygdomme baseret på røntgenbilleder end den nuværende standard, AC24.

Slutteligt kan vi konkludere, at der stadig er et stykke vej til et helt automatisk system til at finde forkalkninger, men at måling af form, størrelse, antal, fordeling og omfang af forkalkninger i aorta kan bidrage til at identificere mennesker, som har den største risiko for at dø af hjertekarsygdomme, og som derfor har mest brug for behandling.

# Preface

The work described in this dissertation was carried out between March 2008 and February 2011 at the Department of Computer Science of the University of Copenhagen, Denmark (DIKU). The project is a collaboration between DIKU and Synarc Imaging Technologies, Denmark, funded by a combination of the Department of Computer Science of the University of Copenhagen, den Danske Forskningsfond, and Synarc Imaging Technologies. The Ph.D. research took place at Synarc Imaging Technologies, while teaching obligations and other duty hours were carried out both at the Department of Computer Science of the University of Copenhagen and at Synarc Imaging Technologies.

The project was supervised by Dr. Sami Brandt (Synarc Imaging Technologies, Denmark), Assoc. Prof. Dr. Marleen de Bruijne (Erasmus Medical Center, The Netherlands and University of Copenhagen, Denmark) and Prof. Dr. Mads Nielsen (University of Copenhagen, Denmark). Three months of the Ph.D. research were carried out at Microsoft Research Laboratory (Cambridge, UK) under the supervision of Dr. Pushmeet Kohli (Microsoft Research, United Kingdom).

The dissertation at hand is based on research papers that have been accepted or are in review in international medical imaging journals and conferences.



Copenhagen, February 2011



# List of Publications

## Papers in international journals

- A Bayesian Framework For The Automated Prediction Of Cardiovascular Disease From Lumbar Standard Radiographs. Petersen, Kersten; Ganz, Melanie; Mysling, Peter; Lillemark, Lene; Crimi, Alessandro; Nielsen, Mads; Brandt, Sami. Currently in preparation.
- Abdominal Aortic Calcified Deposits And Their Relationship To Mortality In Post Menopausal Women. Ganz, Melanie; de Bruijne, Marleen; Dam, Erik B.; Pettersen, Paola; Karsdal, Morten A.; Christiansen, Claus; Nielsen, Mads. Currently in preparation.
- A Texton-Based Generative Shape Model And MDL Model Selection For Statistical Analysis Of Archipelago-Like Textures. Ganz, Melanie; Nielsen, Mads; Brandt, Sami. Currently submitted for publication.
- Distribution, Size, Shape, Growth Potential And Extent Of Abdominal Aortic Calcified Deposits Predict Mortality In Post Menopausal Women. Nielsen, Mads; Ganz, Melanie; Lauze, Francois; Pettersen, Paola P.; de Bruijne, Marleen; Clarkson, Thomas B.; Dam, Erik B.; Christiansen, Claus; Karsdal, Morten A. BMC Cardiovascular Disorders 2010, 10:56
- Patterns Of Abdominal Atherosclerotic Calcified Deposits From Lumbar Lateral X-rays. Lillemark, Lene; Ganz, Melanie; Barascuk, Natasha; Dam, Erik; Nielsen, Mads. The International Journal of Cardiovascular Imaging (formerly Cardiac Imaging), March 2010, Springer.

## Papers in international conferences

- Application Of Renormalization Group Theory To Multi-Scale Energy Minimization Problems. Ganz, Melanie; Kohli, Pushmeet. Currently in preparation.
- Patch-Based Generative Shape Model And MDL Model Selection For Statistical Analysis Of Archipelagos. Ganz, Melanie; Nielsen, Mads; Brandt,

Sami. International Workshop on Machine Learning in Medical Imaging (MLMI) in conjunction with MICCAI 2010, Beijing, China, 2010.

- MACD - An Imaging Marker For Cardiovascular Disease. Ganz, Melanie; de Bruijne, Marleen; Nielsen, Mads. SPIE Medical Imaging, San Diego, CA, USA, 2010.
- Dense Iterative Contextual Pixel Classification Using Kriging. Ganz, Melanie; Loog, Marco; Brandt, Sami; Nielsen, Mads. Proceedings of IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis, Miami Beach, FL, USA, 2009.
- Growth Pattern Of Atherosclerotic Calcifications. Lillemark, Lene; Ganz, Melanie; Dam, Erik; Dandu, Naga; Nielsen, Mads. Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2nd Workshop on Computer Vision for Intra Vascular and Intra Cardiac Imaging (CVII), 2008.

## Published abstracts

- BMD Prediction Of Death Is Encapsulated By The Morphological Atherosclerosis Calcification Distribution (MACD) Index. Ganz, Melanie; Nielsen, Mads; Karsdal, Morten; Christiansen, Claus. Bone, 44, 2009. 36th European Symposium on Calcified Tissues, Vienna, Austria, 2009.
- Analysis Of Individual Abdominal Atherosclerotic Calcified Deposits. Lillemark, Lene; Ganz, Melanie; Nielsen, Mads. Bone, 44, 2009. 36th European Symposium on Calcified Tissues, Vienna, Austria, 2009.
- Superior Assessment Of CVD Death By MACD Index Compared With The Framingham Score Is Highly Associated With Predisposition To Diabetes. Soltyskiska, Ewa; Barascuk, Natascha; Ganz, Melanie; Nielsen, Mads; Karsdal, Morten; Christiansen, Claus. Metabolism and Cardiovascular Risk. Keystone Symposia, nr. S2, Breckenridge, Colorado, USA, 2008.
- Morphological Atherosclerosis Calcification Distribution (MACD) Index Is A Strong Predictor Of Cardio-Vascular Death And Includes Predictive Power Of BMD. Christiansen, Claus; Karsdal, Morten; Ganz, Melanie; Dam, Erik; Nielsen, Mads. 30th ASBMR Annual Meeting, Montreal, Quebec, Canada, 2008.
- Automated Quantification Of The Morphological Atherosclerotic Calcification Distribution On X-Rays Is A Strong Predictor Of Mortality In

Post Menopausal Women. Arteriosclerosis, Thrombosis, and Vascular Biology. Christiansen, Claus; Karsdal, Morten; Lauze, Francois; Dam, Erik; Ganz, Melanie; de Bruijne, Marleen; Sørensen, Mette; Barascuk, Natasha; Nielsen, Mads. In I. Phase, editor, 2008 ATVB Oral Presentation, volume 28, pages 32 to 149. Am Heart Association, 2008.



# Chapter 1

## Introduction

According to the World Health Organization (WHO) [1] the number one cause of death globally that claims more deaths than any other single cause are cardiovascular diseases (CVDs). CVDs are defined as the class of diseases that involve the heart and/or blood vessels [2]. Alone in 2004 an estimated 17.1 million people died from CVDs which equals 29% of all global deaths. Projections by the WHO show that CVDs will remain the single leading cause of death and by 2030 almost 23.6 million people worldwide will die from them. Especially, in the United States (U.S.) CVDs are a major problem and are the leading cause of death (table 1.1) claiming 35.3% of all deaths in 2005 equaling a CVD related death every 37 seconds. In fact, CVDs are such a big problem that the U.S. National Center for Health Statistics [3] states that the life expectancy in the U.S. would increase by almost seven years if all forms of major CVDs were eliminated. In contrast, if all forms of cancer would be eliminated the gain in life expectancy would only amount to 3 years.

Table 1.1: U.S. causes of death in 2005 (based on mortality figures) [4]. Cardiovascular diseases (CVDs) are the leading cause of death.

Cause of Death	Number of victims
CVD	864,480
Cancer	559,312
Accidents	117,809
Alzheimer's Disease	71,599
HIV (Aids)	12,543

## 1.1 Purpose

One of the reasons for CVDs being the number one cause of death is that two thirds of women and half of men who die suddenly from CVDs have no previously recognized symptoms [5–7]. This is the case because traditional prevention strategies fail to recognize that cardiovascular events also occur in subjects in low and intermediate risk groups. Akosah et. al [8] showed that 70% of a group of 222 young adults (men  $\leq 55$  years and women  $\leq 65$  years) were hospitalized for myocardial infarction despite being categorized as low risk candidates by traditional prophylaxis methods. Hence, the current categorization in CVD risk groups seems to assign candidates that are actually at high risk of dying from CVD to a low risk group. Therefore, mortality can not be reduced in low risk groups by traditional methods and there exists a need for new methods especially aimed at prevention.

**An early indicator of the risk of death by CVDs are calcified deposits in the vascular system. This is the reason why in this dissertation I want to investigate methods that can help detect calcified deposits and examine their influence on CVD risk to optimally target timely intervention and to better identify people at risk.**

## 1.2 Outline

The dissertation is divided into two major parts: Part I (chapter 3-5) deals with medical imaging of calcifications and part II (chapter 6-7) focuses on the relevance of especially lumbar aortic calcifications in relation to CVD risk.

In chapter 2 some background knowledge needed to understand the dissertation is provided. The underlying causes of cardiovascular diseases, atherosclerosis, is introduced (section 2.1) as well as an overview of medical imaging of cardiovascular diseases given (section 2.2). Finally, the clinical tools to assess cardiovascular disease risk are described (section 2.3).

Part I of the dissertation consists of research papers that focus on the development of automated methods for the detection of calcifications. Here the contribution to the automated methods, especially in the calcification detection process, is described. The work is divided into two chapters. In chapter 3 an excursion into a method from geostatistics called Kriging leads to a method called "Dense Iterative Contextual Pixel Classification using Kriging". In chapter 4 a texton-based generative shape model using a minimal description length model selection is developed and applied on calcification data as well as other natural image data. Finally, a new methodology based on the concept of renormalization group theory (RGT) is introduced, which is an essential tool in statistical physics. In chapter 5 RGT is used in the context of multi-scale image segmentation.

Part II of the dissertation deals with the clinical relevance of particularly lumbar aortic calcified deposits. In chapter 6 several biomarkers based on lumbar aortic calcified deposits are developed and their clinical relevance tested. Then the biomarkers are via a selection process based on Cox regression combined to form the morphological atherosclerotic calcification distribution (MACD) index. Finally, the newly formed MACD index and its relation to mortality in post menopausal women is examined in chapter 7.

The dissertation concludes in chapter 8 with a short summary, a discussion and perspectives for future research within the field.



# Chapter 2

## Risk Assessment of Cardiovascular Diseases

In order to be able to assess CVD risk as early as possible the underlying causes of CVDs need to be understood. Hence, this chapter provides some background knowledge. First, the pathology of one of the underlying causes of cardiovascular diseases, atherosclerosis, is introduced (section 2.1). Then an introduction to part I of the dissertation is given in section 2.2 where an overview of medical imaging of cardiovascular diseases is given, and the overall project, which this dissertation was part of, is detailed. In the introduction to part II (section 2.3) current clinical tools to assess cardiovascular disease risk are described.

### 2.1 Causes of Cardiovascular Diseases

While CVDs refer to any disease that affects the cardiovascular system the underlying cause of CVDs is usually atherosclerosis [10]. Atherosclerosis is a disease that affects the arteries (figure 2.1), which are the blood vessels that carry oxygen-rich blood to one's heart and the rest of one's body. As its name (athero = soft, pasty material; sclerosis = hardening) suggests it is a process where fatty substances, such as cholesterol, cellular waste material and calcium form a calcified deposit in the media of an artery. The reason for substances entering the media is usually a damage of the endothelium, the innermost lining, of an artery. The endothelium gets damaged if it is for example exposed to elevated lipid levels, high blood pressure or tobacco smoking [11]. The resulting calcified deposit in the media of an artery is then called a plaque (figure 2.1).

Plaques can grow large enough to reduce or even block the blood flow through an artery. On the one hand, plaques can become stable (figure 2.2(c)) with a large necrotic core and a strong calcified cap and subsequently narrow the

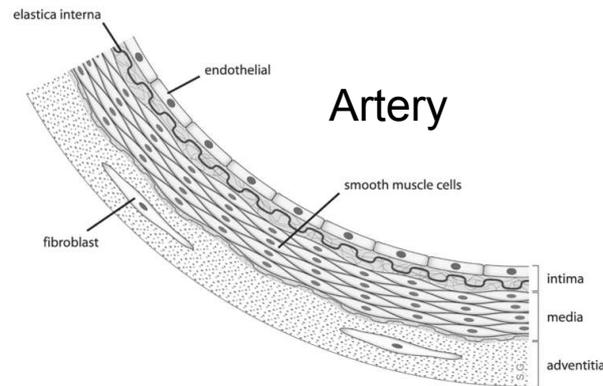


Figure 2.1: The anatomy of an artery showing the intima, made up of endothelial cells as well as the elastica interna, the media and adventitia. (Picture taken from [9], modification and reprint according to Creative Commons Attribution-NonCommercial-ShareAlike 2.5 License).

artery. On the other, they can become instable (figure 2.2(b)) and rupture. Ruptured plaques (figure 2.2(d)) can cause the formation of blood clots that can travel through the arteries and cause a blockage in any other part of the body. If a blood clot blocks a coronary artery it can lead to a heart attack or if it travels to the brain it can cause a stroke. In the case that the clot ends up in the lung it can lead to a pulmonary embolism and even in the extremities a blood clot can cause significant damage and eventually lead to gangrene [12]. The goal is to identify calcified deposits, because they are an expression of atherosclerosis and detectable by medical imaging modalities.

## 2.2 Introduction to Part I: Medical Imaging of Atherosclerosis

Since the aforementioned atherosclerosis develops over decades [14] earlier detection of subclinical atherosclerosis may allow timely intervention and lead to better identification of people at risk. Hereby atherosclerosis imaging plays a large role. There exist a multitude of modalities to evaluate atherosclerosis, both invasive and non-invasive [15, 16].

Invasive methods like quantitative coronary angiography (QCA), intravascular ultrasound (IVUS) and optical coherence tomography (OCT) have all been shown to be able to assess atherosclerosis [15]. But all three modalities are also reserved for small scale studies due to the inherent risks of invasive procedures. Furthermore, use of QCA is only advisable to aid interventional treatment of

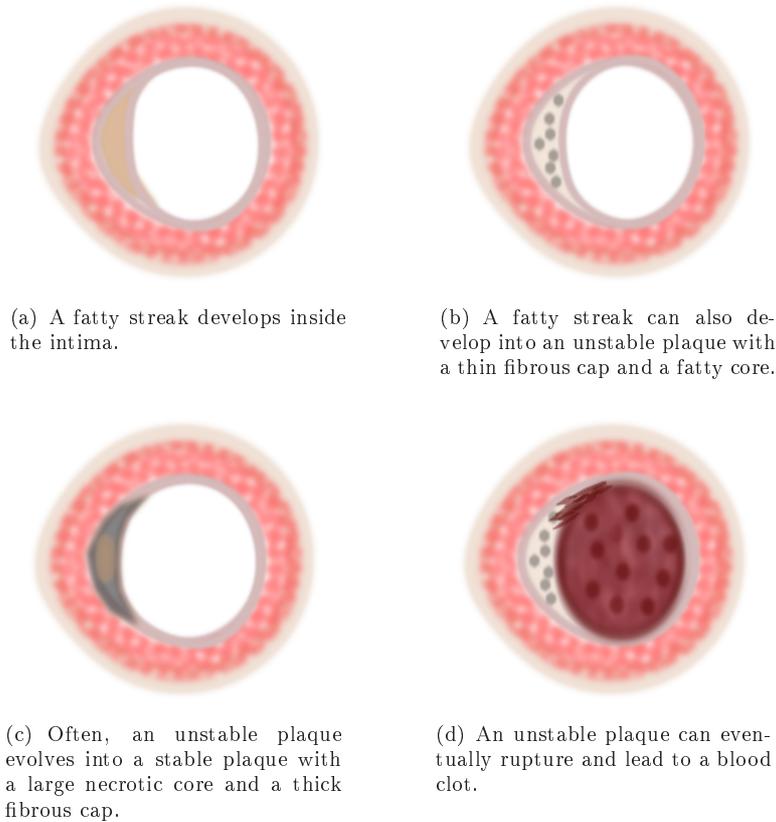


Figure 2.2: A cross section of an artery illustrates the different stages of plaque development: (a) shows a fatty streak, (b) a stable, (c) an unstable and (d) a ruptured plaque.

patients at intermediate risk [17] due to the relatively large exposure to ionizing radiation (16 mSv [18]).<sup>1</sup>

But there exist many alternatives to invasive imaging procedures. Ultrasound (US), computed tomography (CT), magnetic resonance imaging (MRI) and X-ray are all non-invasive imaging techniques that have developed markers that relate to CVD risk.

Ultrasound can be used to visualize the carotid intima-media thickness (IMT) that has been shown to be associated with atherosclerosis [19] and is thus a marker for CVD. Additionally, US is cost-efficient and does not expose patients to harmful radiation. A downside of US is though that IMT is measured in a different vascular bed than the coronary arteries and that it is operator dependent [20].

Imaging of atherosclerosis in the coronary arteries can be done with CT [21].

---

<sup>1</sup>The natural background radiation per year amounts to ca. 3 mSv and a coast to coast flight over the U.S. exposes a person to a radiation dose of 0.03 mSv [18]

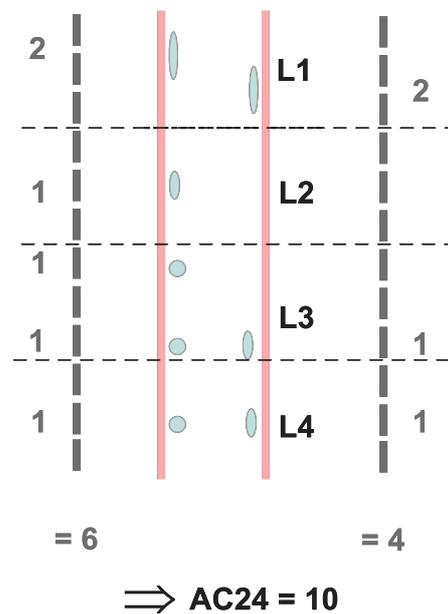


Figure 2.3: The AC24 is constructed by projecting the abdominal aortic calcifications (AACs) to the corresponding aorta wall. The degree of lesion occupation is graded: 0 for no AAC, 1 for AACs occupying less than  $1/3$  of the wall they are projected onto, 2 for AACs occupying more than  $1/3$ , but less than  $2/3$  in the projection, and 3 for a  $2/3$  or more occupation of the wall. Summed up the degrees of lesion occupation lead to an AC24 score ranging between 0 and 24. The schematic view is adopted from [13].

CT of the coronary arteries is able to quantify the degree of coronary artery calcification (CAC) in the Agatston score with good reproducibility [22, 23], which provides a strong measure of cardiovascular risk [24] independently from, and potentially more powerful than, traditional risk factors such as smoking [25]. Atherosclerosis imaging in CT is usually performed with two types of scanners, electron-beam computed tomography and multi-row detector CT. Coronary CT is widely available and its downsides are only its cost [26] and the exposure to moderate levels of radiation (3 mSv [18]).

MRI is a non-invasive modality to assess atherosclerosis in different vascular beds. MRI has been able to quantify atherosclerosis and responses to treatment, but only in the aorta and the carotid arteries [27–29]. The reason for this is that MRI measurements are challenged by the size of the smaller arteries and assessment of the coronary arteries is especially difficult due to cardiac and respiratory motion artifacts. So although its advantage is the lack of exposure to harmful radiation, its disadvantages are the inaccessibility of the coronaries and its cost.

Finally, it has been demonstrated that abdominal aortic calcifications (AACs) detectable by lateral lumbar radiographs are strong predictors of cardiovascu-

Table 2.1: Summary of imaging techniques used in atherosclerosis including some of their advantages and disadvantages.

	<b>Advantages</b>	<b>Disadvantages</b>
QCA	widely available	invasive radiation (16mSv)
IVUS	widely available	invasive
OCT	high resolution	invasive complex
US	cost-efficient no radiation	different vascular bed (carotid arteries) operator dependent
CT	widely available	cost radiation (3 mSv)
MRI	no radiation	coronaries inaccessible cost
X-ray	cost-efficient widely available	different vascular bed (lumbar aorta) radiation (1.5 mSv)

lar morbidity and mortality [10]. They correlate strongly with coronary artery calcifications and can hence predict the risk of coronary artery problems [30]. Therefore in X-rays the state of the art methodology to estimate CVD risk is the abdominal aortic calcification score (AC24) proposed by the Framingham study group [13]. The AC24 is constructed by projecting the abdominal aortic calcifications (AACs) to the corresponding aorta wall (see figure 2.3). So while an advantage of X-rays is the wide availability of the modality, a disadvantage is the exposure to a low level of harmful radiation (0.3 - 1.5 mSv [18, 31]) as well as the fact that the assessment of AC24 is again made in a different vascular bed than the coronaries.

A summary of the mentioned advantages and disadvantages of the different imaging techniques can be seen in table 2.1.

### 2.2.1 Choice of Modality and Region of Interest in the Body

In this dissertation the approach will be to examine abdominal aortic calcifications as observed in standard lumbar radiographs (see figure 2.4). As already discussed AACs are strong predictors of cardiovascular morbidity and mortality [10]. The reasons for choosing X-ray images of the AACs in contrast to for example using MRI are threefold:

- X-ray is still one of the most widely available imaging modality besides ultrasound and compared to imaging of carotid plaques via ultrasound taking a lumbar aortic X-ray is not as operator dependent [20].
- The gold standard for vertebral fracture diagnosis are standard radio-

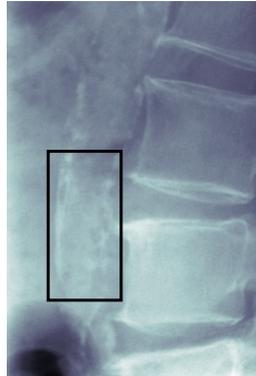


Figure 2.4: Illustrations of abdominal aortic calcifications (AACs) in a lumbar aortic X-ray for ease highlighted by a black square around some of them.

graphs [32]. This means every time it needs to be clinically assessed if a vertebral fracture is present, an X-ray of the spine is taken, which then also can be used for CVD risk assessment.

- In clinical studies for drug development against osteoporosis [33,34] inclusion criteria are, besides other metabolic factors, low BMD as measured by DXA scans or radiographically confirmed vertebral fractures. Furthermore, spinal lateral radiographs are used to measure efficacy, so every subject in osteoporosis trials has at least one standard radiograph taken. And in some adverse effect studies AC24 scorings of lumbar radiographs have been requested to examine the effect of osteoporosis treatment on the cardiovascular system leading to even more subjects in clinical trials where an X-ray of the lumbar spine is taken.

That spinal fracture is assessed by standard radiographs and the availability of large, long duration studies from clinical osteoporosis trials are big advantages when examining abdominal aortic calcifications. For all subjects that have participated in a clinical osteoporosis trial, the AAC scoring can then be performed without additional ionizing radiation exposure or cost as these images are already available. Furthermore, historical data from clinical trials can be used to develop new AAC markers and verify them.

As already stated there exists a need to intervene in the early stages of CVDs and to develop methods to enable CVD risk scoring of large populations either in clinical trials or in a screening setup. But the CVD risk examination of a larger population is only possible if it can be done fast, cheap and is easily accessible. By choosing X-rays we provide a cheap, widely available and inexpensive modality.

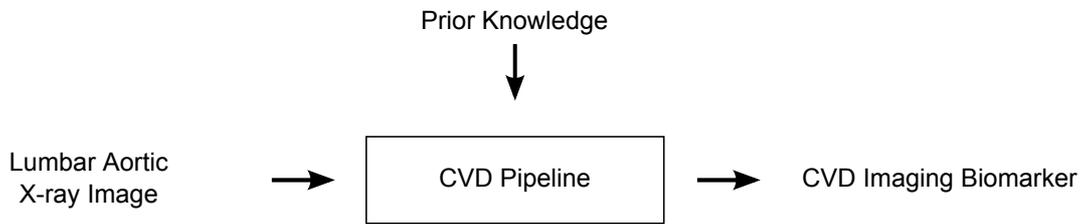


Figure 2.5: This shows how we apply the CVD Pipeline on lumbar aortic X-ray images.

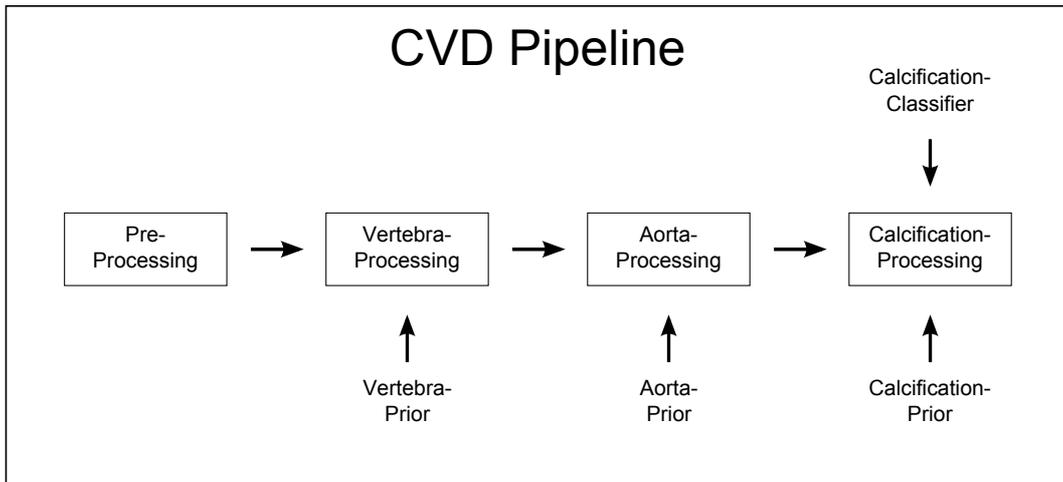


Figure 2.6: A schematic description of the CVD Pipeline that allows for automatic annotation of calcifications in aortic X-ray images.

### 2.2.2 Automating the Measurement of Imaging Biomarkers on Lumbar Aortic Radiographs

When trying to make CVD risk scoring available for larger populations, the real cost arises from the human involvement, e.g. the radiologists that have to spent time on AC24 scorings of lumbar aortic X-rays. Therefore, it is essential to provide automated scoring frameworks to admit the possibility of large population studies in order to identify further risk factors of CVDs.

To meet this challenge we have developed a machine-learning based approach [35] for automatic annotation of calcified deposits in lumbar aortic X-ray images that we call "CVD pipeline" (see figure 2.5). The CVD pipeline is comprised of several steps: a pre-processing, a vertebra-processing, an aorta-processing and a calcification-processing stage. A diagram describing the CVD pipeline can be seen in figure 2.6.

### Pre-Processing

The pre-processing in the CVD pipeline consists of an image normalization which is a prerequisite for achieving robust pixel classification results on images from different X-ray machines. This can be achieved with standard intensity normalization techniques, such as subtraction of the mean and division with the standard deviation. Another possibility is to normalize the intensities according to models from physics that describe the scattering of  $\gamma$ -rays.

### Vertebra- and Aorta-Processing

The second step of the CVD pipeline deals with finding the vertebrae in the image. This is done by generating a vertebrae template that consists of six regions of interest, the five vertebral classes defined in [36] and a background class, and then applying a random forest classifier [37]. The classifier is trained on 100,000 sampled points from the five foreground classes using Gaussian derivative features up to third order on scales 0.18, 0.56 and 1.78 mm [38]. On the basis of the classification results a shape model is used to identify the lumbar vertebrae, L1-L4.

Next the aorta is defined on the basis of the identified vertebra via an aorta position posterior, where we marginalize over all possible vertebra shapes [38]. Once we have found the aorta we use it as our region of interest in the search for calcifications.

### Calcification-Processing

The calcification processing combines two things, the classification output of a pixel classifier and prior knowledge about the location and shape of calcified deposits.

For the classification we employ a random forest classifier [37] with seven features: Gradient Magnitude at the scales 0.56 and 1.78 mm, Hessian Coherence at 1.78 mm, the second Gaussian derivative in y-direction at the scales 0.18 and 0.56 mm as well as the third Gaussian derivative in x- as well as y-direction at a scale of 0.56 mm.<sup>2</sup> After identifying the features the classifier is trained on 100,000 points and a five fold cross validation is used; in this case five fold means  $\frac{4}{5}$  of our data is used for training and  $\frac{1}{5}$  is tested on.

A weakness of a pixel-based classification is that it does not use contextual knowledge. The feature space is not well separated and hence the output from the pixel classification is noisy. Calcification priors can distill contextual information from the image. They enhance the result by using image intensity, shape and statistical information since these properties of the X-rays can help

---

<sup>2</sup>The seven features are found by training the classifier on 100,000 points from manual annotations of the aorta template of a separate data set and then employing a sequential floating forward feature selection (SFFS) [39].

to predict shape and distribution of the calcifications.

Hence, in the CVD pipeline several calcification priors are used: First, we combine the output of the pixel classifier with a calcification likelihood map. The map is constructed by registering all manual annotations of the training set into a common normalized coordinate system [40] and then turning this calcification histogram through normalization into a probability map. The calcification likelihood map ensures that calcifications that are placed in unlikely places are subdued, while calcifications in very likely places in the aorta are emphasized. Second, we learn a shape model of the calcifications from the manual annotations and then impose the same shape statistics of the training data onto our segmentation result.

### **Challenges for the CVD Pipeline**

The CVD pipeline consists of several steps that each face their own obstacles. Furthermore, since we deal with historical clinical trial data, our X-ray images are afflicted by clutter, occlusions and a low signal-to-noise ratio and object boundaries are hardly visible. This naturally challenges the processing as a whole. An example X-ray image can be seen in figure 2.7.

In the vertebra stage the biggest problem is to identify the lumbar vertebrae, L1-L4, correctly. A shape model of the 4 lumbar vertebrae can also falsely detect the thoracic vertebra T12-L3 or L2-L5 and hence predict a wrong aorta. Furthermore, vertebrae are sometimes only partially visible and can then not be identified by the vertebra shape model. Additionally, the vertebrae boundaries can appear smeared or even duplicate, because the X-ray procedure projects the three-dimensional vertebrae into two dimensions.

The conditional shape model of the aorta given the vertebrae depends on the vertebrae, but also on potential calcifications. Therefore it needs to be able to take the vertebrae as well as potential calcifications into account to define a region of interest for further processing. Especially, if the vertebrae prediction step fails, e.g. can not match the vertebrae shape model to the actual vertebrae, it can be seen that the conditional shape model is very sensitive to false positives (vertebrae boundary pixels that look like potential calcifications). Finally, since the biological shape variation of the aortas occurs on a curved manifold, a more complicated shape model might be needed.

Setting aside the problems that arise from trying to detect the vertebra and the aorta, the detection of calcified deposits in X-ray images itself has various challenges. First, the calcified plaques are small objects of low contrast and large variability in shape, size and appearance and therefore affected by background noise, e.g., noise from Compton scattering. Second, because the X-ray is a 2-dimensional projection the calcifications are often totally or partially occluded by other physiological structures, such as the lowest ribs, part of the



Figure 2.7: Here an example X-ray image is shown, where one clearly sees the four lumbar vertebrae. To the left of the vertebrae the aorta is located, but it is hard to see since it is less prominent than the vertebrae. It is inside the aorta that we try to detect calcified deposits.

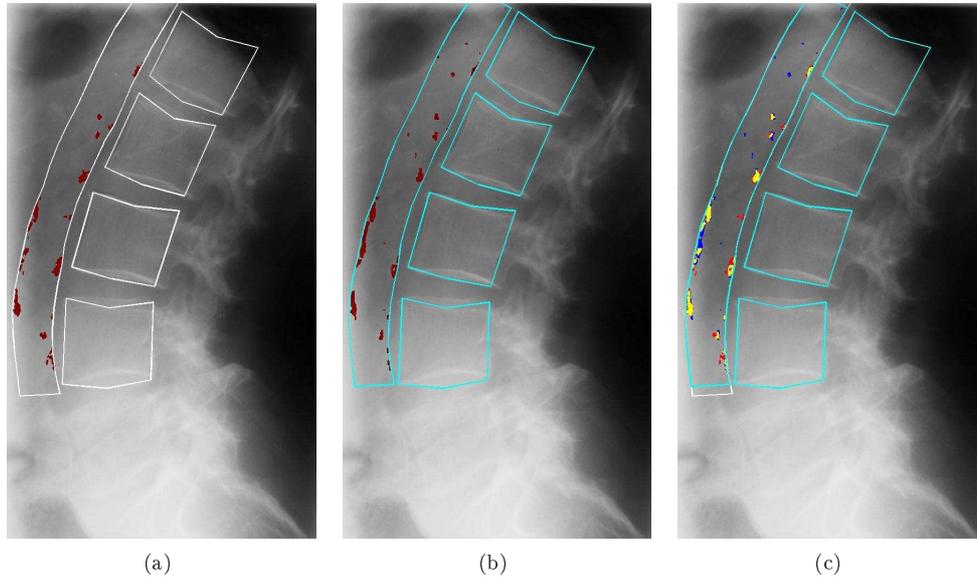


Figure 2.8: Here we can see a typical pipeline result based on the image in figure 2.7. (a) We can see the manual annotation of the vertebrae and the aorta outline (white) as well as the calcifications (red). (b) In the middle the automatically retrieved vertebrae and aorta outline (green) as well as a probability map of the calcifications (yellow to red) are shown. (c) The overlap of the calcification segmentation with the manual annotation is displayed; yellow indicates true positives, blue indicates false positives and red indicates false negatives.

hip or folds of fat in obese persons. Third, other structures, e.g. calcifications in the bowels or parts of vertebral cortical bone, have a similar appearance in the image.

An example of an automatic calcification segmentation of the CVD pipeline is shown in figure 2.8. The picture displays a manual annotation besides the automated segmentation result and also a comparison of the two. The performance and limitations of the CVD pipeline, especially concerning the calcification shape modeling, will be examined in chapters 3 and 4.

### 2.2.3 Application of Renormalization Group Theory in Medical Imaging

In section 2.2.2 the CVD pipeline and its focus on detecting lumbar aortic calcifications has been described. A key feature of the CVD pipeline is the pixel classifier that is used to segment the calcifications inside a given region of interest. But it has its limitations. When dealing with high resolution images the computational time is extensive. And since it classifies each pixel for itself, it can not take spatial interactions between pixels into account and therefore requires us to use calcification shape priors to re-introduce spatial coherency.

Hence, we have investigated methods to replace the pixel classifier by other image segmentation algorithms. To do this we developed a new method for image segmentation based on renormalization group theory (RGT) which is especially interesting for images with a high resolution and where very long spatial interactions play a role (see figure 2.9). A strength of RGT is that it can as described in [41] and [42] evolve energy functions correctly between different scales when employing a hierarchical approach. At every scale the resulting energy function can then be optimized by the optimizer of choice. Hence, for large problems RGT can be applied to general segmentation techniques to construct an appropriate hierarchical version of the problem.

Segmentation problems in general can be solved by 3 different approaches - local optimizing techniques, global optimizing techniques and stochastic methods. Local optimizing techniques are e.g. active contour [43], active appearance [44] or level set models [45]. A problem of active contour or active appearance models is their inability to handle topology changes. Level sets can handle those, but since level sets essentially employ gradient descent methods one runs into problems again. Most of the functionals one deals with in medical imaging are far from convex, so level sets only yield local optima. Global optimizers like graphcut [46], on the other hand, guarantee to find the global optimum. The problem lies here in the computational tractability in terms of time and memory consumption, as soon as the connectivity of the graph becomes dense. But this exactly is the case in medical images, where high

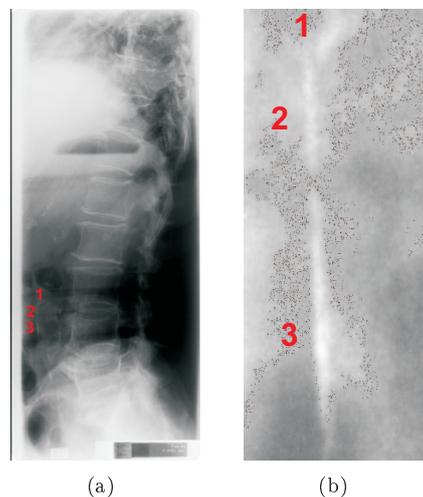


Figure 2.9: First an X-ray of our study population (resolution 570 dpi). Then a small part of the former X-ray, showing a row of three calcifications (800 by 400 pixels). To separate the 3 calcifications correctly from each other, the high resolution as well as the several hundred pixels long interactions inside the calcifications are important.

Table 2.2: Specifications of the EU SCORE card [49] and the Framingham Coronary Heart Disease Risk Score (Framingham score) [50].

<b>EU SCORE</b>	<b>Framingham score</b>
Age	Age
Smoking status	Smoking status
Total cholesterol	Total cholesterol
Systolic blood pressure	Systolic blood pressure
	High-density lipoprotein
	Hypertension treatment status

resolution as well as long interactions ranges play a role. Stochastic methods, markov chain monte carlo (MCMC) methods [47] or simulated annealing [48], also find the global optimum. But similarly they present a computational challenge. Especially if long range interactions are of importance, the convergence time is a problem.

In chapter 5 RGT will be applied to one of the above methods to improve tractability showing how hierarchical processing of large images with long spatial interactions is possible.

## 2.3 Introduction to Part II: Diagnostic Tools for Cardiovascular Disease Risk

From the image processing approach to CVD risk assessment we now move into the clinical evaluation of CVD risk. Besides the imaging modalities introduced in chapter 2.2 to assess CVD risk, there exist also other markers of CVD risk, such as metabolic biomarkers, which are used in clinical practice. If one is able to process large population studies by automating the measurement of imaging biomarkers (see chapter 2.2.2) the performance of the imaging biomarkers needs to be compared to already established clinical measures of CVD risk.

### 2.3.1 Current Metabolic Biomarkers for CVDs

Already in 1981 a list of 246 coronary risk factors had been compiled [51]. Later on, risk scoring systems for use in the clinical management of cardiovascular risk in Europe, the EU SCORE card [49], as well as in the US, the Framingham Coronary Heart Disease Risk Score (Framingham score) [50], have been developed. The EU SCORE is a combination of the age, smoking status, levels

of total cholesterol and systolic blood pressure, while the Framingham score is comprised of the same variables plus the high-density lipoprotein (HDL) and the hypertension treatment status (see table 2.2).

The second part of this dissertation (chapter 6-7) will give an example of a comparison between newly developed markers relating to the geometrical outline of calcified deposits in lumbar aortic X-ray images and the established EU SCORE and Framingham score.

**Overall this dissertation is comprised of a theoretical approach to the automated segmentation of calcified deposits and its practical application on data from existing clinical osteoporosis trials. While part I creates the technical basis for the automated segmentation and the development of new imaging biomarkers, part II evaluates the performance of imaging biomarkers compared to established clinical measures. Together the two parts can hopefully improve the understanding of CVD risk and thereby optimally target timely intervention and better identify people at risk.**

# I

## Medical Imaging of Lumbar Aortic Calcifications



**P**ART I of the dissertation consists of research papers that focus on the development of automated methods for the detection of lumbar aortic calcifications. The contributions to the automated methods were done in the calcification processing stage of the CVD pipeline.

The first chapter gives an overview over a method from geostatistics called Kriging and leads to a "Dense Iterative Contextual Pixel Classification using Kriging". Chapter 4 introduces a textron-based generative shape model using a minimal description length model selection that is applied on the lumbar aortic calcification data as well as natural image texture data. In chapter 5 we make an excursion into the world of statistical physics. A method based on renormalization group theory (RGT) is used to derive a new approach for multi-scale image segmentation that could be an alternative approach to parts of the CVD pipeline.



# Chapter 3

## Dense Iterative Contextual Pixel Classification Using Kriging

---

This chapter is based on the manuscript "Dense Iterative Contextual Pixel Classification Using Kriging" by Melanie Ganz, Marco Loog, Sami Brandt and Mads Nielsen published in the proceedings of the IEEE Computer Society Workshop on Mathematical Methods in Biomedical Image Analysis, 2009.

---

**Abstract** In medical applications, segmentation has become an ever more important task. One of the competitive schemes to perform such segmentation is by means of pixel classification. Simple pixel-based classification schemes can be improved by incorporating contextual label information. Various methods have been proposed to this end, e.g., iterative contextual pixel classification, iterated conditional modes, and other approaches related to Markov random fields. A problem of these methods, however, is their computational complexity, especially when dealing with high-resolution images in which relatively long range interactions may play a role. In the following, a new method based on Kriging is proposed that makes it possible to include such long range interactions, while keeping the computations manageable when dealing with large medical images.

### 3.1 Introduction

In medical applications, image segmentation tasks become ever more important to aid quantitative analysis. In this paper, we focus on the application of medical imaging to aid the diagnosis and prognosis of cardiovascular diseases (CVDs) [52,53]. Images are traditionally segmented by e.g. active contour [54], active appearance [55] or level set models [56]. Competitors to these models, especially in the domain of medical imaging, are pixel-wise classifiers [57]. To achieve even better segmentations, in cooperation with pixel-wise classification,

various methods have been developed that take contextual information in the images into account. Examples are iterated conditional modes (ICM) [58], iterative contextual pixel classification (ICPC) [59] or Markov Random Fields (MRF) [48].

A pure pixel-wise classification takes only the contributions of a neighborhood in the image into account when assigning the class label to a pixel. It does not consider any other class labels in its decision. The technique of ICM on the contrary employs contextual knowledge. It is divided in two steps: First, the pixels are classified by a pixel-wise classifier. Second, the neighboring class labels are included into a label decision. An advanced version of ICM was presented by Loog and van Ginneken [59]. Their ICPC method went a step further than ICM by creating a simultaneous dependency of a class label on surrounding image values and class labels.

A problem of these methods, however, lies in their computational burden when dealing with high resolution images like medical data. Additionally, if also long range interactions are present in the image, the convergence of the methods often becomes an issue.

We propose a new model that can be applied after a pixel-wise classification, Dense Iterative Contextual Pixel Classification (DICPC). It employs the context of all class labels and can take long range interactions into account. We implement this by approximating the contextual interactions in label space with a linear model based on Kriging [60]. With this approximation it becomes feasible to converge to an optimal segmentation in manageable time, even for high-resolution images with a long interaction ranges.

This chapter is organized as follows: Section 3.2 restates the problem and gives an introduction to a statistical interpretation of segmentation and Kriging. Section 3.3 introduces the DICPC algorithm. Section 3.4 introduces the problem on which we exemplify our method. It concerns the difficult task of quantifying aortic calcifications. Information on the study population and the exact classification settings are presented there as well as evaluation methods. The results can be found in section 3.5, while section 3.6 comprises the discussion and conclusion.

## 3.2 Problem Description

Let an image  $I = (I_1, \dots, I_n)$  of the size  $n$  be described by its pixel values  $I_i$ . In a pixel classification scheme, there exists a feature vector  $\vec{f}_i$  for each pixel that consists of one or multiple features, e.g. intensity values or filter responses, at the pixel location  $i$ . The matrix  $F = (\vec{f}_1, \dots, \vec{f}_n)$  is comprised of all the feature vectors. The labels for every pixel  $i$  also shape a vector,  $\vec{c} = (c_1, \dots, c_n)$ , that consists of the class label at each pixel location. Class

labels are part of the set  $\Gamma = (1, 2, \dots, \gamma)$ , where  $\gamma$  is finite.

The problem lies now in finding the optimal segmentation  $C^*$  for the image  $I$ .

### 3.2.1 Statistical Interpretation

To find the optimal segmentation  $C^*$  we pursue a maximum-a-posteriori (MAP) approach<sup>3</sup>. In a MAP estimation the optimal segmentation is given as

$$C^* = \operatorname{argmax}_{C \in \Gamma^n} P(C|I), \quad (3.1)$$

where  $\Gamma^n$  is the set of all possible segmentations.

There exist different approaches to solve (3.1). Assuming conditional independence of  $C$  we can rewrite it to

$$C^* = \operatorname{argmax}_{C \in \mathcal{C}} \prod_i P(C_i|C_{-i}, I), \quad (3.2)$$

where  $C_i$  denotes the label for a pixel  $i$ . Additionally assuming a Markov property [62] holds in the label space this can be transformed into

$$C^* = \operatorname{argmax}_{C \in \mathcal{C}} \prod_i P(C_i|C_{N_i}, I). \quad (3.3)$$

Here  $C_{N_i}$  are the neighborhood labels of a pixel  $i$ . This equation can now be taken as the starting point to describe the different techniques mentioned before.

In order to approximate (3.1) and to move toward the optimal segmentation  $C^*$  one can iteratively update the labeling of single pixels. This is somewhat reminiscent of Besag's iterated conditional models [58] and means that if two alternate segmentations  $C_1$  and  $C_2$  are provided, one can decide pixel by pixel if the segmentation should be updated or not.

A way of optimizing such an iterative procedure has been presented as ICPC in [59]. The ICPC algorithm is based on a simple classification result, so it avoids calculating conditional properties explicitly. Neither is a clique formalism needed as in an MRF approach. A problem of ICPC is though that one can end the optimization in a local optimum, instead of the global optimum. There is no guarantee that one deals with a convex energy function and it can therefore not be guaranteed to reach the global optimum.

Our method, DICPC, is based on the outcome of a classification. It possesses all the benefits of ICPC, but on the contrary to ICPC it provides a convex functional for the common distributions (Gaussian, Poisson) and can therefore be guaranteed to be solved by a second order optimization method.

---

<sup>3</sup>The MAP approach equals a Bayesian minimal risk approach [61] for the risk function of all wrong classifications being equally risky.

### 3.2.2 Kriging

Kriging [63,64] is a geostatistical method that is used to spatially interpolate the value  $z_0$  at any location  $\vec{r}_0$  from irregularly sampled data  $\vec{z}$  at  $N$  points  $\vec{r}_N$ . This is done by the local affine model  $z_0 = z(\vec{r}_0)$  such that

$$z_0 = w_0 + \sum_{i=1}^N w_i z_i = w_0 + \vec{w}^t \vec{z}, \quad (3.4)$$

where  $w_0$  is an offset and  $w_i$  is the weight applied to  $z_i$ . If we regard the  $z_i$  as realizations of random variables  $Z_i$  and request our measure to be unbiased,  $E(Z_0 - \hat{Z}_0) = 0$ , we can define the estimation variance

$$\sigma_E^2 = \text{Var}(Z_0 - \hat{Z}_0). \quad (3.5)$$

Using the linear model we can simplify this to

$$\begin{aligned} \sigma_E^2 &= \text{Var}(Z_0) + \text{Var}(w_0 + \vec{w}^t \vec{Z}) - 2\text{Cov}(Z_0, w_0 + \vec{w}^t \vec{Z}) \\ &= \sigma^2 + \vec{w}^t \mathbf{C} \vec{w} - 2\vec{w}^t \text{Cov}(Z_0, \vec{Z}), \end{aligned} \quad (3.6)$$

where  $\mathbf{C}$  is the covariance matrix of  $\vec{Z}$ .  $\text{Cov}(Z_0, \vec{Z})$  is a column vector of covariances between data points at locations  $\vec{r}_i$  and  $\vec{r}_j$  that can be calculated based on the assumption of spatial stationarity from the entries in  $\mathbf{C}$ . We minimize the estimation variance after the weights  $w_i$  by solving

$$\frac{\partial \sigma_E^2}{\partial \vec{w}} = 2\text{Cov}(\vec{Z}, \vec{Z})\vec{w} - 2\text{Cov}(Z_0, \vec{Z}) = 0. \quad (3.7)$$

This results in the simple Kriging system

$$\text{Cov}(\vec{Z}, \vec{Z})\vec{w} = 2\text{Cov}(Z_0, \vec{Z}), \quad (3.8)$$

which can be solved for the interpolation weights  $\vec{w}$  and is in our case expanded to include the necessary condition  $0 \leq w_i \leq 1$  in order to avoid negative weights.

### 3.2.3 Application of Kriging

The same principle of Kriging can with regularly distributed samples be applied to an image and has been used for image restoration [65]. Kriging is as stated in (3.4) based on a linear estimation model. In the case of a segmentation task where manual segmentations are available, one can learn the weights that minimize the estimation variance,  $\sigma_E^2$ , from the manual segmentations  $S_{\text{man}}$  via the linear model

$$z_{0,\text{man}} = w_{0,\text{man}} + \vec{w}_{\text{man}}^t \vec{z}_{\text{man}}. \quad (3.9)$$

Then we use these weights to compose a linear model for the automated segmentations

$$z_{0,\text{aut}} = w_{0,\text{man}} + \vec{w}_{\text{man}}^t \vec{z}_{\text{aut}}. \quad (3.10)$$

This is possible because we may assume that the covariance structure of the manual segmentations can be transferred to the automated segmentation. The weights of the linear model can then be applied to the automated segmentation  $S_{\text{aut}}$  in a filtering manner to give a kriged estimate of the segmentation

$$K(S_{\text{aut}}) = k * S_{\text{aut}}, \quad (3.11)$$

where  $k$  is a 2D-filter built from the weights  $\vec{w}_{\text{man}}$ . Because our method is based on this type of filtering the computational cost stays low compared to ICM and ICPC. Using this formulation of simple Kriging, we now turn to our application in pixel-based segmentation.

### 3.3 DICPC - Dense Iterative Contextual Pixel Classification

The solution for an optimal contextual segmentation has the form of (3.3),

$$C^* = \operatorname{argmax}_{C \in \mathcal{C}} \prod_i P(C_i | C_{N_i}, I). \quad (3.12)$$

Using Bayes formula this can be transformed into

$$\begin{aligned} C^* &= \operatorname{argmax}_{C \in \mathcal{C}} \prod_i P(I | C_i, C_{N_i}) P(C_i | C_{N_i}) \\ &= \operatorname{argmin}_{C \in \mathcal{C}} \sum_i \left( -\log(P(I | C_i, C_{N_i})) - \log(P(C_i | C_{N_i})) \right). \end{aligned} \quad (3.13)$$

If we assume independence of  $C_{N_i}$  and  $I$ , we can write this as

$$C^* = \operatorname{argmin}_{C \in \mathcal{C}} \sum_i \left( -\log(P(I | C_i)) - \log(P(C_i | C_{N_i})) \right). \quad (3.14)$$

#### 3.3.1 Gaussian Distributions

To solve (3.14) we need to define  $P(I | C_i)$  and  $P(C_i | C_{N_i})$ . For now we assume that both are Gaussian distributed, but we will relax this assumption later. A probability prior  $P(C_i | C_{N_i})$  for the segmentation  $C$  can be formulated as follows

$$\begin{aligned} P(C_i | C_{N_i}) &= G_\sigma(C_i - K(C_i)) \\ &= \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(C_i - K(C_i))^2}{2\sigma_i^2}\right) \end{aligned} \quad (3.15)$$

Furthermore, we assume also separability of the image pixels and therefore define  $P(I|C_i) = \prod_i P(I_i|C_i)$  to be of the form

$$\begin{aligned} P(I_i|C_i) &= G_\sigma(I_i - C_i) \\ &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(I_i - C_i)^2}{2\sigma^2}\right). \end{aligned} \quad (3.16)$$

In the case of a Gaussian prior  $P(C_i|C_{N_i})$  and a Gaussian likelihood  $P(I_i|C_i)$ , the posterior distribution is Gaussian again and in that case a closed form solution exists for (3.14). Plugging in  $P(I_i|C_i)$  and  $P(C_i|C_{N_i})$  into (3.16) leads to

$$E_G = \sum_i \left( a + \frac{(I_i - C_i)^2}{2\sigma^2} + \frac{(C_i - K(C_i))^2}{2\sigma_i^2} \right), \quad (3.17)$$

where  $a = \log(2\pi\sigma\sigma_i)$ . We replace  $K(C_i)$  with  $k * C_i$  according to the definition of (3.11) and get

$$E_G = \sum_i \left( a + \frac{(I_i - C_i)^2}{2\sigma^2} + \frac{(C_i - (k * C_i))^2}{2\sigma_i^2} \right). \quad (3.18)$$

Due to the Parseval theorem the energy is preserved in the Fourier transform, so the cost can be equivalently computed in the Fourier domain. The minimization in the Fourier domain is performed by differentiating the energy functional with respect to the real and imaginary parts of the Fourier coefficients and setting the result to zero. This way we arrive at

$$C^* = \mathcal{F}^{-1} \left( \frac{\tilde{I}_i}{1 + \varpi(\tilde{k}')^*(\tilde{k}')} \right), \quad (3.19)$$

where  $\varpi$  is the ratio between the neighborhood and the global variance,  $\mathcal{F}$  describes the Fourier transform and  $\tilde{I} = \mathcal{F}(I)$  as well as  $\tilde{k}' = \mathcal{F}(k') = \mathcal{F}(k-1)$ .

### 3.3.2 Other Distributions

If  $P(I_i|C_i)$  and  $P(C_i|C_{N_i})$  are not Gaussian, but e.g. Laplace distributed, the energy function of (3.17) changes to

$$E_L = \sum_i \left( a + \frac{|I_i - C_i|}{2\sigma^2} + \frac{|C_i - K(C_i)|}{2\sigma_i^2} \right), \quad (3.20)$$

A solution to (3.20) can be found via variational methods by any approach for total variation minimization. In the case of distributions where the mode of the distribution is the same as its mean this solution is equal to the MAP solution. Thus, for the Laplacian distribution the MAP solution is identical to

the minimum variance solution. In the case of other distributions, it depends on the distribution if the total variation minimization equals the MAP solution. To exemplify this we derive the solution again with the help of a gradient descent method; the same solution can be found with the iterative update equation

$$I^{t+1} = \sum_i \alpha \partial_{I_{N_i}} (-\log(P(C_{N_i}|I_{N_i}^t)) - \log(P(I_{N_i}^t))), \quad (3.21)$$

where  $t$  gives the number of the iteration,  $\alpha$  the step size and  $\partial_{I_{N_i}}$  a partial derivative after  $I_i$ . The solution of the update equation (3.21) is given by

$$I^{t+1} = \sum_i \gamma(\log(2\pi\sigma\sigma_i) - \frac{1}{\sigma^2}(C_i - I_i)^2 + \frac{1}{\sigma_i^2}(I_i - K(I_i))^2) \quad (3.22)$$

We can solve this by first neglecting the constant factor advancing to

$$I^{t+1} \approx \sum_i -\frac{1}{\sigma^2}(C_i - I_i)^2 + \frac{1}{\sigma_i^2}(I_i - K(I_i))^2, \quad (3.23)$$

We get to the optimal segmentation when the image  $I^{t+1}$  is the same as  $I^t$  before the iteration. Therefore the changes between  $I^t$  and  $I^{t+1}$  should vanish:

$$\sum_i (C_i - I_i^t)^2 + \varpi(I_i^t - K(I_i^t))^2 = 0. \quad (3.24)$$

We arrive at the same optimal solution as given in (3.19)

$$C^* = \mathcal{F}^{-1} \left( \frac{\tilde{C}}{1 + \varpi(\tilde{k}')^*(\tilde{k}')} \right). \quad (3.25)$$

### 3.3.3 DICPC algorithm

All in all, the contextual classification approach that is proposed looks as follows:

1. Learn the weights  $\vec{w}_{man}$  from manual segmentations and construct a 2D-filter  $k$ .
2. Define a distribution to be used in the prior  $P(C_i|C_{N_i})$ .
3. Define a distribution to be used in the conditional probability  $P(I_i|C_{N_i})$ .
4. Use a direct or a variational approach to solve (3.14).

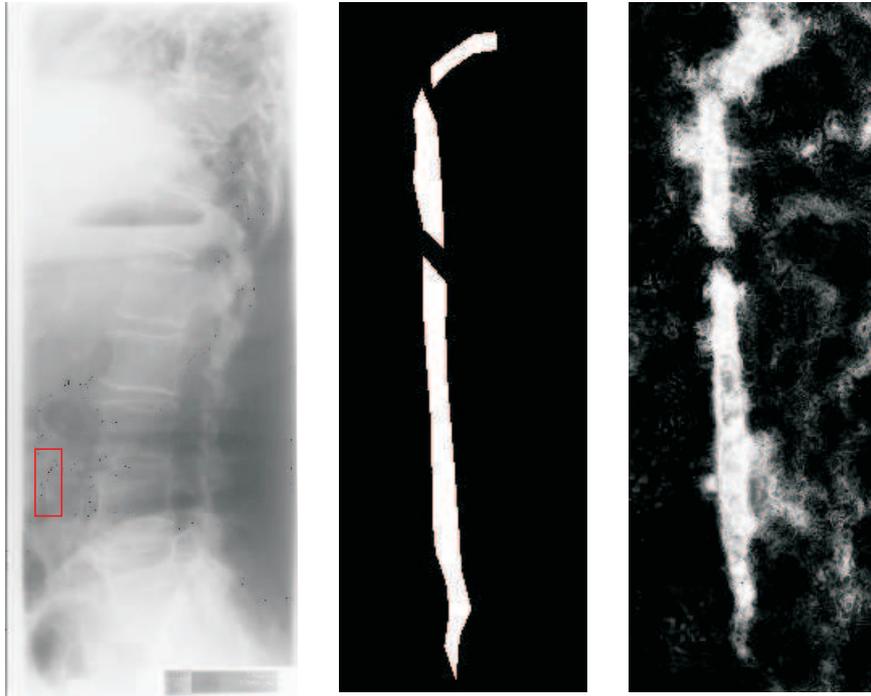


Figure 3.1: The graph displays one of the original images and a manual as well as an automatic segmentation of a calcification. Note how much background noise there is present in the automatic segmentation.

## 3.4 Methods

### 3.4.1 Study population

The study population used in this paper is the EPI follow-up population, which was part of the multi-centered PERF Study [66]. The EPI part of the PERF study was an epidemiological study addressing the role of a number of metabolic risk factors in the pathogenesis of CVD and osteoporosis [67] carried out in Ballerup, Denmark, in 1992 and 2001. Three trained radiologists, unaware of the patients conditions, annotated the vertebrae, the aorta and the calcifications in the digitized X-ray images. They used Sectra radiological reading units and annotation software implemented in MatLab (Mathworks, MA, USA).

Five images were randomly selected and 18 patches with one or multiple calcifications as illustrated in figure 3.1 were acquired. The testing of the different methods was performed on these 18 calcification patches.

### 3.4.2 Evaluation Methods

To measure the compliance of an annotation  $A_1$  and the outcome of a classification  $A_2$ , first the Jaccard index [68] is used. This is a quantification measure for segmented areas. The ratio of the number of pixels present in both segmentations to the total number of pixels in the segmentations is taken. The Jaccard index varies from 0 which equals no overlap to 1 corresponding to complete overlap,

$$r_{\text{Jacc}} = \frac{|A_1 \cap A_2|}{|A_1 \cup A_2|} \quad (3.26)$$

Furthermore we use sensitivity  $r_{\text{Sens}}$  and specificity  $r_{\text{Spec}}$  to evaluate the performance of our contextual classification methods.

Note that these measurements require the images to have the same resolution. Furthermore, the errors of  $r_{\text{Jacc}}$ ,  $r_{\text{Sens}}$  and  $r_{\text{Spec}}$  increase with smaller objects.

### 3.4.3 Classification and Kriging Settings

We use a  $k$ -Nearest-Neighbor classifier [69] with  $k = 25$ . Training pixels are selected with a bias toward calcified pixels, meaning that 30% of the training pixels chosen are calcified pixels which are defined according to the manual annotations. The features used were the intensity, the gradient magnitude, the Hessian trace, the Hessian determinant and the Hessian eigenvalues, an adaption of Koenderink's shape classification measure [70] and the ratio of the difference and sum of the Hessian eigenvalues. All features were calculated at three different scales, corresponding to 2, 5 and 17 mm.

In our method, we learn the Kriging filter for calcifications from manual segmentations by calculating the weights for a seven by seven neighborhood in which we krig to the central pixel. This way we arrive at a seven by seven filter that is used as  $k$  in (3.11). In principle one can use any size neighborhood, since the filtering process is computationally very fast even for large filters.

### 3.4.4 Comparison to other methods

To investigate the performance of our method we compare it to post-processing methods used for the segmentation of lumbar aortic calcifications.

#### Disk Morphology

The morphological operations used are an opening and a closing with a disk of the size of 1 mm. We chose 1 mm in order to remove pixel noise, but not parts of calcifications.

### Biological Morphology

The morphological operations used are an opening and a closing with structuring elements derived from biological findings. Larsen et al. showed in [71] that the size of lumbar aortic calcifications is approx.  $5 \pm 3$  mm in height and  $2 \pm 1$  mm in width. According to these findings the first structuring element was designed to remove everything smaller than a standard deviation of a calcification in size. Therefore we used a disk of the size of 1 mm for the opening. For the closing, we made use of a rectangular structuring element of the size of one standard deviation of a calcification in width (1 mm) and one standard deviation of a calcification in height (3 mm).

## 3.5 Results

The pure pixels classification as well as the pixel classification in correspondence with the three different methods, disk morphology (figure 3.2(a)), biological morphology (figure 3.2(b)) and DICPC (figure 3.2(c)), were evaluated for the 18 calcification patches at the kNN threshold (70% to 30% sampling leads to a threshold of  $\frac{17}{25}$ ). In general we can observe that the Kriging produces results that are much closer to the original pixel classification than the morphological operations. The morphology imprints the shapes of its structuring elements onto the pixel classification result and produces harsh boundaries. The Kriging, on the contrary, makes the pixel classification boundaries finer and even prescinds structures out of the background around the calcifications. The average values for the Jaccard index, sensitivity and specificity for the 18 patches are given in table 3.1. We can observe that the Kriging Prior increases the Jaccard index and the sensitivity compared to all other methods while leaving the specificity unchanged compared to the pure pixel classification. Furthermore the statistical significance of difference between the means of the different methods was tested via a paired one-sided Student's t-test. The t-tests confirmed that the in figure 3.2 observed differences of the Kriging compared to the other methods are significant for the Jaccard index, the sensitivity and the specificity. The results of the tests are shown in figures 3.3(a),3.3(b),3.3(c).

## 3.6 Discussion and Conclusion

When dealing with high resolution medical images that present long range interactions one runs into computational problems when trying to use standard contextual classification techniques like ICM or ICPC. This is why we compare our new method to other techniques, disk and biological morphology, which are common post-processing methods for this application.



(a) The result of the disk morphology. (b) The result of the biological morphology. (c) The result of DICPC.

Figure 3.2: Results for different morphological operations and DICPC. The morphology imprints the shapes of its structuring elements onto the pixel classification result and produces harsh boundaries. The Kriging, on the contrary, makes the pixel classification boundaries finer and even prescinds structures out of the background around the calcifications.

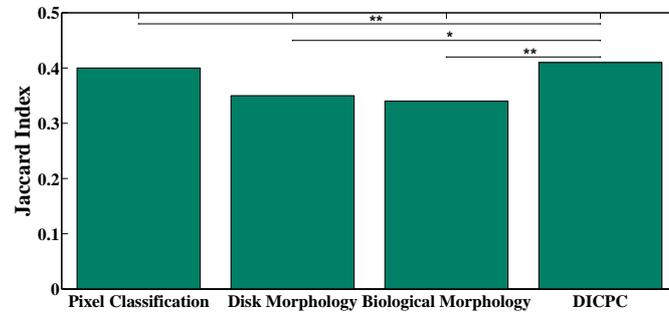
Table 3.1: The area overlap results for the population at the kNN threshold (70% to 30% sampling leads to a threshold of  $\frac{17}{25}$ ). We can observe that the Kriging Prior increases the Jaccard index and the sensitivity compared to all other methods while leaving the specificity virtually unchanged compared to the pure pixel classification.

	Pure Pixel Classification	Disk Morphology	Biological Morphology	Kriging Prior
Jaccard index	40%	35%	34%	41%
Sensitivity	0.60	0.41	0.43	0.62
Specificity	0.96	0.99	0.98	0.95

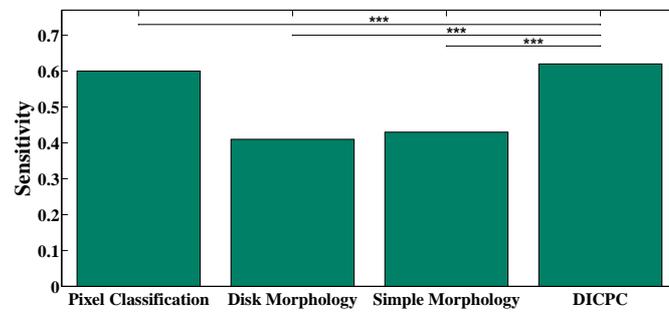
We observe that our new method, DICPC, improves the sensitivity and the Jaccard index, while it leaves the specificity almost unchanged. The other methods, disk morphology and biological morphology, even lower the Jaccard index and the sensitivity while improving the specificity inconsiderably. The lowering of the Jaccard index and the sensitivity is caused by the relatively harsh boundaries that the morphological operations produce in contrast to the Kriging.

Kriging and therefore DICPC is only the first step in the right direction. A weakness of DICPC is the linear model that underlies Kriging. It implies only pairwise interactions. The goal is to develop a contextual method that is as

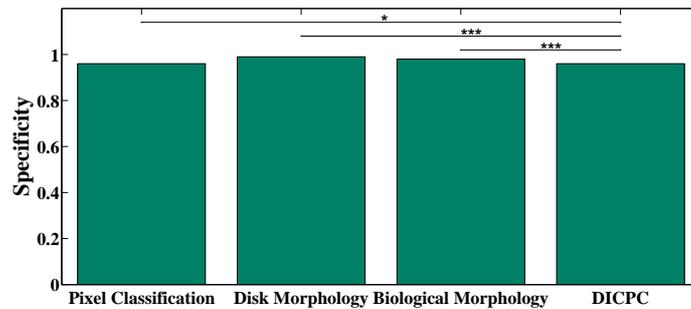
fast and computationally feasible as DICPC, but based on joint probabilities of the class labels. The next chapter introduces such a method based on texton dictionaries.



(a) Comparison of the Jaccard index of the different methods



(b) Comparison of the sensitivity of the different methods



(c) Comparison of the Specificity of the different methods

Figure 3.3: The graph displays the results for the paired Student's t-tests between the different methods for the Jaccard index, sensitivity and specificity. The stars indicate the outcome of a paired one-tailed Students' t-test:  $\star < 0.05$ ,  $\star\star < 0.01$  and  $\star\star\star < 0.001$ . The in figure 3.2 observed differences of the Kriging compared to the other methods are significant for the Jaccard index, the sensitivity and the specificity.



# Chapter 4

## A Texton-based Generative Shape Model and MDL Model Selection for Statistical Analysis of Archipelago-Like Textures

---

This chapter is based on the manuscript "Patch-based Generative Shape Model and MDL Model Selection for Statistical Analysis of Archipelagos" by Melanie Ganz, Mads Nielsen and Sami Brandt published in the proceedings of the International Workshop on Machine Learning in Medical Imaging (MLMI) in conjunction with MICCAI 2010 and the manuscript "A Texton-based Generative Shape Model and MDL Model Selection for Statistical Analysis of Archipelago-Like Textures" by Melanie Ganz, Mads Nielsen and Sami Brandt currently submitted for publication.

---

**Abstract** In the following we propose a statistical generative shape model for archipelago-like structures. These kind of structures occur, for instance, in medical images, where our intention is to model the appearance and shapes of calcifications in X-ray images. The generative model is constructed by (1) learning a patch-based dictionary for possible shapes, (2) building up a time-homogeneous Markov model to model the neighborhood correlations between the patches, and (3) automatic selection of the model complexity by the minimum description length principle. The generative shape model is proposed as a probability distribution of a binary image where the model is intended to facilitate sequential simulation. Our results show that a relatively simple model is able to generate structures visually similar to the training images. Furthermore, we use the shape model as a shape prior in the statistical segmentation of calcifications, where the area overlap with the ground truth shapes improved significantly compared to the case where the prior was not used.

### 4.1 Introduction

In the field of computer vision as well as medical imaging one essential problem is the handling of texture. Textures have a wide-spread nature: they



Figure 4.1: The textures from regular to stochastic (Image courtesy of Wen-Chieh Lin [72]).

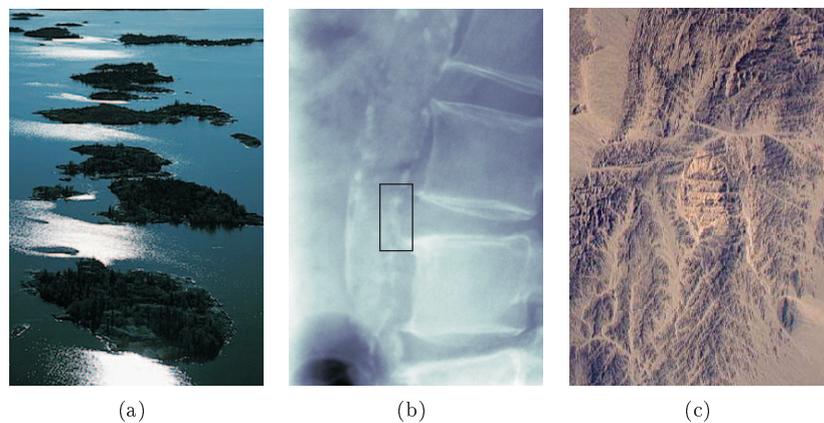


Figure 4.2: Illustrations of archipelago-like structures (a) in nature, (b) in a lumbar aortic X-ray and (c) in geostatistics.

can be classified as either regular or stochastic and there exists a broad spectrum of textures between those two extremes [72] (figure 4.1). Archipelago-like structures (see figure 4.2) would fall in the category of near-stochastic texture. In this case, traditional methods for texture as well as shape modeling fail and other methods, likely based on statistics, are needed. An example of a recent statistical method in signal, image, and video processing is "sparse representations" [73]. Under the assumption that natural images admit a sparse decomposition, a dictionary is learned from training data and can be used for any canonical texture problem: synthesis, classification, segmentation, compression or shape from texture.

We have adopted the sparsity method on two problem areas:

1. Texture synthesis which can be used to solve several practical problems in computer vision, graphics, and image processing from geostatistical modeling of water beds [74, 75] over handling occlusions in 3D reconstruction [76] to inpainting problems [77].
2. Texture segmentation which can be used in many biological segmentation problems that deal with archipelago-like structures, e.g., brain lesions as observed in MRI [78] or calcified deposits in the arteries observed by X-ray [52] or CT imaging methods [79].

In this chapter, we will develop a method for texture synthesis and texture based segmentation and focus on two application areas, medical image segmentation problems and natural image texture synthesis. First, we will define our statistical objective (section 4.2) before describing our generative shape model (section 4.3) and how we intend to use it in the case of segmentation (section 4.4). Then we will present our experiments (section 4.5). On the one hand, the texture synthesis will be examined on natural image data, e.g. water and fire. On the other, in the medical domain an example of lumbar aortic X-ray projections will be used, where our goal is to automatically segment lumbar aortic calcifications that are related to cardiovascular disease (CVD) [10,30,80]. We will finish with a short discussion of the achieved results (section 4.6) and conclude with an evaluation of our proposed algorithm (section 4.7).

## 4.2 Statistical Objective

In our medical imaging application we are interested in a general segmentation problem, namely segmenting calcifications from the background in the lower abdominal aorta in X-ray images. We model the aorta through a latent variable model with the values one and zero for calcifications and background, respectively, and estimate the posterior distribution by combining the likelihood of the pixel data with a shape prior given by our generative shape model. The posterior distribution is then given by

$$p(\mathbf{u}|\mathbf{c}) \propto p(\mathbf{c}|\mathbf{u})p(\mathbf{u}), \quad (4.1)$$

where  $\mathbf{u} = (u_1, u_2, \dots, u_L)$  is our latent variable vector, a vector of unknown binary pixel labels of the image matrix  $\mathbf{I}$ , and  $\mathbf{c} = (c_1, c_2, \dots, c_L)$  is the binary, hard classification vector, which is a function of the image  $\mathbf{I}$ , corresponding to the pixel  $l = 1, 2, \dots, L$ .

The likelihood function is constructed as follows. In addition to the hard classification  $\mathbf{c}$ , the classifier returns pixel classification probabilities  $\gamma_l = P(c_l = 0)$  of the pixel  $l$  being zero. Conversely,  $1 - \gamma_l = P(c_l = 1)$  is the probability of the pixel  $l$  being one. Let  $r_l = |c_l - u_l|$  be the pixel wise residual. Then one can divide the discrete outcomes for the residual in four possible cases:

1. If  $u_l = 0$  and  $\gamma_l \geq 0.5$ , it follows that  $c_l = 0$  and  $c_l = u_l$ . Thus,  $r_l = 0$ .
2. If  $u_l = 0$  and  $\gamma_l < 0.5$ , it follows that  $c_l = 1$  and  $c_l \neq u_l$ . Thus,  $r_l = 1$ .
3. If  $u_l = 1$  and  $1 - \gamma_l \geq 0.5$ , it follows that  $c_l = 1$  and  $c_l = u_l$ . Thus,  $r_l = 0$ .
4. If  $u_l = 1$  and  $1 - \gamma_l < 0.5$ , it follows that  $c_l = 0$  and  $c_l \neq u_l$ . Thus,  $r_l = 1$ .

Now we assume for the residual the discrete distribution  $\{P(r_l = 0), P(r_l = 1)\} = \{\max\{\gamma_l, 1 - \gamma_l\}, \min\{\gamma_l, 1 - \gamma_l\}\}$ , where we use  $\gamma_l, l = 1, 2, \dots, L$ , as the parameters of the residual distribution, that models the noise distribution or the probabilities for correct and misclassification, respectively. This choice yields the likelihood function

$$\begin{aligned} p(\mathbf{c}|\mathbf{u}) &= \prod_l (\gamma_l^{1-c_l} \gamma_l^{c_l})^{1-u_l} ((1 - \gamma_l)^{c_l} (1 - \gamma_l)^{1-c_l})^{u_l} \\ &= \prod_l \gamma_l^{1-u_l} (1 - \gamma_l)^{u_l}. \end{aligned} \quad (4.2)$$

Our goal is to construct the shape prior  $p(\mathbf{u})$  that statistically models the structures of archipelagos shown in figure 4.2.

### 4.3 Generative Shape Model

To construct a prior model for archipelago-like structures, we first build a texton dictionary (section 4.3.1) that contains the patch prototypes in which the structures are represented as building blocks. This dictionary will subsequently be used to define texture models based on texton frequencies learnt from training images. The grammar that models the neighborhood relations between the textons will be constructed via two different Markov mesh random fields (Section 4.3.2). The texton size  $m \times m$  and number of patches  $k$  in the dictionary will be selected by the minimum description length (MDL) principle (section 4.3.4), which completes our prior model for archipelago-like textures.

#### 4.3.1 Texton Dictionary

To construct the texton codebook, we extract  $n$  training textons by sliding a window of the size  $m \times m$  over each training image. Some example training images are shown in figure 4.3. Let the matrix  $\mathbf{X}$  contain the  $n$  training textons each stacked into a column vector. The textons are to be summarized by the  $m^2 \times k$  texton dictionary  $\mathbf{D}$  that contains the binary patch prototypes and that minimizes

$$E = \|\mathbf{X} - \mathbf{D}\mathbf{A}\|_{\text{fro}}^2, \quad (4.3)$$

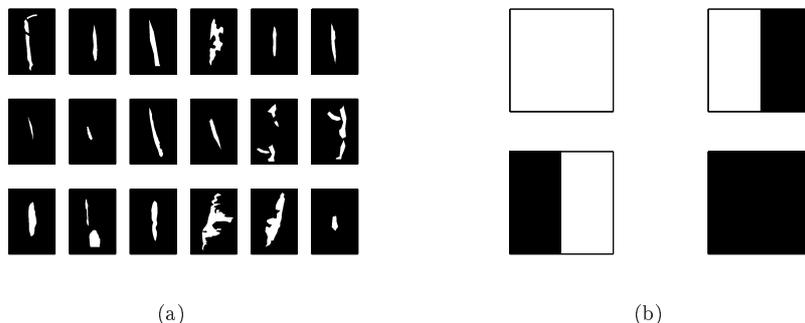


Figure 4.3: (a) Some training images of lumbar aortic calcifications. (b) An example of a dictionary retrieved from training on all training images for a model with the patch size  $m = 2$  and the cluster number  $k = 4$ .

```

1: Initialize the  $k$  cluster centers randomly
2: while Cluster centers change do
3:   Find the nearest center for each data point via the Hamming distance
   Update the cluster centers by calculating the mean of all data points belonging to
   a cluster and projecting it back to the binary manifold4
4: end while

```

**Algorithm 1:** Binary K-Means

where, for a fixed  $j$ ,  $a_{ij} = 1$  for only one  $i = i'$ , while  $a_{ij} = 0$  when  $i \neq i'$  [81] and  $\|\cdot\|_{\text{fro}}$  indicates the Frobenius norm.  $\mathbf{A}$  has the size  $k \times n$  and thus represents the sparse representation of  $\mathbf{X}$  in terms of  $\mathbf{D}$ . In general, we should minimize (4.3) over both  $\mathbf{D}$  and  $\mathbf{A}$ , but because it is a combinatory discrete optimization problem, we are satisfied by approximating the solution. We thus divide the problem into two parts:

1. We find the texton dictionary  $\mathbf{D}$  via clustering the training textons by utilizing a binary version of the  $K$ -means [82] algorithm (Algorithm 1), where the Euclidean distance used in our earlier work [83] has been replaced by the Hamming distance.
2. We find the optimal  $\mathbf{A}$ , given the dictionary  $\mathbf{D}$ , by picking up the prototype for each  $j$  that minimizes 4.3.

Clearly, the dictionary is not globally optimal, but it gives us a fair model class with varying patch sizes  $m \times m$  and number of clusters  $k$ . The model selection, i.e., determining  $m$  and  $k$  will be described in section 4.3.4.

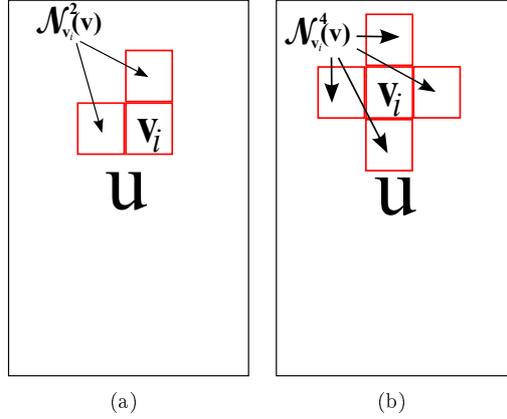


Figure 4.4: The causal neighborhood  $\mathcal{N}_{v_i}$  for the patch  $v_i$  which is a subset of the image represented by the latent variable vector  $u$  (a) in the 2-neighborhood and (b) 4-neighborhood Markov mesh random field model.

### 4.3.2 Texton Grammar Models

Our texton dictionary does not yet describe the archipelago-like structures well, even though we could easily generate a random image that has a similar texton histogram by trivially counting the occurrence of each patch in the training images and generating a random image by drawing random patches from the empirical patch distribution. The problem is that the neighboring patches are not independent, i.e., the neighbor patches significantly constrain the outcome of a patch.

To take these neighbor correlations into account we develop two different causal contextual texton grammar models. The models are called contextual because they are based on spatial interaction and take the neighborhood of a pixel into account. Furthermore, a contextual model can be classified as either causal or non causal depending on the nature of the neighborhood. In causal models, the concept of the past of a pixel is introduced. In this case the past neighborhood of a pixel refers to a neighborhood that has already been fixed and only that neighborhood influences the label of the current pixel. Causal Markov random fields (MRF) are generally called Markov mesh random fields (MMRF) or simply Markov meshes. MMRF models are advantageous, because when designing a prior it needs to be taken in consideration that sampling from it should be feasible. Furthermore, different visiting orders of the patches in the image can be considered. In the following, we will for clarity use a specific simple visiting order instead of a general one, but we will generalize it in section 4.3.3.

In the two MMRF models we will introduce below, the probability distribution

---

<sup>4</sup>This is equivalent to finding the cluster centers in the squared Hamming distance.

of our latent variable vector  $\mathbf{u}$  is dependent on the patches  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N$  we divide the image into

$$p(\mathbf{u}) = p(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N). \quad (4.4)$$

The probability distribution of the patches becomes

$$\begin{aligned} p(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N) &= p(\mathbf{v}_1)p(\mathbf{v}_2|\mathbf{v}_1) \dots p(\mathbf{v}_N|\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{N-1}) \\ &\stackrel{\text{Markov}}{\cong} \prod_{i=1}^N p(\mathbf{v}_i|\mathcal{N}_{\mathbf{v}_i}(\mathbf{v})), \end{aligned} \quad (4.5)$$

where  $\mathcal{N}_{\mathbf{v}_i}$  denotes the causal neighborhood of  $\mathbf{v}_i$ ,  $i = 1, 2, \dots, N$  and  $N$  is the total number of distinct patches of size  $m \times m$  in the image. This construction allows sequential simulation of the patch distribution by first drawing the patch  $\mathbf{v}_1$  from  $p(\mathbf{v}_1)$  at time point 1, then  $\mathbf{v}_2$  from  $p(\mathbf{v}_2|\mathbf{v}_1)$  at time point 2, and so on. The first model we consider is a 2-neighborhood MMRF. By using a 2-neighborhood MMRF we assume that the current patch probability depends only on the two neighbors that have previously been processed (see figure 4.4(a)). Therefore the causal neighborhood used in (4.5) has the form

$$\mathcal{N}_{\mathbf{v}_i}(\mathbf{v}) = \mathcal{N}_{\mathbf{v}_i}^2(\mathbf{v}) \cap \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{N-1}\}. \quad (4.6)$$

The second model is a 4-neighborhood MMRF model where the current patch probability depends on those elements of a 4-neighborhood that have been processed. The neighborhood can be seen in figure 4.4(b). This model is chosen for better spatial symmetry when compared to the 2-neighborhood model. Accordingly  $\mathcal{N}_{\mathbf{v}_i}(\mathbf{v})$  changes to

$$\mathcal{N}_{\mathbf{v}_i}(\mathbf{v}) = \mathcal{N}_{\mathbf{v}_i}^4(\mathbf{v}) \cap \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{N-1}\}. \quad (4.7)$$

The probability distributions introduced above can be estimated from the 3- or 5-dimensional histogram  $H$  of patch labels and their neighborhoods  $\mathcal{N}_{\mathbf{v}_i}^2$  or  $\mathcal{N}_{\mathbf{v}_i}^4$ .

### 4.3.3 Visiting Order

To complete the description of our shape prior, we will in the following introduce the indexing  $i_1, i_2, \dots, i_N$  that defines a patch visiting order. Let us define the indexing for the original patches as shown in figure 4.5(a). The patches can be visited in various orders, such as  $i_1 = 1, i_2 = 2, i_3 = 3, \dots, i_N = N$ , as above, to which we will refer as simple visiting order, or at random, such as  $i_1 = 25, i_2 = 5, i_3 = 9, \dots, i_N = 73$ , see figure 4.5(b).

Since we have chosen to sample sequentially from our prior, instead of using global optimization techniques like e.g. graphcut [46], it is instructive to show

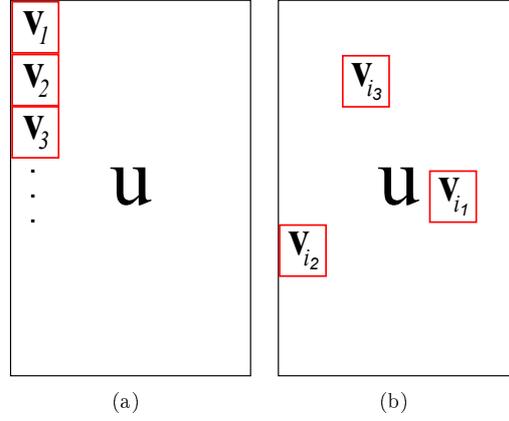


Figure 4.5: Different patch visiting orders of the latent variable vector  $\mathbf{u}$ : (a) The original simple patch order. (b) An example of a different causal visiting order - a random visiting order.

how the visiting order affects the prior distribution. Sequential sampling is based on the factorization

$$\begin{aligned}
 p(\mathbf{u}) &= p(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N) \equiv p(\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \dots, \mathbf{v}_{i_N}) \\
 &= p(\mathbf{v}_{i_1}) \cdot p(\mathbf{v}_{i_2} | \mathbf{v}_{i_1}) \dots p(\mathbf{v}_{i_N} | \mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \dots, \mathbf{v}_{i_{N-1}}) \\
 &\cong p(\mathbf{v}_{i_1}) \cdot p(\mathbf{v}_{i_2} | \mathcal{N}_{\mathbf{v}_{i_2}}(\mathbf{v})) \dots p(\mathbf{v}_{i_N} | \mathcal{N}_{\mathbf{v}_{i_N}}(\mathbf{v})), \tag{4.8}
 \end{aligned}$$

where the visiting order  $i_1, i_2, \dots, i_N$  affects for which patches the approximation, following from the Markov assumption, is made.

We experimented with two visiting orders. In the case of the 2-neighborhood MMRF we choose to use the simple visiting order, whereas in the case of the 4-neighborhood MMRF we apply a random visiting order. In summary, the resulting priors render as

$$p(\mathbf{u}) \cong p(\mathbf{v}_1) \cdot p(\mathbf{v}_2 | \mathcal{N}_{\mathbf{v}_2}(\mathbf{v})) \dots p(\mathbf{v}_N | \mathcal{N}_{\mathbf{v}_N}(\mathbf{v})), \tag{4.9}$$

where in the 2-neighborhood case

$$\mathcal{N}_{\mathbf{v}_i}(\mathbf{v}) = \mathcal{N}_{\mathbf{v}_i}^2(\mathbf{v}) \cap \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{N-1}\} \tag{4.10}$$

and in the 4-neighborhood case

$$\mathcal{N}_{\mathbf{v}_i}(\mathbf{v}) = \mathcal{N}_{\mathbf{v}_i}^4(\mathbf{v}) \cap \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_{N-1}\}. \tag{4.11}$$

#### 4.3.4 Model Selection

To use our proposed model we need to find the optimal cluster number  $k$  and optimal patch size  $m$  and estimate the transition probabilities for our Markov

model. We decided to use MDL [84] for the model selection due to its tangible definition of the model selection problem as the best model is defined to have the minimal lossless transmission code length. MDL exactly fits to our purpose, since we are dealing with a binary problem for which it is easy to construct a compression model. Moreover, MDL provides a natural definition for noise, as noise is considered everything that can not be compressed by the model [85]. Let us first derive the code length for our model using a two-part coding model. The total code length of our model in bits is

$$L = L_{\text{par}} + L_{\text{res}}, \quad (4.12)$$

where  $L_{\text{par}} = L_{\mathbf{D}} + L_{\mathbf{A}}$  is the code length of the model parameters and  $L_{\text{res}}$  the code length of the residual. We choose to code  $\mathbf{D}$  simply as a binary matrix, so one needs  $m^2 \times k$  bits to encode it, hence

$$L_{\mathbf{D}} = m^2 \times k + \underbrace{\lceil \log_2(\max k) \rceil}_k + \underbrace{\lceil \log_2(\max m) \rceil}_m, \quad (4.13)$$

where the latter two terms, code lengths for  $k$  and  $m$ , are constant and can thus be dropped. The content of  $\mathbf{A}$  can be encoded by using the time-homogeneous Markov model as soon as the 3- or 5-dimensional histogram  $H$  of patch labels and their causal neighborhoods is available. The histogram can be encoded either, if sparse, by storing its  $N_{\text{nnz}}$  non-zero bin indices, and the counts in such bins; or otherwise by storing the counts in all the bins. In this way, assuming an ideal coding method,

$$L_{\mathbf{A}} = \underbrace{\min(N_{\text{nnz}} \cdot \lceil \log_2(n) \rceil^2 + \lceil \log_2(N_{\text{nnz}}) \rceil, k^3 \cdot \lceil \log_2(n) \rceil)}_H - \underbrace{\sum_{\text{data}} \log_2(\hat{p}_k)}_{\text{data}}, \quad (4.14)$$

where the conditional probability  $\hat{p}_k = p(\mathbf{v}_{i_k} | \mathcal{N}_{\mathbf{v}_{i_k}}(\mathbf{v}))$  of the patch  $k$  is computed from the histogram  $H$ .

Finally, let us consider the residual encoding, where the residual of our model is  $\epsilon = \mathbf{X} - \mathbf{DA}$  and each pixel can obtain only values  $\{-1, 0, 1\}$ . We can thus code  $\epsilon$  by only transmitting the indices of first the negative and then the positive entries of the residual. In this way the code length for  $\epsilon$  in bits becomes

$$L_{\text{res}} = q \lceil \log_2(N_{\text{pix}}) \rceil + \underbrace{\log_2[q]}_q, \quad (4.15)$$

where  $q$  is the number of non-zero residuals and  $N_{\text{pix}}$  is the number of pixels in the image. The latter term is bounded by  $\log_2 \lceil N_{\text{pix}} \rceil$  and can thus be dropped.

## 4.4 Sampling from the Posterior with the Shape Prior

Let us consider the simulation of the posterior (4.1) to obtain realizations for the latent variable  $\mathbf{u}$  as

$$p(\mathbf{u}|\mathbf{c}) \equiv p(\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \dots, \mathbf{v}_{i_N}|\mathbf{c}). \quad (4.16)$$

If the likelihood function is separable, we may use the same causal construction with which the prior was designed. Hence, the posterior at time point 1 becomes

$$p(\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \dots, \mathbf{v}_{i_n}|\mathbf{c}) = p(\mathbf{v}_{i_1}|\mathbf{c})p(\mathbf{v}_{i_2}|\mathbf{v}_{i_1}, \mathbf{c}) \dots p(\mathbf{v}_{i_n}|\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \dots, \mathbf{v}_{i_{n-1}}, \mathbf{c}), \quad (4.17)$$

where

$$\begin{aligned} p(\mathbf{v}_{i_N}|\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \dots, \mathbf{v}_{i_{N-1}}, \mathbf{c}) &\propto p(\mathbf{c}|\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \dots, \mathbf{v}_{i_{N-1}})p(\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \dots, \mathbf{v}_{i_{N-1}}) \\ &\propto \prod_l \gamma_l^{1-u_l} (1 - \gamma_l)^{u_l} p(\mathbf{v}_{i_1}, \mathbf{v}_{i_2}, \dots, \mathbf{v}_{i_{N-1}}) \end{aligned} \quad (4.18)$$

Here  $l$  denotes the element of the latent variable vector  $\mathbf{u}$  and  $\gamma_l$  and  $1 - \gamma_l$  are again the probabilities of the pixel  $l$  having the label 0 or 1.

We thus assume that the posterior is similarly sequentially simulated by first drawing the patch  $\mathbf{v}_{i_1}$  from  $p(\mathbf{v}_{i_1}|\mathbf{c})$  at time point 1, then  $\mathbf{v}_{i_2}$  from  $p(\mathbf{v}_{i_2}|\mathbf{v}_{i_1}, \mathbf{c})$  at time point 2, etc.

The same sampling strategy will be used for drawing samples from the shape prior in section 4.5.

## 4.5 Experiments

### 4.5.1 Modeling different Shape Distributions

#### Model Selection

In our experiments for the lumbar aortic x-ray data, we used a training set of 18 manually annotated calcifications (figure 4.3(a)). The manual annotations are binary, where the value 1 equals a calcified pixel, while the value 0 corresponds to a background pixel. We parametrized the model class with the set of all pairs of patch sizes and cluster numbers  $\{m, k\}$ ,  $m \in \{2, 4, 6, 8, 10\}$  and  $k \in \{2, 4, 8, 16, 32\}$ , over which we optimized the compression code length for the 2- as well as the 4-neighborhood. The results as per-pixel-normalized code lengths are shown in table 4.1(a) and 4.1(b). MDL selected the models  $m = 2$ ,  $k = 4$  and  $m = 4$ ,  $k = 2$  for the 2- and the 4-neighborhood cases, respectively. The learnt dictionary of patches for the model  $m = 2$  and  $k = 4$  is displayed in figure 4.3(b).

Table 4.1: The code length per pixel in bits. The color spectrum red to blue indicates the size of the code length. One can see that MDL most of the time selects small patch sizes and a small number of clusters.

(a) Calcification, 2-neighborhood.						(b) Calcification, 4-neighborhood.					
m \ k	2	4	8	16	32	m \ k	2	4	8	16	32
2	0.12	0.09	0.09	19.23	n.a.	2	0.51	0.93	3.45	44.29	n.a.
4	0.26	0.17	0.15	0.24	0.78	4	0.44	1.17	0.48	1.25	2.16
6	0.36	0.26	0.21	0.27	1.02	6	0.67	0.62	0.58	0.75	1.59
8	0.50	0.33	0.28	0.31	0.98	8	0.67	0.52	0.51	0.57	1.33
10	0.61	0.48	0.32	0.35	1.10	10	0.72	0.67	0.48	0.54	1.29

(c) Water, 2-neighborhood.							(d) Water, 4-neighborhood.						
m \ k	2	4	6	8	12	16	m \ k	2	4	6	8	12	16
2	1.82	0.28	0.25	0.24	18.29	18.49	2	11.38	7.85	8.05	12.99	31.16	31.35
4	3.15	1.83	1.37	0.82	0.72	0.94	4	4.10	3.82	4.18	3.62	3.72	3.15
6	3.85	2.21	1.98	1.92	1.59	1.60	6	4.75	3.83	3.45	3.15	2.99	2.89
8	4.36	3.38	2.83	2.30	2.16	1.99	8	4.70	3.92	3.66	3.22	2.85	2.74
10	4.70	3.54	2.98	2.87	2.52	2.49	10	4.94	4.00	3.67	3.41	3.02	3.00

(e) Geostatistical Channels, 2-neighborhood.							(f) Geostatistical Channels, 4-neighborhood.						
m \ k	2	4	6	8	12	16	m \ k	2	4	6	8	12	16
2	1.13	0.77	0.83	1.07	16.87	19.57	2	12.32	12.68	8.30	8.58	29.64	32.90
4	2.35	1.85	1.89	1.67	2.72	5.80	4	4.49	3.73	4.45	4.92	5.21	8.11
6	2.75	2.45	2.22	2.31	3.19	5.06	6	3.60	3.37	3.58	3.74	4.61	6.39
8	3.44	2.69	2.64	2.83	3.35	4.99	8	3.74	3.28	3.20	3.31	4.48	5.77
10	3.87	3.43	3.14	3.21	3.85	5.34	10	4.16	3.61	3.58	3.77	4.29	5.79

(g) Fire, 2-neighborhood.							(h) Fire, 4-neighborhood.						
m \ k	2	4	6	8	12	16	m \ k	2	4	6	8	12	16
2	1.20	0.98	0.97	0.82	0.71	17.96	2	3.13	5.66	8.20	4.67	30.28	30.61
4	2.09	1.85	1.65	1.69	1.68	1.83	4	2.61	3.92	4.58	3.42	4.04	4.29
6	2.58	2.24	2.08	2.01	2.05	2.27	6	3.21	3.17	3.19	3.03	3.35	3.27
8	2.82	2.52	2.36	2.26	2.27	2.43	8	3.19	3.07	2.93	2.83	2.87	3.03
10	3.07	2.78	2.58	2.52	2.52	2.66	10	3.18	3.17	3.02	2.92	2.89	3.14

The corresponding experiments were conducted for binary images of natural textures, such as water, geostatistical channel systems and fire (figure 4.7(a)-(c)) with  $k \in \{2, 4, 6, 8, 12, 16\}$ ; the results are displayed in table 4.1(c)-(h). In the case of water, MDL selected the model  $m = 2, k = 8$  and  $m = 8, k = 16$  for the 2- and the 4-neighborhood, respectively. For the image of geostatistical channel systems the MDL optimum was achieved at  $m = 2, k = 4$  and  $m = 8, k = 6$  for the 2- and the 4-neighborhood. Finally, MDL yields the minimal code length for  $m = 2, k = 12$  and  $m = 4, k = 2$  for the 2- and the 4-neighborhood, respectively, for the natural image of fire. In general, we can observe that MDL most of the time selects small patch sizes and a small number of clusters.

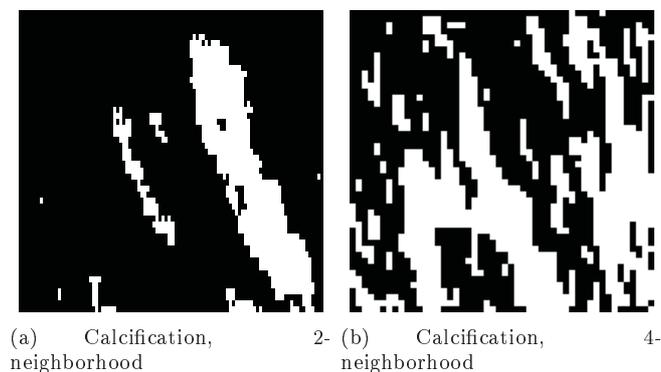


Figure 4.6: Texture synthesis results for calcification data for the 2- and 4-neighborhood MMRF model. In the case of the calcifications, the 2-neighborhood MMRF seems to produce a more adequate simulation than the 4-neighborhood MMRF model.

### Texture Synthesis

One way of testing a generative shape model is to synthesize textures from the learnt model. If the prior model is good, simulated textures should resemble real calcification textures. We simulated the prior with the MDL selected parameters as explained in Section 4.3.2, using textures from real images (c.f. figure 4.2(b) and 4.7(a)-(c)). Figure 4.6 illustrates the results for the calcification shape distribution, whereas the synthesis results for water, geostatistical channel systems and fire are shown in figure 4.7(d)-(i). It can be seen that the shapes are qualitatively similar to original shapes. It can also be seen that the 2-neighbor model with the simple visiting order produced superior results to the 4-neighborhood model with the random visiting order.

### 4.5.2 Statistical Shape Segmentation

To complete the experiments, we apply the generative shape model as a shape prior on a test set of 38 images displaying lumbar aortic X-ray data. In 30 of the 38 images radiologists have identified at least one calcification in the lumbar aorta while 8 images were judged to have no calcifications. We use the shape prior as described in Section 4.4 with a 2-neighborhood model and the simple visiting order together with the likelihood function. The pixel-wise likelihood was constructed from the pixel classification probabilities, as explained in section 4.2, where the initial segmentation is performed by a Random Forests classifier [37] with a set of 7 Gaussian derivative features, trained by manual annotations of calcified lesions.

To measure the performance of our segmentation, we draw  $S = 20$  samples  $\mathbf{u}^{(n)}$ ,  $n = 1, 2, \dots, S$  from the posterior distribution  $p(\mathbf{u}|\mathbf{c})$  and estimate the

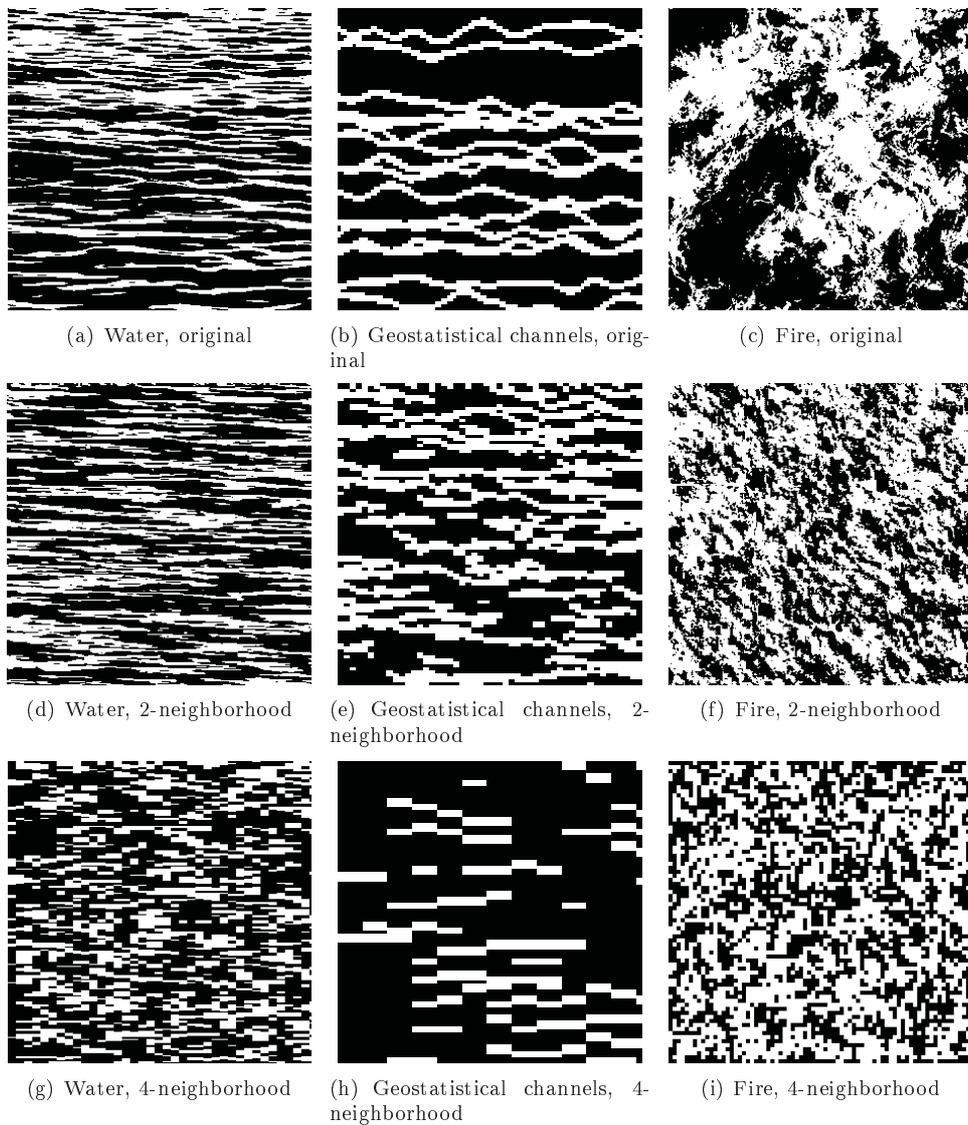


Figure 4.7: Texture synthesis results. In (a)-(c) the original natural image data is shown, whereas the synthesis results for the 2-neighborhood MMRRF model can be seen in (d) - (f) and for the 4-neighborhood MMRF model in (g)-(i). In general the 2-neighborhood MMRF produces better results than the 4-neighborhood MMRF model.

expected value of a chosen scoring function  $f_{\text{eval}}(\mathbf{u}; \mathbf{u}_{\text{ann}})$ , where  $\mathbf{u}_{\text{ann}}$  denotes

the ground truth annotation,

$$\begin{aligned}
 E\{f_{\text{eval}}(\mathbf{u}; \mathbf{u}_{\text{ann}}|\mathbf{c})\} &= \int f_{\text{eval}}(\mathbf{u}; \mathbf{u}_{\text{ann}})p(\mathbf{u}|\mathbf{c})d\mathbf{u} \\
 &\approx \frac{1}{S} \sum_n f_{\text{eval}}(\mathbf{u}^{(n)}; \mathbf{u}_{\text{ann}}) \\
 &= \overline{f_{\text{eval}}(\mathbf{u}; \mathbf{u}_{\text{ann}})}. \tag{4.19}
 \end{aligned}$$

We compare the resulting mean score with the value of  $f_{\text{eval}}(\mathbf{u}_{\text{ref}}; \mathbf{u}_{\text{ann}})$ , where  $\mathbf{u}_{\text{ref}}$  is the classification probability map thresholded at 0.5. As evaluation function  $f_{\text{eval}}(\mathbf{u}; \mathbf{u}_{\text{ann}})$  we use the Jaccard Index [68]

$$f_{\text{eval}}(\mathbf{u}; \mathbf{u}_{\text{ann}}) = \frac{|I_{\mathbf{u}} \cap I_{\text{ann}}|}{|I_{\mathbf{u}} \cup I_{\text{ann}}|}, \tag{4.20}$$

where  $I_{\mathbf{u}}$  and  $I_{\text{ann}}$  are  $\mathbf{u}$  and  $\mathbf{u}_{\text{ann}}$  in image matrix form, respectively. Thus the Jaccard index measures the area overlap between the binary segmentation results and the manual annotation, which we assume to be our ground truth. In the case where no calcifications were present in the manual aorta the area overlap is defined to be 0. The numerical results for  $f_{\text{eval}}(\mathbf{u}; \mathbf{u}_{\text{ann}})$  and  $f_{\text{eval}}(\mathbf{u}_{\text{ref}}; \mathbf{u}_{\text{ann}})$  are given in table 4.2. It can be seen that our method improves the classification results on average by 50% for all 38 as well as for only the 30 images with calcifications present. The improvement vs. simple thresholding is statistically significant according to a pairwise two-sided t-test with the value of  $p = 0.000001$  as well as  $p = 0.0000001$  in the two cases. Visual comparison of the manual annotation compared to our result is displayed in figure 4.8 for one of the subjects with the most improvement (from 0.143 to 0.337) and in figure 4.9 for one of the subjects with the least improvement (from 0.407 to 0.411). In figure 4.10 a detailed look at a segmentation with and without our prior is given.

## 4.6 Discussion

In our experiments, using MDL proved itself to be successful for choosing the optimal patch size and number of clusters. MDL chose small patch sizes supporting the limited amount of training data. Furthermore, it not only produced visually good results, but even though a simple local model was chosen the texture synthesis was able to produce global structures. In the case of more training data and a more complicated MMRF model with many patch prototypes, it is possible for MDL to select a model with a larger patch size  $m$  and cluster number  $k$ . In this case it might also be useful to use a more sophisticated coding method, e.g., universal coding [86].

Table 4.2: Evaluation of the segmentation results using the Jaccard index against the manual annotation. The Jaccard index improves by 50% after applying the proposed shape prior on a test set of 38 and 30 images displaying lumbar aortic X-ray data, respectively.

For all 38 test images	Mean $\pm$ Standard Deviation
$f_{eval}(\mathbf{u}_{ref}; \mathbf{u}_{ann})$	$0.10 \pm 0.11$
$\overline{f_{eval}(\mathbf{u}; \mathbf{u}_{ann})}$	$0.15 \pm 0.14$
For 30 test images with calcifications	Mean $\pm$ Standard Deviation
$f_{eval}(\mathbf{u}_{ref}; \mathbf{u}_{ann})$	$0.13 \pm 0.11$
$\overline{f_{eval}(\mathbf{u}; \mathbf{u}_{ann})}$	$0.19 \pm 0.13$

An alternative to MDL would have been to use cross validation. But because we deal with a very limited amount of training data, cross validation is not an optimal choice for us since in order to derive parameters cross validation divides the data further in a training, validation and testing subset. Conversely, MDL can take all training data into account at the same time when trying to optimize parameters.

Another observation was the remarkable difference between the 2- and the 4-neighborhood synthesis results. The 2-neighborhood produced much more realistic structures. Reasons for this could be that the simple visiting order used with the 2-neighborhood model supports the Markov constraint, while the 4-neighborhood with the random visiting order lays down global constraints that are not in accordance with the Markov assumption. In practice, fixing two distant patches constrains all selections of allowed patches in between the two patches, while the Markov assumption constrains only neighboring patches. Additionally, the 2-neighborhood model is also in better accordance with the limited amount of training data and therefore outperforms the 4-neighborhood model.

In general, our model performs well in the case of texture synthesis in comparison to earlier models based on patch MRF models for near stochastic texture as for example described in [87]. But a thorough evaluation of our algorithm in comparison to other near stochastic texture synthesis algorithms as done in [72] for near regular texture has still to be made.

The segmentation experiments show that our prior improves the Jaccard index

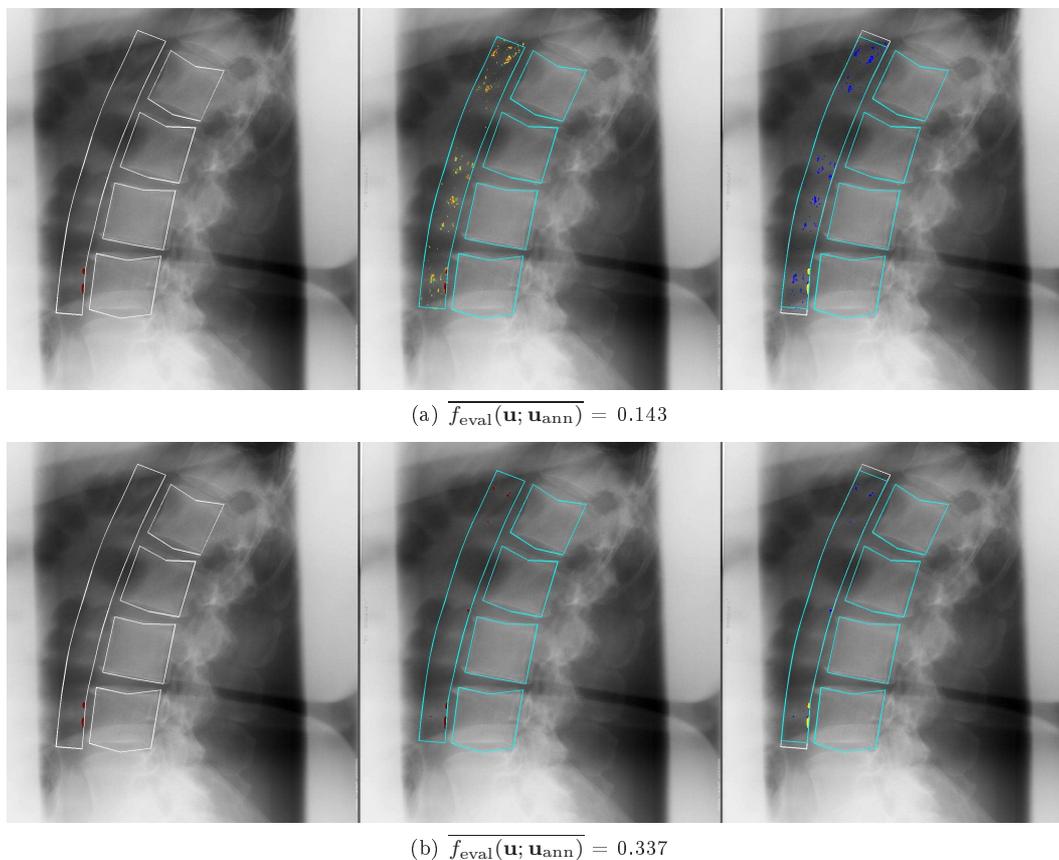


Figure 4.8: One of the subjects where the spatial prior improves the segmentation the most (from 0.143 to 0.337). In each image on the left the manual annotation of the vertebrae and the aorta outline (white) as well as the calcifications (red), in the middle the automatically retrieved vertebrae and aorta outline (green) as well as a probability map of the calcifications (yellow to red) and on the right the overlap of the two where yellow indicates true positives, blue indicates false positives and red indicates false negatives are shown.

by 50% (see table 4.2) and that the improvement is statistically significant. This is the case for all 38 images, of which 8 images have no calcifications present and therefore the Jaccard index is zero, as well as for the 30 images that all have calcifications present.

But Table 4.2 also indicates that the mean of the Jaccard index is low and its standard deviation high.

The main reason that the mean of the Jaccard index on our data set is low is that the fully automatic classification of calcifications, where first the vertebrae need to be detected and then the aorta region of interest defined before segmenting calcifications, is a very hard task. Figures 4.8 and 4.9 display two results from our dataset that exemplify how difficult the problem is. Further-

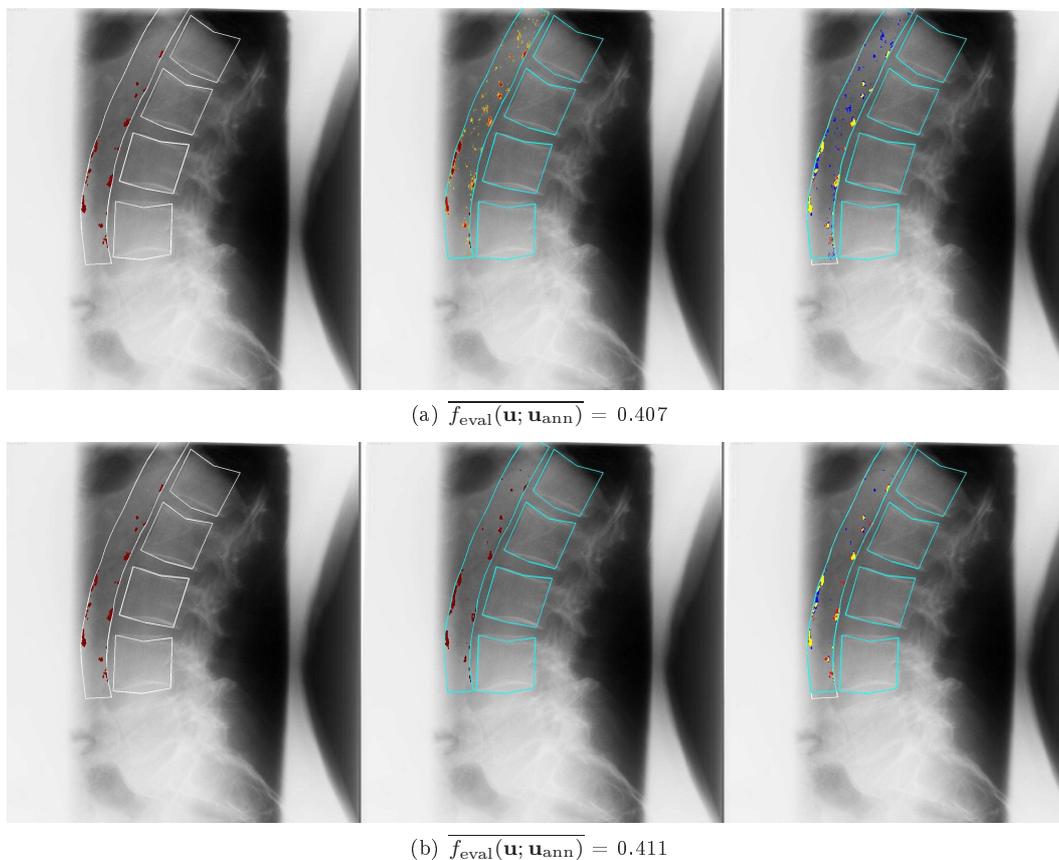


Figure 4.9: One of the subjects where the spatial prior improves the segmentation the least (from 0.407 to 0.411). In each image on the left the manual annotation of the vertebrae and the aorta outline (white) as well as the calcifications (red), in the middle the automatically retrieved vertebrae and aorta outline (green) as well as a probability map of the calcifications (yellow to red) and on the right the overlap of the two where yellow indicates true positives, blue indicates false positives and red indicates false negatives are shown.

more, two experts only have an Jaccard index of approx. 0.50 when scoring the same image underlining the fact that the segmentation of calcifications is a hard task. Finally, the Jaccard index is also low, because it is a very harsh measure to use in our problem. Our Jaccard index values can not be compared to values achieved in, e.g., brain imaging where two large areas are compared to each other. Alternatively, the number of correctly classified pixels or Cohen's  $\kappa$  [88] could be used to measure the inter-rater agreement for categorical items such as pixels. However, these measures will be dominated by the very large class of non-calcified pixels, and in the case of Cohen's  $\kappa$  individual pixel scorings cannot be considered statistically independent. Similar reasons apply to the question why the variance of our segmentation

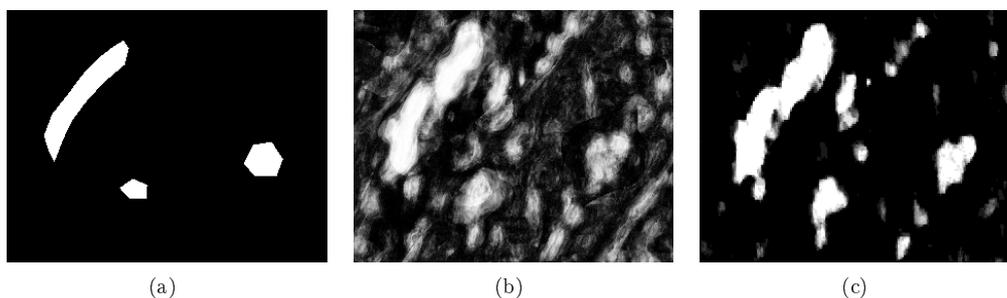


Figure 4.10: Calcification segmentation result: (a) manual annotation (ground truth), (b) the corresponding pixel-wise classification probabilities, (c) conditional mean  $\bar{\mathbf{u}} = \frac{1}{N} \sum \mathbf{u}^{(n)}$  of the posterior. Compare especially (b) and (c) to see how much background noise the proposed patch-based prior removes.

results is so high. Not only is the Jaccard index as described above a harsh measure for our problem, but on images with only a few or very small calcifications even a small number of false positives or false negatives leads to a large difference in the Jaccard index. Hence, the Jaccard index is in our application very sensitive to small changes and leads to a large standard deviation of our results. Another reason for the large standard deviation is the large variation in our images. A large biological variation over different subjects as well as large differences in image quality, make the preconditioning steps - finding the vertebrae and defining the aorta region of interest - more challenging and conversely lead to variable segmentation results and hence variable Jaccard indexes.

It is difficult to compare our work to earlier approaches for segmentation of calcifications in X-rays as was done in [89] and [90].

On the one hand, the results of [89] can not be directly compared with our results, since the preconditions are different. While in [89] the aorta is derived based on manually given landmark points at the corners and midpoints of the first four lumbar vertebrae, in our setup the vertebrae are found automatically without employing manual annotations. Since our experiments have shown that small errors in the vertebrae segmentation influence the calcification segmentation results greatly, we can not compare our results to the results given in [89]. On the other hand, the Jaccard index for the classification of calcifications is not given in [89] and the number of correctly classified pixels and Cohen's  $\kappa$  that are given are dominated by a large background class, which as mentioned skews the values of the given measures.

Comparison of the calcification segmentation results to [90] is also not straight forward, since it is also not a fully automatic approach, but assumes that the corner points of the vertebrae have been indicated and hence also differs in the preconditions from our setup. Furthermore, the Jaccard index is given, but

not for complete images but for 48 areas that were selected from 10 X-rays. This can not be directly compared to the performance on full images where all areas are taken into account.

## 4.7 Conclusion

In this paper, we have proposed a generative model and MDL model selection for structures resembling archipelagos. The model is based on a patch-based description of the shapes combined with a Markov Mesh Random Field model that takes patch correlations into account. Our selection for the dictionary, the binary  $K$ -means-clustered texton prototypes, seems reasonable even though it is not strictly optimal in the Frobenius norm. However, searching for the optimal codebook is itself a combinatorial optimization problem and less important in practice. As far as the MMRF model is concerned, our synthesis results showed that a simple local model based on only a few training examples can capture complex global structures and generate visually sound results. The visiting order was shown to influence the results greatly, whereas we concluded that a visiting order which is in accordance with the Markov assumption should be used. Our segmentation results, based on the proposed prior shape distribution, were promising and improved the area overlap by 50%. This indicates that our shape model can be successfully used as a prior distribution in statistical segmentation of calcifications on X-ray image data. An interesting direction in the future could be to introduce an appropriate multi-resolution extension of the generative model, which could take even longer interactions between patches into account.



# Chapter 5

## Application of Renormalization Group Theory to Multi-Scale Image Segmentation Problems

---

This chapter is based on work done under supervision of Pushmeet Kohli during the stays abroad at Microsoft Research Cambridge, UK and is currently in preparation.

---

**Abstract** We propose an approach for image segmentation that is based on deducting energy potentials from images in a multi-scale fashion with Renormalization Group Transformations (RGT) in order to make exact segmentations of large images possible. By combining RGT with state of the art segmentation techniques we present a novel approach to multi-scale image segmentation algorithms. Our experiments show that RGT leads to improved segmentations compared to other common multi-scale image segmentation techniques.

### 5.1 Introduction

In the last years image processing applications such as segmentation or restoration have become increasingly important. While segmentation problems used to be solved by approximate techniques such as simulated annealing [48] or iterated conditional modes [91], today they are mostly solved via graphical approaches [46] or inference algorithms, e.g. belief propagation [92]. As the resolution of the images being processed increases [93], the need for fast and efficient multi-scale energy minimization techniques arises. We propose an approach for deducting energy potentials from images in a multi-scale fashion based on Renormalization Group Transformations (RGT) in order to make exact segmentations of large images possible.

**Related Work** There are many multi-scale segmentations methods available and the related work is nicely reviewed in [94]. Note that most multi-scale approaches assume that the energy functions of the image that is being segmented is of the same form at different scales, while in fact one should try to evolve the energy function correctly over different scales. We will try to do that by employing RGT.

RGT is a method from statistical physics and most of the work related to it has been done in statistical or solid state physics. RGT has been used to explain phase transitions in physical systems [95], e.g. magnets or fluids which are alike in having fluctuations in structure over a vast range of sizes, but has also been applied in quantum electrodynamics [96]. Two of the most prominent protagonists of RGT in solid state physics were Leo P. Kadanoff [95] and Kenneth G. Wilson [41], who applied RGT to the Kondo effect [97].

The basis of especially Kadanoff's and Wilson's RGT calculations of phase transitions were 2D Ising models [98], which are a simpler form of a Markov Random field (MRF) [99]. RGT can approximate the analytical solution of a two dimensional (2D) Ising model and describe phase transitions of the material that is modeled. This suggests itself to be applied on images, which can be interpreted as 2D Ising models.

In computer vision there have been few papers that have tried to apply RGT on images. The first to transfer RGT from solid state physics to images was Basilis Gidas [42, 100]. Further work was done later by Petrou et al. [101–103]. When Gidas and Petrou presented the renormalization group approach their calculations were computationally hardly tractable, because of the loss of locality of the model at the coarser stages. By combining RGT with state of the art segmentation techniques we present in the following a computationally efficient and novel approach to multi-scale image segmentation problems.

In section 5.2 we introduce our notation and the energy function that we choose to transform, while in section 5.3 we introduce the reader to the general concept of multi-scale image segmentation. Then we show how to transform the energy function with the help of RGT and give details of RGT for a specific choice of mapping and grid in section 5.4. Finally, section 5.5 covers the experiments and section 5.6 offers a discussion and conclusions and describes future work.

## 5.2 Notation

In this section, we will use a random field model to formulate the image segmentation problem. The random field  $\mathbf{X}$  has a set of  $n$  nodes  $\mathcal{V} = \{1, 2, 3, \dots, n\}$  as well as a neighborhood system  $\mathcal{N}$  that connects pairs of nodes. Each random variable  $x_i \in \mathbf{X}$  is associated with one lattice point  $i \in \mathcal{V}$  and takes one value from the label set  $\mathcal{L} = \{l_1, l_2, \dots, l_k\}$ . All possible label sets,  $\mathbf{x}$ , form the set

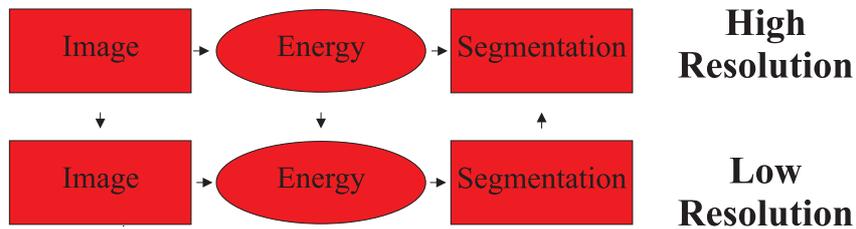


Figure 5.1: A description of our problem: We want to segment e.g. a gigapixel image at high resolution. We could do this following the high resolution path left to right, but often this is not feasible. Therefore in this section we exploit several options to employ low resolution versions of the image or the energy to segment the high resolution image.

of labelings  $\mathbf{L} = \mathcal{L}^n$ .

Transferred to image segmentation,  $x_i$  represents the assigned label of the  $i$ th image pixel, while  $\mathcal{V}$  corresponds to the set of all image pixels and  $\mathcal{N}$  to the set of all edges in a given neighborhood. An optimal image segmentation,  $\mathbf{x}^*$ , based on the data,  $\mathbf{D}$ , is then given by the maximum-a-posteriori (MAP) estimate

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathbf{L}} P(\mathbf{x} | \mathbf{D}) = \arg \min_{\mathbf{x} \in \mathbf{L}} E(\mathbf{x}), \quad (5.1)$$

where the energy function  $E : \mathcal{L}^n \rightarrow \mathbb{R}$  can in principle be defined freely. But for most computer vision problems and in our case, the energy function is of the form of an Ising model [98],

$$E(\mathbf{x}) = \sum_{i \in \mathcal{V}} \phi_i + \sum_{i, j \in \mathcal{N}} \phi_{ij}, \quad (5.2)$$

where the label set consists of only two labels, foreground (fg) and background (bg). In our application the unary potentials  $\phi_i$  are derived from the color of the pixels and appearance models that are constructed from user marked brush strokes as shown in [104, 105], while the pairwise interaction terms  $\phi_{ij}$  are constructed from edge features  $g(i, j)$  based on the color difference of neighboring pixels and defined as  $\phi_{ij} = g(i, j)$  for  $i \neq j$  and 0 otherwise.

Further we define two mappings: A mapping  $\mathcal{M} : \mathcal{V} \rightarrow \mathcal{V}^l$  that maps the pixels of the original image  $I$  to the set of pixels  $\mathcal{V}^l$  in the low resolution version  $I^l$  of the image and a mapping  $\mathcal{K} : \mathbf{L} \rightarrow \mathbf{L}^l$  that connects the space of labelings  $\mathbf{L}$  of  $\mathcal{V}$  to the space of labelings  $\mathbf{L}^l$  of  $\mathcal{V}^l$ .

### 5.3 Multi-scale Image Segmentation

The principle behind multi-scale image segmentation can be seen in figure 5.1. One starts out with a high resolution image  $I$  and wants to end up with a high

resolution segmentation  $\mathbf{x}^*$ . For images of reasonable size one can do that simply via a high resolution energy  $E$  that is minimized according to eq. 5.1, but for large images this gives rise to a large scale optimization problem which is computationally extremely expensive or even unfeasible to solve. Therefore it is better to first solve the problem at low resolution and to obtain a coarse labeling of the high resolution problem. The low resolution problem can be constructed in a multitude of ways. Two of the most common are:

**Low Resolution Energy Function from a Low Resolution Image** Traditional multi-scale methods as [106, 107] define the low resolution energy function  $E^l$  by using a low resolution version  $I^l$  of the image  $I$ . The coarse energy function  $E^l$  is hereby simply of the same form as (5.2), but based on  $\mathcal{V}^l$  instead of  $\mathcal{V}$ .

**Low Resolution Energy Function from the Original Energy Function** Another approach is to define the coarse grid energy  $E^l$  directly from the potentials of the original energy function  $E$ . This is done by summing over the unary variables that comprise the coarse node  $\mathcal{V}^l$  and likewise over the pairwise potentials defined between the nodes  $\mathcal{V}$  that  $\mathcal{V}^l$  contains.

The coarse energy function can then be minimized to extract a partial labeling of the original image  $I$ , which can in turn be refined to a solution for the whole image  $I$  by another optimization on a small subset of pixels only. The partial labeling based on the coarse energy  $E^l$  can be computed via different approaches, either via a Boundary Band Approach [106] or via a min-marginal approach [108]. We will in the following use a hybrid approach made up of a combination of the two that was also introduced and detailed in [108].

Finally, the coarse labeling is projected back to the high resolution problem and a solution for the original image obtained.

## 5.4 Using RGT for Multi-scale Image Segmentation

In general, RGT deals with energy functions of physical systems. It can be used as described in [41] and [42] to evolve energy functions between different representational scales, e.g. a coarse and a fine scale of a physical system. In the following we introduce a new way of defining the low resolution energy function  $E^l$  on the basis of the original energy function.<sup>5</sup>

### 5.4.1 Energy Function Model

The energy function  $E$  of a two spin state Ising Model is in general given as in (5.2). For simplicity we base the following derivation on a simple spin model

---

<sup>5</sup>The approach presented is due to [109]. Other approaches are explained in [110], e.g. the method of neglect and the method of potential moving [111].

( $\phi_i = hs_i$  and  $\phi_{ij} = ks_i s_j$ ) leading to

$$E = h \sum_i s_i + k \sum_{\langle ij \rangle} s_i s_j, \quad (5.3)$$

where  $s_i$  is the spin state (+1 or -1) of the spin located at the site  $i$  and  $h$  and  $k$  are the unary and pairwise interaction factors. Furthermore we define the block spin  $S_i$  to be made up of several spins  $s_i, s_j$ , etc. Then the mapping  $\mathcal{M} : s \mapsto S$  maps the spins  $s$  at high resolution to the block spins  $S$  at low resolution, where the mapping  $\mathcal{K} : \mathbf{L} \rightarrow \mathbf{L}^l$  connects the two labeling spaces,  $\mathbf{L}$  and  $\mathbf{L}^l$ , at high and low resolution.

Statistically, the probability of the labelings  $\mathbf{L}^l$  of the coarse energy function  $E^l$  can be described by the sum of the probabilities of the original labelings  $\mathbf{L}$ ,

$$e^{-E^l} = \sum_{a \in \mathbf{L}} e^{-E(a)}, \quad (5.4)$$

where  $a$  is a labeling  $\in \mathbf{L}$ .

We can derive the functional form of  $E^l$  via linear operator perturbation theory [112], which divides our coarse energy in a linear term,  $E_0$ , and a higher order term,  $V$ ,

$$E^l = E_0 + V. \quad (5.5)$$

The linear term of the coarse energy function should look as similar as possible to the interaction term of our original energy function and describe the interaction between spins inside one block,  $S_I$ , so we set it to

$$E_0 = k \sum_I \sum_{i,j \in I} s_i s_j. \quad (5.6)$$

The higher order term,  $V$ , is then outlining the interaction between spins in different blocks,

$$V = k \sum_{I \neq J} \sum_{i \in I, j \in J} s_i s_j. \quad (5.7)$$

Now we can change equation 5.4 to

$$e^{-E^l} = \sum_{a \in \mathbf{L}} e^{-E_0(a)} \cdot e^{-V(a)} \quad (5.8)$$

$$= \langle e^{-V} \rangle_0 Z_0(k)^m, \quad (5.9)$$

where

$$\langle e^{-V} \rangle_0 = \frac{\sum_{a \in \mathbf{L}} e^{-E_0(a)} \cdot e^{-V(a)}}{\sum_{a \in \mathbf{L}} e^{-E_0(a)}} \quad (5.10)$$

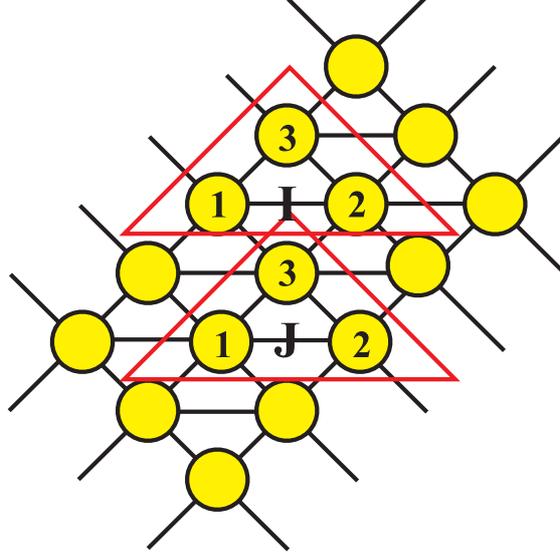


Figure 5.2: The graph shows the triangulation used by Kadanoff.

is the expectation value of  $e^{-V}$ . In the expression above

$$Z_0(k)^m = \sum_{a \in \mathbf{L}} e^{-E_0(a)} \quad (5.11)$$

describes the partition function for  $m$  blocks of spins  $s_i$  that are mapped to  $S_I$ .  $Z_0$  is a normalizing constant.

Now all we need is an expression for  $\langle e^{-V} \rangle_0$ . Via an expansion in orders of  $V$  and with the formula  $\log(x) = (x - 1) - \frac{1}{2}(x - 1)^2 + O(x^3)$  [113] we get

$$\langle e^{-V} \rangle_0 = e^{-\langle V \rangle_0 + \frac{1}{2}(\langle V^2 \rangle_0 - \langle V \rangle_0^2) + O(V^3)}. \quad (5.12)$$

Our final expression for the coarse energy function is then

$$E^l = -m \log(Z_0(k)) + \langle V \rangle_0 - \frac{1}{2} (\langle V^2 \rangle_0 - \langle V \rangle_0^2) + O(V^3). \quad (5.13)$$

#### 5.4.2 Kadanoff's Block Spin Method

The mapping  $\mathcal{M}$  used to downsample the image can be chosen in a multitude of ways. It can be an interpolation based down sampling algorithm like [114], a majority vote mapping or even a mapping where one specific spin in each block determines the value of the larger block  $S_I$ . An optimal representation is one of the key points and choices in the application of RGT.

Furthermore, one can choose many different type of grids, that then determine how the sums over the spins  $i$  and  $j$  in the energy function are calculated. Obvious choices here are a standard pixel grid or a triangular lattice. In the

following let us present a more detailed calculation based on a majority vote mapping and a triangular lattice.

We assume a block spin configuration on a triangular lattice by Kadanoff [115] as shown in figure 5.2, where block spins are defined by a mapping  $\mathcal{M}$  that groups three spins together and computes the block spin  $S_I$  via a majority rule:

$$S_I = \text{sign}(S_I^1 + S_I^2 + S_I^3), \quad (5.14)$$

where  $S_I^j$  is the  $j$ th spin in the  $I$ th block. Then we define a set of spins that make up  $S_I$

$$\sigma_I \equiv \{S_I^1, S_I^2, S_I^3\}. \quad (5.15)$$

For a set of 3 spins with 2 states each, the labeling space  $\mathbf{L}$  is made up of  $2^3 = 8$  possible configurations. The block spin  $S_I$  has only two possible labelings, but 4 different labelings of the original spins are mapped to each coarse labeling  $\mathbf{L}^l$

$$\mathcal{L}_1^l = +1 \quad \mathcal{L} = \{\{+1, +1, +1\}, \{-1, +1, +1\}, \{+1, -1, +1\}, \{+1, +1, -1\}\} \quad (5.16)$$

$$\mathcal{L}_2^l = -1 \quad \mathcal{L} = \{\{-1, -1, -1\}, \{+1, -1, -1\}, \{-1, +1, -1\}, \{-1, -1, +1\}\}$$

reflecting the original 8 configurations.

We can apply this choice of mapping and grid to our coarse energy function given in (5.13). To do this we need to find the expression for  $\langle V \rangle_0$ . Since according to (5.7)  $V$  couples nearest neighbor blocks, we can determine it from figure 5.2 to be

$$V_{IJ} = k(S_J^3)(S_I^1 + S_I^2) \quad (5.17)$$

and therefore

$$\begin{aligned} \langle V_{IJ} \rangle_0 &= 2k \langle S_I^1 S_J^3 \rangle_0 \\ &= 2k \langle S_I^1 \rangle_0 \langle S_J^3 \rangle_0, \end{aligned} \quad (5.18)$$

since the blocks comprising  $S_I$  and  $S_J$  are independent of each other. One can evaluate the average for all configurations given in (5.16) and one gets

$$\langle V_{IJ} \rangle_0 = 2k\phi(k)^2 \sum_{I \neq J} S_I S_J, \quad (5.19)$$

where  $\phi(k) \equiv \frac{e^{-3k} + e^k}{e^{-3k} + 3e^k}$  and the coarse energy pairwise coupling  $k' = 2k\phi(k)^2$ . So the coarse energy function is to first order given by

$$E^l(S_I) = -m \log(Z_0) + k' \sum_{I \neq J} S_I S_J + O(V^2) \quad (5.20)$$

The same expansion can be conducted for one order higher, giving

$$E^l(S_I) = -m \log(Z_0) + k' \sum_{I \neq J} S_I S_J - \frac{1}{2} k'' \sum_{I \neq J} S_I^2 S_J^2, \quad (5.21)$$

where

$$k'' = k^2 \left( 2\phi(k)^2 + 2\phi(k) \frac{e^{-3k} - e^k}{e^{-3k} + 3e^k} - 4\phi(k)^4 \right) \quad (5.22)$$

is the modified coupling parameter for the higher order interactions at the coarser scale.

The unary term in  $E$  can be derived in the same way by expansion and is at coarse resolution changed to

$$E^l = h' \sum_I s_I, \quad (5.23)$$

where  $h' = 3\phi(k)$ .

## 5.5 Experiments and Results

### 5.5.1 Segmentation Quality

In section 5.4 we have derived the form of an RGT low resolution energy. In this section we examine the achieved segmentation quality of the different methods for constructing the smaller energy minimization problem. Therefore, we compare segmentation results of the low resolution energy function derived via RGT ( $S_{RGT}$ ) with a) a low resolution energy function derived from a low resolution image ( $S_{LRI}$ ) and b) a low resolution energy function calculated from the high resolution energy function ( $S_{LRE}$ ). The downsampling function used in the three methods is thereby the same, the difference is only that we add appropriate couplings in front of the downsampled energy ( $S_{RGT}$ ) or that the image ( $S_{LRI}$ ) or the energy ( $S_{LRE}$ ) is being downsampled. We also present the segmentation result of the original high resolution image ( $S_{HRI}$ ). The segmentations were made via the graphcut algorithm [116], which can optimize a submodular energy functions as in our case ( $k = -1$ ) exactly in polynomial time.

First, we present the four segmentations for three different images, in figures 5.3, 5.4 and 5.5. One can see slight differences between the three low resolution segmentations, especially in the fine structures as for example around the bag that the man holds in figure 5.3. If one measures the area overlap between the manual annotation and the three segmentations via the Jaccard index  $J$  [68], where the area overlap between two binary segmentations  $S_1$  and  $S_2$  is given

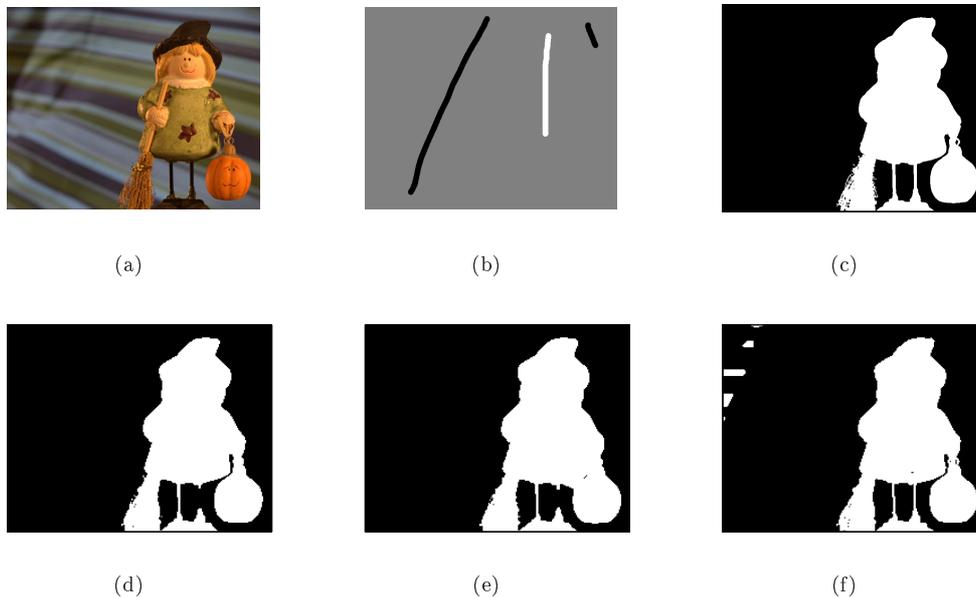


Figure 5.3: We display (a) the original image, (b) the user marked brush strokes used to initialize the segmentation and the four segmentations, (c)  $S_{HRI}$ , (d)  $S_{LRI}$ , (e)  $S_{LRE}$  and (f)  $S_{RGT}$ . One can see slight differences between the three low resolution segmentations, especially in the fine structures as for example around the bag that the man holds.

by

$$J(S_1, S_2) = \frac{|S_1 \cap S_2|}{|S_1 \cup S_2|}, \quad (5.24)$$

the segmentations have basically the same area overlap as shown in table 5.1. But if we compare the quality of partial labelings generated from different coarse energy functions, we can see differences between the three low resolution segmentation approaches as can be observed in figure 5.6. The traditional approach of constructing the energy from the low resolution outperforms the approach of constructing the energy from the original energy function. Further, using RGT to construct the coarse energy outperforms both of the other methods.

### 5.5.2 Segmentation Time

Besides the segmentation quality we also analyze the computational time needed for the different segmentations. The results can be seen figure 5.7 and they show that the smaller energy minimization problem constructed by downsampling a high resolution energy is the slowest approach, while our proposed RGT based method to minimize the energy minimization problem is the

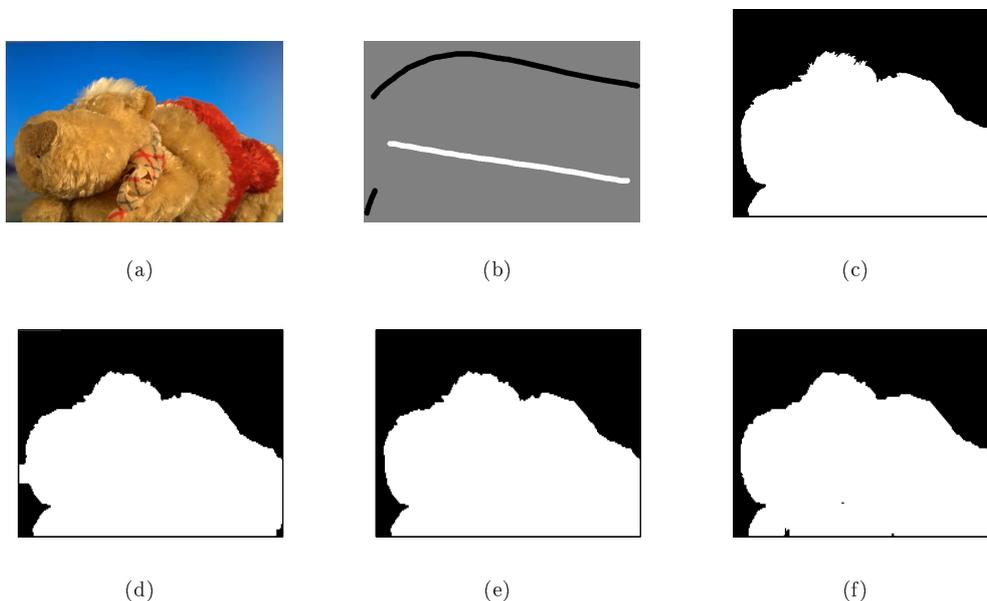


Figure 5.4: Again we display (a) the original image, (b) the user marked brush strokes used to initialize the segmentation and the four segmentations, (c)  $S_{HRI}$ , (d)  $S_{LRI}$ , (e)  $S_{LRE}$  and (f)  $S_{RGT}$ . In this picture one can not see large differences between the different segmentation algorithms, but it is still interesting to look at the bottom boundary of the bear, where the algorithms differ.

Table 5.1: Evaluation of the segmentation results using the Jaccard index against the manual annotation. Notice that the improvements can not be captured by the Jaccard index.

(a) To figure 5.3			(b) To figure 5.4			(c) To figure 5.5		
	Mean	Median		Mean	Median		Mean	Median
$S_{LRI}$	0.97	0.97	$S_{LRI}$	0.97	0.96	$S_{LRI}$	0.92	0.92
$S_{LRE}$	0.97	0.97	$S_{LRE}$	0.96	0.96	$S_{LRE}$	0.92	0.92
$S_{RGT}$	0.97	0.97	$S_{RGT}$	0.96	0.96	$S_{RGT}$	0.92	0.92

fastest.

## 5.6 Discussion and Conclusion

Our experiments show that RGT can be applied to multi-scale image segmentation problems. Using multi-scale image segmentation techniques we can reduce the size of an image segmentation problem and make it tractable and solvable. RGT offers to conduct the multi-scale approach by correctly evolving energy functions over different scales instead of assuming that the energy

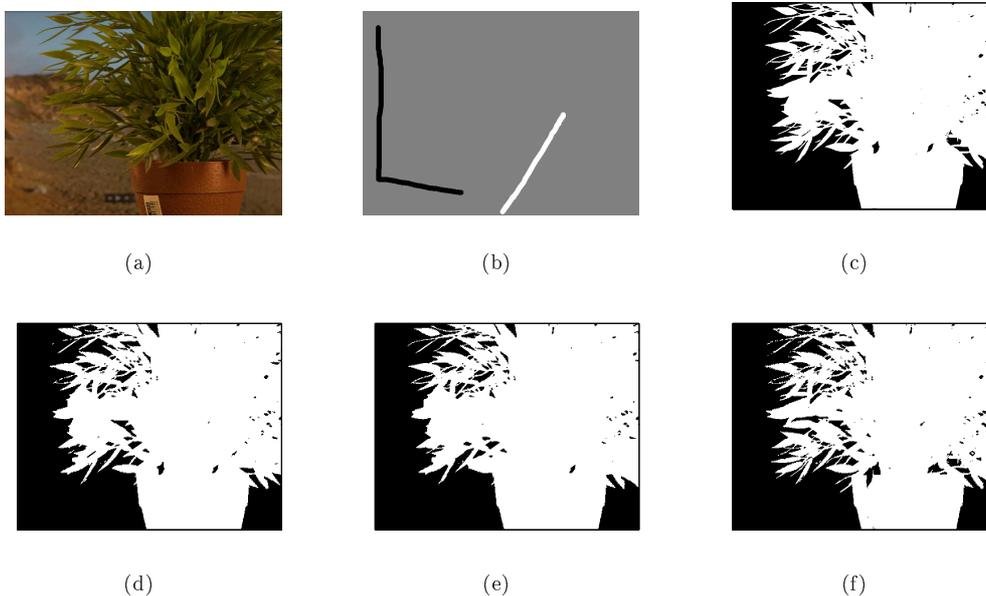


Figure 5.5: Yet again we display (a) the original image, (b) the user marked brush strokes used to initialize the segmentation and the four segmentations, (c)  $S_{HRI}$ , (d)  $S_{LRI}$ , (e)  $S_{LRE}$  and (f)  $S_{RGT}$ . In the plant image it is especially interesting to look at the left side of the plant. RGT seems to be able to separate the leaves on the lower left of the plant much better than any other method.

functions are of the same form at every scale, an assumption that underlies the other two methods. The mappings  $\mathcal{M}$  and  $\mathcal{K}$  can be chosen dependent on the problem and the need for exact or approximate inference.

The goal for the further development of the RGT transformation on images is to proceed in a more structured manner. Instead of choosing an arbitrary mapping and to see what effect this has, one could attempt to choose the optimal mapping and method of coarsening. A principle component type of analysis of different cluster variables could help to choose the optimal mapping and method of coarsening.

In conclusion, we need to experiment with further applications of RGT to be able to demonstrate clearly that it can be the solution to a wide range of problems, but our preliminary results support the notion that RGT provides an interesting solution to multi-scale energy minimization problems.

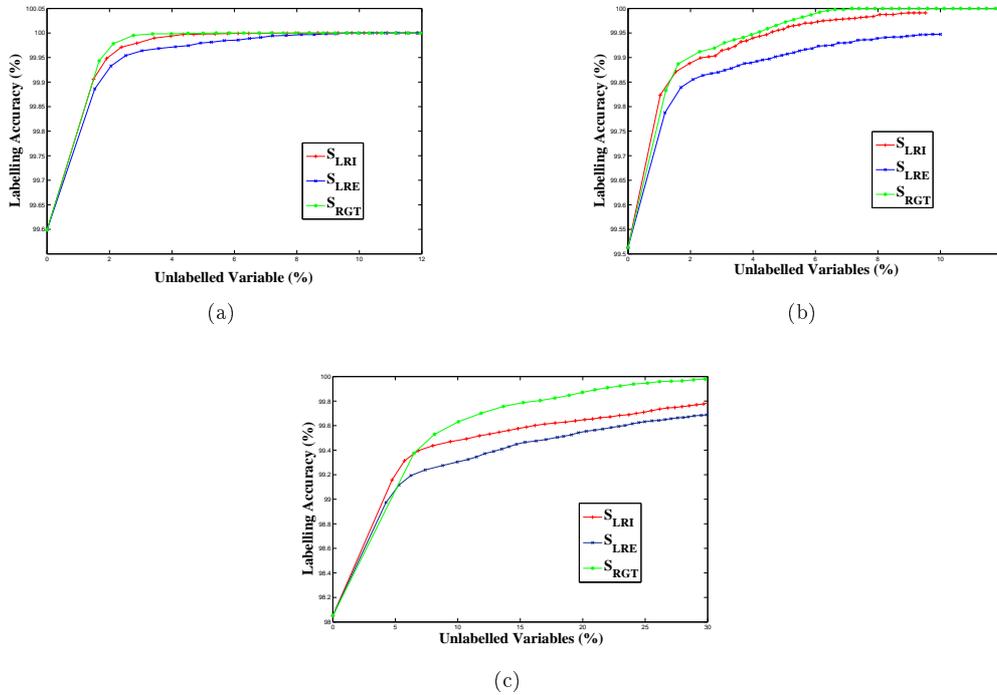


Figure 5.6: We display for the images shown in figures 5.3, 5.4 and 5.5 how the accuracy of the partial solutions derived at low resolution changes as we increase the number of unlabeled variables. The RGT approach achieves better labeling accuracy, especially if the number of unlabeled variables is increased.

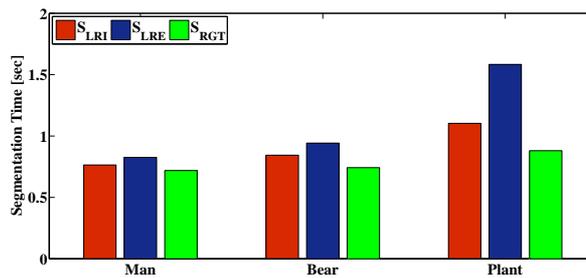


Figure 5.7: We display the computational time for the different methods of constructing the smaller energy minimization problem for the images of the man, the bear and the plant shown in figures 5.3, 5.4 and 5.5. RGT has the lowest segmentation time of the three segmentation methods.

## II

# Clinical Application of the Imaging of Lumbar Aortic Calcifications



**PART II** of the dissertation consists of research papers that focus on the clinical relevance of lumbar aortic calcifications.

In chapter 6 several biomarkers are developed and their clinical relevance tested. Then the biomarkers are combined to form the morphological atherosclerotic calcification distribution (MACD) index and the MACD index and its relation to mortality in post menopausal women is examined in chapter 7.



# Chapter 6

## Abdominal Aortic Calcified Deposits and their Relationship to Mortality in Post Menopausal Women

---

This chapter is based on the manuscript "Abdominal Aortic Calcified Deposits and their Relationship to Mortality in Post Menopausal Women" by Melanie Ganz, Marleen de Bruijne, Erik B. Dam, Paola Pettersen, Morten A. Karsdal, Claus Christiansen and Mads Nielsen currently in preparation and on the manuscript "MACD - an Imaging Marker for Cardiovascular Disease" by Melanie Ganz, Marleen de Bruijne and Mads Nielsen published in the proceedings of the SPIE Medical Imaging conference, San Diego, CA, USA, 2010.

---

**Abstract** Abdominal aortic calcifications (AACs) correlate strongly with coronary artery calcifications and can be predictors of cardiovascular mortality. We investigated whether size, shape, and distribution of AACs relate to mortality and how such prognostic markers perform compared to the state-of-the-art AC24 marker introduced by Kauppila et al. For 308 post-menopausal women we quantified the number of AACs and the percentage of the abdominal aorta that the lesions occupied in terms of their area, simulated plaque area, thickness, wall coverage and length. We analyzed inter-/intra-observer reproducibility and predictive ability of mortality after 8-9 years. This was done via Cox regression with and without adjustment for biological risk factors leading to hazard ratios (HR). The coefficient of variation was below 25% for all markers. The strongest individual predictors were the number of calcifications (HR=2.4,  $p < 0.001$ ) and the simulated area percentage (HR=2.96,  $p < 0.001$ ) of a calcified plaque and, unlike AC24 (HR=1.66,  $p < 0.001$ ), they were predictive of mortality after adjusting for traditional risk factors. In a combined Cox regression model the strongest complementary predictors were the number of calcifications and the percentage of the aorta area they occupied. Hence, morphometric markers of AAC quantified from radiographs may be a useful tool for screening and monitoring risk of CVD mortality.

## 6.1 Introduction

Cardiovascular diseases (CVDs) are the most prevalent cause of death in Europe [117] and the United States [118]. This is the case despite general acceptance that a healthy lifestyle and risk factor management can prevent the development of CVDs [6]. Furthermore, two-thirds of women who die suddenly from CVDs have no previously recognized symptoms. Thus, effective and broadly applicable indicators of cardiovascular risk may prompt timely intervention.

Current non-invasive modalities for imaging atherosclerosis are radiographs, ultrasound, computed tomography (CT) and magnetic resonance imaging (MRI) [119]. Ultrasound is used to visualize the carotid intima-media thickness (IMT), because carotid IMT has been shown to be associated with atherosclerosis [19], and is thus a marker for CVD. Multi-slice CT is able to quantify the degree of coronary artery calcification (CAC) with good reproducibility [22,23], which provides a strong measure of cardiovascular risk [24] independently from, and potentially more powerful than, traditional risk factors such as smoking [25]. MRI is a non-invasive modality to assess atherosclerosis in different vascular beds. However, MRI measurements are challenged by the size of the smaller arteries and assessment of the coronary arteries is especially difficult due to cardiac and respiratory motion artifacts [120].

An alternative to examining coronary arteries for calcification is to assess the abdominal aorta. Although calcifications of the coronary arteries can also be assessed by radiographs, we focus on radiographic examination of the abdominal aorta, because abdominal aortic calcifications (AACs) are strong predictors of cardiovascular morbidity and mortality [10], correlate strongly with coronary artery calcifications, and may hence predict the risk of coronary artery diseases [30,80]. The state of the art methodology to estimate CVD risk from lumbar aortic radiographs is the abdominal aortic calcification score (AC24) proposed by the Framingham study group [13]. A big advantage is that such AAC scoring can, for example in the case of post-menopausal women, be performed without additional ionizing radiation exposure or cost as these images are already performed in osteoporosis trials [33,34].

We investigated if the morphometric aspects of calcifications quantified from plain radiographs could become useful novel markers of AACs. Due to the semi-quantitative grading of the AC24 score, such markers from abdominal aortic examinations could potentially be more sensitive, in particular to assess the potential significance of smaller calcifications. We outlined the boundaries of the calcified deposits in the lumbar aortic region on radiographs of post menopausal women and quantified the number of calcified deposits as well as the percentage of the abdominal aorta covered by calcifications in terms of area, simulated-plaque area, thickness, wall coverage, and length. These po-

tential AAC markers were evaluated for precision and their ability to predict CVD-related mortality.

## **6.2 Materials and Methods**

### **6.2.1 Study Population**

In 1992-93, 686 post menopausal women living in the Copenhagen area in Denmark were recruited via a household postal survey to participate in the EPI study [67] addressing the role of a number of metabolic risk factors in the pathogenesis of CVD and osteoporosis and were examined radiologically. The follow-up of the EPI study was the Prospective Epidemiological Risk Factors (PERF) study [66], which was performed after 8.5 years. PERF was initiated to obtain further insight into the epidemiology and pathogenesis of menopause-related diseases, particularly osteoporosis.

Of the original 686 participants, we chose those whose interval between their first and second clinic visit was 8-9 years, with known alive/mortality status, who were post menopausal and whose lumbar aorta was visible on a single radiograph at baseline and at follow-up. This left us with 308 subjects. Information about causes of death was obtained from the Central Registry of the Danish Ministry of Health and the death causes were grouped into three groups: CVD, cancer and other death causes.

### **6.2.2 Metabolic and Physical Measurements**

At baseline, demographic information and CVD risk parameters such as age, weight, height, body mass index (BMI), waist and hip circumferences, systolic and diastolic blood pressure (BP), treated hypertension, treated diabetes, smoking, regular alcohol and daily coffee consumption, and weekly fitness activity were collected. Using a blood analyzer (Cobas Mira Plus, Roche Diagnostics Systems, Hoffman-La Roche, Basel, Switzerland), measurements of fasting glucose and a lipid profile (total cholesterol, triglycerides, low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), apolipoprotein A and B (ApoA and ApoB respectively) and lipoprotein(a) (Lp(a))) were obtained. On the basis of these measurements, the composite risk markers, systemic coronary risk evaluation (SCORE) [49] and Framingham score [50], were calculated. The SCORE is a combination of the age, smoking status, levels of total cholesterol and systolic blood pressure, while the Framingham score is comprised of the same variables plus the HDL-C and the hypertension treatment status.

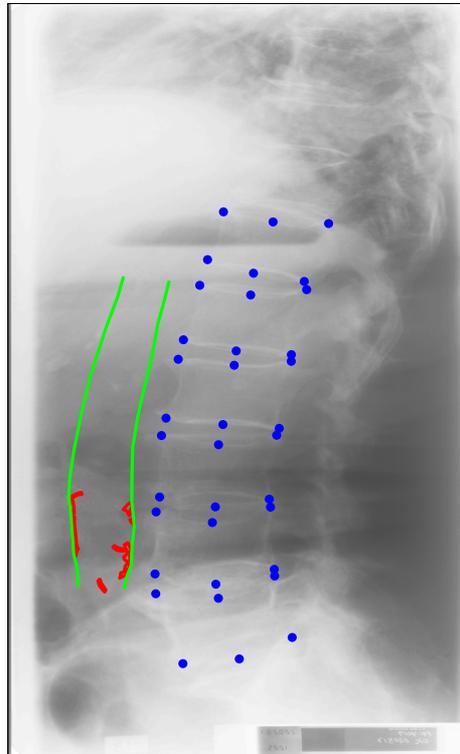


Figure 6.1: A manual annotation of an X-ray: In blue we see distinct vertebra points, in green the aorta wall and in red the calcifications.

### 6.2.3 Radiographic Analysis

The lateral X-ray images of the lumbar aorta (L1-L4) were acquired on film in 1992 and again in 2001, and digitized in 2007/2008 using a DosimetryProAdvantage scanner (Vidar, Herndon, USA), providing an image resolution of 570 dpi on a 12-bit gray scale. Three trained radiologists without prior knowledge of the patients' conditions manually annotated the corners and mid points of the vertebrae (L1-L4), the corresponding abdominal aorta walls, and calcifications in the digitized images. The three radiologists had ten, eight and five years of experience. They used radiological reading units (Sectra, Linköping, Sweden) and annotation software specifically implemented for that task in Matlab (The MathWorks, Natick, USA), which allowed them to change brightness and contrast, zoom in and out, and to edit outlines, as seen in figure 6.1.

The AC24 [15] was constructed by projecting the AACs to the corresponding aorta wall. Then the aortic sections adjacent to each vertebra L1-L4 were graded by the degree of lesion occupation: 0 for no AACs, 1 for AACs occupying less than 1/3 of the wall they were projected onto, 2 for AACs occupying

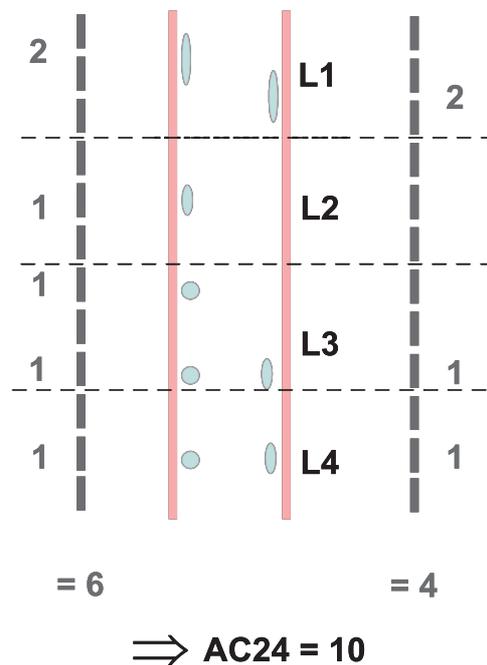


Figure 6.2: A schematic overview of the AC24 scoring adopted from [13].

more than  $1/3$ , but less than  $2/3$  in the projection, and 3 for a  $2/3$  or more occupation of the wall. The degree of lesion occupation of the projections to the anterior and posterior aorta wall for all four lumbar vertebrae was summed, leading to an AC24 score ranging between 0 and 24. In addition to the AC24 scores provided by the radiologists, the outlines of the calcifications were used in an alternative computer-based computation of the AC24. The AC24 scoring is illustrated in figure 6.2.

For all images with calcifications, annotations were performed by one of the three radiologists. For a subset of 8 images, annotations by two radiologists were made twice in order to evaluate inter- and intra-observer precision. Re-outlining was performed blinded to earlier outlines and separated by approximately six to eight weeks.

#### 6.2.4 AAC Markers

Besides measuring the AC24 there is more information to gain from the AACs. Therefore several potential severity scores relating to the geometrical outline of the calcified deposits in the lumbar aortic region were examined. The proposed AAC markers were automatically computed from the radiologist's computer-assisted outlines of calcified deposits in the radiographs:

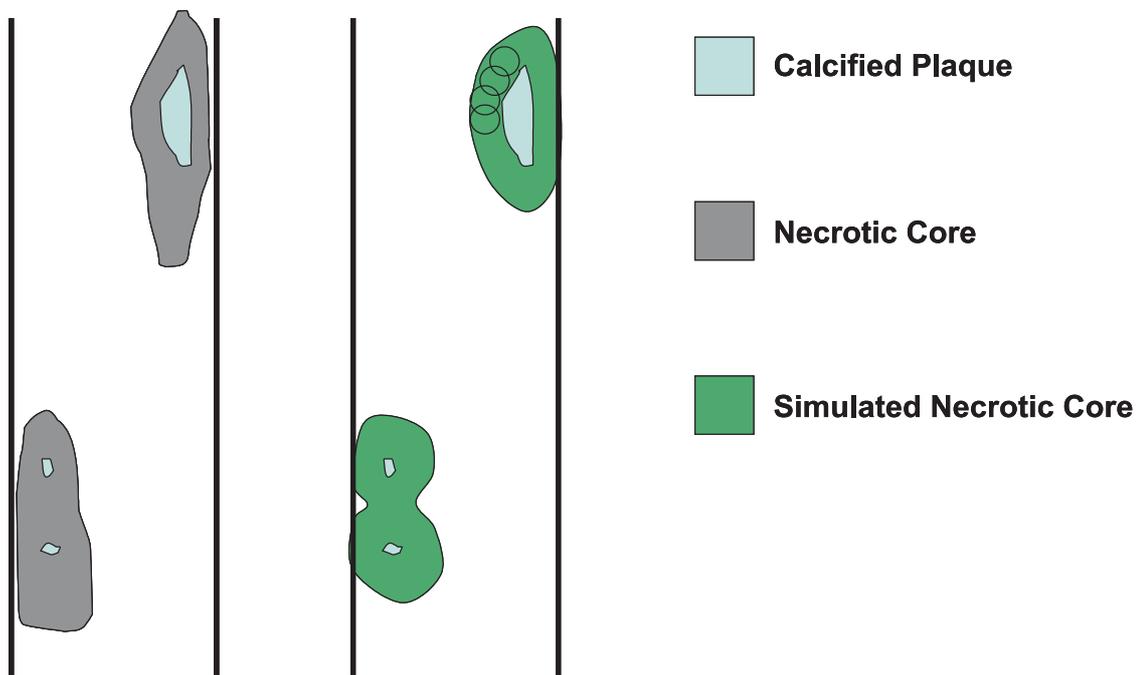


Figure 6.3: Left: A schematic visualization of a plaque similar to what can be seen in histology. The calcified plaque is surrounded by an area of necrotic tissue. Right: The simulated area tries to imitate the area of necrotic tissue as seen in histology by a morphological dilation of the calcified plaque.

- Area percentage (Area %): The percentage of the area of the lumbar aorta adjacent to L1-L4 occupied by AACs.
- Simulated area percentage (Sim. area %): We estimated the size of the underlying atherosclerotic inflammation from the area and shape of the observed AACs since X-ray analysis can only visualize the calcified core of the AACs. The extent of the atherosclerotic inflammation was simulated by a morphological dilation [121] with a circular structuring element of radius 200 pixels (approx. 8.9 mm). The size of the structuring element was derived by a parameter study on a subset of the data and it was confirmed to be biologically sensible by comparing with histology and image analysis observations which estimated the size of the atherosclerotic inflammation surrounding the calcified plaque to be between 3 mm [122] and 5-10 mm [123]. An illustration of this computer-based simulation of the full plaque area is given in figure 6.3. The simulated area percentage is the percentage of the lumbar aorta covered by the simulated plaques, including both calcified core and simulated inflamed area.
- Thickness percentage (Thickness %): The average thickness of the AACs

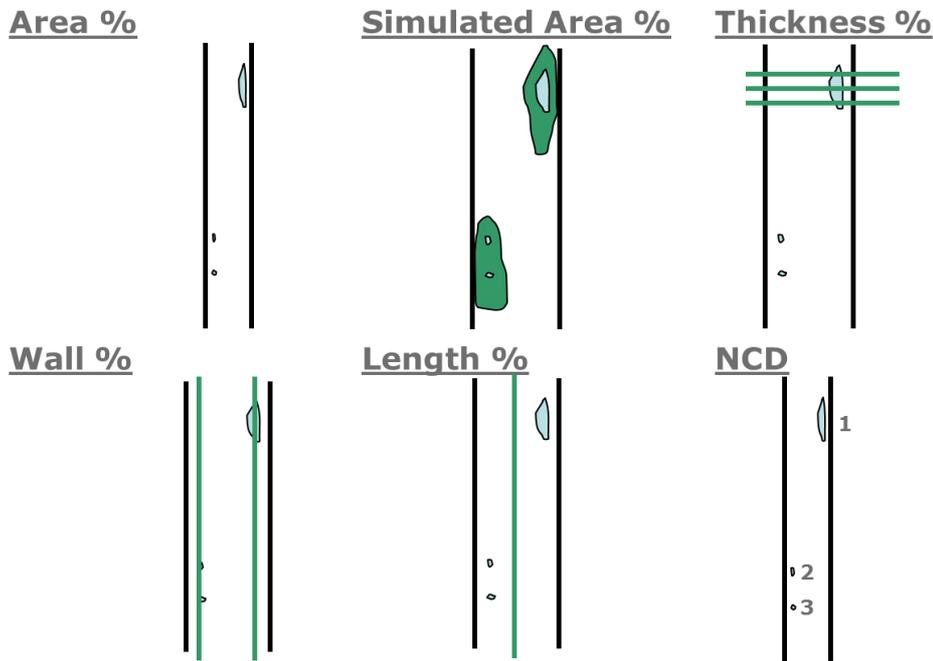


Figure 6.4: A schematic view of the different AAC markers

along the aorta wall relative to the aorta width.

- Wall percentage (Wall %): The percentage of the anterior and posterior lumbar aorta wall covered by projections of AACs.
- Length percentage (Length %): The fraction of the length of the aorta where AACs were present at any position (anterior, posterior or internal).
- Number of calcified deposits (NCD): The number of distinct AACs visible in each radiograph.

A schematic view of the six proposed markers can be seen in figure 6.4. We examined the degree to which these markers could be reliably established on the basis of manual annotations of X-ray images, and evaluated their association to mortality, also when adjusted for metabolic or physical markers.

### 6.2.5 Statistical Analysis

Kendall's coefficient of concordance  $\tau$  [124] was used to assess the level of agreement between AC24 scorings of calcified images made by radiologists directly on the original X-rays and AC24 scorings by the computer, based on the radiologist's annotated outlines. To measure the inter- and intra-observer

variability of the manual annotations of radiologists on the eight images allocated specifically for this purpose, we used the Jaccard Index A [68]. We computed the ratio of the area identified as calcified in two outlines, divided by the area identified as calcified in at least one outline:

$$A = \frac{|A_1 \cap A_2|}{|A_1 \cup A_2|}$$

where  $A_1$  and  $A_2$  are binary annotations. The Jaccard Index varies from 0 for no agreement to 1 for complete agreement. Typically, Cohen's  $\kappa$  [88] would be used to measure the inter-rater agreement for categorical items such as pixels. However, the statistics will be dominated by the very large class of non-calcified pixels, and individual pixel scorings cannot be considered statistically independent.

The inter- and intra-observer variability of the AAC markers computed from the radiologists' outlines was analyzed on the eight images by the mean coefficients of variation (CV).

The predictive power of mortality in terms of hazard ratio (HR) per standard deviation change of the individual AAC scorings was calculated with the help of a Cox regression analysis, also known as survival analysis. A basic overview over survival analysis is given in [125], while a thorough coverage of the subject can be found in [126].

The basis of the Cox regression model is the examination of the behavior of the hazard function with respect to  $q$  different environmental parameters. The hazard function is given by

$$h(t; z_i) = h_0(t) \exp(z_i^t \beta), \quad (6.1)$$

where  $t$  is the time and  $z_i$  with  $i = 1, \dots, n$  are the  $q$ -dimensional environmental parameter vectors for each individual patient  $i$  in the study,  $h_0$  is an unknown baseline hazard function used to model the hazard without environmental influences and  $\beta$  is a  $q$ -dimensional vector giving the coefficient estimates of a Cox regression of the result status (e.g. dead or alive) to the predictors in  $z_i$ . The Cox regression tries to estimate the regression parameters  $\beta$  and measure their significance.

In all Cox regression analyses, we use the marker values for the complete population and vary the binary outcome variable (e.g. CVD dead = 1, alive or other dead = 0) according to the group of interest we focus on.

First we used Cox regression analysis on the image markers to test their individual prognostic power. In the Cox regression the outcome variable was the time of death and survivors were right censored. This analysis was performed on unadjusted markers as well as markers adjusted with three different sets of biological variables: a) a model consisting of age, smoking status and triglyceride levels, b) the SCORE [49] and c) Framingham score [50] (Model

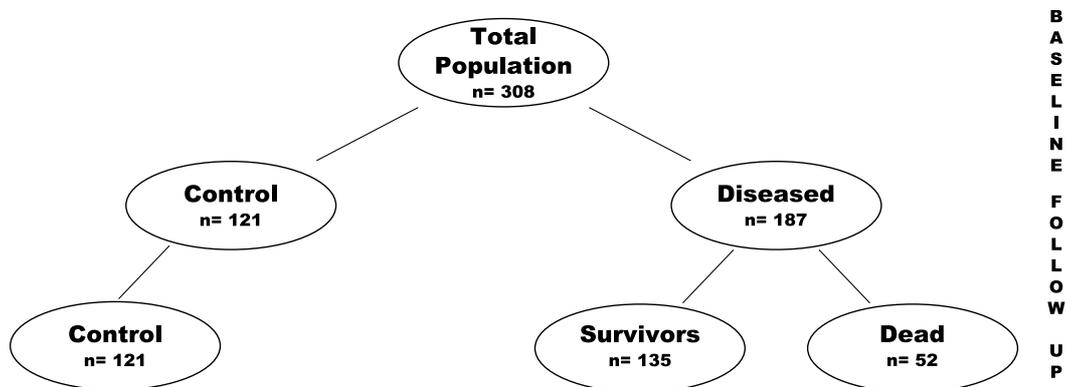


Figure 6.5: A schematic overview of the study population.

a) consists of age, smoking status and triglyceride levels, because in a combined model including all physical/metabolic parameters only age, smoking, and triglyceride level persisted after elimination of insignificant contributions.). We adjusted by combining the biological variables of each set into one new variable by a linear weighting with their  $\beta$ -weights derived by a Cox regression. This new variable was then included in another Cox regression model for the imaging marker we adjusted. The resulting  $\beta$ -weight for the imaging marker determined the biologically adjusted prognostic power.

Furthermore, to analyze the complementarity of the AAC markers we build a Cox regression model with all variables and then successively deleted the least significant marker until only significant markers were left. Hereby significance of the marker  $q$  was given as the model weight  $\beta_q$  being significantly different from zero ( $p < 0.05$ ). This way, single markers that complemented each other and gave supplementary information were identified.

### 6.3 Results

The data consisted of baseline images taken in 1992 of 308 subjects. Of these, 121 subjects had no calcifications at baseline or follow-up. Of the remaining 187 subjects, 52 had died before follow-up due to cancer ( $n=27$ ), CVD ( $n=20$ ) or other causes ( $n=5$ ), and 135 surviving subjects had varying degrees of abdominal aortic calcification at baseline or follow-up. A schematic overview of the study population is given in figure 6.5, while an overview of the physical and metabolic measurements is given in table 6.1.

The radiologist and computer-based AC24 scores for the images of calcification in the 135 subjects calcified images were in excellent agreement (Kendall's  $\tau = 0.97$ ,  $p < 0.0001$ ).

Table 6.1: The mean and standard deviation of the measured metabolic and physical markers.

Physical/Metabolic markers	Population (n=308)	Survivors (n=256)	Deceased (n=52)
Age (years)	60.3 ± 7.5	59.3 ± 7.1	65.6 ± 7.0
Waist (cm)	80.7 ± 10.9	80.2 ± 9.9	83.1 ± 12.4
Waist-to-hip ratio	0.80 ± 0.08	0.80 ± 0.08	0.83 ± 0.10
Body Mass Index ( <i>frackgm</i> <sup>2</sup> )	24.7 ± 3.9	24.7 ± 3.8	25.1 ± 4.6
Smoking (%)	37	33	58
Systolic BP (mm Hg)	127 ± 21	125 ± 20	136 ± 26
Diastolic BP (mm Hg)	77 ± 10	76 ± 10	77 ± 11
Hypertension (%)	16	15	17
Glucose (mmol/L)	5.44 ± 1.27	5.37 ± 0.99	5.79 ± 2.17
Total cholesterol (mmol/L)	6.44 ± 1.19	6.36 ± 1.14	6.85 ± 1.33
Triglycerides (mmol/L)	1.24 ± 0.75	1.15 ± 0.56	1.69 ± 1.25
LDL-C (mmol/L)	2.89 ± 0.82	2.85 ± 0.80	3.07 ± 0.93
HDL-C (mmol/L)	1.77 ± 0.48	1.77 ± 0.44	1.74 ± 0.62
ApoB/ApoA	0.57 ± 0.18	0.56 ± 0.17	0.64 ± 0.23
Lp(a) (mg/dL)	21.4 ± 21.7	21.9 ± 22.0	18.4 ± 19.8
EU SCORE	2.60 ± 2.58	2.16 ± 2.12	4.73 ± 3.45
Framingham	14.75 ± 3.54	14.21 ± 3.46	17.31 ± 2.74

From the eight images with four annotations each, the mean Jaccard Index between the two radiologists' AAC outlines was 51% area overlap for the inter-observer variation and 56% area overlap for the intra-observer variation (see figure 6.6 for an example). Separately, the two radiologists had an intra-observer variability of 53% and 59% area overlap, respectively. The CV values for the AAC marker precision on the same set of eight images were between 12.5% and 24.9% (table 6.2).

The mean values and respective standard deviations of each of the AAC markers is reported in table 6.3. One can easily see that they all have a relation to identifying the people that are dying of CVD or CVD/cancer, because all markers show elevated values in the CVD-death and cancer-death group compared to the survivors.

Table 6.4 shows that the simulated area percentage and number of calcifications (NCDs) have the largest individual predictive power (HR=2.96, p<0.001 and HR=2.44, p<0.001) for CVD-mortality. The hazard ratios (HRs) for the simulated area percentage and NCD were between 2.0 - 2.96 and 1.76 - 2.44

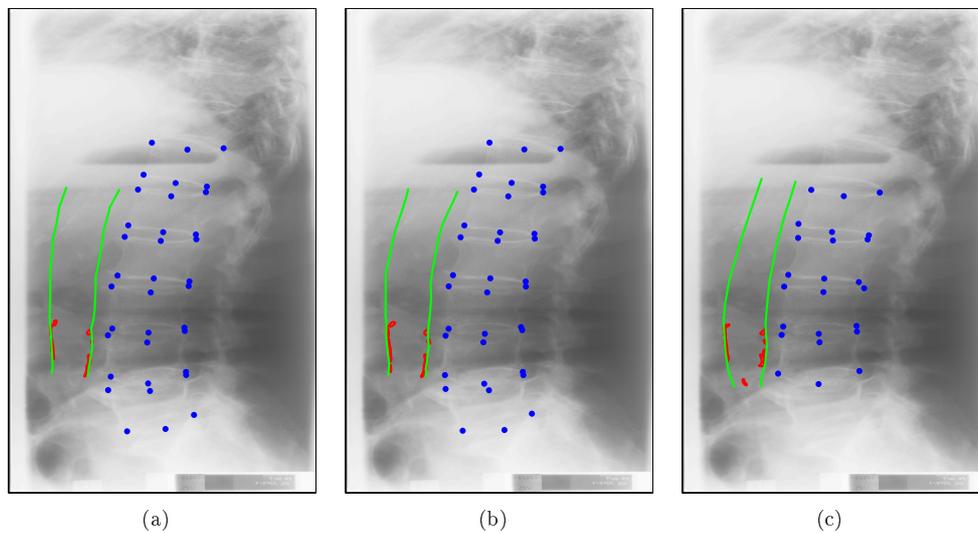


Figure 6.6: An X-ray of a participant in the EPI follow-up population: (a) an annotation by a radiologist, (b) a second annotation by the same radiologist and (c) an annotation done by another radiologist. Notice how although the outlines of the annotated calcifications do not vary much, the first radiologists misses several calcifications completely. This is often what leads to a low area overlap when comparing annotations from different radiologists.

Table 6.2: The inter- and intra-observer mean coefficients of variation  $\pm$  their standard error for the AAC markers based on the inter-intra-observer test population. The coefficients of variation are all below 25% and therefore we can measure the AAC markers with reasonable accuracy.

Inter-Intra-Observer Population	Inter-Observer CV % $\pm$ standard error	Intra-Observer CV % $\pm$ standard error
Area	$23.9 \pm 4.7$	$24.7 \pm 4.9$
Sim. area %	$24.9 \pm 5.4$	$20.4 \pm 5.3$
Thickness %	$17.1 \pm 3.3$	$16.1 \pm 3.6$
Wall %	$12.1 \pm 2.1$	$12.9 \pm 2.7$
Length %	$12.1 \pm 2.0$	$12.9 \pm 2.7$
NCD	$19.4 \pm 3.1$	$16.6 \pm 3.5$

respectively for the CVD-death group and between 1.68 - 2.37 and 1.69 - 2.28 respectively for the combined CVD/cancer-death group. All HRs were significantly different from unity ( $p < 0.01$ ) both before and after adjusting for three different biological models, so the ratio of the probability of dying in the CVD or CVD/cancer death group versus the rest is significantly larger than 1.

Table 6.3: The mean  $\pm$  one standard deviation of all the imaging markers stratified for the different subsets of patients. One can see large differences in the AAC markers measurements in the survivor group and the CVD, cancer and CVD/cancer groups of deceased. Note especially that the 5 subjects that died of other causes are can not be discerned from the survivors with the help of the AAC markers.

	All (n=308)	Survivors (n=256)	CVD (n=20)	Cancer (n=27)	CVD/Can. (n=47)	Other (n=5)
AC24	1.67 $\pm$ 2.55	1.35 $\pm$ 2.34	3.50 $\pm$ 2.35	3.41 $\pm$ 3.23	3.45 $\pm$ 2.86	1.35 $\pm$ 2.36
Area %	0.6 $\pm$ 1.2	0.5 $\pm$ 1.1	1.0 $\pm$ 0.9	1.6 $\pm$ 1.8	1.3 $\pm$ 1.5	0.5 $\pm$ 1.1
Sim. area %	11 $\pm$ 17	8.9 $\pm$ 15.7	24 $\pm$ 16	25 $\pm$ 24	25 $\pm$ 21	8.7 $\pm$ 15.5
Thickness %	11 $\pm$ 20	9.0 $\pm$ 19	17 $\pm$ 16	25 $\pm$ 28	21 $\pm$ 24	8.7 $\pm$ 19
Wall %	1.03 $\pm$ 1.83	0.79 $\pm$ 1.64	2.08 $\pm$ 1.70	2.51 $\pm$ 2.68	2.33 $\pm$ 2.30	0.80 $\pm$ 1.63
Length %	7.5 $\pm$ 12.8	6.0 $\pm$ 11.7	15.4 $\pm$ 11.2	17.3 $\pm$ 17.6	16.5 $\pm$ 15.1	5.9 $\pm$ 11.6
NCD	3.8 $\pm$ 7.7	2.6 $\pm$ 6.4	8.5 $\pm$ 6.5	11.6 $\pm$ 13.4	10.3 $\pm$ 11.0	2.6 $\pm$ 6.3

AC24s unadjusted individual predictive power was lower (HR=1.66,  $p < 0.001$ ). After adjustment for the three different biological models the significance of the HRs for AC24 was reduced and in some cases removed, leading to a HR between 0 and 1.66 for the CVD-death group and between 1.29 and 1.64 for the CVD/cancer-death group.

The results of the combined predictive power of the seven imaging markers can be seen for the CVD and the CVD/cancer group in table 6.5. First the non-adjusted hazard ratios from table 6.4 are stated again and then two elimination models are shown. When combining the markers in a Cox regression model, only area percentage and NCD remained significant ( $p_{\text{area}} < 0.001$ ,  $p_{\text{NCD}} < 0.001$ ) with negative and positive regression coefficients respectively.

## 6.4 Discussion

We evaluated whether a radiologist's manual scoring of the AC24 correlated with a computer-based scoring of the AC24 derived from a radiologist's manual outline of the calcifications on a digitized radiograph. The Kendall's coefficient of concordance showed the two scorings were in excellent agreement. We also evaluated inter- and intra-observer variability of manual annotations using the Jaccard Index and coefficients of variation of the AAC markers, including the AC24. Although the Jaccard Index showed that the variation in the outlined calcified deposits was high, the coefficients of variation for the AC24 and the other AAC markers based on the outlines were relatively low. These results

---

<sup>6</sup>In parentheses the 95% confidence intervals of the relative risk is shown and the symbols \*, \* and † denote the significance corresponding to  $p < 0.05$ ,  $p < 0.01$  and  $p < 0.001$ , respectively.

Table 6.4: The hazard ratio per standard deviation increase in marker values stratified into death cause and adjusted for physical/metabolic markers, EU SCORE and Framingham score respectively.<sup>6</sup> Note only the simulated area and NCD are significant for the CVD group after any adjustments.

	Hazard Ratio Not adjusted	Hazard Ratio Bio adj.	Hazard Ratio SCORE adj.	Hazard Ratio Fram. adj.
AC24				
CVD	1.66 (1.25-2.19) <sup>†</sup>	NS	1.38 (1.02-1.86)*	NS
CVD/Can.	1.64 (1.35-2.00) <sup>†</sup>	1.31 (1.06-1.63)*	1.40 (1.13-1.72)*	1.29 (1.02-1.63)*
Area %				
CVD	1.60 (1.16-2.20) *	NS	NS	NS
CVD/Can.	1.68 (1.36-2.09) <sup>†</sup>	1.32 (1.04-1.66)*	1.47 (1.16-1.86)*	1.34 (1.04-1.72)*
Sim. area %				
CVD	2.96 (1.76-4.99) <sup>†</sup>	2.00 (1.15-3.49)*	2.46 (1.41-4.27)*	2.27 (1.26-4.09)*
CVD/Can.	2.37 (1.73-3.25) <sup>†</sup>	1.68 (1.20-2.34)*	1.96 (1.40-2.73) <sup>†</sup>	1.79 (1.26-2.54)*
Thickness %				
CVD	NS	NS	NS	NS
CVD/Can.	1.45(1.20-1.75) <sup>†</sup>	NS	1.27 (1.04-1.55)*	NS
Wall %				
CVD	1.50 (1.16-1.95) *	NS	NS	NS
CVD/Can.	1.60 (1.34-1.91) <sup>†</sup>	1.26 (1.04-1.53)*	1.42 (1.17-1.73) <sup>†</sup>	1.30 (1.05-1.62)*
Length %				
CVD	1.55 (1.18-2.04) *	NS	NS	NS
CVD/Can.	1.61 (1.34-1.95) <sup>†</sup>	1.26 (1.03-1.55)*	1.42 (1.16-1.73) <sup>†</sup>	1.29 (1.03-1.62)*
NCD				
CVD	2.44 (1.72-3.48) <sup>†</sup>	1.76 (1.20-2.60)*	2.20 (1.48-3.26) <sup>†</sup>	2.04 (1.34-3.12) <sup>†</sup>
CVD/Can.	2.28(1.79-2.90) <sup>†</sup>	1.69 (1.30-2.21) <sup>†</sup>	2.00 (1.53-2.62) <sup>†</sup>	1.86 (1.40-2.47) <sup>†</sup>

demonstrated that even though the outlining of the individual plaques is a challenging task, the resulting markers based on the annotations provided reasonably precise measurements.

In the course of the 8-9 years of the study 52 people died, of whom 20 died from CVD-related causes and 27 from cancer. The Cox regression models showed similar correlations to CVD and CVD/cancer mortality for the different markers. The simulated area percentage and the number of calcified deposits could individually predict CVD and CVD/cancer death and contained additional information for CVD mortality even after adjustments for age, triglycerides and cholesterol and the SCORE model and Framingham score. Hence, in this post hoc study, the simulated area percentage and the number of calcified deposits predicted CVD mortality independently from traditional risk factors, in contrast to AC24. A reason for this could be that the AC24 does not discriminate between severity and spread of individual calcifications.

Our experiments with Cox regression elimination models showed that the haz-

Table 6.5: The individual hazard ratios per standard deviation for the markers in the CVD and the CVD/cancer group as well as two Cox regression elimination models. First the non-adjusted hazard ratios from table 6.4 are stated again and then two elimination models are shown, where  $\beta$  denotes the logistic regression coefficient of the given marker in the combined elimination models.<sup>6</sup> Only area percentage and NCD are left after a backward marker selection.

	CVD: $\beta \cdot std$	CVD elim.: $\beta \cdot std$	CVD/cancer: $\beta \cdot std$	CVD/cancer elim.: $\beta \cdot std$
AC24	1.66 <sup>†</sup>	-	1.64 <sup>†</sup>	-
Area %	1.60*	b = -3.84 <sup>†</sup>	1.68 <sup>†</sup>	b = -2.39 <sup>†</sup>
Sim. area %	2.96 <sup>†</sup>	-	2.37 <sup>†</sup>	-
Thickness %	1.32	-	1.45 <sup>†</sup>	-
Wall %	1.50*	-	1.60 <sup>†</sup>	-
Length %	1.55*	-	1.61 <sup>†</sup>	-
NCD	2.44 <sup>†</sup>	b = 2.76 <sup>†</sup>	2.28 <sup>†</sup>	b = 1.88 <sup>†</sup>

ard of dying of CVD or CVD/cancer was proportional to the number of calcifications and inversely proportional to their area. Therefore, our hypothesis is that many small plaques play a relevant role for the hazard of dying of CVD or CVD/cancer. Biologically this can be explained by the fact that the risk of death due to myocardial infarct (MI) may be related to the number of active plaques [127]. During plaque development, smaller plaques develop into larger complicated lesions that either rupture or become stable plaques. Smaller lipid-laden plaques with high turnover have been identified as those most likely to rupture and result in MI [128, 129]. Thus, a large number of smaller calcifications may indicate a higher risk of rupture than few large, stable, calcifications in the same area. This higher emphasis on the number of calcifications, rather than the total calcium burden, may reflect aspects of vulnerability that help improve the CVD-mortality prediction, as observed in this work.

The sample size is a limitation of the present study. The relatively small population with only 20 CVD deaths, a limited representation of ethnicity and gender and a mixture of death causes may limit the utility of generalizing our results. Therefore, the presented findings need to be validated in larger, independent studies. Although the current analysis is based on manual annotations by trained radiologists, the annotation procedure can in principle be automated. A first step toward automated detection and segmentation of aortic calcifications from radiographs has been provided by de Bruijne [89], Lauze, F. et al. [90] and Petersen, K. et al. [35].

Compared to markers of CVD obtained with other imaging modalities, such as carotid IMT or CAC, a clear advantage of using standard radiographs is the availability of large, long duration clinical studies about osteoporosis [33, 34]. The radiographic images derived from osteoporosis trials of post menopausal women could then be used at the same time for cardiovascular risk stratification in these women. This way the clinical applicability of AAC markers could be increased. Furthermore, historical trial data can be used to verify the developed AAC markers and can improve understanding of CVD death risk factors.

## 6.5 Conclusion

As shown above, simple statistical modeling can help to identify potential imaging markers. While AC24 captures essential information about AACs, our results demonstrate that some of these novel morphometric markers of AACs identified in this study, i.e., the number and the simulated area percentage of a calcified plaque in the abdominal aorta, may capture complementary information. Further steps can be taken by building combined biological and imaging markers or by developing even more AAC markers and repeating the same procedure as above. Although it is left to show the clinical applicability and reproducibility of the newly identified markers, this statistical approach for marker development seems to be a step in the right direction and the proposed radiographic AAC markers may enable improved monitoring of CVD mortality risk.



# Chapter 7

## Distribution, Size, Shape, Growth Potential and Extent of Abdominal Aortic Calcified Deposits Predict Mortality in Post Menopausal Women

---

This chapter is based on the highly accessed manuscript "Distribution, Size, Shape, Growth Potential and Extent of Abdominal Aortic Calcified Deposits Predict Mortality in Post Menopausal Women" by Mads Nielsen, Melanie Ganz, Francois Lauze, Paola C. Pettersen, Marleen de Bruijne, Thomas B. Clarkson, Erik B. Dam, Claus Christiansen and Morten A. Karsdal published in the journal BMC Cardiovascular Disorders 2010, 10:56.

---

**Abstract** Aortic calcification is a major risk factor for death from cardiovascular disease. We investigated the relationship between mortality and the composite markers of number, size, morphology and distribution of calcified plaques in the lumbar aorta. 308 post menopausal women aged 48-76 were followed for  $8.3 \pm 0.3$  years, with deaths related to cardiovascular disease, cancer, or other causes being recorded. From lumbar X-rays at baseline the number (NCD), size, morphology and distribution of aortic calcification lesions were scored and combined into the Morphological Atherosclerotic Calcification Distribution (MACD) index. The hazard ratio for mortality was calculated for the MACD and for three other commonly used predictors: the EU SCORE card, the Framingham Coronary Heart Disease Risk Score (Framingham score), and the gold standard Aortic Calcification Severity score (AC24) developed from the Framingham Heart Study cohorts. For the 10% subjects at highest risk of CVD death the SCORE card and the Framingham score resulted in a mortality hazard ratio of 4.9 ( $p < 0.01$ ) and 0, respectively. The MACD index revealed the best predictive power for identification of patients at 10% highest risk of mortality, with a hazard ratio of 15.6 ( $p < 0.001$ ). This study indicates that the MACD index may provide a more sensitive predictor of mortality from aortic calcification than the commonly used AC24 and SCORE/Framingham score systems.

## 7.1 Background

Cardiovascular diseases (CVDs) remain the most common cause of death in the developed world, even though vast epidemiological and interventional studies have demonstrated significant declines in CVD prevalence with adherence to a healthy lifestyle, and the identification and management of risk factors [6]. Since two thirds of women who die suddenly from CVD have no previously recognized symptoms [6], it is essential to find effective indicators of cardiovascular risk that may prompt timely intervention.

Biomarkers and biochemical markers are receiving increased attention for their potential prognostic value, and for identification of those patients in most need of intervention [130]. An extensive list of more than 200 potential CVD risk factors has been compiled [51] and multivariate analysis models, such as the EU SCORE card [49] and the Framingham Coronary Heart Disease Risk Score (Framingham score) [50], have been developed to estimate the risk of CVD death. However, more information may be provided by in-depth analysis of already-established risk factors.

Recently, several interesting findings have been reported on abdominal aortic calcifications as a CVD risk factor: i) Premature parental CVD has been associated with abdominal aortic calcification [131]. ii) Abdominal aortic calcium levels were significantly related to coronary calcium levels independent of the usual risk factors [132, 133]. iii) In type II diabetes patients, abdominal aortic calcification was shown to constitute an independent risk factor of clinical vascular disease [134]. iv) An increased total-to-high density lipoprotein (HDL) cholesterol ratio increased the risk of presence of aortic calcification [135]. v) Lumbar aortic calcifications in bone densitometer images have been shown to constitute an independent risk factor of CVD [136]. Hence, abdominal aortic calcification is an important risk factor for CVD.

Further investigations have indicated that it is rather the number of active lipid-laden remodeling, growing, plaques, rather than the total burden of calcified plaques, including stable plaques, that is related to cardiovascular death [137]. Also the number, distribution and size of calcified plaques have been shown to relate to mortality [138]. As the aortic calcification severity score (AC24) assesses, in terms of lesions, only the extent of calcification in the aorta, we developed a broader morphological atherosclerotic calcification distribution (MACD) index specifically to score the number, length, width, shape, and distribution of abdominal aortic calcifications (AAC) found in lumbar X-rays of post menopausal women. This index was created to further understand the composition of the plaque burden in relation to cardiovascular death. Low dose computed tomography might have been used to evaluate coronary calcifications for screening purposes [139], however its cost is a limiting factor.

We evaluated whether each risk included in the composite MACD marker per-

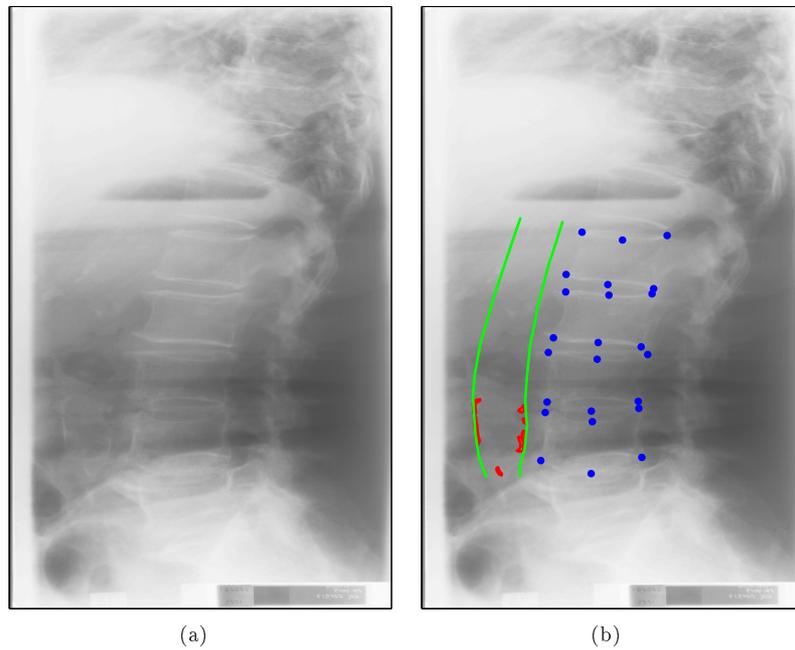


Figure 7.1: Lateral lumbar X-ray with calcifications in the lower region without (a) and with (b) computer-mediated annotations performed by a radiologist.

sisted after correction for generalized risk assessments used in the SCORE card [49], the Framingham score [50] or individual risk factors, such as smoking, cholesterol or triglycerides levels.

## 7.2 Methods

### 7.2.1 Subjects

In 1992-93, 686 post menopausal women living in the Copenhagen area in Denmark were recruited via a household postal survey to participate in a study addressing the role of a number of metabolic risk factors in the pathogenesis of CVD and osteoporosis [67].

Follow-up was performed after 8.5 years and information about all 95 individuals who died in the observation period was obtained from the Central Registry of the Danish Ministry of Health.

### 7.2.2 Markers

At baseline, information was collected on demographics and known risk parameters such as age, weight, height, body mass index (BMI), waist and hip cir-

cumferences, systolic and diastolic blood pressure (BP), treated hypertension, treated diabetes, smoking, regular alcohol and daily coffee consumption, and weekly fitness activity. Using a blood analyzer (Cobas Mira Plus, Roche Diagnostics Systems, Hoffman-La Roche, Basel, Switzerland), fasting glucose levels and lipid profiles, consisting of total cholesterol, triglycerides, LDL-cholesterol (LDL-C), HDL-cholesterol (HDL-C), apolipoproteins (ApoA and ApoB) and lipoprotein a (Lp(a)), were obtained.

On basis of these measurements, the composite risk SCORE card [49] and Framingham score [50] were both calculated based on the gender, age, systolic blood pressure, total cholesterol, and smoking status; and the Framingham score also based on HDL-C.

Lateral X-rays of the lumbar aorta (L1-L4 vertebrae) were taken at baseline and at follow-up. The images were digitized using a Vidar DosimetryPro Advantage scanner providing an image resolution of 570 dpi on a 12-bit gray scale. Trained, blinded radiologists annotated the digitized images on a Sectra radiological reading unit using annotation software developed in Matlab (Mathworks, MA, USA) (figure 7.1). The radiologists were instructed to annotate the 6 points used for vertebral height measurements on L1-L4 [140], to delineate the aorta, and finally to outline every individual calcified deposit visible in the lumbar aorta and note their possible association to the anterior and/or posterior wall. The software enabled digital zooming and editing [53]. The inter- and intra- observer variability was tested by three radiologists annotating the same 16 randomly selected images.

Geometric data relating to the calcified deposits in the L1-L4 region was quantified as follows:

- Area percentage (Area %): The percentage of the aorta lumen area occupied by calcified deposits.
- Simulated area percentage (Sim. area %): As X-rays only capture the calcified core and not the biological extent of atherosclerotic lesions, we implemented a statistically validated method [138], in which the atherosclerotic plaque size was estimated from the area and form of the observed calcified lesion, and the resulting area percentage was recorded. The estimation was done using a grass-fire equation based on a morphological dilation [121] with a circular structuring element of radius 200 pixels corresponding to 8.9 mm. The biological extent of atherosclerotic lesions around an elongated calcified lesion was estimated to be larger than the biological extent of atherosclerotic lesions around a circular calcification of similar size. Thus, equal areas of calcification but of different shapes were given different scores (see figure 7.2).
- Thickness percentage (Thickness %): The average thickness of the calcified deposits along the aorta wall, expressed as a percentage of the aorta

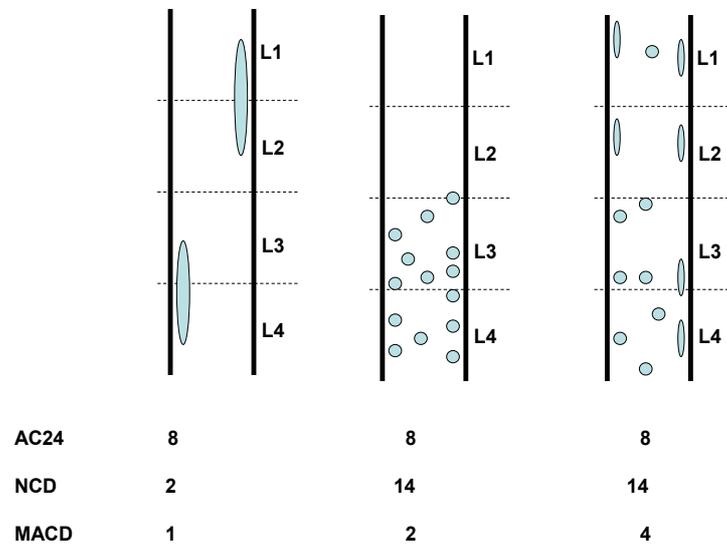


Figure 7.2: For a given amount of calcified tissue, one can see schematically how the AC24, the NCD and the MACD can be influenced differently by variations in calcification morphology and distribution.

width.

- **Wall percentage (Wall %):** The percentage of the aorta wall covered by calcified deposits.
- **Length percentage (Length %):** The percentage of the length of the aorta in which a calcified deposit was present, in any position (anterior, posterior or internal).
- **Number of Calcified Deposits (NCD):** The number of distinct calcified deposits.

The relationship between these individual markers and CVD mortality in this cohort has already been demonstrated [138]. Furthermore, two novel composite markers were created:

1. **Morphological Atherosclerotic Distribution (MAD) factor:**  
The simulated plaque area divided by the area estimates the portion of the biological atherosclerotic process which is not detected by X-rays.
2. **Morphological Atherosclerotic Calcification Distribution (MACD) index:**  
The NCD multiplied by the MAD factor. Biologically that can be un-

derstood as the number of plaques multiplied by the disease potential described by the MAD factor.

### 7.2.3 Statistical Analysis

Patients were stratified into survivors and deceased. The latter were substratified into CVD-related, cancer-related and other-cause deaths. Since cancer and CVD have many risk factors in common, an additional group containing all cancer or CVD deaths was created to increase numbers and improve statistical significance.

To test the prognostic power, metabolic and physical parameters and AAC markers were used in separate Cox-regression models with the time of death as the outcome variable while right-censoring survivors. Significance was tested as the model weight being significantly different from zero. To test if one marker carried additional prognostic power compared to the remaining markers, a model including all elementary metabolic/physical parameters was sequentially stripped for the insignificant markers until significance persisted for all markers. To test if an AAC marker carried prognostic power in addition to the other AAC markers and/or metabolic/physical markers, each marker was compared in combined stripped models. Separate models for CVD, CVD/cancer and all-cause death were created.

As CVD and CVD/cancer death rates were 6.5% and 15.2% respectively, a 10% percentile cut-off was used to separate subjects at high risk from those (90%) at normal risk. Hazard ratios were computed, adjusted for the influence of other risk parameters by combining all other risk factors into Cox-regression models.

## 7.3 Results

Of the 686 post menopausal women enrolled in the original study in 1992-93, 95 died prior to follow-up with 52 (55%) of them having baseline X-ray examinations in which the full lumbar (L1-L4) aorta was visible on a single radiograph. Of these 52 deaths, 20 (38%) were due to CVD, 27 (52%) to cancer and 5 (10%) to other causes. Another 129 women had relocated from the Copenhagen area or did not want to participate in the follow-up study and provided no clinical data for it. Of the 462 women completing the follow-up visit, lumbar aorta from 256 (55%) were visible on a single radiograph (figure 7.3). This compares with the aorta visibility percentage reported in earlier studies [136]. Therefore in total, 308 (52 plus 256) women were included in the current analysis. Baseline demographics and risk parameters showed no

---

<sup>7</sup>The symbols \*, \* and † denote the significance corresponding to  $p < 0.05$ ,  $p < 0.01$  and  $p < 0.001$ , respectively.

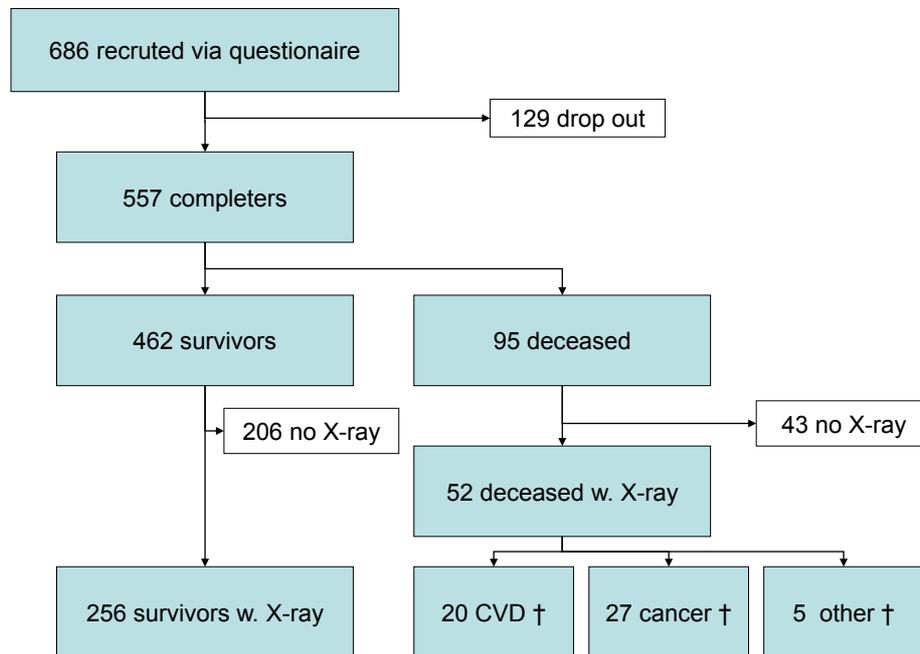


Figure 7.3: Of 557 post menopausal women who completed an 8.5 year follow-up study, 55% of those alive at follow-up and 55% of those who were deceased had useful X-rays with the full abdominal aorta visible in a single X-ray. Thus, the study population included in this analysis consisted of a total of 308 women: 256 survivors and 52 deceased.

difference between the discontinued women and those completing the study. Observer reproducibility, assessed by three radiologists scoring the same 16 X-rays, three times each, resulted in an inter- and intra- observer area overlap [68] of 56% and 60% respectively, showing good reproducibility [40]. These annotations were used to compute the AC24 that ranges from 0 to 24 based on the length of the vertebral sections affected by calcified deposits [13]. Most of the physical and metabolic markers provided prognostic separation of the groups of survivors and deceased as depicted in tables 7.1 and 7.2. In a combined model including all physical/metabolic parameters only age, smoking, and triglyceride level persisted after elimination of insignificant contributions. All three parameters were positively associated with death. These were combined into one parameter denoted "combined metabolic/physical parameter" (HR per SD = 2.94 (2.18-3.95),  $p < 0.001$ ) for further analysis. All imaging-based AAC markers showed higher values in the CVD, cancer, and combined CVD/cancer groups than in the survivor group (table 7.3) and independently and significantly predicted death in the CVD and combined CVD/cancer groups (table

Table 7.1: Population characteristics of the study population stratified into survivors and deceased (all-cause) expressed as mean±standard deviation. Most of the physical and metabolic markers provide a subtle separation of the groups of survivors and deceased.

Physical/Metabolic markers	Population (n=308)	Survivors (n=256)	Deceased (n=52)
Age (years)	60.3 ± 7.5	59.3 ± 7.1	65.6 ± 7.0
Waist (cm)	80.7 ± 10.9	80.2 ± 9.9	83.1 ± 12.4
Waist-to-hip ratio	0.80 ± 0.08	0.80 ± 0.08	0.83 ± 0.10
Body Mass Index ( $\text{frackg}\text{m}^2$ )	24.7 ± 3.9	24.7 ± 3.8	25.1 ± 4.6
Smoking (%)	37	33	58
Systolic BP (mm Hg)	127 ± 21	125 ± 20	136 ± 26
Diastolic BP (mm Hg)	77 ± 10	76 ± 10	77 ± 11
Hypertension (%)	16	15	17
Glucose (mmol/L)	5.44 ± 1.27	5.37 ± 0.99	5.79 ± 2.17
Total cholesterol (mmol/L)	6.44 ± 1.19	6.36 ± 1.14	6.85 ± 1.33
Triglycerides (mmol/L)	1.24 ± 0.75	1.15 ± 0.56	1.69 ± 1.25
LDL-C (mmol/L)	2.89 ± 0.82	2.85 ± 0.80	3.07 ± 0.93
HDL-C (mmol/L)	1.77 ± 0.48	1.77 ± 0.44	1.74 ± 0.62
ApoB/ApoA	0.57 ± 0.18	0.56 ± 0.17	0.64 ± 0.23
Lp(a) (mg/dL)	21.4 ± 21.7	21.9 ± 22.0	18.4 ± 19.8
EU SCORE	2.60 ± 2.58	2.16 ± 2.12	4.73 ± 3.45
Framingham	14.75 ± 3.54	14.21 ± 3.46	17.31 ± 2.74

7.4, column 2). This significance persisted for simulated Area, NCD, MAD factor, and MACD also when adjusted for the combined metabolic/physical parameter, EU SCORE, or Framingham score. AC24, wall% and length% all maintained a significant prediction under adjustment in the CVD/cancer group, but did not have sufficient statistical power in the smaller CVD group (table 7.4).

In a combined elimination model using all elementary calcification markers, only the number of calcified deposits (NCD) (positive association to death) and area % (negative association to death) persisted in the CVD group and the CVD/cancer group. The composite marker MACD showed highest predictability in all tests and also higher predictability (but not significantly so) than the combined elimination models of the elementary calcification markers. In the CVD deaths group, the highest 10% of NCD or MACD scores were significantly associated with death. This did not hold for AC24 or area% values in the same group (table 7.5). This relation persisted but with decreasing hazard

Table 7.2: Hazard ratios of all-cause death (HR) per standard deviation of metabolic/physical markers and their 95% confidence interval (CI) based on a Cox regression model as well as for a sequentially stripped model including all metabolic/physical markers.<sup>7</sup> Not all of the metabolic/physical markers have a significant hazard ratio and in a stripping model only age, smoking and triglycerides prevail in this population.

Physical/ Metabolic markers	HR per SD	HR per SD
	[95% CI] Alone	[95% CI] Comb.
Age	2.25 <sup>†</sup> (1.67-3.03)	2.41 <sup>†</sup> (1.75-3.31)
Waist	1.29* (1.01-1.65)	NS
Waist-to-hip ratio	1.37* (1.12-1.67)	NS
Body Mass Index	NS	NS
Smoking	1.37* (1.08-1.73)	1.50* (1.17-1.94)
Systolic BP	1.53 <sup>†</sup> (1.20-1.94)	NS
Diastolic BP	NS	NS
Hypertension	NS	NS
Glucose	1.23* (1.03-1.46)	NS
Total cholesterol	1.44* (1.12-1.86)	NS
Triglycerides	1.51 <sup>†</sup> (1.29-1.76)	1.46 <sup>†</sup> (1.22-1.75)
LDL-C	NS	NS
HDL-C	NS	NS
ApoB/ApoA	1.45* (1.14-1.83)	NS
Lp(a)	NS	NS
EU SCORE	1.79 <sup>†</sup> (1.51-2.13)	Not Incl.
Framingham	2.63 <sup>†</sup> (1.87-3.71)	Not Incl.

ratios when adjusted by standard composite metabolic/physical markers (EU SCORE or Framingham score) or the combined metabolic/physical parameter in the elimination model from table 7.2. Similar results were obtained in the CVD/cancer group with slightly lower hazard ratios and higher significance levels due to the larger population.

## 7.4 Discussion

We investigated whether more information could be obtained from calcified deposits in the abdominal aorta to better predict CVD death than the gold standard AC24 score, which was developed from the Framingham Heart Study cohorts. We hypothesized that the presence of many small, spatially distributed,

Table 7.3: Stratification of abdominal aortic calcification marker values according to cause of death shown as mean  $\pm$  standard deviation. One can observe quite a difference between the values of the AAC markers of the survivors and the CVD, cancer and CVD/cancer group. Conversely, there is virtually no difference between the survivors and the 5 subjects that died of other causes.

	All (n=308)	Survivors (n=256)	CVD (n=20)	Cancer (n=27)	CVD/Can. (n=47)	Other (n=5)
AC24	1.67 $\pm$ 2.55	1.35 $\pm$ 2.34	3.50 $\pm$ 2.35	3.41 $\pm$ 3.23	3.45 $\pm$ 2.86	1.35 $\pm$ 2.36
Area %	0.6 $\pm$ 1.2	0.5 $\pm$ 1.1	1.0 $\pm$ 0.9	1.6 $\pm$ 1.8	1.3 $\pm$ 1.5	0.5 $\pm$ 1.1
Sim. area %	11 $\pm$ 17	8.9 $\pm$ 15.7	24 $\pm$ 16	25 $\pm$ 24	25 $\pm$ 21	8.7 $\pm$ 15.5
Thickness %	11 $\pm$ 20	9.0 $\pm$ 19	17 $\pm$ 16	25 $\pm$ 28	21 $\pm$ 24	8.7 $\pm$ 19
Wall %	1.03 $\pm$ 1.83	0.79 $\pm$ 1.64	2.08 $\pm$ 1.70	2.51 $\pm$ 2.68	2.33 $\pm$ 2.30	0.80 $\pm$ 1.63
Length %	7.5 $\pm$ 12.8	6.0 $\pm$ 11.7	15.4 $\pm$ 11.2	17.3 $\pm$ 17.6	16.5 $\pm$ 15.1	5.9 $\pm$ 11.6
NCD	3.8 $\pm$ 7.7	2.6 $\pm$ 6.4	8.5 $\pm$ 6.5	11.6 $\pm$ 13.4	10.3 $\pm$ 11.0	2.6 $\pm$ 6.3

radiographically visible calcified deposits of varying shape in the lumbar aorta had a stronger relation to CVD death than the AC24 segment-wise scoring of the extent of calcified deposits on the aortic wall.

The AC24 score [13] quantifies the burden of calcified plaques in the aorta by segment-wise scoring of the calcified deposit coverage of the aortic wall. We investigated whether additional aspects of the outline of the individual plaques may be associated with the progression and/or prognosis of atherosclerosis. We analyzed the area %, thickness %, wall % and length % of the abdominal aorta covered by calcification and the number of distinct calcified deposits. Furthermore, we calculated the simulated plaque area in which the atherosclerotic plaque size was estimated from the area and form of the observed calcified lesion. Lastly, two composite markers were created: i) The morphological atherosclerotic distribution (MAD) factor was constructed by dividing the simulated plaque area with the absolute plaque area. ii) The morphological atherosclerotic calcification distribution (MACD) index is given by the NCD multiplied by the MAD factor.

In the present cohort, eight different markers (AC24, area %, simulated area %, wall%, length%, NCD, MAD and MACD) exhibited a significant hazard ratio per standard deviation increase for death in the combined CVD/cancer group when adjusted for physical/metabolic markers, the EU SCORE, and the Framingham score respectively. However, only four markers (simulated area %, NCD, MAD and MACD) had sufficient power in risk segregation of CVD mortality when adjusted by physical/metabolic markers, the EU SCORE and the Framingham score. The composite MAD factor showed increased sensitivity to CVD compared to cancer mortality. The reason for this may be that the MAD factor essentially scores how small and widely distributed the indi-

Table 7.4: Hazard ratio per standard deviation increase in marker value stratified into death cause and adjusted for physical/metabolic markers, the EU SCORE, and the Framingham score respectively.<sup>7</sup> Only the simulated area, NCD, MAD and MACD persist after adjusting for physical/metabolic markers, the EU SCORE, and the Framingham score.

	Hazard Ratio Not adjusted	Hazard Ratio Bio adj.	Hazard Ratio SCORE adj.	Hazard Ratio Fram. adj.
AC24				
CVD	1.66 (1.25-2.19) <sup>†</sup>	NS	1.38 (1.02-1.86)*	NS
CVD/Can.	1.64 (1.35-2.00) <sup>†</sup>	1.31 (1.06-1.63)*	1.40 (1.13-1.72)*	1.29 (1.02-1.63)*
Area %				
CVD	1.60 (1.16-2.20) *	NS	NS	NS
CVD/Can.	1.68 (1.36-2.09) <sup>†</sup>	1.32 (1.04-1.66)*	1.47 (1.16-1.86)*	1.34 (1.04-1.72)*
Sim. Area %				
CVD	2.96 (1.76-4.99) <sup>†</sup>	2.00 (1.15-3.49)*	2.46 (1.41-4.27)*	2.27 (1.26-4.09)*
CVD/Can.	2.37 (1.73-3.25) <sup>†</sup>	1.68 (1.20-2.34)*	1.96 (1.40-2.73) <sup>†</sup>	1.79 (1.26-2.54)*
Thickness %				
CVD	NS	NS	NS	NS
CVD/Can.	1.45 (1.20-1.75) <sup>†</sup>	NS	1.27 (1.04-1.55)*	NS
Wall %				
CVD	1.50 (1.16-1.95) *	NS	NS	NS
CVD/Can.	1.60 (1.34-1.91) <sup>†</sup>	1.26 (1.04-1.53)*	1.42 (1.17-1.73) <sup>†</sup>	1.30 (1.05-1.62)*
Length %				
CVD	1.55 (1.18-2.04) *	NS	NS	NS
CVD/Can.	1.61 (1.34-1.95) <sup>†</sup>	1.26 (1.03-1.55)*	1.42 (1.16-1.73) <sup>†</sup>	1.29 (1.03-1.62)*
NCD				
CVD	2.44 (1.72-3.48) <sup>†</sup>	1.76 (1.20-2.60)*	2.20 (1.48-3.26) <sup>†</sup>	2.04 (1.34-3.12) <sup>†</sup>
CVD/Can.	2.28 (1.79-2.90) <sup>†</sup>	1.69 (1.30-2.21) <sup>†</sup>	2.00 (1.53-2.62) <sup>†</sup>	1.86 (1.40-2.47) <sup>†</sup>
MAD				
CVD	3.37 (1.83-6.21) <sup>†</sup>	2.44 (1.22-4.89)*	3.02 (1.55-5.86)*	2.85 (1.44-5.64)*
CVD/Can.	2.19 (1.58-3.04) <sup>†</sup>	1.58 (1.11-2.26)*	1.83 (1.29-2.59) <sup>†</sup>	1.74 (1.22-2.48)*
MACD index				
CVD	5.22 (2.40-11.36) <sup>†</sup>	3.17 (1.48-6.78)*	4.36 (1.97-9.66) <sup>†</sup>	4.22 (1.79-9.97) <sup>†</sup>
CVD/Can.	2.99 (2.05-4.35) <sup>†</sup>	2.01 (1.37-2.95) <sup>†</sup>	2.43 (1.64-3.59) <sup>†</sup>	2.27 (1.51-3.41) <sup>†</sup>

vidual calcified plaques appear. When the MAD factor was combined with the number of calcified plaques, which as an individual parameter alone was shown to be a strong predictor of mortality, the resulting MACD index displayed superior predictive power over any other marker. The MACD index produced hazard ratios  $>4$  per standard deviation increase in the CVD death group, even after adjustment for metabolic/physical factors.

In trying to identify which tool would be most useful in clinical practice to identify CVD patients at highest risk of death, we found, from applying the

Table 7.5: Hazard ratio for high risk subjects based on 90% threshold in the CVD deaths group.<sup>7</sup> Here one can directly compare the performance of the different markers. NCD and MACD are best at identifying high risk subjects in this population.

AC Marker Adjusted by	Hazard Ratio Alone	AC24	Area%	NCD	MACD index
None		NS	NS	10.9 (4.4-27) <sup>†</sup>	15.6 (6.3-38) <sup>†</sup>
EU SCORE	4.9 (1.9-13)*	NS	NS	8.5 (3.2-23) <sup>†</sup>	13.2 (4.9-35) <sup>†</sup>
Framingham	NS	NS	NS	10.8 (4.1-28) <sup>†</sup>	15.7 (6.1-40) <sup>†</sup>
All metabolic/ physical	10.1 (4.1-25) <sup>†</sup>	NS	NS	7.2 (2.8-18) <sup>†</sup>	9.8 (3.7-26) <sup>†</sup>

various scoring systems to post menopausal subjects who had died from CVD, that the MACD index is potentially a better predictor of mortality. For the 10 % highest risk subjects the MACD index produced a hazard ratio for death of 10 and more even after adjustment for metabolic and physical markers, while the hazards ratios for the AC24 and the Framingham score were both insignificant, and the EU SCORE, had a value of the hazard ratio of 5.

## 7.5 Conclusion

In conclusion, assessment of the shape, size, number, distribution, and extent of lumbar aortic calcifications may aid in identifying patients at risk of CVD death and thus most in need of treatment. Since atherosclerosis is a systemic disease in which lumbar aortic calcifications occur, increasing attention has been devoted to the correlation between the number of lumbar aortic calcifications in radiographs and coronary calcifications [141]. [10, 141, 142] suggest that radiographs provide equally valuable information on CVD and offer the advantage of simplicity for in-office quantification. Some studies even suggest the number of lumbar aortic calcifications is an independent predictor of CVD events [141]. Importantly, only the calcified core of an atherosclerotic lesion is detected in X-rays whereas the surrounding necrotic tissue and region of high remodeling and fibrosis are not detectable. Hence, the actual pathologically involved area is underestimated in radiographs. Consequently, the morphological enlargement of plaques (used in the MAD factor and thereby the MACD index) may carry information related to the projected area of the inflammatory processes and indirectly indicate an increased risk. This additional information may result in a better prediction of mortality risk than the current state-of-the-art, the AC24 radiographic scoring of atherosclerotic plaques.

The present study has its limitations. Its findings are only valid for a follow-

up period of 8.5 years and may not necessarily apply to shorter follow-ups. For short follow-up times, the predictive power could possibly be based only on the total plaque burden as described by the AC24 score. Furthermore, the present population is restricted in size, geographical and ethnic content to post menopausal Danish women. Therefore, the present study needs validation in other populations and longer term clinical settings.



# Chapter 8

## Discussion and Conclusions

### 8.1 Summary

The main goal of the presented work was to make CVD preventive diagnostics available for clinical studies or even screening programs through low cost imaging. Our proposed framework for automated segmentation of calcified deposits including the contributions to the calcification shape priors and the evaluation of the developed biomarkers on clinical trial data was the first step in that direction.

In the first part of the dissertation we focused on the development of different calcification shape priors that could be used within a fully-automatic framework for detection of lumbar aortic calcifications, called the CVD pipeline. Two different approaches were proposed: In chapter 3 we used a method from geostatistics, called Kriging, to try to include knowledge learned from training data about the spatial correlations between different pixels in the output of a pixel classifier. We observed that our new method, DICPC, improved the sensitivity and the Jaccard index on our test set of calcification data. But DICPC was only the first step in the right direction and its weakness was the linear model that underlies Kriging which implies only pairwise interactions. Hence, going one step further, in chapter 4 we proposed a new way of not only learning some spatial characteristics from training data, but a whole patch representation of the training data. Such a representation can capture the statistical nature of a given texture and enforce it on testing data. Applied by itself it can be used in texture synthesis, in combination with a pixel classifier it can improve the spatial coherence of classifier output. Our segmentation results, based on the proposed shape prior, were promising and improved the Jaccard index by 50%. This indicates that the shape model can be successfully used as a prior in statistical segmentation of calcifications on X-ray image data.

In chapter 5 we made an excursion into the world of statistical physics. A

method based on renormalization group theory (RGT) was used to derive a new approach for multi-scale image segmentation that could be an alternative to parts of the CVD pipeline. Our experiments showed that RGT can be applied to multi-scale image segmentation problems. Furthermore, by using RGT we reduced the size of the image segmentation problem and made it tractable and solvable by a global image segmentation algorithm such as graphcut. But this still needs to be applied on our calcification data to show its applicability on medical data.

Part two of the dissertation dealt with the development of imaging biomarkers based on manual annotations or the classification output of the CVD pipeline. In chapter 6 a number of markers were measured and then a special type of backward selection via Cox regression was used to identify the most prominent combination of measurements that relate to death by CVDs. Our results pointed out the fact that simple statistical modeling can help to identify potential imaging markers and demonstrated that some of these novel morphometric markers of abdominal aortic calcifications may capture complementary information when assessing CVD risk.

Finally, in chapter 7 the statistical modeling lead to the formation of the morphological atherosclerotic calcification distribution (MACD) index. MACD could be shown to significantly relate to the risk of death and to outperform standard metabolic CVD biomarkers such as the EU SCORE [49] and the Framingham score [50] as well as to outperform the current gold standard on radiographs, the AC24 score, when identifying patients especially at risk.

## 8.2 Discussion and Conclusions

*"The scientific method is based on the principle that observation is the judge of whether something is so or not. All other aspects and characteristics of science can be understood directly when we understand that observation is the ultimate and final judge of the truth of an idea."<sup>8</sup>*

In this dissertation different image processing methods have been applied to test if automatic segmentation of calcifications in lumbar aortic X-rays is possible. Furthermore, the hypothesis that image analysis of calcified deposits in the lumbar aorta can quantify CVD risk of death has been tested. The results have provided new insights into the relation between imaging biomarkers of the lumbar aorta and CVD risk of death, but have also highlighted some remaining challenges.

For example, the CVD pipeline is still suffering from overall low Jaccard index values compared to radiologists' readings. There are several reasons for

---

<sup>8</sup>From "The Meaning of It All: Thoughts of a Citizen Scientist" by Richard P. Feynman

this. First, the fully automatic detection of calcifications is very challenging as already described in section 2.2.2. The fact that there are no ground truths available and that two radiologists have only ca. 50% area overlap with each other does not make the task easier. Second, the area overlap measure we use, the Jaccard index, is a very strict measure. For example, if an x-ray has no calcifications present in the manual annotation and the CVD pipeline finds one calcification, the Jaccard index drops to zero. Hence, one can not compare the Jaccard index we achieved to index values achieved in e.g. brain segmentation where two large areas are compared to each other. Third, the scores that relate to biological findings e.g. AC24 do not suffer as much as the area overlap and are actually quite robust [35] which indicates that the CVD pipeline is on the way to make fully automatic segmentation possible.

Despite the obstacles we have met in the CVD pipeline project, the new scientific results produced within it until now, e.g. a new way of modeling textures [83] and the static SMC sampler [38], show that one can not be ambitious enough. Even if at present applicability in clinical trials is out of question, the results achieved until now show that it is not impossible to get there. Especially, once digital X-rays of more recent studies (and hence of higher quality) and from different populations, that help to increase e.g. the generalization ability of the shape models, are in place, the CVD pipeline can be tested again and re-evaluated versus the performance of radiologists. If these steps are taken, then clinical trial availability is maybe only some years away.

Furthermore, while the application of RGT on image segmentation and energy minimization is clearly still in its fledgling stages, the results encourage to explore the subject more deeply. And the mere existence of the RGT approach shows how fertile inter-disciplinary work can be. Just as biologists have been involved in the development of the new biomarkers and have helped to interpret the statistical findings, transferring RGT onto images was only possible because of the very different background an education in theoretical physics offers. In the end working in the field of medical imaging is exciting and interesting, because many disciplines – biology, chemistry, computer science and physics – meet and something new is created by combining knowledge from all fields.

Finally, the work regarding the development of new biomarkers, e.g. MACD, has confirmed the fruitfulness of a statistical approach. When trying to gain new insights about e.g. a disease, it can be beneficial to not make a hypothesis first and then perform statistical tests until the same hypothesis is confirmed, but to let the data guide the search for a new hypothesis. Of course statistics alone can not explain biology, but it can maybe lead to us to new insights about our problem.

### 8.3 Future Work

While the application of Kriging and our proposed dense iterative contextual pixel classification (DICPC) was a first try at modeling spatial relations between classification output, the modeling of textures via a generative shape model based on a sparse texton dictionary was more successful. An interesting future direction for the texton approach could be to apply the textons also for non-binary images. In this case a different basis of the dictionary, e.g. ridgelets [143] or wedgelets [144], might be necessary. Furthermore, an appropriate multi-resolution extension of the generative model could be introduced that would be able to take even longer interactions between patches into account and as a consequence improve the segmentation further.

With respect to the CVD pipeline as a whole, it is clear that the performance of the calcifications shape priors depends greatly on the performance of the pixel classifier. This could be improved by either trying to select better features or by choosing a different classifier. But regardless of the performance that can be achieved on the EPI [67] and PERF [66] data sets available to us, it is left to show the applicability of the CVD pipeline for clinical trials by applying it to different data sets to be sure not to overfit to the data the pipeline was developed on. This can be done by conducting transfer experiments where we train the pipeline on one data set and then show its efficiency on another data set of a different population.

The RGT approach could be extended in several ways. On the one hand, different grid structures and mappings should be tested. On the other hand, to really get a broad spectrum of people interested in the method it should be extended to multi-labels, so rather a Potts than an Ising model.

The possibilities for future work in the realm of biomarkers are extensive. On top of the list is the development of new imaging biomarkers, possibly even on CT images acquired in the Danish Lung Cancer Screening Trial [145] instead of on X-ray images. Second, all newly developed markers including MACD need to be validated again, since the current findings are only based on a population that is restricted in size, geographical and ethnic content to post menopausal Danish women. A validation of our findings in other populations and longer term clinical settings would substantiate them further for use in clinical trials. Third, an option could be to combine the imaging and metabolic biomarkers for CVD risk assessment or to even try to relate to sth. else than CVD death when building risk assessment models.

**Through the presented research I have made progress in identifying the relations of CVD risk to abdominal aortic calcifications as imaged by X-rays, but also identified new questions that need to be answered. So the work presented in this dissertation is meant to be part of a bigger effort by the medical imaging community to pre-**

vent cardiovascular diseases by developing even better biomarkers for CVD risk and by making CVD preventive diagnostics through low cost imaging available.



## Bibliography

- [1] World Health Organization: Cardiovascular diseases (CVDs) (Fact sheet Nr. 317). <http://www.who.int/mediacentre/factsheets/fs317/en/index.html> (January 2011)
- [2] Maton, A.: Human biology and health. Prentice Hall (1997)
- [3] National Center for Health Statistics: U.S. Decennial Life Tables for 1989 to 1991. Eliminating Certain Causes of Death, 1989 to 1991 **1** (September 1999)
- [4] National Center for Health Statistics: Compressed mortality file: underlying cause of death, 1979 to 2005. <http://wonder.cdc.gov/mortSQL.html> (November 2010)
- [5] American Heart Association (AHA): Heart disease and stroke statistics - 2010 update. <http://www.americanheart.org/presenter.jhtml?identifier=3000090> (2010)
- [6] Mosca, L., Appel, L., et al.: Evidence-Based Guidelines for Cardiovascular Disease Prevention in Women. *Circulation* **109**(5) (2004) 672–693
- [7] Murabito, J., Evans, J., Larson, M., Levy, D.: Prognosis after the onset of coronary heart disease. An investigation of differences in outcome between the sexes according to initial coronary disease presentation. *Circulation* **88**(6) (1993) 2548
- [8] Akosah, K., Schaper, A., Cogbill, C., Schoenfeld, P.: Preventing myocardial infarction in the young adult in the first place: how do the national cholesterol education panel iii guidelines perform?\* 1. *Journal of the American College of Cardiology* **41**(9) (2003) 1475–1479
- [9] Ghesquiere, S.: The role of phospholipases in atherosclerosis. UPM, Universitaire Pers Maastricht (2006)
- [10] Wilson, P., Kauppila, L., O'Donnell, C., Kiel, D., Hannan, M., Polak, J., Cupples, L.: Abdominal aortic calcific deposits are an important predictor of vascular morbidity and mortality. *Circulation* **103**(11) (2001) 1529
- [11] George, S., Johnson, J.: Atherosclerosis: molecular and cellular mechanisms. Wiley-VCH (2010)
- [12] American Heart Association (AHA): Atherosclerosis. <http://www.americanheart.org/presenter.jhtml?identifier=4440> (January 2011)
- [13] Kauppila, L., Polak, J., Cupples, L., Hannan, M., Kiel, D., Wilson, P.: New indices to classify location, severity and progression of calcific lesions in the abdominal aorta: a 25-year follow-up study. *Atherosclerosis* **132** (1997) 245–250
- [14] Libby, P., Theroux, P.: Pathophysiology of coronary artery disease. *Circulation* **111**(25) (2005) 3481
- [15] Lindsay, A., Choudhury, R.: Form to function: current and future roles for atherosclerosis imaging in drug development. *Nature Reviews Drug Discovery* **7**(6) (2008) 517–529
- [16] Kaul, P., Douglas, P.: Atherosclerosis Imaging: Prognostically Useful or Merely More of What We Know? *Circulation: Cardiovascular Imaging* **2**(2) (2009) 150

- [17] Greenland, P., Abrams, J., Aurigemma, G., Bond, M., Clark, L., Criqui, M., Crouse III, J., Friedman, L., Fuster, V., Herrington, D., et al.: Prevention Conference V: Beyond secondary prevention: identifying the high-risk patient for primary prevention: noninvasive tests of atherosclerotic burden: Writing Group III. *Circulation* **101**(1) (2000) 16–22
- [18] RadiologyInfo.org developed jointly by American College of Radiology and Radiological Society of North America: Radiation exposure in x-ray and ct examinations. [http://www.radiologyinfo.org/en/safety/index.cfm?pg=sfty\\_xray](http://www.radiologyinfo.org/en/safety/index.cfm?pg=sfty_xray) (2010)
- [19] Mukherjee, D., Yadav, J.: Carotid artery intimal-medial thickness: indicator of atherosclerotic burden and response to risk factor modification. *American Heart Journal* **144**(5) (2002) 753–759
- [20] Korcarz, C., Hirsch, A., Bruce, C., DeCara, J., Mohler, E., Pogue, B., Postley, J., Tzou, W., Stein, J.: Carotid intima-media thickness testing by non-sonographer clinicians: the office practice assessment of carotid atherosclerosis study. *Journal of the American Society of Echocardiography* **21**(2) (2008) 117–122
- [21] Achenbach, S., Raggi, P.: Imaging of coronary atherosclerosis by computed tomography. *European Heart Journal* **31**(12) (2010) 1442
- [22] Raman, R., Raman, B., Napel, S., Rubin, G.: Semiautomated Quantification of the Mass and Distribution of Vascular Calcification with Multidetector CT: Method and Evaluation1. *Radiology* **247**(1) (2008) 241
- [23] Rutten, A., Isgum, I., Prokop, M.: Coronary Calcification: Effect of Small Variation of Scan Starting Position on Agatston, Volume, and Mass Scores1. *Radiology* **246**(1) (2008) 90
- [24] Haberl, R., Becker, A., Leber, A., Knez, A., Becker, C., Lang, C., Bruening, R., Reiser, M., Steinbeck, G.: Correlation of coronary calcification and angiographically documented stenoses in patients with suspected coronary artery disease: results of 1,764 patients. *Journal of the American College of Cardiology* **37**(2) (2001) 451–457
- [25] Arad, Y., Spadaro, L., Goodman, K., Lledo-Perez, A., Sherman, S., Lerner, G., Guerci, A.: Predictive value of electron beam computed tomography of the coronary arteries: 19-month follow-up of 1173 asymptomatic subjects. *Circulation* **93**(11) (1996) 1951
- [26] Bateman, T.: Business aspects of cardiovascular computed tomography: tackling the challenges. *JACC Cardiovascular Imaging* **1**(1) (2008) 111
- [27] Corti, R., Fuster, V., Fayad, Z., Worthley, S., Helft, G., Smith, D., Weinberger, J., Wentzel, J., Mizsei, G., Mercuri, M., et al.: Lipid lowering by simvastatin induces regression of human atherosclerotic lesions: two years' follow-up by high-resolution noninvasive magnetic resonance imaging. *Circulation* **106**(23) (2002) 2884
- [28] Corti, R., Fuster, V., Fayad, Z., Worthley, S., Helft, G., Chaplin, W., Muntwyler, J., Viles-Gonzalez, J., Weinberger, J., Smith, D., et al.: Effects of Aggressive Versus Conventional Lipid-Lowering Therapy by Simvastatin on Human Atherosclerotic Lesions: A Prospective, Randomized, Double-Blind Trial With High-Resolution Magnetic Resonance Imaging. *Journal of the American College of Cardiology* **46**(1) (2005) 106–112
- [29] Ayaori, M., Momiyama, Y., Fayad, Z., Yonemura, A., Ohmori, R., Kihara, T., Tanaka, N., Nakaya, K., Ogura, M., Sawada, S., et al.: Effect of bezafibrate therapy on atherosclerotic aortic plaques detected by MRI in dyslipidemic patients with hypertriglyceridemia. *Atherosclerosis* **196**(1) (2008) 425–433
- [30] Wittman, J., Kok, F., van Saase, J., Valkenburg, H.: Aortic calcification as a predictor of cardiovascular mortality. *The Lancet* **2**(8516) (1986) 1120–2
- [31] Compagnone, G., Baleni, M., Pagan, L., Calzolaio, F., Barozzi, L., Bergamini, C.: Comparison of radiation doses to patients undergoing standard radiographic examinations with conventional screen-film radiography, computed radiography and direct digital radiography. *British Journal of Radiology* **79**(947) (2006) 899

- 
- [32] Lenchik, L., Rogers, L., Delmas, P., Genant, H.: Diagnosis of osteoporotic vertebral fractures: importance of recognition and description by radiologists. *American Journal of Roentgenology* **183**(4) (2004) 949
- [33] Silverman, S., Christiansen, C., Genant, H., Vukicevic, S., Zanchetta, J., de Villiers, T., Constantine, G., Chines, A.: Efficacy of Bazedoxifene in Reducing New Vertebral Fracture Risk in Postmenopausal Women With Osteoporosis: Results From a 3-Year, Randomized, Placebo-, and Active-Controlled Clinical Trial. *Journal of Bone and Mineral Research* **23**(12) (2008) 1923–1934
- [34] Black, D., Delmas, P., Eastell, R., Reid, I., Boonen, S., Cauley, J., Cosman, F., Lakatos, P., Leung, P., Man, Z., et al.: Once-yearly zoledronic acid for treatment of postmenopausal osteoporosis. *New England Journal of Medicine* **356**(18) (2007) 1809
- [35] Petersen, K., Ganz, M., Mysling, P., Lillemark, L., Crimi, A., Nielsen, M., Brandt, S.: A Bayesian Framework for the Automated Prediction of Cardiovascular Disease from Lumbar Standard Radiographs
- [36] De Bruijne, M., Nielsen, M.: Image segmentation by shape particle filtering. In: *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*. Volume 3., IEEE (2004) 722–725
- [37] Breiman, L.: Random forests. *Machine learning* **45**(1) (2001) 5–32
- [38] Petersen, K., Nielsen, M., Brandt, S.: A Static SMC Sampler on Shapes for the Automated Segmentation of Aortic Calcifications. *Computer Vision-ECCV 2010* (2010) 666–679
- [39] Pudil, P., Novovicova, J., Kittler, J.: Floating search methods in feature selection. *Pattern recognition letters* **15**(11) (1994) 1119–1125
- [40] Lillemark, L., Ganz, M., Barascuk, N., Dam, E., Nielsen, M.: Growth patterns of abdominal atherosclerotic calcified deposits from lumbar lateral X-rays. *The International Journal of Cardiovascular Imaging (formerly Cardiac Imaging)* (2010) 1–11
- [41] Wilson, K.: Problems in physics with many scales of length. *Scientific American* **241**(2) (1979) 140–57
- [42] Gidas, B.: A renormalization group approach to image processing problems. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**(2) (1989) 164–180
- [43] Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. *International journal of computer vision* **1**(4) (1988) 321–331
- [44] Cootes, T., Edwards, G., Taylor, C.: Active appearance models. *Computer Vision—ECCV’98* (1998) 484
- [45] Sethian, J.: *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*. 1999. Cambridge Monographs on Applied and Computational Mathematics (1999)
- [46] Greig, D., Porteous, B., Seheult, A.: Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society. Series B (Methodological)* **51**(2) (1989) 271–279
- [47] Hastings, W.: Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**(1) (1970) 97
- [48] Geman, S., Geman, D.: Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* **6** (1984) 721–741
- [49] De Backer, G., Ambrosioni, E., Borch-Johnsen, K., Brotons, C., Cifkova, R., Dallongeville, J., Ebrahim, S., Faergeman, O., Graham, I., Mancia, G., Cats, V., Orth-Gomer, K., Perk, J., Pyorala, K., Rodicio, J., Sans, S., Sansoy, V., Sechtem, U., Silber, S., Thomsen, T., Wood, D.: European guidelines on cardiovascular disease prevention in clinical practice. third joint task

- force of european and other societies on cardiovascular disease prevention in clinical practice (constituted by representatives of eight societies and by invited experts). *Atherosclerosis* **173** (2004) 381–391
- [50] Wilson, P., D'Agostino, R., Levy, D., Belanger, A., Silbershatz, H., Kannel, W.: Prediction of coronary heart disease using risk factor categories. *Circulation* **97** (1998) 1837–1847
- [51] Hopkins, P., Williams, R.: A survey of 246 suggested coronary risk factors. *Atherosclerosis* **40**(1) (1981) 1
- [52] Conrad-Hansen, L., de Bruijne, M., Lauze, F., Tanko, L., Nielsen, M.: Quantizing calcification in the lumbar aorta on 2-D lateral X-ray images. *Computer Vision for Biomedical Image Applications* (2005) 409–418
- [53] Conrad-Hansen, L., De Bruijne, M., Lauze, F., Tanko, L., Pettersen, P., He, Q., Chen, J., Christiansen, C., Nielsen, M.: Quantifying calcification in the lumbar aorta on X-ray images. In: *Proceedings of the 10th international conference on Medical image computing and computer-assisted intervention*, Springer-Verlag (2007) 352–359
- [54] Michael Kass, A.W., Terzopoulos, D.: Snakes: Active contour models. *International Journal of Computer Vision* **1**(4) (1988)
- [55] Cootes, T., Edwards, G., CJ, T.: Active appearance models. *Lecture Notes in Computer Science, Computer Vision ECCV 1998* **1407** (1998)
- [56] Osher, S., Sethian, J.: Fronts propagating with curvature dependent speed: Algorithms based on hamilton-jacobi formulations. *Journal of Computational Physics* **79** (1988) 12–49
- [57] van Ginneken, B., Stegmann, M., Loog, M.: Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database. *Medical Image Analysis* **10**(1) (2006) 19–40
- [58] Besag, J.E.: Discussion of a paper by P. Switzer. *Bull. Int. Stat. Inst.* **50**(3) (1983) 422–425
- [59] Loog, M., van Ginneken, B.: Supervised segmentation by iterated contextual pixel classification. *Pattern Recognition, 2002. Proceedings. 16th International Conference on* **2** (2002) 925–928
- [60] Krige, D.G.: A statistical approach to some basic mine valuation problems on the witwatersrand. *OR* **4**(1) (1953)
- [61] Berger, J.: *Statistical decision theory: foundations, concepts, and methods*. Springer-Verlag (1980)
- [62] Besag, J.: Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)* **32**(2) (1974) 192–236
- [63] Journel, A., Huijbregts, C.: *Floating search methods in feature selection*. Academic Press (1978)
- [64] Isaaks, E., Srivastava, R.: *Applied Geostatistics*. Oxford University Press (1989)
- [65] Pham, T.: Image restoration by ordinary kriging with convexity. In: *ICPR '00: Proceedings of the International Conference on Pattern Recognition, Washington, DC, USA, IEEE Computer Society* (2000) 3334
- [66] Bagger, Y., Tanko, L., Alexandersen, P., Hansen, H., Qin, G., Christiansen, C.: The long-term predictive value of bone mineral density measurements for fracture risk is independent of the site of measurement and the age at diagnosis: results from the Prospective Epidemiological Risk Factors study. *Osteoporos Int* **17**(3) (2006) 471–7
- [67] Tanko, L., Bagger, Y., et al.: Enlarged Waist Combined With Elevated Triglycerides Is a Strong Predictor of Accelerated Atherogenesis and Related Cardiovascular Mortality in Postmenopausal Women. *Circulation* **111**(15) (2005) 1883–1890

- 
- [68] Jaccard, P.: The distribution of the flora in the alpine zone. *New Phytologist* **11**(2) (1912) 37–50
- [69] Arya, S., Mount, D.M., et al.: An optimal algorithm for approximate nearest neighbor searching fixed dimensions. *J. ACM* **45**(6) (1998) 891–923
- [70] Koenderink, J.J., van Doorn, A.J.: Surface shape and curvature scales. *Image Vision Comput.* **10**(8) (1992) 557–565
- [71] Larsen, L., Ganz, M., Dam, E., Nielsen, M.: Growth pattern of atherosclerotic calcifications. Proceedings of the Second International Workshop on Computer vision for intracardiac and intravascular imaging, held in New York, USA, September 10, 2008 in conjunction with the 11th MICCAI conference (2008)
- [72] Lin, W., Hays, J., Wu, C., Liu, Y., Kwatra, V.: Quantitative evaluation of near regular texture synthesis algorithms. 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1 (CVPR'06) (2006)
- [73] Bruckstein, A., Donoho, D., Elad, M.: From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM review* **51**(1) (2009) 34–81
- [74] Zhang, T., Switzer, P., Journel, A.: Filter-based classification of training image patterns for spatial simulation. *Mathematical Geology* **38**(1) (2006) 63–80
- [75] Strebelle, S.: Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology* **34**(1) (2002) 1–21
- [76] Debevec, P., Taylor, C., Malik, J.: Modeling and rendering architecture from photographs: A hybrid geometry-and image-based approach. In: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, ACM (1996) 11–20
- [77] Criminisi, A., Perez, P., Toyama, K.: Object removal by exemplar-based inpainting. In: Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on. Volume 2., IEEE (2003)
- [78] Johnston, B., Atkins, M., Mackiewicz, B., Anderson, M.: Segmentation of multiple sclerosis lesions in intensity corrected multispectral MRI. *Medical Imaging, IEEE Transactions on* **15**(2) (2002) 154–169
- [79] Isgum, I., Rutten, A., Prokop, M., van Ginneken, B.: Detection of coronary calcifications from computed tomography scans for automated risk assessment of coronary artery disease. *Medical physics* **34** (2007) 1450
- [80] Bolland, M., Wang, T., van Pelt, N., Horne, A., Mason, B., Ames, R., Grey, A., Ruygrok, P., Gamble, G., Reid, I.: Abdominal aortic calcification on vertebral morphometry images predicts incident myocardial infarction. *Journal of Bone and Mineral Research* **25**(3) (2010) 505–512
- [81] Mairal, J., Bach, F., Ponce, J., Sapiro, G., Zisserman, A.: Discriminative learned dictionaries for local image analysis. In: Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, IEEE (2008) 1–8
- [82] Duda, R., Hart, P., Stork, D.: Pattern classification. Volume 2. Citeseer (2001)
- [83] Ganz, M., Nielsen, M., Brandt, S.: Patch-Based Generative Shape Model and MDL Model Selection for Statistical Analysis of Archipelagos. *Machine Learning in Medical Imaging* (2010) 34–41
- [84] Rissanen, J.: Minimum-Description-Length Principle
- [85] Rissanen, J.: MDL denoising. *IEEE Transactions on Information Theory* **46**(7) (2000) 2537–2543
- [86] Rissanen, J.: Fisher information and stochastic complexity. *Information Theory, IEEE Transactions on* **42**(1) (2002) 40–47

- [87] Wang, L., Liu, J.: Texture classification using multiresolution Markov random field models. *Pattern recognition letters* **20**(2) (1999) 171–182
- [88] Cohen, J.: A coefficient of agreement for nominal scales. *Educational and psychological measurement* **20**(1) (1960) 37–46
- [89] de Bruijne, M.: Shape particle guided tissue classification. *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis (MMBIA)* (2006) 64–64
- [90] Lauze, F., de Bruijne, M.: Toward automated detection and segmentation of aortic calcifications from radiographs. *Medical Imaging 2007: Image Processing* **6512**(1) (2007) 651239
- [91] Besag, J.: On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society. Series B (Methodological)* **48**(3) (1986) 259–302
- [92] Pearl, J., Shafer, G.: *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Morgan Kaufmann San Mateo, CA (1988)
- [93] Kopf, J., Uyttendaele, M., Deussen, O., Cohen, M.: Capturing and viewing gigapixel images. *ACM Transactions on Graphics* **26**(3) (2007) 93
- [94] Perez, P., Heitz, F.: Restriction of a Markov random field on a graph and multiresolution image analysis. *Rapport de recherche - Institut national de recherche en informatique et en automatique* (1994)
- [95] Kadanoff, L.P.: Scaling laws for ising models near  $t_c$ . *Physics* **2** (1966) 263
- [96] Gell-Mann, M., Low, F.E.: Quantum electrodynamics at small distances. *Physical Review* **95** (1954) 1300
- [97] Wilson, K.G.: The renormalization group: Critical phenomena and the kondo problem. *Reviews of Modern Physics* **47**(4) (1975) 773
- [98] McCoy, B., Wu, T.: *The two-dimensional Ising model*. Harvard University Press Cambridge (1973)
- [99] Besag, J.: Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)* **36**(2) (1974) 192–236
- [100] Gidas, B.: Efficient methods for computing coarse-grid clique parameters for the renormalization group algorithm. Technical report, Division Applied Mathematics, Brown University (1987)
- [101] Nicholls, G., Petrou, M.: A generalization of the renormalization group method for multiresolution image analysis. *Proceedings of the 11th International Conference on Pattern Recognition* **1** (1992) 567–570
- [102] Petrou, M., Nicholls, G.: On multiresolution image restoration. *Proceedings of the 12th IAPR International Conference on Signal Processing* **3** (1994) 63–67
- [103] Petrou, M.: Accelerated optimization in image processing via the renormalization group transformation. In Titterton, D., ed.: *Complex Stochastic Systems and Engineering*. Clarendon Press (1995) 105–120
- [104] Blake, A., Rother, C., Brown, M., Perez, P., Torr, P.: Interactive image segmentation using an adaptive gmmrf model. *Computer Vision-ECCV 2004* (2004) 428–441
- [105] Rother, C., Kolmogorov, V., Blake, A.: Grabcut: Interactive foreground extraction using iterated graph cuts. In: *ACM SIGGRAPH 2004 Papers*, ACM (2004) 314
- [106] Lombaert, H., Sun, Y., Grady, L., Xu, C.: A multilevel banded graph cuts method for fast image segmentation. In: *Tenth IEEE International Conference on Computer Vision, 2005. ICCV 2005. Volume 1*. (2005)
- [107] Sinop, A., Grady, L.: Accurate banded graph cut segmentation of thin structures using laplacian pyramids. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2006* (2006) 896–903

- 
- [108] Kohli, P., Lempitsky, V., Rother, C.: Uncertainty Driven Multi-scale Energy Minimization. Microsoft Research Technical Report (April 2010)
- [109] Niemeijer, T., Van Leeuwen, J.: Wilson theory for spin systems on a triangular lattice. *Physical Review Letters* **31**(23) (1973) 1411–1414
- [110] Kadanoff, L.P.: *Statistical Physics: Statics, Dynamics and Renormalization*. World Scientific (1999)
- [111] Migdal, A.: Recursion equations in gauge field theories. 30 years of the Landau Institute: selected papers **69** (1996) 114
- [112] Kato, T.: *Perturbation theory for linear operators*. Springer Verlag (1995)
- [113] Bronstein, I., Semendjajew, K., Musiol, G., Muehlig, H.: *Taschenbuch der Mathematik*. Harri Deutsch Verlag (2008)
- [114] The Mathworks: `imresize`. <http://www.mathworks.com/access/helpdesk/help/toolbox/images/imresize.html>
- [115] Goldenfeld, N.: *Lectures On Phase Transitions And The Renormalization Group*. Addison-Wesley (1993)
- [116] Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2004) 1124–1137
- [117] European Heart Network (EHN): European cardiovascular disease statistics 2008. <http://www.ehnheart.org/cdv-statistics.html> (2008)
- [118] American Heart Association (AHA): Cardiovascular disease statistics. <http://www.americanheart.org/presenter.jhtml?identifier=4478> (2008)
- [119] Jacoby, D., Mohler, E., Rader, D.: Noninvasive atherosclerosis imaging for predicting cardiovascular events and assessing therapeutic interventions. *Current Atherosclerosis Reports* **6**(1) (2004) 20–26
- [120] Ouhloud, M., Lethimonnier, F., Dippel, D., van Sambeek, M., van Heerebeek, L., Pattynama, P., van der Lugt, A.: Evaluation of a dedicated dual phased-array surface coil using a black-blood FSE sequence for high resolution MRI of the carotid vessel wall. *Journal of Magnetic Resonance Imaging* **15**(3) (2002) 344–351
- [121] Serra, J.: *Image Analysis and Mathematical Morphology*. Academic Press (1982)
- [122] De Weert, T., De Monye, C., Meijering, E., Booij, R., Niessen, W., Dippel, D., Van Der Lugt, A.: Assessment of atherosclerotic carotid plaque volume with multidetector computed tomography angiography. *The International Journal of Cardiovascular Imaging (formerly Cardiac Imaging)* **24**(7) (2008) 751–759
- [123] Cai, J., Hatsukami, T., Ferguson, M., Kerwin, W., Saam, T., Chu, B., Takaya, N., Polissar, N., Yuan, C.: In vivo quantitative measurement of intact fibrous cap and lipid-rich necrotic core size in atherosclerotic carotid plaque: comparison of high-resolution, contrast-enhanced magnetic resonance imaging and histology. *Circulation* **112**(22) (2005) 3437
- [124] Kendall, M., Smith, B.: The problem of m rankings. *The Annals of Mathematical Statistics* **10**(3) (1939) 275–287
- [125] Bull, K., Spiegelhalter, D.: Tutorial in biostatistics survival analysis in observational studies. *Statistics in Medicine* **16**(9) (1997) 1041–1074
- [126] Collett, D.: *Modelling Survival Data in Medical Research*. Chapman and Hall (2003)
- [127] Madjid, M., Zarrabi, A., Litovsky, S., Willerson, J., Casscells, W.: Finding vulnerable atherosclerotic plaques: Is it worth the effort? *Arteriosclerosis, Thrombosis, and Vascular Biology* **24**(10) (2004) 1775

- [128] Kavurma, M., Bhindi, R., Lowe, H., Chesterman, C., Khachigian, L.: Vessel wall apoptosis and atherosclerotic plaque instability. *Journal of Thrombosis and Haemostasis* **3**(3) (2005) 465–472
- [129] Ohwaki, K., Bujo, H., Jiang, M., Yamazaki, H., Schneider, W., Saito, Y.: A secreted soluble form of LR11, specifically expressed in intimal smooth muscle cells, accelerates formation of lipid-laden macrophages. *Arteriosclerosis, thrombosis, and vascular biology* **27**(5) (2007) 1050
- [130] Karsdal, M., Henriksen, K., Leeming, D., Mitchell, P., Duffin, K., Barascuk, N., Klickstein, L., Aggarwal, P., Nemirovskiy, O., Byrjalsen, I., et al.: Biochemical markers and the FDA Critical Path: how biomarkers may contribute to the understanding of pathophysiology and provide unique and necessary tools for drug development. *Biomarkers* **14**(3) (2009) 181–202
- [131] Parikh, N., Hwang, S., Larson, M., Cupples, L., Fox, C., Manders, E., Murabito, J., Massaro, J., Hoffmann, U., O'Donnell, C.: Parental occurrence of premature cardiovascular disease predicts increased coronary artery and abdominal aortic calcification in the Framingham Offspring and Third Generation cohorts. *Circulation* **116**(13) (2007) 1473
- [132] Wu, M., Chern, M., Chen, L., Lin, Y., Sheu, M., Liu, J., Chang, C.: Electron beam computed tomography evidence of aortic calcification as an independent determinant of coronary artery calcification. *Journal of the Chinese Medical Association* **69**(9) (2006) 409–414
- [133] Schousboe, J., Claffin, D., Barrett-Connor, E.: Association of Coronary Aortic Calcium With Abdominal Aortic Calcium Detected on Lateral Dual Energy X-Ray Absorptiometry Spine Images. *The American journal of cardiology* **104**(3) (2009) 299
- [134] Reaven, P., Sacks, J.: Coronary artery and abdominal aortic calcification are associated with cardiovascular disease in type 2 diabetes. *Diabetologia* **48**(2) (2005) 379–385
- [135] Allison, M., Pavlinac, P., Wright, C.: The differential associations between HDL, non-HDL and total cholesterol and atherosclerotic calcium deposits in multiple vascular beds. *Atherosclerosis* **194**(2) (2007) 87–94
- [136] Schousboe, J., Taylor, B., et al.: Abdominal aortic calcification detected on lateral spine images from a bone densitometer predicts incident myocardial infarction or stroke in older women. *Journal of Bone and Mineral Research* **23**(3) (2008) 409–416 PMID: 17956153.
- [137] Shi, Y., Witte, R., O'Donnell, M.: Identification of vulnerable atherosclerotic plaque using IVUS-based thermal strain imaging. *Ultrasonics, Ferroelectrics and Frequency Control, IEEE Transactions on* **52**(5) (2005) 844–850
- [138] Ganz, M., de Bruijne, M., Nielsen, M.: MACD: an imaging marker for cardiovascular disease. *Proceedings of SPIE* **7624** (2010) 76240P
- [139] Jacobs, P., Prokop, M., van der Graaf, Y., Gondrie, M., Janssen, K., de Koning, H., Isgum, I., van Klaveren, R., Oudkerk, M., van Ginneken, B., et al.: Comparing coronary artery calcium and thoracic aorta calcium for prediction of all-cause mortality and cardiovascular events on low-dose non-gated computed tomography in a high-risk population of heavy smokers. *Atherosclerosis* **209**(2) (2010) 455–462
- [140] Genant, H., Wu, C., van Kuijk, C., Nevitt, M.: Vertebral fracture assessment using a semi-quantitative technique. *Journal of bone and mineral research* **8**(9) (1993) 1137–1148
- [141] O'Donnell, C., Chazaro, I., Wilson, P., Fox, C., Hannan, M., Kiel, D., Cupples, L.: Evidence for heritability of abdominal aortic calcific deposits in the Framingham Heart Study. *Circulation* **106**(3) (2002) 337
- [142] Wolffe, J., Siegal, E.: X-ray of the abdominal aorta in detection of atherosclerosis. *Clinical Medicine* **69** (1962) 401
- [143] Candes, E., of Statistics, S.U.D.: Ridgelets: theory and applications. Stanford University (1998)
- [144] Donoho, D.: Wedgelets: Nearly minimax estimation of edges. *The Annals of Statistics* **27**(3) (1999) 859–897

- [145] Pedersen, J., Ashraf, H., Dirksen, A., Bach, K., Hansen, H., Toennesen, P., Thorsen, H., Brodersen, J., Skov, B., Dossing, M., et al.: The Danish randomized lung cancer CT screening trial-overall design and results of the prevalence round. *Journal of Thoracic Oncology* 4(5) (2009) 608



## Acknowledgments

I owe a great deal to my co-authors, colleagues, friends and members of my family who have accompanied me during the work on this dissertation, and who, through their own research, comments and questions have encouraged and supported me.

Foremost, I would like to thank my supervisors, Sami S. Brandt, Marleen de Bruijne and Mads Nielsen, who have shared their knowledge and ideas about image analysis with me during the last three years. I would also like to acknowledge the debt I owe to all my colleagues at Synarc Imaging Technology, especially to Kersten Petersen and Lene Lillemark, who have worked together with me on the CVD pipeline, and Natasha Barascuk, who has taught me a lot about the biology of CVDs. Thanks especially to Lene, who read my dissertation during the draft stages and gave me a number of great comments to improve it. And since I also was employed at the University of Copenhagen I owe thanks to my colleagues there, particularly Jon Sparring, Aasa Feragen, Kenny Erleben, Francois Lauze and Marco Loog, for numerous interesting discussions about my academic work now and in the future. Three months of my Ph.D. I spent at Microsoft Research Laboratory in Cambridge, UK, and I want to express my gratitude to the people I was allowed to work with there, especially Pushmeet Kohli and Sebastian Nowozin.

In the end I want to thank all my friends, in particular Mads Danker Danielsen and Franziska Meier who have helped me to improve this dissertation through a view from the outside. Additionally, a huge thanks goes to my family for supporting me from far away. Without their trust and support I would not have made it this far. Finally, I have to say thanks to Claus, who had to live in the same apartment with me while I was working on and writing this dissertation and who has supported me in everything I chose to do.

My apologies if I have inadvertently omitted anyone to whom acknowledgment is due. Without doubt there will be errors left in this work, for which I take absolute responsibility, but I hope that the dissertation nevertheless can inspire some people to take a look at CVD risk assessment in X-rays.

