Institut für Signalverarbeitung und Prozessrechentechnik
Universität zu Lübeck
Direktor: Prof. Dr.-Ing. Alfred Mertins

Cardiovascular Biomedical Image Analysis: Methods and Applications

Inauguraldissertation
zur Erlangung der Doktorwürde (Dr.-Ing.)
der Universität zu Lübeck
– Technisch-Naturwissenschaftliche Fakultät –

vorgelegt von
Diplom-Ingenieur
Alexandru Paul Condurache
aus Lübeck

Lübeck, im November 2006

Referent: Priv.-Doz. Dr.-Ing. Erhardt Barth
Koreferent: Prof. Dr.-Ing. Til Aach
Vorsitz des Prüfungsausschusses: Prof. Dr.-Ing. Erik Maehle
Tag der mündlichen Prüfung: 23. Februar 2007

gezeichnet: Prof. Dr.rer.nat. Enno Hartmann
– Dekan der Technisch-Naturwissenschaftlichen Fakultät der Universität zu Lübeck –

Cardiovascular Biomedical Image Analysis: Methods and Applications

Alexandru Paul Condurache

Preface

My interest in seeing and thinking machines was ignited mainly during my high school years by novels such as those written by Isaac Asimov and Arthur C. Clark, which however, found the ground already prepared by Jules Verne. My study at the "Electronica" Faculty of the "Polithenica" University of Bucharest was then the consequence. The passage from science-fiction to something more "earthly" happened during the last years as a student, while I was specializing in "Imagini, Forme si Inteligenta Artificiala (IFIA)"[1] under the guidance of my professor Vasile Buzuloiu and the group he lead at the "Laboratorul de Analiza si Prelucrarea Imaginilor (LAPI)"[2]. From this group – to whom my gratitude is directed as a whole – I would like to mention Dr. Mihai Ciuc and Dr. Constantin Vertan, which found the best way to make us students the acquaintance with Image Processing and other related fields in a clear and enjoyable way.

During my activity at LAPI I met Dr. Erhardt Barth, who gave me the possibility to prepare my diploma thesis at the Institute for Signal Processing of the then Medical University of Lübeck, within a project related to vision-based quality control. He has my gratitude and full appreciation for being not only a teacher but also a very good friend.

During my time in Lübeck – an excellence center in medical technologies – I was first acquainted with medical image processing, an acquaintance which shaped my career from then on trough the additional specialization year in medical imagery, again at the "Politehnica" University and then during my later stay as a research associate at the Institute for Signal Processing (ISIP) of the University of Lübeck. This doctoral thesis has grown in a natural way from my interest in the field. It benefited largely from the physical proximity between ISIP and the Lübeck campus of the University Clinic of Schleswig-Holstein (UKSH), as well as from the innovative and open spirit of its medical staff, from which I would like to mention those with whom I collaborated more closely during the last years: Dr. Stephan Grzybowsky, Dr. Peter Radke, Dr. Axel Kaiser, Dr. Martin Misfeld and especially Dr. Hans-Günther Machens.

My special gratitude goes to Professor Til Aach, head of ISIP until 2004 and the main "guiding-light" during my time spent researching for this thesis. The numerous discussions we had as well as his critical, well-balanced points have been a constant source of scientific inspiration just as much as his basketball skills, proven during the many games we've played together, have helped establishing a relaxed, friendly and therefore – in my opinion – optimal working atmosphere.

Towards my colleagues and friends at ISIP: Andre Folkers, Uli Hofmann, Christian Kier, Kerstin Menne, Volker Metzler, Ingo Stuke and Daniel Toth goes my thanks not only for the scientific debates we held together, but also for teaching me german and then enduring my

---

[1]Images Patterns and Artificial Intelligence
[2]The Image Processing and Analysis Laboratory (IPAL)

endless speeches with stoicism. I also thank them for making me feel very well during my entire stay in Germany until now.

For their support and understanding I thank my family: my parents Anca and Serban Condurache who were there for me whenever I needed them and most specially to my wife Andreea whose beautiful smile was always near to provide me with the emotional support I looked for.

I would like to end this Preface with my personal conclusion for these last five years: *"Small opportunities are often the beginning of great enterprises"* – Demosthenes

To my Family

Contents

Abstract

The medical field as a whole benefited largely during the past decades from the introduction and large-scale availability of various non-invasive imaging methods. The introduction of computers provided the ground on which seeing machines appeared to assist the medical staff at various tasks. The realization of such machines is characterized by a strong interdisciplinarity and algorithmic complexity and therefore their true potential surfaced only recently, together with the appearance of more powerful hardware.

Cardiovascular medicine was one of the first areas of application for non-invasive imaging methods: X-ray angiograms were used to enhance diagnostic and evaluation capabilities. Therefore the rate of acceptance for a machine-vision-enhanced routine is relatively high among physicians in the field.

This work profits from such synergies and brings together machine-vision algorithms and research with cardiovascular applications, and can be therefore considered an attempt to establish or – following a more conservative rhetoric – to cement a new branch of medical image analysis: that of cardiovascular biomedical image analysis. This work is not however a plea for replacing the human specialist by a machine, but rather a prove that machine-vision support can enhance the abilities of physicians to the benefit of the patient.

We concentrate on non-invasive 2D imaging methods, which are largely used in cardiovascular medicine. Starting from a specific application, machine-vision-based solutions are searched and novel methods are developed – when needed – to solve the problem at hand.

We describe how to improve the navigation abilities of an operating physician during Percutaneous Transluminal Coronary Angioplasty (PTCA) by presenting him a dynamic roadmap of the coronary vessel tree instead of the static one used until now. The dynamic roadmap moves synchronous with the heart and the surgical tools, like e.g., guidewire tip and balloon are superimposed. We also show how to measure automatically the Myocardial Blush Grade (MBG) which is used to estimate revascularization of the cardiac muscle. The MBG helps evaluate the success of a PTCA which was performed previously as treatment to a heart attack.

We develop novel automatic and semiautomatic vessel segmentation algorithms. The main purpose of segmentation is here the quantification of measures of the investigated vasculature, like e.g., the area and length. Such measures applied, e.g., to the retina vessel-tree, help prevent blindness by facilitating a timely diagnostic of the diabetic retinopathy. In the case of transplanted skin tissue, they are used, e.g., in medical research to facilitate the investigation of the effects of different angiogenesis-inducing drugs. In the case of peripheral angiography vessel segmentation is useful during the diagnosis phase where it provides the grounds for, e.g., enhanced blood-flow analysis.

We describe also how to compute automatically new quality measures mainly for tricuspid xenograft heart-valve implants. The examination of heart-valve implants is done preoperatively in a dedicated test-bed. The purpose of such quality control is to ensure that only the best

possible implants reach the patient.

In this work, we describe machine-vision-based supportive technologies for the cardiovascular medicine. We introduce algorithmic solutions spanning almost the entire spectrum of domains, starting from research to medical routine and within this routine from the preventive to the interventional medicine including the diagnostic, preoperative, operative and evaluation phases.

Zusammenfassung

In den vergangenen Jahrzehnten hat die gesamte Medizin enorm von der Einführung und breiten Verfügbarkeit nicht-invasiver Bildgebungsmethoden profitiert. Die Entwicklung von Computern ermöglichte es, sehende Maschinen zu konstruieren, die das medizinische Personal bei einer Vielfalt von Aufgaben unterstützen. Systeme dieser Art setzen eine hohe Interdisziplinarität der Entwicklung voraus und sind im Allgemeinen sowohl algorithmisch als auch in der Rechenzeit sehr anspruchsvoll. Das wahre Potenzial dieser Methoden kommt deshalb erst in letzter Zeit unter anderem durch den Fortschritt in der Hardwareentwicklung zum Vorschein.

Im Bereich der kardiovaskulären Medizin kam mit der Röntgen-Angiografie eine der ersten nicht-invasiven Bildgebungsmethoden zum Einsatz. Die damit arbeitenden Mediziner haben dieses Verfahren schätzen gelernt und als selbstverständlich akzeptiert, so dass die Akzeptanz einer maschinengestützten klinischen Routine in diesem Feld relativ hoch ist.

Diese Arbeit nutzt solche Synergieeffekte durch die Vereinigung von Algorithmen des maschinellen Sehens (machine vision) mit kardiovaskulären Anwendungen. Sie kann deshalb als Ansatz aufgefaßt werden, das neue Feld der kardiovaskulären biomedizinischen Bildanalyse zu zementieren. Das Ziel dieser Arbeit ist dabei nicht, den meschlichen Spezialisten durch eine Maschine überflüssig zu machen, sondern zu zeigen, dass die Möglichkeiten des Arztes durch die Unterstützung mit maschinellem Sehen erweitert werden – zum Vorteil des Patienten.

Der Fokus liegt auf nicht-invasiver 2D Bildgebung, die in der kardiovaskulären Medizin häufig genutzt wird. Beginnend mit einer spezifischen Anwendung werden machine-vision-basierte Methoden gefunden und neue Verfahren entwickelt, um das vorliegende Problem zu lösen.

In der Arbeit wird eine dynamische Karte des koronaren Gefäßsystems vorgestellt, der die Navigationsmöglichkeiten des operierenden Arztes während einer perkutanen transluminalen Koronarangioplastie (PTCA) im Vergleich zu dem bisher genutzten statischen Karte erheblich verbessert. Die dynamische Gefäßkarte bewegt sich synchron zum Herz während die chirurgischen Werkzeuge überlagert werden. Des Weiteren wird eine automatische Bestimmung des Myocardial Blush Grade (MBG) erreicht. Der MBG ist ein Maß für die (veränderte) myokardiale Durchblutung und wird genutzt, um den Erfolg einer PTCA als Behandlung eines Herzinfarkts zu beurteilen.

Ein weiterer Schwerpunkt der Arbeit liegt auf automatischen und semiautomatischen Gefäßsegmentierungsalgorithmen mit dem Ziel der Quantifizierung verschiedener Gefäßeigenschaften, wie z.B. Fläche oder Länge. Ein Anwendungsfall dieser Algorithmen ist z.B. die Untersuchung des Gefäßbaums der Netzhaut, wo es darum geht die Erblindung eines Diabetes-Patienten durch eine frühzeitige Diagnostik der Retinopathia diabetica zu verhindern. Für den Fall des transplantierten Hautgewebes helfen diese Algorithmen z.B. der medizinischen Forschung, um den Einfluss verschiedener Medikamente auf die Angiogenese zu bestimmen. Bei der peripheren Angiographie unterstützt die Gefäßsegmentierung die Diagnostik z.B. durch

verbesserte Blutflussanalyse.

Darüber hinaus wird in dieser Arbeit die automatische Berechnung neuer Qualitätsmerkmale hauptsächlich für von einem artfremden Spender stammende Herzklappenimplantate vorgestellt. Die Untersuchung der Implantate erfolgt präoperativ in einer speziellen Testumgebung zu dem Zweck, einem Patienten nur die bestmöglichen Implantate einzusetzen.

In dieser Arbeit, werden machine-vision spezifische, algorithmische Lösungen für die Unterstützung des Arztes in der kardiovaskulären Medizin vorgestellt. Dabei werden alle Anwendungsbereiche einbezogen, beginnend von der Forschung bis zur klinischen Routine und innerhalb der Routine von der präventiven bis zur interventionellen Medizin. Dies schließt die diagnostische, präoperative, operative sowie die Evaluierungsphase mit ein.

Chapter 1

Introduction

"Never say »it can't be done«
but begin with »lets's see about this«"
– Nicolae Iorga

Humans have invented machines to assist and support them in their work and daily life. While trying to construct better ones it became quickly clear that they should be given the possibility to adapt to the task they need to accomplish, i.e., they should be able to receive and process information. As a large part of the information a human processes comes as a visual input, the acquiring and interpretation of images – i.e. seeing – is central to a multitude of human activities and therefore seeing machines are needed. First the introduction of computers made such machines possible.

Machine vision has since then evolved beyond human vision in some domains, like e.g., sensory input. Machine vision systems are able with the help of some dedicated hardware to extract information about a scene not only by projection from 3D to 2D of items that reflect or absorb some portions of the visible spectrum of the electromagnetic radiation, but also from other parts of this spectrum, as well as information transmitted otherwise than by electromagnetic waves – e.g. ultrasound imaging – and from scenes where items are separated by other properties than the reflection/absorption of some radiation.

However, irrespective of the way image information about a scene is sensed, the perception of such information, which is the main goal of vision, is still related to the description and interpretation of objects and in this respect machine vision still has something to learn from its human counterpart.

For the purposes of this work, we define a seeing machine as one which is able to emulate the human vision for certain tasks. Therefore, such a machine can receive images, i.e., 2D-representations of scenes, and return an analysis of these images.

1.1   Biomedical cardiovascular image analysis

Seeing machines are currently being used to support the human activity in a multitude of domains from quality inspection [187] to medical image analysis [55]. Presently they can successfully replace humans in a multitude of tasks usually involving repetitive processes, but they can only support them at tasks requiring higher adaptability.

1

The medical field in particular requires a highly adaptive and practically infallible approach. This is why in this case, their main role is to assist the medical staff usually by taking away from them painstaking repetitive work and therefore allowing them to concentrate on integrating information and taking decisions. Usually, in this case, the task of such machines is to provide an analysis of the images and sequences of images acquired, thereby, e.g., enhancing the human visual perception of digital medical images and improving the human-analysis of medical data.

Since its beginnings some three decades ago, medical image analysis has grown – particularly in the last years – from being a collection of general image-processing algorithms applied to an interesting data-set to being a discipline in its own right [73] defined over specific imaging methods [40] and a particular problematic which clearly needs particular solutions.

The cardiovascular medicine, which handles with the cardiovascular system comprising the heart and the blood vessels, makes no exception from this rule. Computer-based supportive systems are already encountered in many areas related to the field covering almost all aspects from research and diagnostic to intervention and postoperative evaluation. Being convinced that such seeing machines are able to enhance the capabilities of physicians and researcher in the field to the benefit of the patient, the main purpose of this work is to introduce such systems to new areas of application. For this purpose, new algorithms have been developed when needed and already established methods have been adapted to solve the particular problem at hand.

Cardiovascular image analysis represents currently an active research field within the larger frame of medical image analysis. This is proved by the large number of publications in different dedicated conferences, workshops and journals [82].

1.2 An enhanced catheter intervention

The coronary artery disease (CAD) is defined as the buildup of plaque on the inner walls of the arteries supplying blood to the heart-muscle and one of its most serious consequences is a myocardial infarction. This occurs if the plaque ruptures and suddenly blocks completely a coronary artery. Treatment of the CAD and of the myocardial infarction is facilitated by a catheter intervention, i.e., the Percutaneous Transluminal Coronary Angiography (PTCA). The purpose of PTCA is to relieve the impaired artery.

Our aim is to help the physician during PTCA by making it easier for him to navigate through the coronary vessel tree so that it proceeds faster and safer. Presently for navigation a single image showing the entire vessel tree – i.e. a complete coronary angiogram – is displayed on a screen next to the one showing live images. This vessel roadmap is inherently static and inaccurate. In this thesis we have developed algorithms to show the physician a dynamic roadmap based on complete coronary angiograms, and where his tools are superimposed. We have already published and patented some of the aspects related to the dynamic roadmap in [6], [48], [50], and [47], [49] respectively. We intent to present an overview of the method in [46].

If PTCA is performed as treatment for myocardial infarction, it is important for the evaluation of the postoperative evolution of the patient to establish how much (if any) blood is reaching the heart muscle after the artery has been reopened. For this purpose a dedicated measure is used: the Myocardial Blush Grade (MBG). The MBG estimates the darkening – caused by the presence of blood with contrast agent in the heart tissue – of a target region in X-ray images acquired in the end-phase of a catheter intervention.

Currently, the MBG is established by a visual analysis of this darkening. Within this work

we show how to automatize this procedure, thereby improving from a what in clinical routine is termed a "semi-quantitative" analysis to a "quantitative" analysis with less inter- and intra-observer variability. We have described some aspects related to the automatic estimation of the MBG in [54].

1.3 Vessel segmentation and quantification in 2D-projection images

Vessel imaging and analysis is required in a multitude of medical applications. For the imaging part, 2D projection imaging such as angiography or retinal photography is often the method of choice as it returns results of sufficient quality without the overhead required by 3D techniques. To analyze the vessel images, in many cases they need to be segmented, e.g., as a prerequisite for the quantification of the observed vessel structures [52] or as integrating part of a more complex computer-supported analysis of the image data [176].

Here we describe novel general-purpose, automated vessel-segmentation methods, which typically reach a solution in a two step approach: first the vessels are enhanced and then they are segmented. We introduce here also the hysteresis classification paradigm with application to vessel segmentation. This yields several new parametric, unsupervised and supervised classification methods, some of which we have already published in [53], [43] and [44]. Although these algorithms are introduced here for vessel segmentation we believe that their applicability extends over this initial problematic. A detailed approach to hysteresis classification is intended for publication in [45].

To evaluate the effectiveness of different angiogenesis-inducing substances, the surface and length of newly grown vessels has to be quantified in microangiograms. Microangiograms are X-ray images of blood vessels of a diameter less than $20\mu m$ of tissue transplanted on the back of laboratory animals.

In this work we introduce a novel semi-automatic framework, designed specifically to precisely segment vessels in microangiograms but also in other types of 2D-projection vessel-images for the purpose of quantifying their surface and length. Within this framework both human operator and machine are given the roles they can play best: the machine has the tedious work of segmenting all vessel structures pixel-wise such that no vessel is missed and shows the human operator a segmentation proposition, which he can then accept or modify at a larger scale until it fits its expectations. We have published some aspects related to the semi-automatic vessel segmentation framework in [52], [51].

1.4 Quality control for heart valves

Patients suffering from a heart valve deficiency are often treated by replacing the ill valve with artificial or biological implants. Particularly in the case of biological implants, their quality is first evaluated such as to ensure an optimal postoperative evolution of the patient. That implant which fulfills in an optimal way a set of quality criteria is then chosen for transplant. Currently, this evaluation is done mainly by visual inspection by trained personnel, using a movie showing the evolution of the valve over a heart beat in an operatory. Usually, only the most important quality criterion is estimated using some measures which are computed manually.

Typically a relatively large number of such biological implants is available and needs to be inspected. Slow manual quality measurement means that the implant is chosen from a only small number of candidates and potentially does not represent a real optimum. Also, manual measurements are hardly reproducible which renders comparisons between current and previously inspected valves rather difficult. This represents at the same time a brake in the path of research as, e.g., a comparative study on the postoperative evolution of some patients can be conducted only if all received implants are of similar quality.

In this work we describe how to measure automatically several quality criteria therefore supporting an enhanced analysis. We also provide more accurate and reproducible results and at the same time relieve the human operator of the tedious work of estimating quality measures thus allowing him to concentrate on choosing among several candidate valves. Automatic quality measurements in comparison to manual ones, permit the evaluation of a larger number of xenograft implants in the same period of time, such that the chances increase of discovering the truly best valve. A few initial aspects related to the automatic measuring of quality criteria for heart valves have been published in [95]. A more detailed article was published in [57]. A final contribution is intended for publication in [56].

Chapter 2

Machine vision for cardiovascular medicine: a tutorial

*"Everything should be made
as simple as possible but not simpler."*
– Albert Einstein

Intuitively vision deals with the acquisition and interpretation of images. Therefore, to construct an artificial – as in not-human – vision, specifically designed for the needs of cardiovascular medicine, one needs some dedicated hardware and software. Some of the hardware deals with the acquisition of images[1] and some, together with the software, deals with the interpretation of images.

Here we describe the hardware needed to sense the images such that its limitations and the distortions which it introduces in the acquired-images are clear making thus in turn clear the need for specific algorithmic-steps to deal with these problems (Section 2.1 and Section 2.2).

At the same time we introduce and localize the classes of algorithms which make such cardiovascular machine vision possible (Section 2.3 and Section 2.4) and show how to file a valid algorithm-performance claim (Section 2.5).

2.1   From human to machine-vision

Human vision.   Humans apprehend the surrounding reality by sampling the environment with the senses and processing the received information. Once aware of the surrounding reality one can act or react to serve his own purposes. Of the five senses we have at our disposal – i.e. vision, hearing, smell, taste and touch – vision is by far the most important bringing the largest amount of information [194].

As visible light can pass largely unattenuated through the Earth's atmosphere, it is the main source of energy near the surface of Earth. From an evolutionary point of view it is a very interesting coincidence that many organisms living in such an environment and mammals in particular have developed over the time the ability to sense[2] and perceive[3] it, i.e., vision[4].

---

[1]Here are discussed only those 2D projection imaging techniques which acquire data from electromagnetic radiations. For a discussion on other techniques see [75].

[2]The sensation is a passive process of bringing information into the brain.

[3]The perception is an active process of interpreting sensed-information.

[4]In a broader sense vision can be defined as the ability to sense and perceive electromagnetic radiation –
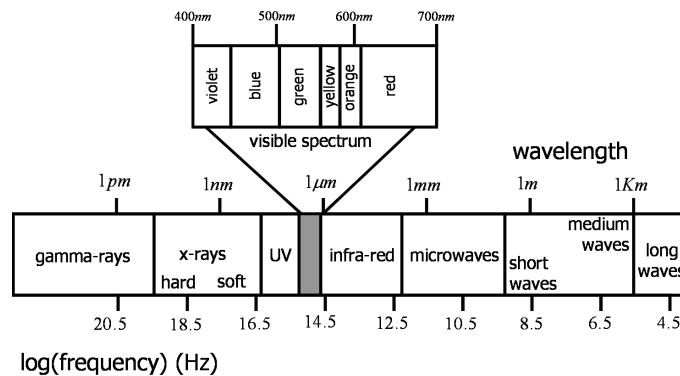
Figure 2.1: Visible light in the electromagnetic spectrum.

Light is one of the principal modalities by which the sun emits energy. Nuclear fusion reactions involving hydrogen, which take place in the core of the sun, heat its surface – i.e. the photosphere – to a temperature of around 6000K ($\approx$ 5700°C). Assuming the sun is a perfectly radiating (black) body, most of its energy will be emitted in a very small band of the electromagnetic spectrum with wavelength between 400 and 700nm. This portion of the electromagnetic spectrum is called light or more specific visible light (See Figure 2.1). It has been calculated that 40% of the sun's energy is emitted as visible light (wavelengths between 400 and 700nm), 50% is emitted on waves of a lower frequency, i.e., infrared (wavelengths of 750nm to 1 mm), 9% is emitted on waves of a higher frequency, i.e., ultraviolet (wavelengths of 10nm to 380nm) and 1% as X-rays (wavelength of 0.01nm to 10nm[5]).

The human visual system consists of eye including retina, optic nerve, optic chiasm, optic tract, lateral geniculate nucleus, optic radiations and the visual cortex (see Figure 2.2).

The information of the 3D scene which is viewed is projected by an optical system – the eye – on a light-sensitive 2D surface – the retina – so that rays emitted from a point in the scene are collected into one point in the retina, forming an image. An image can be thus defined as a 2D representation of a scene, and a scene as a collection of items to be viewed. Photons which travel along these rays are converted in the retina (by a chemical process) in electrical signals which are sent down the optical nerves, through the optic chiasm and the optic tract into the lateral geniculate nucleus. From here the visual information travels through the optic radiations to the visual cortex where is processed and presented in an abstract form to higher centers of the brain (see Figure 2.2).

The eye including the retina is the light sensing organ and although some processing of the visual information is done already in the retina, the process of visual perception takes place mainly in the lateral geniculate nuclei and especially in the layers of the cortex and it is not yet fully understood. Assuming that the visual reality consists of objects and background, i.e., items of high and low concern respectively, one of the main outputs of the visual system is the description of objects and their interpretation based on this description. Both this interpretation

---

typically in the band in which the sun radiates. For the purpose of giving an initial definition to machine vision – a human invention – a restriction to human vision is made which is called simply vision.

[5]Electromagnetic radiation of wavelength smaller than 0.1nm is called hard X-rays but also Gamma-rays. There is no physical difference between X-rays and Gamma-rays. Some sources differentiate them by the way they are produced: X-rays appear as a consequence of matter interaction with high speed electrons and Gamma rays as a consequence of nuclear transitions.
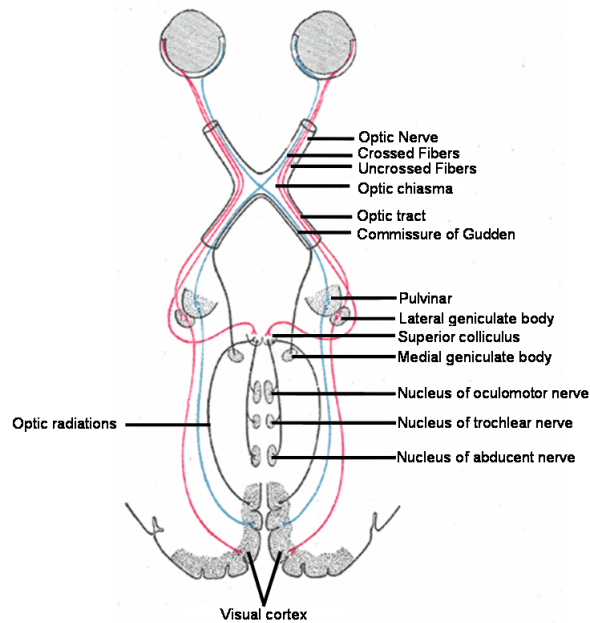
Figure 2.2: Schematic representation of the human visual system – after [91].

and the action/reaction that may follow can be both voluntary and involuntary.

Machine vision. In time, humans have invented machines to take away some of the work they had to do. For more demanding tasks these machines had to have the ability to perceive the environment in which humans leave in a similar manner as they and thus needed something like vision.

An uncontroversial definition of machine vision is very difficult if not impossible. For the purposes of this work, a machine vision system recovers useful information about a scene from its 2D projections [115], i.e., from images. For this purpose it emulates the human vision and in doing so it uses methods from other related fields, the most important being:

- **Imaging** – which deals with the acquisition of images.

- **Image processing** – used to enhance particular information and suppress noise.

- **Pattern recognition** – used to classify/recognize data.

Thus, a machine vision system accepts as input images and returns as output an interpretation of objects present in that images. For this purpose it uses specific hardware to acquire and analyze images, e.g., a video camera and a computer.

2.2 From photons to discrete signal

In the case of machine vision, processing and analysis of images is usually done in a computer, which is a discrete system and needs discrete inputs. Between a scene and a discrete representation the information contained in an image is processed at several intermediary levels. Digital imaging includes aspects related to these levels: the formation of the image in front of a sensor,

the transformation of image information into electrical signals by the sensor (see Section 2.2.1 for visible light and Section 2.2.2 for X-rays), and the discretization of the electrical signal for storage and processing in a computer (see Section 2.2.3). At each level the original information is modified in a certain way and potentially artefacts are introduced (see Section 2.2.4). The purpose followed here is to describe these artefacts such that it becomes clear how they appear and what type of post-processing is necessary to eliminate them or reduce their influence.

Within the frame of this work only data carried by electromagnetic waves and projected before a 2D discrete-sensor is considered. In such a case intensity images are obtained, which encode the amount of electromagnetic radiation leaving from a certain point of the scene.

Depending on the frequency of the radiation which should be sensed, the physical parameters of the image acquisition system – which play a major role in image formation – vary strongly. Two cases are described in more detail, i.e., visible light and X-rays. They represent the main 2D projection imaging modalities encountered in biomedical cardiovascular imaging.

2.2.1   Image formation and detection with visible light

When working with visible light, image formation is the process where rays emitted from scene items are collected to form an image before a plane. This plane is called the image plane, and it typically contains a device able to convert photons associated with the rays in an electrical signal.

Image formation

Image formation with visible light is best approximated by an additive imaging model, where the influences of different items in the scene are added to obtain the image.

The pinhole camera.   Image formation includes in this case, *geometric*, *optic* and *radiometric*[6] aspects. The former relate the position of items in 3D space to their position in the image plane and the latter analyze how much of the radiation emitted by the scene-items is collected by the imaging system and received by the sensor.

Humans have always tried to "save" images of scenes in their surroundings. At the beginning these images were simple drawings that lacked "realism" as it was difficult to find the right perspective. A major improvement in this field, which allowed the drawing of realistic images with relative ease, was the development of the camera obscura. Initially this was a real room[7], a confined space with walls of wood or dark curtains where no light was allowed in with the exception of the light coming from a tiny opening in that wall of the room which was facing the scene to be imaged. On the opposing wall an image of the scene appeared in the right perspective. Everything the "artist" should do to achieve a realistic image was to draw over what he saw on that wall.

Although the image acquisition techniques have evolved dramatically since the camera obscura, a device able to acquire images is still called a camera. The most simple camera is a pinhole camera (see Figure 2.3). This is a size-reduced model of the camera obscura, the light producing the image passes through an infinitesimal small hole (aperture) in one of the walls

---

[6]Radiometry is the measurement of optical radiation, i.e., electromagnetic radiation with wavelength between 10nm and 1mm thus including infrared, visible and ultraviolet but excluding X-rays.

[7]Camera is the latin word for a vaulted room.

and the image is obtained on the opposing wall, i.e., the image plane. Due to the small opening only one light ray per scene point reaches a point in the image plane.
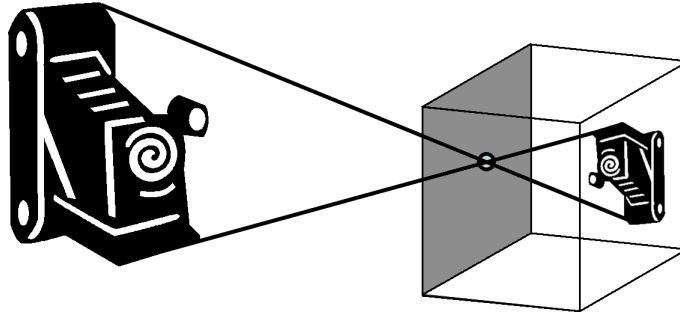
Figure 2.3: Image formation by a pinhole camera.

The perspective camera model. The geometric model which describes image formation by a pinhole camera is called *perspective model*. It consists of the image plane π and a point O called the center or focus of projection. The coordinate system attached to the pinhole camera has its origin in the center of projection. The image plane is situated at a distance d_i from the center of projection (See Figure 2.4).

The relation between a 3D scene point $[X, Y, Z]$ and the 2D coordinates in the image plane $[x, y, -d_i]$ in camera coordinates is:

$$x = -\frac{d_i X}{Z}, y = -\frac{d_i Y}{Z} \tag{2.1}$$

The equations (2.1) are non linear due to the factor $\frac{1}{Z}$. The image coordinates contain only ratios of the world coordinates from which neither true size nor distance between items can be inferred without additional knowledge. Perspective projection does however map lines into lines.

Under the approximation that Z is constant, i.e., the *weak-perspective model*, which becomes valid if the scene's depth is much smaller than the average distance form the camera to the scene (typically by a factor of 20), the Equations (2.1) describe a sequence of two transformations: an orthographic projection in which world points are projected along rays parallel to the optical axis on the image plane and a scaling by a factor $-\frac{d_i}{Z}$.
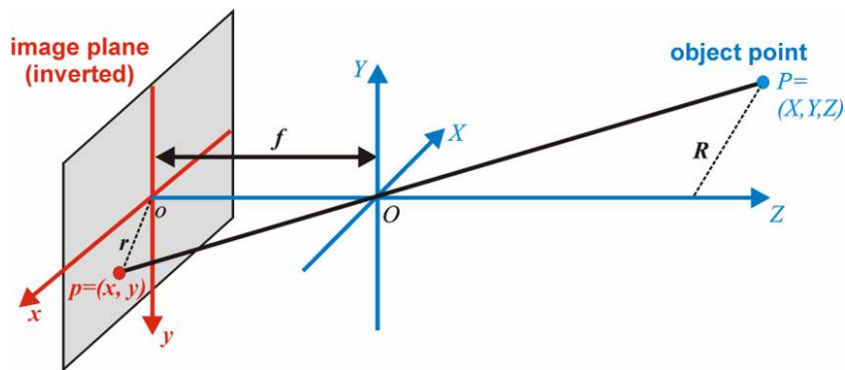
Figure 2.4: The perspective camera model – with permission from [2].

The relations between world and image points can be precisely described in matrix-form to the point that image-based inferences on distances and size can be made. Further details can be found in [188].

Particularities of the pinhole camera. The fact that the light enters the pinhole camera through a tiny aperture has several important effects on the image which is obtained:

- The opening acts as a low-pass filter and attenuates high frequencies.

  – Light is both a wave and a particle beam ; as a wave is also subject to specific phenomena like diffraction at the opening. Thus the image is blurred due to diffraction to a size of at least the order of the wavelength of light. In this case a specific model can be assumed, known as Fraunhofer diffraction, which allows the computation of the cut-off frequency of the low-pass filter modeling the blurring. [109].

- Irrespective of the distance between the center of projection and the imaged scene or between the center of projection and the image plane, the image obtained will always be "sharp" or "in focus".

  – If the opening of the camera is larger, then several of these rays would go through the opening to form the image of a point, meaning that each point will be projected on a small disc in the image plane when this is situated at a certain distance from the center of projection. This is equivalent to a low-pass filtration of the image. By varying the position of the scene or that of the image plane from the center of projection the area of the disc can also vary. Assuming that the image is inherently band-limited and that the distance to the scene and to the image plane from the center of projection are fixed, there is a minimal cut-off frequency of this filter or a maximal size of the disc above which this effect is disturbing, affecting the image information.

  – In the case of the pinhole camera, because the opening is infinitesimal small one and only one ray of light links each point of the scene with a point in the image plane. Thus the image will always be correctly formed in the image plane regardless of the position of the scene along the optical axis. This means that the pinhole camera has an infinite depth of field. Conversely, the image will be correctly formed irrespective o the position of the image plane along the optical axis. This means that the pinhole camera has an infinite depth of focus (see Figure 2.3 and Figure 2.6).

- A light sensitive device placed in the image plane will need to be exposed to the scene for a long period of time before acquiring all the needed information.

  – Assuming that a light-sensitive sensor is placed in the image plane of a pinhole camera and is supposed to use the energy given by the photons traveling along the ray of light which links scene points to image points to modulate a certain value, e.g., electrical load, then, to gather enough energy for a correct modulation, the sensor should be exposed to the image for a time which is for most applications impractically long. This exposure time is roughly inversely proportional to the square of the aperture diameter. The solution to this problem is to enlarge the opening and permit

more light from a scene-point to reach the image plane, and then use an optical system of lenses to focalize the rays of light coming from a scene-point into a single point on the image plane, so that no blurring appears.

Image optics. Optical systems can be adjusted to work under under a wide range of illumination conditions and exposure times – the exposure time being controlled by a shutter – and thus make cameras suitable to practical applications. An optical system is supposed to preserve the geometrical relations between world and image coordinates. However this is true only for the simplest optical systems: the thin lens. This model is often encountered in practice and it can be applied whenever the distance from the viewed object to the lens and from the lens to the image plane is very large in comparison to the thickness of the lens. A thin lens model is enough for a discussion on the principles of optical systems which is the purpose followed here. A more detailed description can be found in [169].

The thin lens model.   The optical behavior of a thin lens is characterized by two elements, i.e., the optical axis and the focal length. Considering that the image trough the lens of a point P, belonging to an object situated at distance D from the center of the lens O, is formed at p in an image plane situated at distance d, the optical axis goes through O and is perpendicular to the image plane. The focal length f is the distance between O and two points situated on each side of the lens, named right F_r and left F_l focus respectively. This is shown in Figure 2.5. The formation of an image through a thin lens is subject to two basic rules:

- A ray of light parallel to the optical axis on one side goes through the focus on the other side.

- A ray of light coming from the focus on one side emerges parallel to the optical axis on the other side.

In such a setup the relation between the distance measured from the focus on one side to the imaged item and from the focus on the other side to the image plane can be deduced from pairs of similar triangles as:

$$Zz = -f^2 \tag{2.2}$$

which becomes

$$\frac{1}{\hat{z}} = \frac{1}{f} + \frac{1}{\hat{Z}} \tag{2.3}$$

for $\hat{z} = z + f$ and $\hat{Z} = Z + f$ and considering a coordinate system with O as origin.

Particularities of image optics.   The characteristics of the optical system influence strongly the image which is obtained. Therefore such a system has to be carefully designed. The most important parameters of an optical system are:

- The *depth of field* and the *depth of focus*, which are related to the blurring of the image when D and d respectively vary.

  - If the item viewed is not situated precisely at a distance $\hat{Z}$ from the lens, then the image of a certain point gets blurred in the image plane and it appears as a disk. The
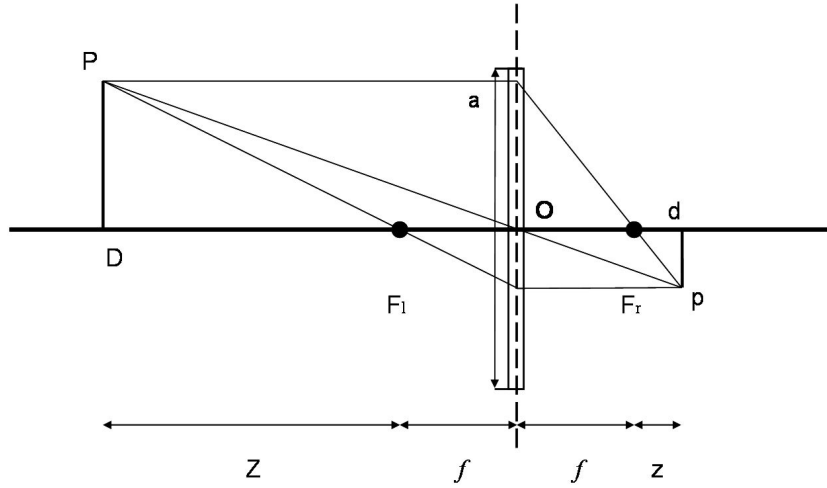
Figure 2.5: Imaging by a thin lens.

(a)   (b)
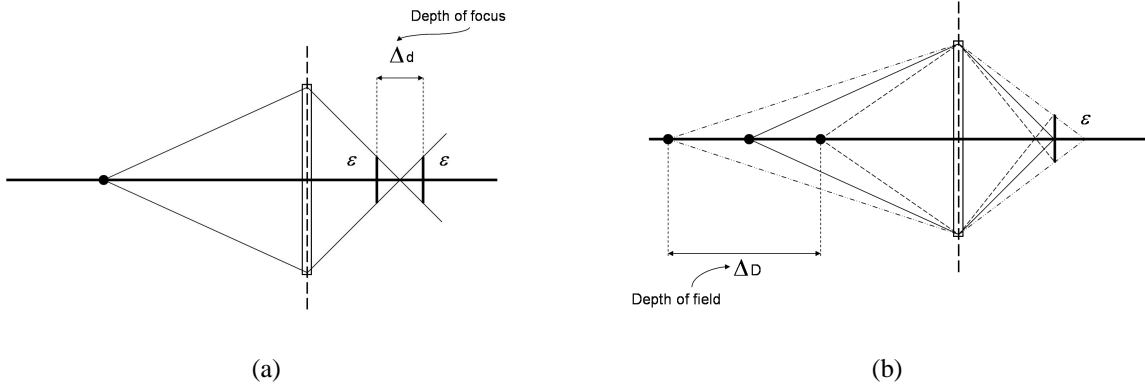
Figure 2.6: Illustration of depth of focus (a) and depth of field (b) – after [109].

same thing happens if the image plane is not situated exactly at a distance $\hat{z}$ from the lens. This is show in Figure 2.6. Again considering that the viewed scene is band limited such blurring is not disturbing up to a certain limit. The maximal size of the disk beyond which the blurring becomes disturbing, i.e., the blur size ϵ, gives the *depth of field* and the *depth of focus*. The *depth of field* is the distance interval from the viewed item to the lens so that the diameter of the blur is less than the blur size. The *depth of focus* is the distance interval from the lens to the image plane so that the diameter of the blur is less than the blur size.

- The imaging setup: *normal*, *telecentric* or *hypercentric*, which is related to the position of the aperture stop a in relation to the lens and the right focus F_r.

   – In a *normal imaging* system the aperture stop a is positioned in the same plane as the thin lens. Then, the viewed item appears larger if it is positioned closer to the lens and smaller conversely besides being blurred.

   In *telecentric imaging* the aperture stop is placed in the same plane with the right focus F_r. In such a case varying the position of the viewed item causes only blurring

while the size remains constant. The disadvantage of *telecentric imaging* is that the lens should be at least as large as the viewed item.

In the case of *hypercentric imaging* the aperture stop is situated between the right focus and the image plane, than the viewed item appears larger if it is positioned further away from the lens and smaller conversely.

A surface parallel to the optical axis will be imaged in a normal and hypercentric setup but not in a telecentric setup. A normal setup will project the surface-side facing the optical axis and the hypercentric setup the other surface-side.

- The *field of view*, which defines which portion of the scene is projected onto the image plane.

  – The aperture allows only a certain part of a potentially infinitely stretching scene to be imaged. The field of view gives an angular measure of the portion of the scene actually seen by the camera. The field of view is defined as half of the angle subtended by the aperture as seen from the left focus [188]:

  $$w = \arctan \frac{a}{2f} \qquad (2.4)$$

- The *radial distortions*, which are related to geometrical transformation of the image due to the lens.

  – Real optical systems are not perfect systems. Diffraction effects at the aperture of the lens, similar to that discussed for the pinhole camera, introduce distortions. Distortions appear also if we ignore such effects, and thus restrict ourself to geometric optics. Assuming the position of the item viewed ,as well as that of the image plane, are fixed and that the item is "in focus", a point on that item is in most cases imaged in a small disk in the image plane – i.e. blurring/distortion. Such aberrations are a consequence of the diffraction and reflection in the lens. The most important for B/W imaging are: spherical aberration, coma, astigmatism and field curvature which cause blurring and radial distortion which affects the shape of the image. Radial distortion can be modeled as:

  $$\vec{x} = \vec{x_d}(1 + k_1 r^2 + k_2 r^4) \qquad (2.5)$$

  with $\vec{x}$ the corrected image coordinates, $\vec{x_d}$ the distorted image coordinates, $k_{1,2}$ parameters nd $r^2 = x_d^2 + y_d^2$.

Further details and a good description of all lens aberrations can be found in Jähne [109].

Radiometry.   Light coming from light sources or reflected from items in a scene reaches the image plane through the optic system of a camera and forms an image there. This image is then acquired by a light sensitive device. To fully describe the entire process one needs also the relation among the amounts of light energy emitted from light sources, reflected by items and reaching the image plane. Radiometry is the part of image formation concerned with such aspects. It is defined as the measurement of optical radiation, i.e., electromagnetic radiation with wavelength between 10nm and 1mm thus including infrared, visible and ultraviolet but excluding X-rays [109], [188].

Radiometric image formation.   The main radiometric task is establishing the relation between the power of light per unit area and at each point of the image plane – i.e. image irradiance – and the power of light per unit area ideally emitted by each point of a scene –i.e. scene radiance. For this purpose, a model of the way a surface reflects incident light is needed. Typically the Lambertian model is used. It assumes that each point appears equally bright from all viewing directions. Then the relation between image irradiance $E(\vec{p})$ and scene radiance $L(\vec{P})$ is given by the fundamental equation of radiometric image formation:

$$E(\vec{p}) = L(\vec{P})\frac{\pi}{4}\left(\frac{a}{f}\right)^2 \cos^4 \alpha \qquad (2.6)$$

with a aperture, f focal length and α the angle between the ray from the scene point $\vec{P}$ to the image point $\vec{p}$ and the optical axis. In practice, the dependency on α is ignored. Then with respect to the camera, the image irradiance depends on the quantity $\frac{f}{a}$ which is called the F-number of the optical system. It influences how much light reaches the sensor and is typically mentioned on lenses in a scale containing only powers of two. Increasing the F-number by one unit on this scale, halves the quantity of light reaching the image plane. For further details see [188].

Image sensing

Once the image has been formed in the image plane, it is acquired for storage and processing by a light-sensitive device which is able to use the energy of light to modulate some chemical or electrical process. Here are discussed only imaging sensors able to covert radiative energy into an electrical signal and specifically quantum detectors – which convert photons into electrons – as they are commonly used in cardiovascular biomedical imaging.

Quantum detectors.   The functioning of quantum detectors is based on the photoelectric effect.

The photoelectric effect.   First discovered by Hertz in 1887, then explained by Einstein in 1905 and finally confirmed by the experiments of Millikan in 1916, the photoelectric effect is the extraction of electrons –i.e. photoelectrons from a material as a consequence of shading light over it or electromagnetically radiating it [169].

Einstein has postulated that a beam of light consists of any small quantities of energy called light-quanta or photons. Each photon has a certain energy which is proportional to its frequency ν after the formula:

$$E = h\nu = \frac{hc}{\lambda} \qquad (2.7)$$

where h and c are universal constants. $h = 6.63 \cdot 10^{-34} Js$ is called Planck's constant and $c = 3 \cdot 10^8 m/s$ is the speed of light. Whenever a photon hits an electron an "all or nothing" change of energy takes place, meaning that either the electron takes the entire energy of the photon or none of it.

As postulated by Bohr, electrons occupy only certain energetic levels in atoms and the value of this energetic levels is specific for each material. Electrons can leave an atom only if a specific quantity of energy – related to the energetic level where the electron is situated – is received. If

a photon gives an electron more than the binding energy of his level, then the electron is set free with a certain kinetic energy E_c:

$$E_c = \frac{mv^2}{2} = h\nu - \Phi \tag{2.8}$$

with Φ the energy needed to leave the atom.

The internal photoelectric effect. Photoelectrons can be generated from any materials, provided the energy of the incoming photons surpasses a certain threshold which depends on the particular energetic levels of each atom. However, of interest in this case are only photoelectrons generated in semiconductors and which do not leave the semiconductor body. Such photoelectrons are said to be generated by the internal photoelectric effect, as opposed to electrons which leave the body from which they are generated, in this case by the external photoelectric effect.

The sensor-element. Photoelectrons generated by the internal photoelectric effect can either be captured in potential wells – e.g. in the case of metal-oxide-semiconductor (MOS) capacitor – or they can contribute to current going through the semiconductor – e.g. in the case of the photodiode. Such a quantum detector which transform incident light in electrical charge represents a sensing site and usually constitute the main part of a sensor-element. A sensor-element is characterized by [109]:

- **Detectable wavelength range**, which is given by the interval between the minimal and the maximal detectable photon frequency (or energy according to Equation (2.7)). For sensors working in visible light, e.g., this between 450nm and 1000nm.

- **Quantum efficiency**, which is the percentage of photons hitting a photoreactive surface that will produce an electron-hole pair. This value varies between 10% for a photographic film and 95% for high-end sensors. As shown in Figure 2.7 it varies as a function of the wavelength.

- **Dark current**, which is the signal generated by the detector in the absence of light due to thermal energy.

- **Electron capacity**, which is the maximal number of photoelectrons that can be stored in a sensor element.

- **Saturation exposure**, which is the image irradiance which generates a number of photoelectrons equal to the Electron capacity in a time interval equal to the exposure time

- **Dynamic range**, which is the ration between the maximum output signal when the sensor is saturated and the signal corresponding to the dark current

- **Responsivity**, which is the ration between the flux of light incident to the detector area and the resulting signal. Responsivity depends mostly on the wavelength of thee incident light, but also on the angle at which the light hits the sensor, temperature, aging, etc.

- **Noise-equivalent power**, which gives the minimum radiative flux that can be measured by a detector in a given frequency band
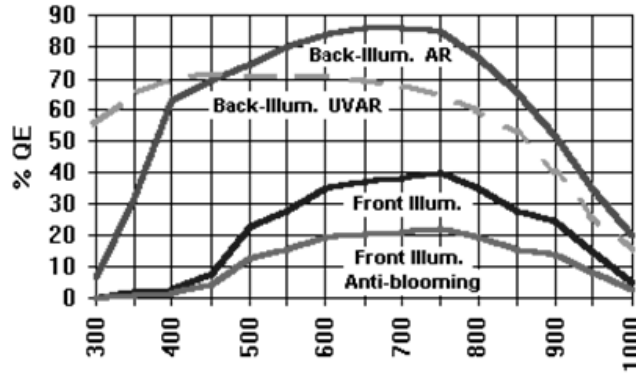
Figure 2.7: Quantum efficiency as a function of the wavelength for a CCD sensor for different types of illumination and coating – after [109].

Image sensors.   To sample an image, several such sensing sites are gathered in a matrix setup, to form an image sensor together with other elements, like e.g., storage sites and control circuitry. Major quality criteria for image-sensors are the fill factor – representing the percentage of the area of a sensor site which is actually sensible to light – and the frame rate – which represents the number of complete images acquired in the time interval (usually one second). There are two main image sensor architectures: the Charge Coupled Device image sensors (CCD) and the Complementary Metal Oxide Semiconductor image sensors (CMOS).

CCD sensors.   In CCD image sensors, when exposure – i.e. the time interval when each sensing site converts light into charge – is complete, the sensor elements are read row by row by moving their charge downwards on columns into a shift register. This sends the information through a processing chain – which is separated from the image sensor (see Figure 2.8 (a)) – where the charge-signal for each row is transformed into a voltage signal and amplified (e.g. by a transistor functioning in the same way as a bipolar transistor in common-base connection) and then converted in digital format by an analog-digital converter.

The electronic devices used to transfer the charge through and from the image sensor are called Charge Couple Devices and give also the name of the sensor. A typical CCD contains an array of linked MOS capacitors. By correctly timing the voltage pulses on the grill connectors $P_i, i = i \ldots N$ of the capacitors, potential wells can be sequentially generated which transport the charge through the device as shown in Figure 2.9. When the data appears at the input of the CCD, the voltage on I decreases, the voltage on G_1 increases so that a channel is formed under it and the charge can reach the potential well P_1 where the voltage was also increased. To move the charge into the potential well P_2 the voltage on P_1 decreases slowly while the voltage on P_2 is high. It is important that the decrease of P_1 is slow otherwise charge will be lost. The transfer is repeated between P_2 and P_3 and finally the charge-signal reaches the output with a delay $t = 3t_o$ where t_o is the time the voltage on the connector of a potential well is constant.

There are also several architectures for CCD image sensors. There are *frame transfer* and *interline transfer* sensors. As shown in Figure 2.10 (a), for the frame transfer sensor, the sensing sites function also as shift registers and constitute thus alone a sensor-element. At the end of the exposure time the entire image is transfered in an optically isolated storage area and from
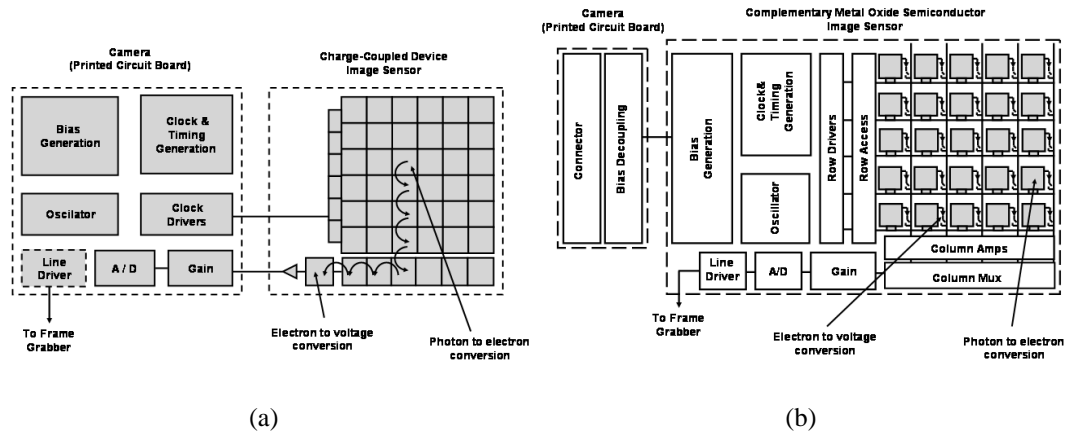
Figure 2.8: Schematics and image acquisition for a CCD sensor (a) and a CMOS sensor (b) – after [130].
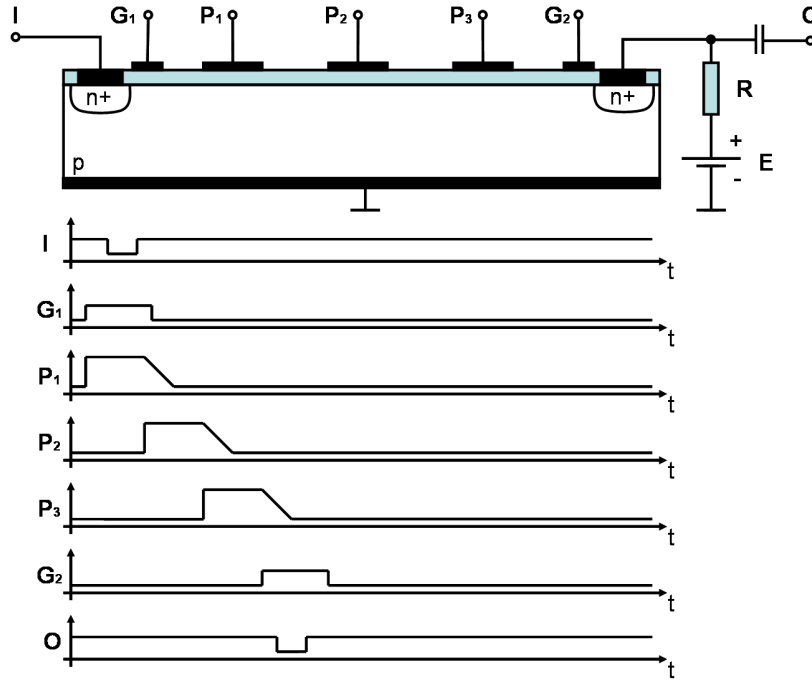
Figure 2.9: The schema of a three phase CCD and the corresponding voltage pulses – after [62].

there is read-out row by row with a horizontally shift register. The interline transfer sensor – shown in Figure 2.10 (b) – has storage-sites next to the sensing sites. They build together a sensor-element. At the end of the exposure time the charge is transfered first in the storage sites and then row by row in the horizontal shift register.

For a certain sensor-size, the frame-transfer architecture provides more sensitive area than the interline transfer sensor and can achieve thus better fill factors. To the limit – e.g. for large-area CCD-sensors – one can renounce to the storage area for the frame transfer in which case the image is directly read. Such a sensor is called *full-frame transfer sensor* and it can exhibit optimal fill factors. The time needed for this operation is sensibly longer than the time needed
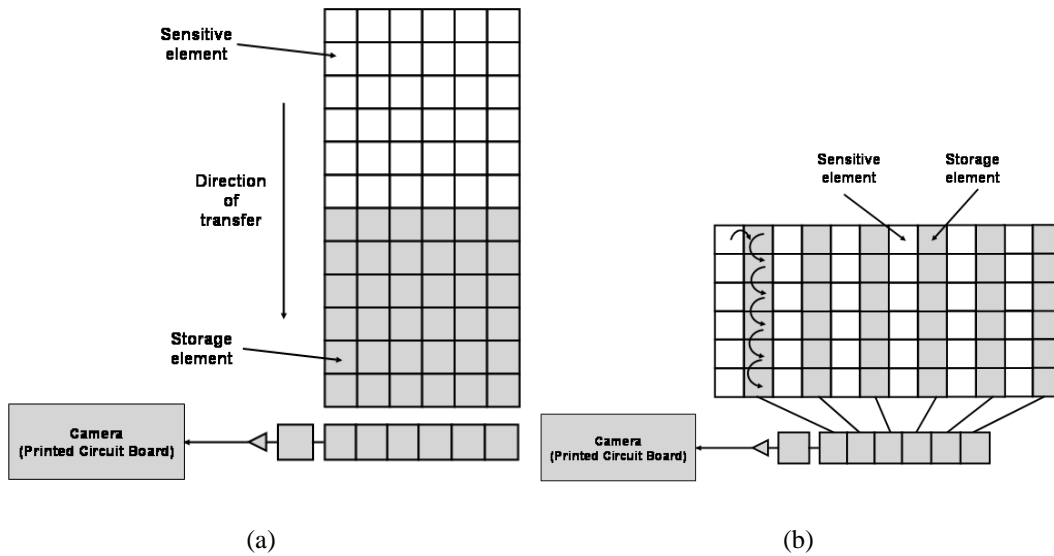
Figure 2.10: Schematics of a frame transfer (a) and interline (b) CCD sensor.

to shift the image from the active site into the storage site –which is equal to the time needed to read one row for a square sensor. This means that such sensors have to pay for their large size and fill factor by smaller frame rates.

To achieve a correct image, the CCD sensor has to be protected from light during read-out, otherwise, artefacts from well illuminated sites in the image can appear, like e.g., smearing. This is a must in particular for full-frame transfer sensors.

The sensor can be protected from light by a *mechanical shutter*. It provides good protection from light but has an inherent latency which limits the frame rate. If higher frame rates are desired, than an *electronic shutter* can be used. The electronic shutter functions by draining the accumulating charges at the beginning of the exposure time. With electronic shutters exposure times of 64 μs can be achieved, but they reduce the fill factor.

CMOS sensors.   CMOS sensors have a different architecture than CCD sensors. A CMOS sensor contains an array of individually -addressable sensing-elements. Thus, the charge obtained at the end of exposure at a sensing site can be gathered directly and it does not have to travel through the entire sensor.

For a long time such an architecture offered only images of a lower quality in comparison to the CCD image sensors. This was mainly due to the parasite capacitance of the lines linking a sensor site to the output amplifier. This was compensated by bringing the first transistor of the amplifier near the sensing site to build a sensor element called Active Pixel Sensors (APS). APS technology is at the heart of modern CMOS image sensors, and currently permits the acquisition of images of a quality close to that obtained by CCDs.

In a CMOS image sensor each sensing site is controlled individually (see Figure 2.8 (b)). The charge to voltage conversion and amplification takes place near the sensing site and each such sensor-element builds an APS. The analog to digital conversion and the clocking circuitry is also integrated on one chip with the sensor. This high integration is achieved by manufacturing the circuits in the CMOS technology, hence the name of the sensor.
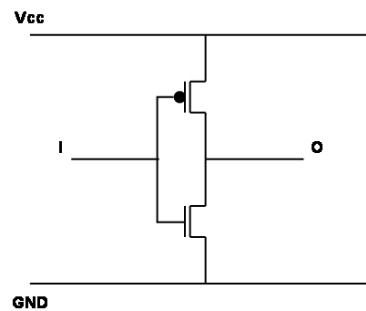
Figure 2.11: Schematic of a CMOS inverter

CMOS technology is actually thought for building large numbers of transistors on a small area. Transistors can be used to implement logical functions. For this purpose they have to act as switches and ideally would have to dissipate power only when switching between two states. With most types of transistors this is very difficult if not impossible to achieve. They usually dissipate power even if their state is constant. This leads to large power consumption and over-heating, which is a major hindrance to integrating many transistors on a small chip. However, two MOS transistors in a complementary setup (CMOS), i.e., one with a p-channel[8] and one with a n-channel[9], can switch between two states and practically dissipate power only when switching (see Figure 2.11). This allows the construction of chips integrating large numbers of transistors per unit area and gives the name for the entire production technology.

CMOS image sensors pay for the high degree of on-chip integration of imaging functions by reduced fill factors in comparison to CCD. This problem is at least partially solved by the introduction of *micro-lenses* which gather light over a larger angle than that allowed by the fill factor.

It is widely accepted that in the near future the image quality will be similar for both image sensors. Then, considering the advantages a CMOS sensor brings, like e.g., cheap production costs, high density, individual access to each sensing site, low power dissipation, it is believed that CMOS image sensors will replace CCD image sensors in most applications.

2.2.2   Digital imaging with X-ray

In the case of X-rays, image formation is the process where rays emitted from a point source are absorbed by items in the scene and the remaining radiation is collected to form an image before a plane. This image plane contains devices able to convert photons associated with the rays in an electrical signal. This is an absorbing type of imaging, and it can be modeled as each point in the viewed scene is generating a single ray which links it directly to a point in the image plane. This is then termed direct imaging [109].

Image formation includes mainly geometric aspects which relate the position of items in 3D space to their position in the image plane. The perspective projection of a pinhole camera can be used to describe the geometric image formation with X-rays also (see Chapter 2.2.1). This is shown in Figure 2.12. X-rays are not affected by lenses and although image irradiance can be defined, scene radiance or an analysis of image formation from a point of view similar to

---

[8]Positively doped semiconductor.
[9]Negatively doped semiconductor.
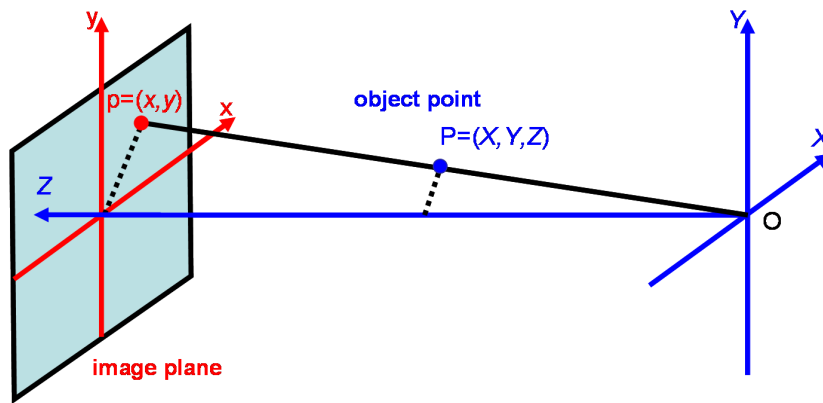
Figure 2.12: Image formation with penetrating X-rays

radiometry in the case of visible light, is pointless.

Image formation with X-rays is best approximated by a multiplicative imaging model, where the influences of different items in the scene are multiplied to obtain the image.

X-ray production and interaction with matter

X-rays were first discovered by the german physicist Wilhelm Röntgen in 1895. Their property of "seeing" through solid objects made them from the very beginning interesting for the medicine as a method of non-intrusive investigation – a role which they still maintain to this date.

X-ray is an electromagnetic radiation but of higher energy than visible light, as the wavelength characteristic for the X-rays typically lie in the interval of 0.01nm to 10nm (see Formula 2.7) this being the reason why is not completely absorbed in some solid bodies as opposed to light.

X-ray production. X-ray are produced by striking a target material with electrons traveling at a high speed. Such electrons interact with the atoms of the target material. Most high-speed electrons kick-out other electrons from the external energetic levels of the atoms in the target material. This type of interaction is most often encountered and accounts for over 99% of the input energy of the beam of high-speed electrons. It generates new electron radiation called delta radiation and heat. The rest of the interaction produces X-ray photons by one of three possible mechanisms:

- **Deceleration of high-speed electrons.** The high-speed electron travels through the target material until it encounters the electrical field of a nucleus, by whom it is attracted. Consequently it modifies its trajectory and decreases its speed, thus giving-up energy in the form of a X-ray photon. This type of electromagnetic radiation is called Bremsstrahlung [10]. Its spectrum covers the entire X-ray band as it depends on how close the high-speed electron goes by the nucleus.

---

[10]"Bremse" is the german word for brake.
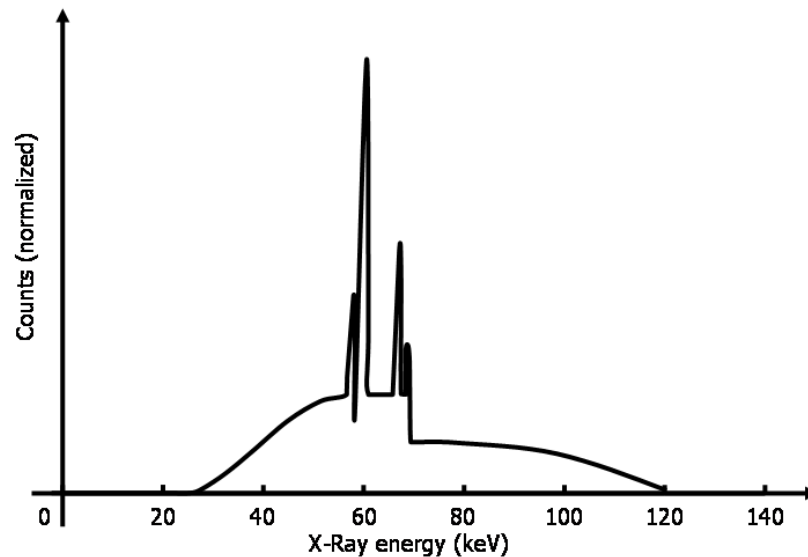
Figure 2.13: X-ray spectra – after [103]

- **Collision with the nucleus.** If the high-speed electron directly collides with the nucleus, than it gives-up its entire energy as Bremsstrahlung. High-energy X-ray quanta are produced this way.

- **Collision with an electron of the inner layers of the atom.** Electrons are allowed to occupy only certain energetic levels in an atom. If a high-speed electron collides with an electron from one of the internal energetically stable layers, than it sets that electron free. The place left behind by the electron is immediately occupied by another electron from an outer layer. For this, it gives-up a certain amount of energy which is specific to the atom. X-ray photons which are created like this make the so called characteristic radiation. The number of X-ray quanta produced by this mechanism is large in comparison with the other two, and causes distinctive peaks in the X-ray spectra (see Figure 2.13)

Medical X-rays.   Typically the energy of X-rays is given in electron-volts (eV). One eV is the amount of kinetic energy an electron receives if it is accelerated across a potential difference of one V: $1eV = 1.602 \cdot 10^{-19}J$. Using the Formula (2.7), one can calculate that an electron has to be accelerated across 124V to 124kV to obtain X-rays. As X-rays with wavelength of 0.1nm to 10nm have not enough energy to penetrate thicker bodies, they are called soft X-rays and are of no use in medical imaging. Therefore hard X-rays are needed with energies from 12.4keV to around 124keV, thus corresponding to a wavelengths of 0.01nm to 0.1nm [103], which are obtained from electrons accelerated by potential differences of 12.4kV to 124kV. A typical X-ray spectrum is shown in Figure 2.13.

The X-ray tube.   Usually X-rays are produced in an X-ray tube. Such a device contains a cathode, a filament and a rotating anode in a vacuum envelope. Electrons are produced by heating the filament (see Figure 2.14). They are then accelerated in vacuum between the cathode
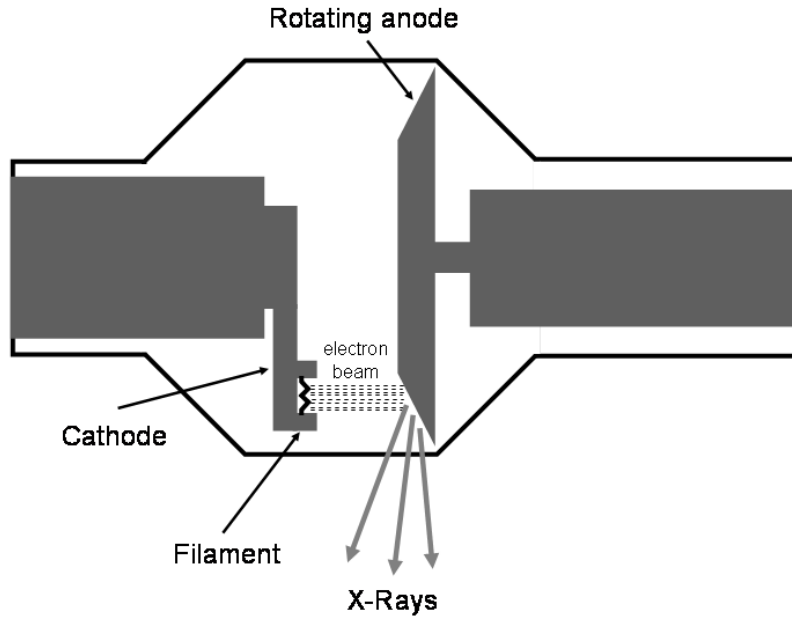
Figure 2.14: Schematic of a X-ray tube

and the anode and hit the anode to produce X-rays. The material from which the anode is made can be chosen so that it suits the target application, e.g., molibden is frequently used in anodes of mammography X-ray tubes because of its 20keV characteristic peaks [27]. The anode is rotating to prevent the melting of the target material during the bombardment with electrons. Ideally X-rays are created by a point-source as a larger source will blur the final image. To model a point source as good as possible, the rotating anode makes a certain angle with the vertical axis through the X-ray tube. This allows a better heat dissipation and a smaller X-ray generating spot visible from the image plane. This is shown in Figure 2.14.

X-ray interaction with matter.   After the X-rays are produced into the tube, they go through the imaged scene before reaching the image plane. X-rays may interact with matter in three fundamental ways: by the *photoelectric effect*, by the *Compton effect* and by *scattering*. This types of interactions may stop, deflect, or generate new X-ray photons —but not necessarily with the same direction and sense as the original one— thus reducing the intensity of the incident radiation. This attenuation can be expressed by an exponential relationship. For a mono-energetic X-ray beam and a material of uniform density and atomic number, this is:

$$I = I_o e^{-\mu L} \tag{2.9}$$

where μ is an attenuation coefficient and L the length of the path traveled by the X-ray through the matter. The probability of such interactions and thus μ, depends mostly on the type of material of which the imaged items are made. Therefore items of different materials will appear with different contrast in the final image.

Yet another characteristic of X-ray imaging is that if the energy of the incident X-ray radiation is too high, the contrast of the obtained images will be reduced.

Interaction by the photoelectric effect.   When an electron which is situated on one of the energetic layers of an atom is hit by a photon with sufficient energy to set it free, then it becomes a free electron. This mechanism is called the photoelectric effect (see Section 2.2.1). The photon transfers its entire energy to the electron and ceases to exist after this interaction. If the electron was ejected from one of the internal layers of the atom, an electron situated on one of the distant layers will occupy the free place and at the same time give up energy to produce a characteristic photon[11]. Thus the photoelectric effect typically produces, an ion, a free electron and a new photon but of less energy than the incident photon. This characteristic photon does not travel very far before being absorbed in the material.

The photoelectric interaction is most likely to occur if the energy of the incident photon is just grater than the binding energy of the electron with which interacts. Considering the types of interactions which appear when a human body is imaged using X-rays, it has been shown that the majority of photons with energies less than 25keV interact by the photoelectric effect [103] (see Figure 2.15).

The probability of an interaction is also proportional to the cubic of the atomic number of characteristic for the material of the imaged body and inversely proportional to the cubic of the energy of the incoming X-ray photon.

Compton effect.   The Compton effect was first observed in 1923 by the american physicist Arthur Compton. The photoelectric effect occurs most often when the energy of the photon is comparable with the binding energy of the electron. If the energy of the photon is considerably larger than that of the electron, than the two interact by the Compton effect. In this case, the incident photon does not disappear, but is scattered with loss of energy. Thus, after such an interaction, an ion, a free electron, a scattered photon and a characteristic photon appears. Similar as in the case of the photoelectric effect, the latter is then quickly absorbed.

The scattered photon may be deflected at any angle from 0 to 180 degrees and it retains most of its energy after collision. Lower energy photons are backscattered while higher energy photons are normally forward scattered. The probability of a Compton interaction does not depend on the atomic number of the material, and thus such interaction does not contribute directly to the imaging process but is rather a noise factor. The vast majority of high-energy X-ray photons – energies larger than 50keV – interact with matter by the Compton effect (see Figure 2.15). However, because the incident photon keeps most of its energy after such an interaction, the radiation dosage absorbed by the items in the scene – i.e. in our case the body of the patient – is less than the dosage absorbed by the photoelectric effect.

Rayleigh scattering.   The photoelectric effect and the Compton effect describe plastic interactions between atoms and X-ray photons. After such an interaction, the atom is changed to a positive ion and the photon either gives its entire energy or recoils on another direction but with less energy. The Rayleigh scattering is an elastic interaction between X-ray photons and atoms. After the photon hits the electron, this does not leave the atom but enters an excited state where it vibrates with the same frequency as the incident photon. After a very short time, the electron relaxes again and sets free/emits the photon without energy loss but on another direction and/or sense. The scattering occurs mainly in the forward direction, and thus only broadens

---

[11]The production of a characteristic photon is one way by which an excited atom relaxes. It can do this also by emission of Auger electrons. However, the probability of such an event is small.
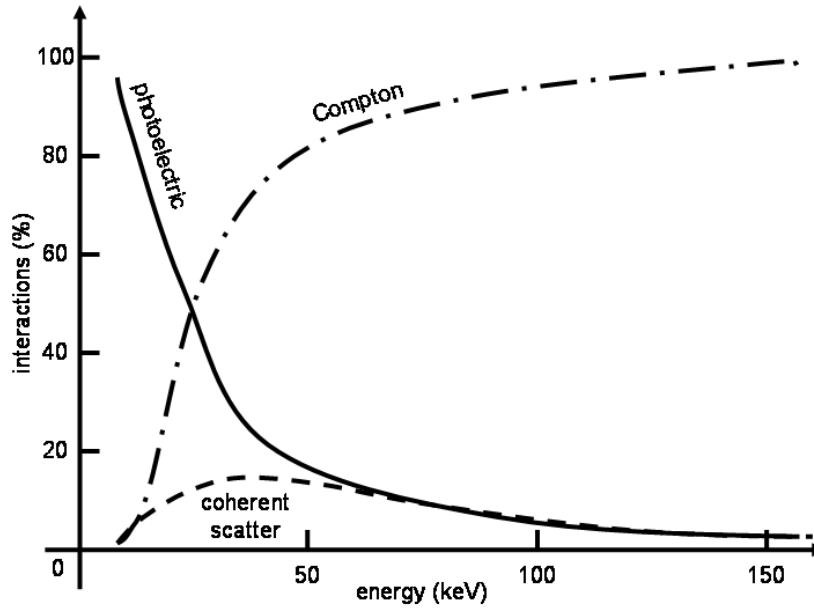
Figure 2.15: Percentage of different types of interactions as a function of energy in water – after Hsieh [103]. Approximately 60% of the human body consists of water.

the X-ray beam, which is equivalent to a low-pass filtration of the image. In comparison to the photoelectric effect or to the Compton effect, scattering occurs rather rarely (see Figure 2.15).

   Contrast reduction for high-energy X-rays.   The attenuation of X-rays (Formula (2.9)) takes into consideration all this effects, and thus μ can be computed as

$$\mu = \theta + \sigma + \sigma_r$$

, with θ, σ and σ_r the attenuation coefficients for the photoelectric effect, the Compton effect and the Rayleigh scattering respectively. Depending on the energy of the incident radiation, either the photoelectric or the Compton effects are responsible for most of the interactions, thus for high energies of the incident radiation there is a reduction in contrast, as most interactions occur according to the Compton effect.

X-ray detection systems

Similar to visible light imaging, X-ray imaging has evolved from taking simple snapshots, to movies in an analog format and then to digital signals. We concentrate on discussing digital imaging techniques which use either a combination of X-ray image intensifier and CCD/CMOS cameras or flat panel detectors.

   In the case of visible light, usually a scene is viewed, through an optical system. The role of the optical system is to gather information from the viewed scene as carried by the light an project it on the image plan while keeping point correspondence between the scene and the image. Thus the optical system makes a transformation from the scene to the image. It contains lenses which can influence the path of electromagnetic radiation with wavelength in an interval which includes the near infrared, the visible and the ultraviolet – i.e. the optical spectrum – but

	Detection method	Conversion
Direct	flat panel detector (Selenium)	X-rays to image
Indirect	flat panel detector (scintillator + photodiode)	X-rays to light to image
Indirect	Image Intensifier + camera	X-rays to light to image

Table 2.1: Digital X-ray detection methods

excludes X-rays. In the case of X-rays, a system with a similar role can be used to form an image before a light image-senor, i.e., CCD or CMOS. Such a system is called X-ray image intensifier. It not only scales the image but it also transforms it so that the information is now carried over visible-light waves. This is then called *indirect imaging* as opposed to *direct imaging* where an image sensor receives directly X-rays which it then converts to an electrical signal.

Indirect imaging contains a step which transforms the image information carried by X-rays into image information carried by visible light before obtaining an electrical signal. Direct imaging can transform directly X-rays to an electrical signal (see Table 2.1)

Intensifier-digital camera combination.   In a X-ray image intensifier – as shown in Figure 2.16 – a *scintillator* is used to transform X-rays into visible light and an *image-intensifier* is used to obtain then a brighter and smaller image– hence the name of the entire system [11], [161].

The scintillator.   A scintillator is a substance – e.g. cesium-iodide (CsI) – which absorbs high-energy electromagnetic radiation or particle radiation end emits visible light after a very short time called decay time. This separates it from a phosphor[12] whose decay time is very long and from a fluorescent substance which receives optical high-energy electromagnetic radiation – i.e. ultraviolet light – and emits visible electromagnetic radiation.

The image-intensifier.   Usually the light flux from the scintillator is very low, as in most cases, the X-ray dosage is reduced to a minimum to avoid damage to the patient's health. To obtain a brighter image an *image-intensifier* is used. The intensifier consists of a photocatode which is a cathode coated in a photoemissive material able to convert photons into electrons by the photoelectric effect at a high efficiency. The electrons are then accelerated and at the same time focused by special grids – called also electron optics – to strike a small scintillation screen at the other end of the image intensifier such that a light-image is again obtained. The kinetic energy with which the electrons strike is linked directly to the intensity of this image which is than viewed by a camera and projected before a CCD or CMOS sensor which transforms it into an electrical signal. Producing a smaller image is important as it can be handled more efficiently by the lenses of the camera used to acquire it.

Flat Panel Detectors.   Image intensifiers have some disadvantages: they are bulky, they have a rather limited life span, they return images with geometrical distortions, as e.g., the Earth's magnetic field influences the electrons' path, they have a rather small dynamic range, they also show brightness variations and loss of contrast at the edges. All this directly influences the

---

[12]The chemic element phosphorus is not a phosphor.
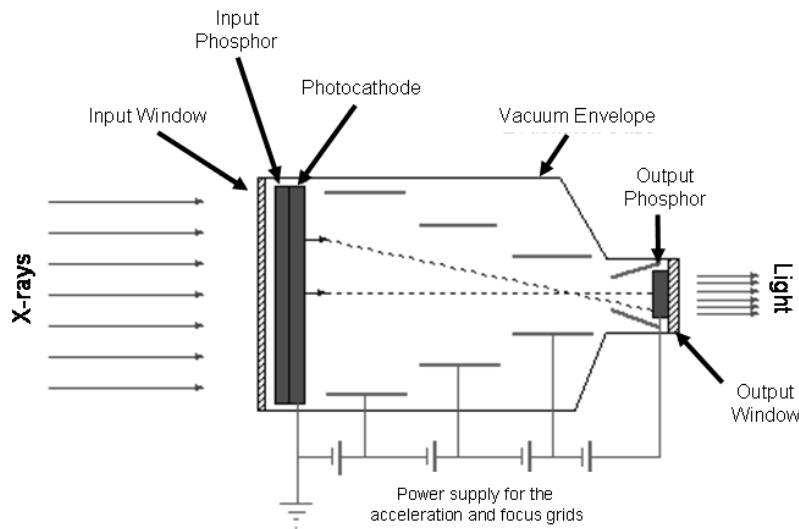
Figure 2.16: Schematic of a X-ray image-intensifier.

quality of the images which are then acquired by the camera, which in turn distorts also the image (see Section 2.2.4).

This disadvantages, which in most cases are inherent to the intensifier, led to the development of Flat Panel Detectors (FPD). A FPD can acquire X-ray images without an image intensifier. The construction of FPDs was made possible by advances into integrated circuit technology and more specifically by the development of amorphous-Silicon Thin-Film-Transistor (TFTs) arrays[13] which is a technology that allows the fabrication of large-area arrays of transistors.

As already pointed out, focusing the X-rays to a smaller surface is difficult and leads to distortions. Alternatively one should build sensors large enough to directly acquire the X-ray-image. Building a CCD or CMOS image sensor of sufficient size is in this case practically impossible. However, using TFTs where each transistor is attached to a sensor-element[14] allows the construction of image sensors with large areas. Such a construct – i.e. transistor plus sensor-element in TFT technology – makes a picture element. Several picture-elements arranged in an array make a TFT-active-matrix. This is shown in Figure 2.17.

A FPD consists of a sensory part and an electronic part. The sensory part is made of a sheet of glass covered with a thin layer of silicon in an amorphous state containing a TFT-active-matrix and a detector-element. The electronic part is needed to read the sensory array and discretize the information. It is either built on the sides of the sensory part or behind it, behind a led screen.

FPD-technologies. Currently two types of FPDs can be encountered in practice, one is used for indirect and one for direct imaging (see Figure 2.18 and Table 2.1):

- **Indirect-imaging FPDs.** If the sensor-element of the TFT active-matrix is a normal photodiode, this is sensible only to visible light. Therefore an additional conversion from

---

[13]There are also Microchannel Plate Detectors [177], which can be used similar as FPDs but this technology is just beginning to find its way into the X-ray medical imaging.

[14]In this case a distinction should be made between a sensor and a detector. A detector only converts X-rays to visible-light or to charge. A sensor can then gather the light-photons or the charge to yield a signal.
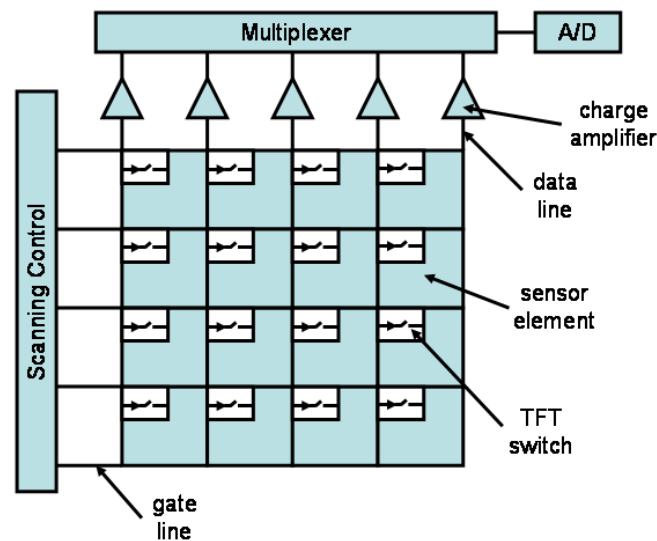
Figure 2.17: Schematic of a TFT-active-matrix – after [75].

X-ray to light-photons is needed before the image information is obtained as charge at the output. For this purpose a scintillator is used which constitutes thus the detector-layer of such FPDs.

- **Direct-imaging FPDs.** A photoconductor[15], as e.g., amorphous Selenium can directly convert X-ray photons to charge. Such a detector layer can be placed in front of a TFT active-matrix whose sensor-element is a capacitor to obtain another type of FPD. As in this case X-ray photons are converted directly to charge, this is called a direct-imaging FPD.

Comparison between FPD-technologies. Indirect-imaging FPDs differ according to the type of scintillator used, however, the large majority uses CsI. CsI scintillators have several advantages, they can be directly deposited on the TFT to give the best possible optical coupling efficiency in comparison with other scintillator screens which are mechanically forced into close contact. When CsI is deposited on the TFT it grows in a columnar structure thus reducing the amount of light scattering and reflection. By comparison, other scintillators induce a larger dispersion of light which leads to images which are more blurry. This is shown in Figure 2.19 – indirect imaging. Direct-imaging FPDs have the advantage of returning images of even higher quality, as there is no spreading of light in this case. This is shown in Figure 2.19.

Comparison to image-intensifiers. In comparison to X-ray image intensifiers, modern FPDs are small, have longer operational lifetimes and larger dynamic ranges. FPDs have usually a better SNR than intensifiers and it can be further improved by binning, i.e., connecting

---

[15]A photoconductor is a solid-state detector, which electrically is a semiconductor or an insulator. They are used to obtain localized charge from localized incident X-ray radiation. A metal is not a good photoconductor as the charge is not localized due to the excess in free carriers.

Figure 2.18: The path from X-ray-photons to charge in indirect imaging FPDs (a) and direct imaging FPDs (b) – after [75].

Figure 2.19: Scattering for different detector-layers used at indirect and direct imaging respectively – after [75].

e.g. a 2×2 neighborhood of pixels to a super-pixel. It is expected that they will completely replace imaging systems using image intensifier and camera in the near future.

2.2.3   Image digitization

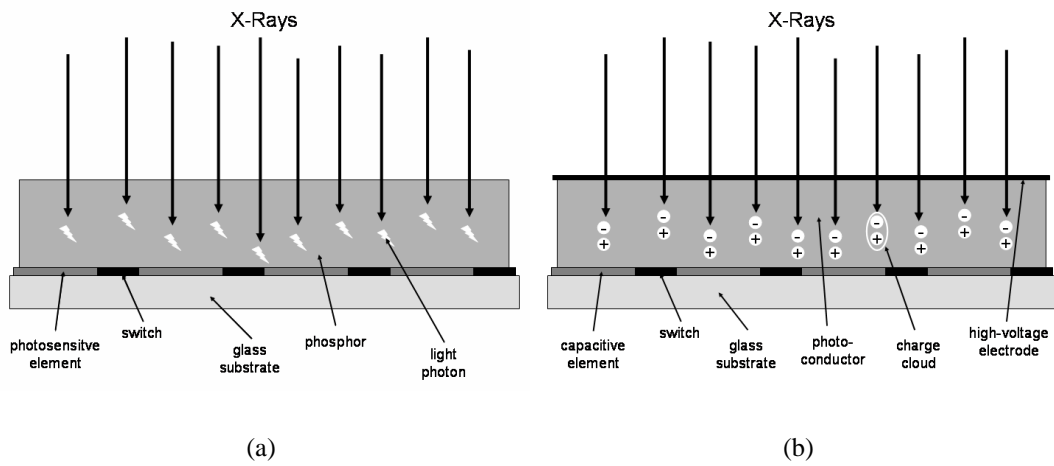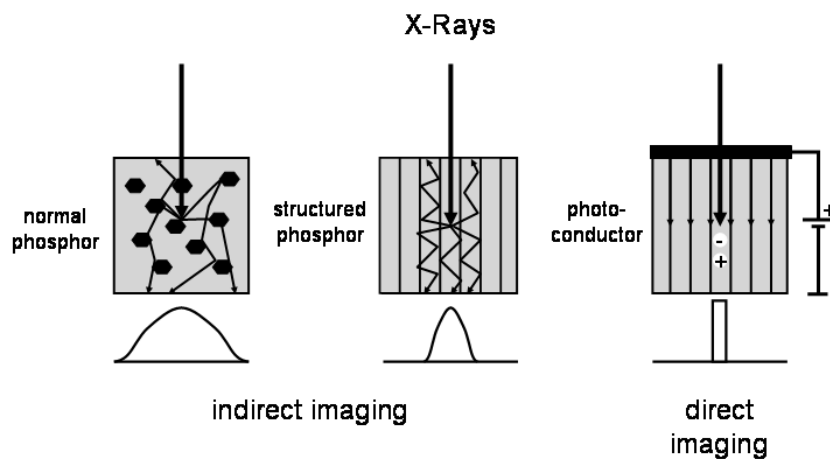After image information is acquired by the sensor the signal is usually converted to voltage and digitized before being processed. Discretization implies a sampling and a quantization step. For CCD and CMOS sensors as well as for FPDs, the image is inherently sampled as each sensor-site acquires one image sample. The quantization can then take place in the imaging device – even on the same chip as in the case of CMOS sensors – or on a dedicated card in a computer.

Sampling.   Ideally, the image is sampled by a train of Dirac pulses δ which in 1D are defined as:

$$f(0) = \int_{-\infty}^{+\infty} f(t)\delta(t)dt \tag{2.10}$$

Mathematically this is a functional, also called Dirac's Delta distribution.

Sampling can be described as the result of a multiplication between the image and the train of Dirac pulses. The distance between the pulses of the network should be chosen according to the Nyquist theorem, which requires that the sampling frequency should be at least double the highest frequency in the image spectrum.

If the Nyquist theorem is not respected, then image information in the high-frequency domain is partially lost during sampling and alias appears. The alias is the distortion of the reconstructed signal, due to incomplete image information. Sampling makes the spectrum of the original-signal periodic and alias can be observed as an overlap between such periods [110], [108] .

If the sampling frequency is chosen to be exactly the double of the highest frequency in the image spectrum, then it is said that the image is critically sampled. However, this is not advisable, as alias may appear if the image is rotated before being sampled, as a rotation of the image leads to a rotation of the spectrum.

Influences of sampling on the signal.   Dirac's Delta distribution is a mathematical abstraction. Practically, it is approximated by a rect function of a small width and high amplitude:

$$\delta \cong \frac{1}{T_0}rect\left(\frac{t}{T_0}\right) \tag{2.11}$$

This corresponds then to modeling a sensor-site as a small surface over which the image information is integrated. This model is also conform with the reality, as sensors-sites gather photons over their entire active area which they then transform to an electrical charge. In this real case, sampling can be described as a multiplication between the image and a network of pulses obtained by filtering the Dirac pulses with a low-pass filter (a short-time integrator). As the convolution is associative this is equivalent to first applying a low-pass filter to the image before ideally sampling it. The cut-off frequency of the low-pass filter depends on the effective area of the sensor-site and thus on the fill factor. A larger fill factor gives a better sensitivity but at the same time a more powerful low-pass filtration. The typical size of objects in an image is such that they are not affected by the sensor-site-area low-pass, thus in most applications, the quality of a sensor increases with the fill factor.
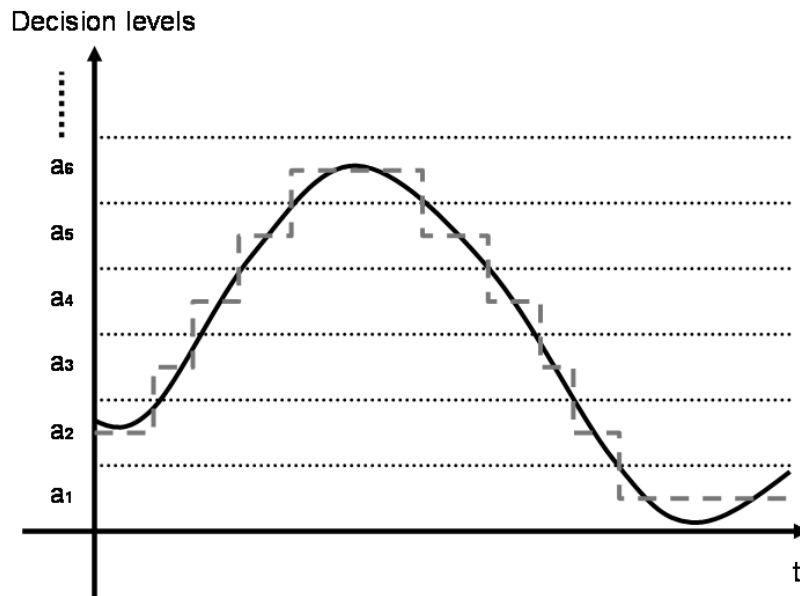
Figure 2.20: Example of a quantizer. As soon as the amplitude of the input signal (continuous-line) leaves the range of a decision-level (between two dotted-lines), the output modifies in discrete steps (dashed-line).

Quantization.   The step subsequent to sampling in digitization is quantization. Quantization transforms a continuous sample value in a discrete one. For this purpose it maps a intervals to constant values, i.e., quantization-levels, as shown in Figure 2.20. These quantization levels are then coded into binary numbers by modulating the length of a series of pulses of constant amplitude. A typical quantization circuit includes an operational comparator, a D/A converter, a digital counter (made with JK latches) and an AND circuit. The D/A converter can be made, e.g., with an operational amplifier and several resistors [61].

Uniform and non-uniform quantization.   If the distance between two quantization-levels, i.e., the quantization step, is always equal, than this is called *uniform quantization*. Uniform quantization is best suited if the signal is uniformly distributed (see Figure 2.21). However, in most cases the signal is rather Gauss distributed as small values appear more often than large ones. In this case, the overall approximation error will decrease if the quantization step is small for small values and large everywhere else. This is shown in Figure 2.21 (b). Such a quantization is called *non-uniform quantization*.

Practically a non-uniform quantizer is achieved by first applying a non-linear compression to the signal followed by uniform quantization. The non-linear compression curve is chosen depending on the interval where most values of the signal are found. When reconstructing the signal, the inverse transform is applied, expanding it back to the original range of values. A typical compression and the corresponding expansion curve – corresponding to Gauss-distributed sample-values – is given in Figure 2.22.

Influences of quantization on the signal.   Quantization means actually signal approximation. The maximum approximation error for each sample is equal to one quantization step. The

Figure 2.21: Uniform (a) and non-uniform (b) quantization curves with approximation errors e at different levels l – after [61].

Figure 2.22: Compression curve (continuous) and expansion curve (dashed) – after [61].

overall approximation error depends clearly on the number of quantization levels – i.e. the more quantization-errors the smaller the error. However, it can be improved if one has knowledge on the distribution of the sample-values. The number of quantization levels is usually expressed in the binary base and gives the number of bits per sample.

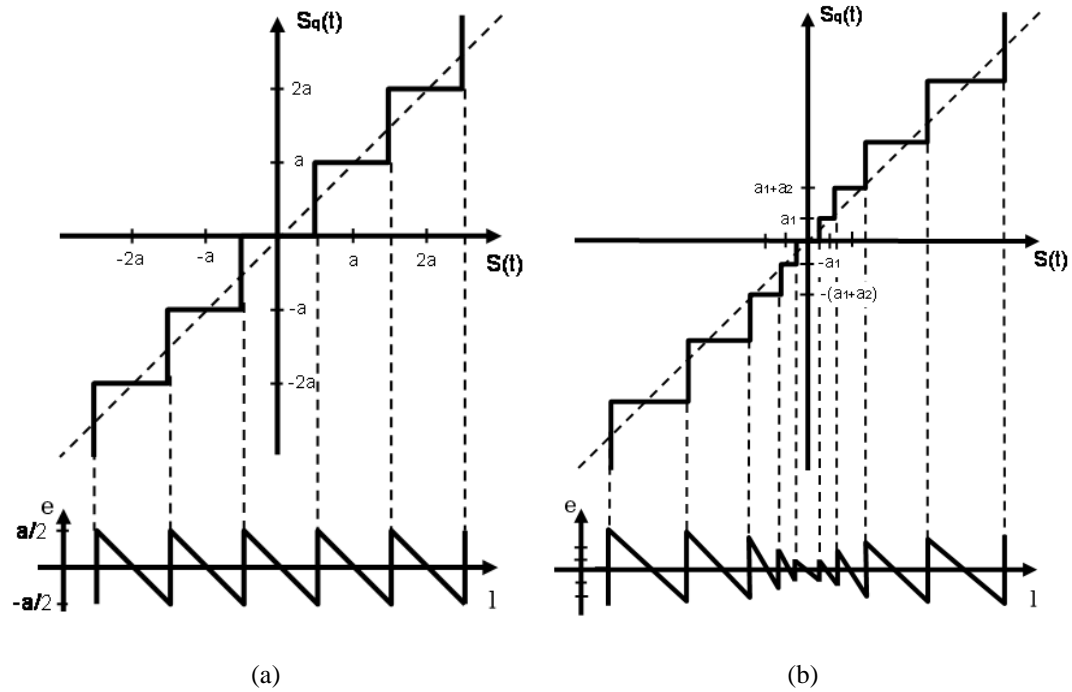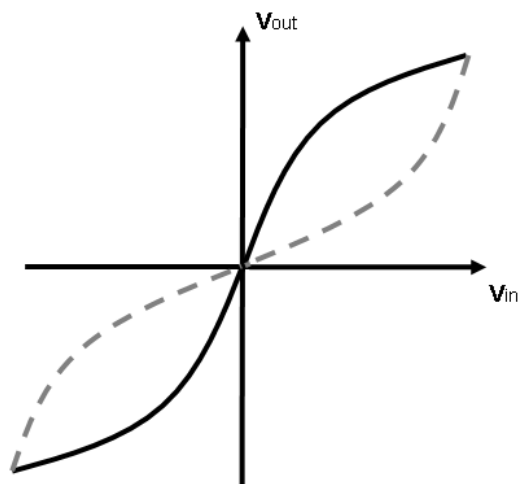Besides quantization noise, an image can also be distorted during quantization due to the fact that the dynamic range of the input signal is larger than that of the quantizer. There are also additional noise factors linked to the quality of the electronic devices which implement the quantization.

2.2.4   Image noise and image distortions

The image which was acquired in a digital format is only a distorted 2D representation of the viewed scene. The information which is contained in a viewed scene goes through a series of transformations before being stored in a digital format on a specific storage medium. Each transformation introduces specific distortions which need to be eliminated, or at least one has to be aware of their existence, before processing and analyzing the image information.

A distinction is made between two types of distortions: blurring and noise. The former is usually appears at the level of the image and the latter at the level of an image element. There are also geometrical distortions due to, e.g., optics and distortions due to the architecture of the image sensor as well as distortions due to digitization. Digitization errors affect all images, irrespective of the primary information carrier, i.e., visible light or X-rays. They include sampling-alias and the sampling low-pass and quantization noise (see Section 2.2.3). Other errors are specific for the acquisition of visible light and X-ray images respectively, hence the two classes are discussed separately.

Image acquisition with visible light

Noise and distortions due to the scene.   The number of photons which leave a certain point on an item of the viewed scene and hit the surface of a detector in a specified time interval and under constant flow of electromagnetic radiation is not constant but a random variable with a Poisson distribution. This is called *photon noise* and is linked to the physical processes which take place when light is reflected, emitted, etc. When the number of photons hitting the sensing element increases, the Poisson distribution becomes similar to a Gaussian distribution, but with signal-dependent variance.

Noise and distortions due to the camera.   After leaving the scene, to obtain an image, the light needs to go through the aperture of the camera. At the aperture diffraction phenomena occur which result in blurring the image. This *camera-blurring* depends on the size of the aperture and on the distance to the image plane. This is shown in Figure 2.23.

Noise and distortions due to the optic.   An optical system introduces a series of distortions in the image (see Section 2.2.1) which include also diffraction phenomena at the opening of the lenses. The cumulated effect of all this distortions can be analyzed using linear systems theory as the formation of an image by an optical system is a linear process. Therefore it is sufficient to know how one point is imaged by the optical system. As the optical system has a low-pass behavior, it smears the point spreading it over a larger surface, hence the name of Point Spread
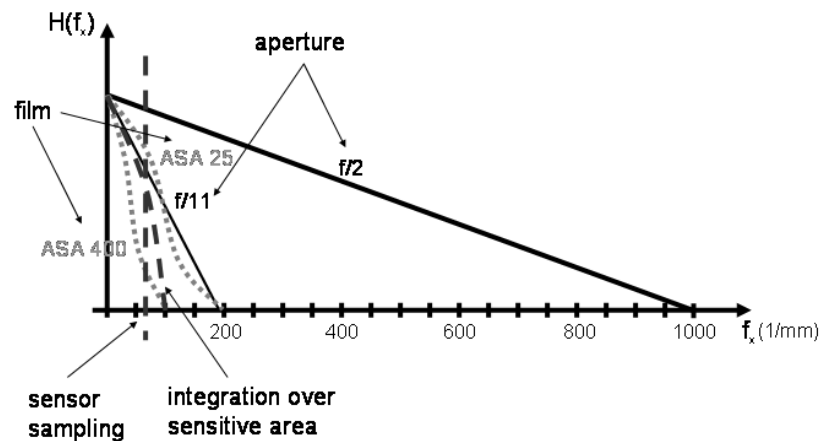
Figure 2.23: Blurring induced by different phenomena in a camera for an usual electronic sensor and for photographic films of different sensibilities.

Function (PSF). What is normally used to describe optical system is the Fourier representation of the PSF called the Optical Transfer Function (OTF) (see Figure 2.23).

The optical system can also introduce *geometrical-distortions* in the image which can be normally compensated by transforms as described in Section 2.2.1.

Noise and distortions due to the sensor. When a visible-light photon hits the surface of a sensor, than it will set an electron free with a certain probability. This electron in turn will contribute to the charge of the detector with a certain probability. The number of detected electrons for a certain number of incident photons is a random variable which follows also a Poisson distribution. Together with the photon-noise this *detection-noise* generates most of the noise in an image. Given the typical number of incident photons, this noise is considered Gaussian distributed. It has a flat power spectral density and it is uncorrelated, therefore is also called white.

A detector generates a signal even when the light flux is zero, due to thermal agitation, this is called dark current. The dark current is not constant but it varies also, thus it is different from sensor-site to sensor-site. As the minimum radiative flux that can be measured at a sensor site is linked to the variation of the dark current, this contributes clearly to the noise in the image. This is called then *thermal-noise*.

Some CCD sensors in particular, show distortions due to their specific architecture and manufacturing process. If a powerful light source is imaged, than charge can spill from one sensor site to another causing thus a *blooming-effect*. Also, while transporting the charge corresponding to the bright spot, it will leave a bright trace in the image, causing thus a *smearing-effect*. This happens for both frame-transfer sensors and interline-transfer sensors. In the latter case this is due to the fact that the storage elements are not perfectly screened from light. CMOS sensors are normally robust with respect to blooming and smearing. Other problems may appear if an interlaced scanning pattern is used, than moving objects will show additional distortions at the borders.

As already pointed out, quantum detectors integrate light over a certain area. In the case of an image sensor, whose sensitive area can be seen as a collection of quantum detectors in

Figure 2.24: Depth of field for a camera with a discrete sensor – with permission from [2] .

a matrix setup, this is equivalent to applying a low-pass filter on the image acquired (see Figure 2.23). If the cut-off frequency of this low-pass filter, f_{qd}, is lower than the cut-off frequency of the optics-low-pass, the optics-in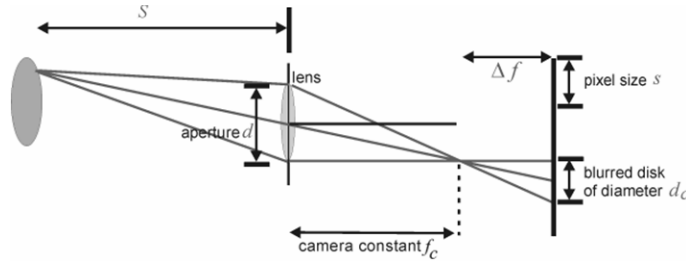duced attenuation of frequencies above f_{qd} will not affect the quality of the recorded image. Nevertheless, the band limitation by the sensitive area is not sufficient to eliminate alias. The optical system has to be adapted also to the distance between two quantum detectors, i.e., the sampling frequency. This allows for a certain maneuver room when choosing or designing an optical system in relation to a certain image sensor.

As shown in Figure 2.24 the sensitive area of a quantum detector gives also the limits of the depth of field of a digital camera (see also Section 2.2.1).

Image acquisition with X-rays

Noise and distortions due to the scene.   X-rays interact with matter by several mechanisms (see Chapter 2.2.2) and with a certain probability. Scattering produces the most important image distortions. Compton scattering leads to spurious X-ray photons hitting any position in the image plane with a certain probability and thus modifying the number of quanta per detector surface contributing thus to the photon noise. Rayleigh scattering leads to a *broadening* of the X-ray beam which is equivalent to a low-pass filtration of the image.

Noise and distortions due to the intensifier.   If an image intensifier plus digital camera combination is used, than the image intensifier introduces some geometrical distortions of the image which need to be compensated by transformations (see Section 2.2.2).

Noise and distortions due to the sensing element.   Similar to visible light, X-ray sensing elements exhibit also noise. An X-ray quanta will generate a response be it a photon for scintillators or charge for photoconductors only with a certain probability. This probability depends on the energy of the incident X-ray radiation, as shown in Figure 2.25 for a CsI scintillator and a Se photoconductor. Even if a response is generated, it will be affected by some physically motivated fluctuations. All this fluctuations are known as gain-fluctuation noise [75].

Scintillators in particular are also responsible for blurring the image. Depending on the incidence angle of the X-rays, a X-ray photon may be absorbed at different depth. Then, depending on the absorption and thickness of the detector-layer, the response varies with the absorption depth. The photon which is generated after an interaction in the scintillator does not follow a direct path on the way to the detector. Because it goes through several interactions before leaving the sensory-layer, it registers a lateral diffusion. If, e.g., an electron from one of the internal
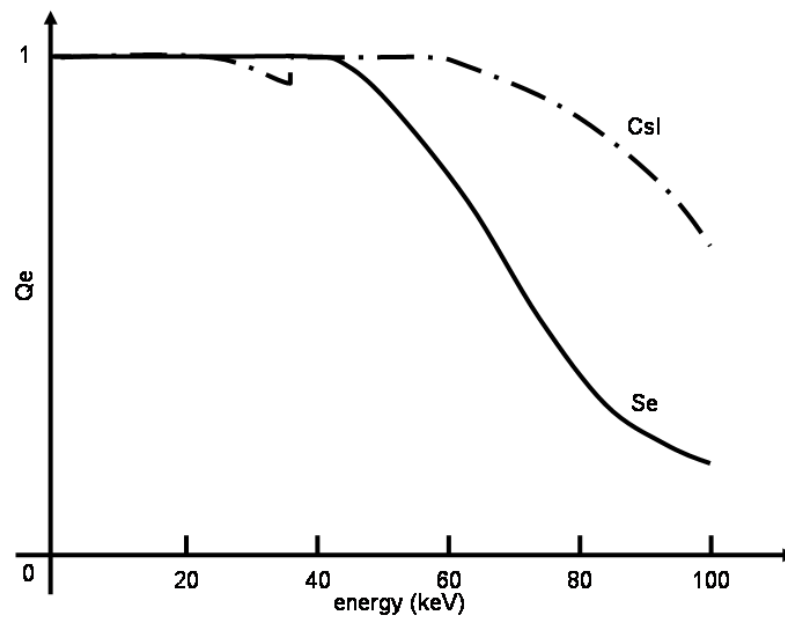
Figure 2.25: Quantum detection efficiency for a 1 mm thick layer of Se (continuous line) and CsI (dash-dotted line) – after [75].

layers of an atom is hit, than it becomes free and a characteristic X-ray photon is produced simultaneously which can again set free another electron and a photon in the vicinity and so on until a light-photon is produced and it leaves the material. Such *blurring* can be analyzed using linear systems theory as in the case of optics for visible light. To describe it the Modulation Transfer Function (MTF) is computed, which is actually the absolute value of the Fourier transform of the corresponding PSF. One of the advantages of photoconductors is that their MTF does not depend on the thickness of the sensory layer, but it does depend on the energy of the incident variation. This is shown in Figure 2.26.

2.3 Image processing

At the end of the imaging chain, the image exists as an array of numbers in binary format stored on a digital media. The image information can now be processed by a computer to achieve a myriad of purposes. The methods used for processing are gathered under the term Digital Image Processing. Such methods can be applied to any 2D data, like e.g., CT slices and are thus not restricted to projection-images.

Image processing covers the following basic classes of problems [110]:

1. Image representation and modeling

2. Image enhancement

3. Image restauration

4. Image analysis

5. Image reconstruction
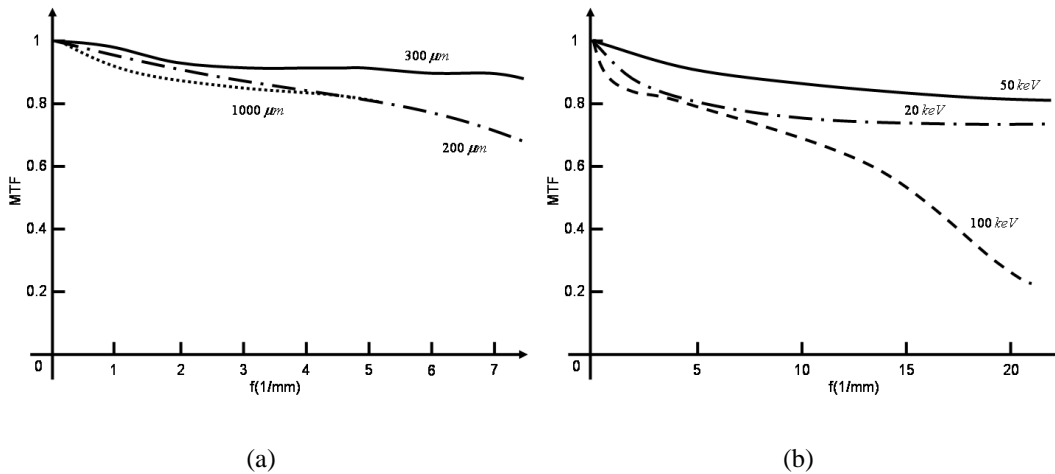
Figure 2.26: MTF for a Se photodetector for different thickness (a) and energies (b) at $500\mu m$ thickness – after [75].

6. Image compression

In the context of machine vision, we make a further distinction between *image analysis* and *image conversion* problems which includes all other categories described above and from which we are interested only in image enhancement. *Image conversion* methods accept as input an image and return a modified version of the input, so in this case, there are images both at the input and at the output. *Image analysis* methods accept as input an image – typically the result of image enhancement – and return a description of items in the image.

2.3.1 Image enhancement

The *purpose of image enhancement* is to make objects conspicuous so that they can be easily separated from the rest of the items present in an image. Typically some information on the representation of the objects in the image is available and should be used to support the choosing of a method or set of methods for image enhancement. For example in an edge detection approach, the objects are the edges, which are known to be regions exhibiting sharp transitions. Then a linear high-pass filter can be used to enhance them.

Image enhancement techniques can be divided into three large classes [110]:

- **Point operations**, which include methods where each gray level of the image is modified according to a certain function, like e.g., contrast stretching and histogram equalization.

- **Spatial operations**, including methods where the value of each element of the image is recomputed using information from a set of neighbors situated in a certain neighborhood. Examples are: filtering with linear time invariant systems, median filtering, morphological processing, etc.

- **Transform operations**, which include methods whose typical implementation requires first the representation of the image in a transformed normally orthonormal space, followed by processing and inverse transformation. This category includes, e.g., some non-linear filtration techniques like homomorphic filtration.

2.3.2 Image analysis

The *purpose of image analysis* is to provide a description of objects providing thus the basis for a consecutive image understanding step, where objects are identified and eventually the relationship between them is established.

Depending on the application an intermediary *segmentation* step may be necessary where objects are separated from other items in the image, i.e., object detection. To conduct this step, one assumes some prior knowledge on the physical-properties of objects which is then related to certain image attributes. The description of objects is then based on single or on combinations of image attributes which are called *features*.

Therefore, one can conclude that image analysis is usually based on some a priori knowledge with respect to objects and more precisely on specific expectations which arise from this knowledge.

Image features

Features can be defined as meaningful detectable attributes of the image. In machine vision the term image feature refers to two possible entities [188]:

- **Global or direct features**, which are general image related attributes, which could also be defined for images showing no objects.

- **Local or composed features**, which typically include a segmentation step to induce a local behavior and contain particular, object-related attributes. These can not be defined for images showing no objects.

Composed features which are arguably more often used than global features, can be further divided into:

- **Low-level features**, which are computed independently using only information extracted from the image function, like e.g., edge-points

- **Mid-level features**, which are defined based on a raw relationships among low-level features, like e.g., perimeter and area.

- **High-level features**, which are usually computed starting from mid-level features and describe object properties, like e.g., its similarity to a circle, its elongation, etc.

Image analysis includes methods to extract and select image features. Feature extraction can be based, e.g., on the amplitudes of the image signal, on the histogram, it may include a transformation and/or a segmentation step and in general they may be as different as the objects they should describe.

Image segmentation

Image segmentation is in many applications an important step of image analysis. The binary image obtained after separating the objects, can be used to perform different measurements, which are usually linked to the geometric shape of each object. Depending on the application, such measurements can be used as such or as features to support a consecutive recognition step.

There are pixel-based methods, which use only the gray-level of a pixel to decide if it belongs to an object, region-based methods, which consider gray-levels in a certain neighborhood, edge-based methods, which detect the contours delineating the objects and model-based/template matching methods, where it is assumed that the geometric shape of the sought object is known and this is then searched for iteratively in the entire image.

Pixel-based segmentation.   A pixel-based segmentation is usually achieved by thresholding the image such that the object and background pixels are divided into two classes. Each class contains pixels which are similar in some respect. These similarities, which motivate the threshold, can be associated to some physical properties of objects and background, like e.g., those related to the way they interact with electromagnetic waves – i.e. how they reflect or emit light. When background and object differ in this respect, accordingly, their pixel gray-levels in the image are different and the two can be therefore separated.

For example, knowing that an objects reflects well light leads to the expectation of finding the object among lighter items in the image. Building upon these expectations a threshold with a high value should select only pixels of object-like structures. Thus, the knowledge with respect to the reflectance-properties of objects has been used to obtain a new binary representation of the imaged scene where only object-like items appear.

Region-based segmentation.   Region-based segmentation uses not only the pixel gray-level, but also the connectivity of objects to achieve a segmentation by, e.g., clustering pixels of similar gray-level and situated at small distances from one another to detect thus an object or at least part of its surface. In the latter case in an additional step several surfaces are linked together to form the object. This entire clustering process is usually controlled by specific similarity functions.

In principle, region-based segmentation uses additional object information – i.e. objects are represented by spatially-connected structures – to achieve a segmentation and thus in some situations represents an improvement over simple pixel-based techniques.

Edge-based segmentation.   In machine vision there is often the case that objects can be characterized by borders, which separate them from other items. Borders can be defined in different ways, usually they are made of points where the image amplitude exhibits a sharp transition, but they can also be points where, e.g., the orientation of pixel neighborhoods changes. After detecting such transition points – which is in itself a feature extraction step, where the points are considered to be the objects – they need to be gathered in border-sets. This is implicitly a segmentation step, because after the border-sets are established, the objects are defined typically as image surface enclosed by the border.

Active contours.   Boundary extraction is a very important step in image analysis and it may result in a segmentation of the image. Boundary extraction is about finding a relation among edge points and it can be seen as consisting from two problems [188]: grouping, i.e., which points belong to a boundary and model fitting, i.e., which boundary curve best interpolates the available edge points.

If precise knowledge on the topography of sought border is available, then a set of methods which try to find instances of the border in the image can be used. Such methods include,

e.g., template matching [109] and Hough transform [15] , [106], [102]. Conversely, adaptive methods are needed like the active contour models [119], [136], [171].

Active contours are self-organizing curves able to lock on the borders of objects with different shapes. They move on the image under the influence of internal energies, which are defined within the curve and external energies, which are computed from the image to a position of minimum energy, which corresponds to the sought border. Internal energy typically model elasticity and stiffness and external energies are attraction energies and are defined in relation with the nature of the borders.

The curve can be open or closed, here we restrict to discussing closed curves, which can be used for boundary detection. By their energy formulation, they behave like elastic strings with specific bending and stretching properties. Two types of active contours can be distinguished [75]: parametric ones, where explicit relations are defined between neighboring curve-model points [119] and geometric ones, where these relations are implicit allowing for improved flexibility [171]. For example, geometric contours are able to split and merge with relative ease in comparison with parametric models and can thus be used, e.g., to detect simultaneously several items in the image.

Model-based segmentation. If enough information about the objects is available to construct an object-model, then segmentation can be achieved by choosing from the image only those points which verify the model, e.g., points where the similarity between the model and the input data is maximal.

2.3.3 Morphological image processing

Morphological image processing can be divided, as function of its input, into *binary morphology* and *gray-level morphology*. Usually, the former is used primarily for image analysis and the latter primarily for image enhancement.

Binary morphology

Binary morphology can be used, e.g., to improve a segmentation by smoothing out object outlines, filling small holes and other similar techniques. It facilitates the search and selection of objects from among several candidates. It can be also used to find borders of objects or to generate features, as it provides the means for an easy quantification of form-based object attributes.

After segmentation, only information about the geometrical structure of the object is available as all other information has been used to achieve the segmentation. Morphological image processing has been initially developed to analyze the geometrical structure within binary images. The basic idea being to quantify the manner in which a structuring element-probe fits or it does not fit within the image [67]. In other words: describe a house – object – starting from the bricks – structuring elements – that make it.

Gray-level morphology

Gray-level morphology has evolved from binary morphology. To understand gray-level morphology, we must remember that with two adjacent gray levels, the brighter one is considered to

be the object (the equivalent of '1' in a binary image), and the darker one is the background (the equivalent of '0'). It is usually used to select and preserve particular intensity patterns, while attenuating others.

Currently morphological image processing represents a collection of methods applicable not only to binary images but also to gray-level images, to accomplish a multitude of tasks and it joins traditional image processing in a complementary way, sometimes even replacing it [144].

2.4   Pattern recognition

Image analysis returns a *feature-based description of objects*. Based on this description, objects are recognized and relations among them can be established. This is then called *image understanding* and it is the final purpose of machine vision. Using such vision-information, decisions can be made, like e.g., moving a robot arm to a certain position to grasp the object. In the case of medical machine vision, such information can be used, e.g., to enhance the diagnostic ability or to improve treatment.

Image understanding includes *pattern recognition*. In this case, the patterns are image objects in a compact, feature-based representation and the recognition problem is usually being posed as a classification task. The corresponding classes can either be predefined in the case of *supervised classification* or learned based on the similarity of patterns in *unsupervised classification*. In this sense, image understanding is usually a supervised classification problem. Of course, unsupervised methods can be used as well provided that in an additional step the correspondences between the learned classes and the sought objects are established.

The five best known approaches to pattern recognition are [111], [29]:

1. the statistical approach [184], [156], [87]

2. the syntactical approach [149], [28]

3. neural networks [97], [21]

4. support vector machines [38], [60], [178]

5. template matching [185]

They can be further divided into:

- **Bottom-up methods**, which in the case of machine vision start from elementary image components to return a classification of objects (normally the approaches 1–4), like e.g., the detection of certain types of cells in a microscopic image

- **Top-down methods**, which use a set of predefined prototypes – based on some prior knowledge or learned – to search for instances of objects in the image (approach 5), like e.g., atlas-based registration of hand bones.

Segmentation as a pattern recognition problem.   If the analyzed image shows only realizations of the same object and background, than detection and identification of objects, i.e., segmentation and image understanding, become the same thing. Pattern recognition methods are not used exclusively at image understanding, they can be used also for segmentation as well, because segmentation can be seen, e.g., as a two-class object-non-object classification of image pixels.

2.4.1 Statistical Classification

A pattern recognition algorithm needs to establish a class correspondence for each input feature vector. In the statistical paradigm, the *feature vectors* are considered *random variables* and usually, their class labels are established based on a posteriori probabilities, such that the *probability of error is minimized*. The mathematical expression of this fundamental strategy is the *Bayes decision rule*.

There are also *other strategies* to design a statistical classifier, then to minimize the total error. For example, depending on the application, certain errors can be more critical than others and accordingly one would like to avoid the critical errors, even if this means accepting some more uncritical errors.

The Bayes decision rule

If a classification algorithm needs to decide on the class label for a certain pattern in a two-class problem, than the best it can do to minimize the total error, is to decide in favor of the class for which the a posteriori probability is maximal:

$$\frac{Pr\left[\omega_1|\vec{x}\right]}{Pr\left[\omega_2|\vec{x}\right]} \overset{\omega_1}{\underset{\omega_2}{\gtrless}} 1 \tag{2.12}$$

As the posterior probabilities are usually difficult to estimate, one can use the *Bayes theorem*:

$$Pr[A|B] = \frac{Pr[B|A]Pr[A]}{Pr[B]} \tag{2.13}$$

to rewrite (2.12) based on the class-conditional probability density function and prior probabilities:

$$\frac{p\left(\vec{x}|\omega_1\right)}{p\left(\vec{x}|\omega_2\right)} \overset{\omega_1}{\underset{\omega_2}{\gtrless}} \frac{Pr\left[\omega_2\right]}{Pr\left[\omega_1\right]} \tag{2.14}$$

Both the class-conditional probability density function $p\left(\vec{x}|\omega_i\right)$ and prior probabilities $Pr\left[\omega_i\right]$ can usually be estimated using some training-data or they may be known a priori.

The large majority of statistical pattern recognition methods – irrespective if they decide among two or several classes – are based in some sense on the Bayes decision rule, which minimizes the probability of error. Depending on how much information is available at the beginning of the design phase of a statistical classifier, the Bayes rule is applied directly or some additional steps are needed.

Estimating the class-conditional densities. If the class-conditional densities and the priors are known in advance, then the Bayes decision theory can be applied directly. If they are not known, then they may be learned, either in a supervised, or in an unsupervised manner. Provided the mathematical form of the sought *pdf* is known, then the specific parameters, like e.g., mean and variance can be estimated to obtain a parametric estimation of the *pdf*. If the mathematical form is not known, then specific estimation methods, like e.g., Parzen estimation and k-nearest neighbor should be used to obtain the *pdf* in a nonparametric approach.

Bayes classifiers.   A supervised parametric approach gives so called *plug-in rules*, which have been previously computed and depend on the parameters of the known distribution functions. The estimated parameters are then directly replaced to obtain the sought classifier, like e.g., for the linear and quadratic Gaussian classifiers [184].

The supervised nonparametric approach returns classifiers which use Equation (2.14) with the class-conditional *pdfs* estimated nonparamterically, or they use some derivate rules like in the case of the k-nearest neighbors classifier [87].

The unsuparvised parametric approach results in a *mixture model analysis*, like e.g., expectation-maximization [21], where typically the number of modes is unknown. However, it can be determined from the data by specific methods, like e.g., minimum message length [152].

The unsupervised nonparametric approach yields *cluster analysis methods*, like e.g., the k-means algorithm [112].

Other decision criteria

Sometimes, e.g., for errors with different relevance degrees or when one does not know the class conditional *pdfs*, the Bayes decision rule can not be applied to design a statistical classifier. In such cases criteria other than the minimization of the total error should be used. We discuss here shortly a few such examples, which are relevant for this work.

Bayes risk minimization.   The decision rule in Equation 2.14 is formulated in terms of the class prior probabilities. To account for a different degree of relevance among errors of different classes, one can establish a decision rule to minimize the Bayes risk. This risk is defined in relation to a cost value previously assigned to each type of decision. The cost C_{ij} will be larger for wrong decisions ($i \neq j$) than for correct decisions ($i = j$) and one can assign a larger cost for that type of error which is more critical. The two-class decision rule to minimize the Bayes risk will be [184]:

$$\frac{p\left(\vec{x}|\omega_1\right)}{p\left(\vec{x}|\omega_2\right)} \overset{\omega_1}{\underset{\omega_2}{\gtrless}} \frac{Pr\left[\omega_2\right] \cdot \left(C_{12} - C_{22}\right)}{Pr\left[\omega_1\right] \cdot \left(C_{21} - C_{11}\right)} \tag{2.15}$$

The decision rule in Equation 2.15 is strongly related to the Bayes decision rule, as it still considers the total error. By imposing certain cost values and experimenting with them, one can compute decision rules such that certain desired error probabilities for certain classes are achieved.

Neyman-Pearson theory.   If from the beginning a certain error probability is desired for a certain class and the class conditional *pdfs* are known, then the Neyman-Pearson decision rule can be used. For a decision among two classes, this rule is computed – by fixing one class error probability, say ϵ_2, and seeking to minimize the other (ϵ_1) – as [184]:

$$\frac{p\left(\vec{x}|\omega_1\right)}{p\left(\vec{x}|\omega_2\right)} \overset{\omega_1}{\underset{\omega_2}{\gtrless}} \lambda \tag{2.16}$$

where λ is a constant, characteristic for the decision rule. One can show that by varying λ, a curve in the space of class error probabilities $\epsilon_1 \epsilon_2$ is computed, i.e., the *receiver operating*

characteristic (ROC)[16] of the decision rule. Using the ROC, one can also formulate a decision in terms of the desired class error probabilities, rather than in terms of class prior probabilities.

Significance test.   If two mutually excluding classes need to be separated and only the conditional *pdf* of a single class ω_k is known, then one can use for classification a significance test. The null hypothesis H_0 of this test is that the investigated sample belongs to ω_k. Over the significance α one can set the probability of error given H_0 and thus establish a decision rule.

Separability analysis.   If the class-conditional densities are not known, then one can find the parameters of a decision rule by optimizing a separability criteria, rather than minimizing class errors (see also Appendix A.6). The Fisher linear classifier (see Section 4.3.2 at page 143) represents a well-known example for this classifier-design strategy.

2.5   Issues of algorithm performance

The purpose of vision is to provide an interpretation of a scene. The interpretations given by machine vision algorithms are subject to *performance claims*, which should show that they are conform with a certain target interpretation[17]. In the cases of interest here, this target interpretation is normally obtained as the result from an analysis by a trained human observer.

The quality of a machine vision algorithm can be evaluated either by analytical or by empirical methods [35], [202]. *Analytical methods* examine the algorithm by analyzing its principles and properties and thus establish if it can reach the desired interpretation. *Empirical methods* examine the algorithm by analyzing its output. They can be further classified into two classes: goodness methods and discrepancy methods. *Goodness methods* use the algorithm's response to a certain input to evaluate some quality measure [127], [128], [166] and *discrepancy methods* compare the algorithm's output against a ground truth [18], [200], [203].

Algorithms for *cardiovascular biomedical image analysis* are usually supposed to take over some visual analysis tasks from human specialists. For this purpose, they need to perform similar to a human observer and thus the best way to test them is by *discrepancy methods*.

2.5.1   Performance metrics

In the following, it is assumed that the data sets used for evaluation are *large enough* to inspire confidence and different performance claims are made on *similar data*, which is *representative* for the clinical problem. We discuss different performance metrics. In this context a distinction between three different categories of problems, each normally requiring different performance metrics, can be made [75]: measurement problems, image-classification problems and image-segmentation problems.

---

[16]This is actually an "operating characteristic", the term "receiver operating characteristic" has been pinned by the traditional usage of such decision-making strategy in the field of signal detection for telecommunications and radar.

[17]Performance claims with respect to the speed at which an interpretation is available are beyond the purpose of this discussion.

Measurement problems

A machine vision algorithm designed to solve such problems should return a scalar or vector quantity, which normally describes an object from the analyzed image. Its performance can be evaluated by comparing the result against an independent measure of the true value. For this purpose, the mean relative error (MRE), but also the mean squared error (MSE) can be used. The MRE is defined as:

$$MRE = E(\frac{|r - t_v|}{t_v}) \tag{2.17}$$

and the MSE is defined as:

$$MSE = E((r - t_v)^2) \tag{2.18}$$

with r the result, t_v the true value and $E(\cdot)$ the expectation operator. For vectorial quantities one obtains an error vector. In this case, the magnitude of this vector can be used as a scalar measure of the error.

Image classification problems

In this case it should be specified if the analyzed image does or does not belong to a certain category. Although the number of categories can be large, in essence this is a binary decision, which has only *two outcomes*: true if the classification was correct or false conversely. The performance of algorithms designed to solve such problems can be evaluated based on the number of correct and incorrect decisions, provided some pre-classified data set, called the test set, is available.

Multi-class classification.   In the case of a multi-class classification problem, the percentage of correct classifications per class and also the percentage of correct classifications on the entire test set can be used as performance metrics.

Binary classification.   In the case of a binary classification problem, the rates of true and false positives (TP/FP) and that of true and false negatives (TN/FN) can be used. A true classification is a correct decision and a false classification is an incorrect decision, either for the positive class or for the negative class. Usually, the positive class is the object-class and therefore, sometimes TP are also called correct classifications (CC). It is customary to specify only the TP and FP rates. Typically in an object no object scenario, a decision in favor of the object is a positive decision. If the decision was correct, then a TP was registered, otherwise a FP was registered.

Composed metrics.   From the TP/FP analysis, several additional performance metrics can be derived, like the sensitivity, the specificity, the precision, the accuracy or metrics based on the ROC (see Section 2.4.1).

The *sensitivity* is defined as:

$$S_e = \frac{N_{TP}}{N_{TP} + N_{FN}} \tag{2.19}$$

the *specificity* is defined as:

$$S_p = \frac{N_{TN}}{N_{TN} + N_{FP}} \tag{2.20}$$

the *precision* is defined as:

$$Pr = \frac{N_{TP}}{N_{TP} + N_{FP}} \tag{2.21}$$

and the *accuracy* is defined as:

$$Ac = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{FN} + N_{TN} + N_{FP}} \tag{2.22}$$

with N_{TP}, N_{TN}, N_{FP}, N_{FN} the number of TPs, TNs, FPs and FNs respectively.

An ideal detection will have: sensitivity, specificity, precision and accuracy equal to one (or 100%). With respect to sensitivity and specificity, it is pointless to specify only one of the two, as perfect sensitivity is achieved by an algorithm deciding always for the object class and perfect specificity can be achieved by an algorithm deciding always against the object class. An algorithm with perfect precision decides always correctly for the object class, but not necessarily for all members of this class. An algorithm of perfect accuracy decides always correctly for and against the object class, for all members of the two classes.

Normally, there is a tradeoff between sensitivity and specificity which is controlled by a certain parameter in the algorithm. In such a case, the algorithm's performance can be evaluated using the *ROC* which gives the TP against the FP rate when the parameter varies such that the FP rate goes from zero to 100%. Using the ROC, performance metrics such as the *area under the ROC* (AROC) can be defined. Clearly, the larger the area, the better the performance, but sometimes it is better to use only a certain *part of the AROC* (pAROC), e.g., when there is a certain maximum acceptable FP rate or a minimum acceptable FN rate or a combination of the two.

Image segmentation problems

This type of problems involve dividing the image in a series of distinct regions. In this case, it is not an entire image, which should receive an interpretation, which can be erroneous or not, but only some subsets of the image. Thus, for one and the same image one can achieve both correct and false results. Performance metrics for such cases are more difficult to design, they are based, e.g., on the number of elementary regions (e.g. pixels) that have to change their class so that the achieved result is in agreement with the ground truth, or on the mutual overlap between ground truth and result. Further details can be found in [75] and [35].

In the case of binary segmentation, the image should be divided in only two regions, i.e., object and background. Then, for each pixel of the image only two outcomes are possible and performance metrics based on TP/FP rates can be used as described previously.

2.5.2   Measuring performance in practice

Performance claims are supposed to either permit a reliable comparison between algorithms or a reliable prediction of performance in clinical practice. There are several factors which influence the credibility of a performance claim and can thus cast a shadow on the reported success of a certain machine vision algorithm.

Optimist bias for a performance claim

If a performance claim is based on *synthetic data*, then it typically shows that the algorithm matches the assumptions made when generating that data. However, it says nothing about how the algorithm will behave under natural conditions. The natural conditions may be different enough to make any attempt to precisely model them futile. Thus synthetic data should be used *only to design* the algorithm and *not during testing*.

An optimist bias can be obtained also if the data set on which the performance claim is made is chosen in such a way that it *particularly supports* a certain type of modeling, which in turn led to the design of the algorithm.

If a machine vision algorithm learns how to interpret some data from a set of examples, the performance claim should be made on a data set different than that used for learning. Usually there is only one data set available for designing the algorithm. To make a performance claim, this set is partitioned so that *training* and *testing* the algorithm is done on *different data*. There are several ways to achieve such a partition including, e.g., dividing the set half-half. Further details on methods such as cross validation and leave one out can be found in [75].

If a performance claim is defined based on some subjectively chosen criteria, then these criteria should be properly chosen such that neither an optimist nor a pessimist bias appears.

Pessimist bias for a performance claim

Performance claims made on *particularly difficult* data sets are pessimistically biased only provided that data set is really an *exception*. Data sets which contain *outliers* can lead to pessimistically biased performance claims. Such outliers may be generated, e.g., due to the imaging procedure and end up playing a role because the data set was inadvertently cleaned.

Most often a pessimist bias appears due to the *incorrect ground truth*. For example, it may lack the detail which is attained by the algorithm or some ground truth is simply wrong. Normally, the reason for this is the human operator, who is supposed to establish this ground truth and who may get bored or tired and make mistakes.

2.5.3   Designing a test to sustain a performance claim

Each algorithm contains usually a set of parameters which influence its behavior and make it thus adaptable to different data. Before making a performance claim, this set of parameters needs to be defined. For this purpose, a sample of real data is needed together with some *objective performance measures* defined over it. The performance measure is then computed for each value of the parameters in a certain variation range, obtaining thus an *error-curve*. The parameter set which yields an *extremum* of the error curve – minimum or maximum depending on the performance measure – should then be chosen.

The variation range of the parameters should be *discretized* in steps small enough to properly sample the behavior of the algorithm and large enough to permit a timely computation of the error curves. If the performance claim is supposed to be used to compare two algorithms, then both parameter sets need to be defined during tests using *the same data*.

Chapter 3

Enhanced catheter intervention

A catheter intervention represents currently the state-of-the-art treatment procedure for the artery disease[1]. Depending on which part of the body is supplied with blood by the affected vessel, the artery disease becomes: the coronary artery disease, the cerebral artery disease or the peripheral artery disease.

We concentrate on machine vision-based improvements for catheter interventions aimed at treating the coronary artery disease and the myocardial infarction (Section 3.1). Dividing a catheter intervention into three phases: diagnostic, intervention and evaluation, we describe methods to support the last two phases. We show how to improve the perception of visual information during the intervention by presenting the physician a dynamic vessel-roadmap where his tools are shown superimposed (Section 3.2) and how to extend his evaluation abilities by automatically and precisely evaluating measures of revascularization of the cardiac muscle (Section 3.3) following treatment of a heart attack.

3.1 Introduction

3.1.1 The coronary artery disease

The coronary artery disease occurs when the arteries that supply blood to the heart muscle (see Figure 3.1) become hardened and narrowed. This is due to the buildup of plaque on the inner walls or lining of the arteries, which is called atherosclerosis.

There are several pathologically relevant consequences of this disease. For example, the blood flow to the heart is reduced resulting in cardiac ischemia and therefore in the decrease of oxygen supply to the heart muscle[2]. Left untreated over a longer period of time, cardiac ischemia may contribute to arrhythmias and heart failures.

One of the most dire consequences of the coronary artery disease is a heart attack. This may happen if a blood clot suddenly cuts off the entire blood supply to a part of the heart muscle. Cells in the heart muscle that do not receive enough oxygen begin to die. This can cause permanent damage to the heart muscle and possibly directly lead to death.

Treatment of the coronary artery disease includes a change in habits for the ill person[3],

---

[1]Sometimes also called the vascular disease.

[2]One of the early symptoms of cardiac ischemia is angina pectoris.

[3]The ill person should, e.g., avoid prolonged stress, eat a healthy diet to prevent or reduce high blood pressure and high blood cholesterol, and maintain a healthy weight.
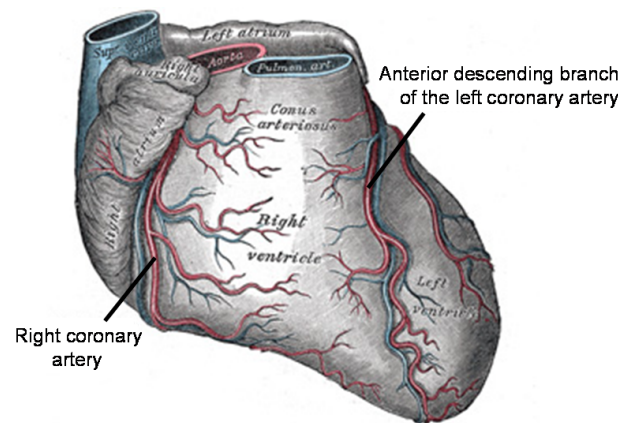
Figure 3.1: Human heart with coronary arteries – after [91].

medication[4] and surgical procedures. Usually, significant narrowing of the coronary arteries requires surgical treatment. If multiple blocked coronary arteries are found during a diagnosis phase, then a coronary artery bypass graft is recommended. If only one or two coronary arteries are ill, then a Percutaneous Transluminal Coronary Angioplasty (PTCA) is effectuated.

During a coronary artery bypass graft healthy blood vessels taken from elsewhere in the body are joined to the affected coronary arteries in such a way that the clogged areas are by-passed.

In PTCA, a balloon is inserted through the skin into a blood vessel and maneuvered to the site of the narrowing. There it is threaded into the blockage and inflated, compressing thus the plaque against the arterial walls. Finally, in most cases a stent is placed to keep the artery open. Such procedures can improve blood flow to the heart, relieve chest pain and possibly prevent a heart attack.

3.1.2 Percutaneous Transluminal Coronary Angioplasty

Percutaneous means through the skin, transluminal, within the artery, coronary, refers to the artery being treated and angioplasty means that the artery is remodeled. 657,000 PTCAs were performed in the US in 2002 and their number has surpassed the 500,000 limit in the last years in Europe and it still maintains an increasing trend [134].

PTCA has evolved from the catheterization procedure [146]. During such a procedure a catheter – i.e. hollow flexible tube – is inserted into the body. If the catheter is supposed to reach the heart than this is called cardiac catheterization. During PTCA, the catheter is advanced only to the ostium – i.e. entrance – to the coronary vessel tree. After the catheter is positioned, a guidewire with a easily steerable tip is advanced to the place of the lesion. Then, a balloon is advanced on the guidewire, through the catheter and then through the coronary arteries, until it reaches the position of the lesion where it is inflated.

PTCA in its current form was made possible by a *series of developments* both in the medical science but also in the field of non-invasive medical imaging. The modern PTCA depends

---

[4]Medication may include drugs that dilate the blood vessels of the heart, drugs which control the high blood pressure, cholesterol-lowering drugs, anticoagulants to reduce the risk of a blood clot being formed and thrombolytics to dissolve the clots that can occur during heart attacks.

largely on *high-performance imaging equipment* and although it is a mature procedure, conducted on a daily basis in many hospitals, there is still room for improvements. Particularly the introduction of *computers* to optimize standard medical equipment offers new almost *limitless possibilities for improvements*. *Medical image analysis* is at the heart of many such improvements.

A short history of PTCA

Initially the cardiac catheterization was used to measure the intra-cardiac pressure in animals. In 1929 the first documented cardiac catheterization on humans was performed with the purpose of finding a way to efficiently administer drugs to the heart. The procedure was documented by taking an X-ray picture of the catheter in the body when it reached the heart. In 1941 the cardiac catheterization began to be used for diagnostic purposes, e.g., to measure the cardiac output.

In 1958 the coronary angiography was introduced as visualization method for the coronary arteries. While observing the heart-region under X-Rays during a cardiac catheterization, a contrasting dye is introduced into the coronary arteries with the help of the catheter and thus one is able to observe the anatomy of the coronary vessel tree in a *coronary angiogram*.

The transluminal angioplasty – i.e. the concept of mechanically remodeling an obstructed vessel – was first proposed in 1964. In 1967 it was elaborated the "femoral approach" to coronary angiography, i.e., introducing the catheter via a groin puncture rather than through the arm. Then in 1977 it was developed the first balloon catheter whose controlled dilations were able to eliminate stenosis.

All these developments, together with the introduction of fluoroscopic image intensifiers, led to physicians performing the first PTCA on an awake patient in 1977. The precise positioning and guiding of these initial balloons was quite difficult, consequently new guiding catheters are developed in 1980 and then in 1982 over-the-wire coaxial balloon systems are introduced and steerable guidewires are developed.

Although the stenosis could be successfully eliminated by inflating the balloon, the physicians were still confronted with two problems. First, there was the collapse of the treated artery after the intervention, in such cases an additional bypass graft intervention was necessary. Second, there was the restenosis, i.e., the vessels close up again short time after the intervention. To sustain the artery after the intervention and stop the restenosis, a stent – i.e. a metallic mesh tube – was used for the first time in 1986. The stent is expanded by a balloon at the position of the blockage and is thus able to sustain the artery and eliminate many complications, but it only reduces the restenosis rate.

The latest advances in this field have seen the introduction of drug-eluting stents as of 2003. Such a stent has a drug coating which interferes with the restenosis process.

The PTCA procedure

PTCA is performed either in common cases after the coronary artery disease has been dignositcized or in emergency cases when a heart attack has occurred [120].

Typically, a patient accusing the symptoms of the coronary artery disease, like e.g., chest pain first goes through a series of non-invasive tests. If these tests cannot rule out the possibility of a blocked or narrowed artery, an angiographic exam is performed. Such an exam is an

invasive test and is performed in the catheter laboratory under X-ray supervision. The catheter laboratory is a special room in the hospital outfitted with X-ray imaging equipment and monitors to directly see the X-ray images (See Figure 3.2). Next to this room there is an observation room outfitted with computers where various software tools, aimed at helping the physician during the operation, are installed and an anteroom where the patients are prepared for the intervention and brought in for intensive care directly after the intervention. There are also several more dependencies whose purpose varies from hospital to hospital.

The imaging equipment. PTCA relies on digital X-ray imaging to facilitate a precise positioning of interventional tools. An imaging chain consist of a X-ray generation unit, a X-ray detector (either a FPD or an image intensifier-digital camera pair – see Section 2.2.2) and an image processing unit, i.e., a host computer where dedicated software is installed. The X-ray generation unit and the X-ray detector are mounted on a C-arm (see Figure 3.2), which can be easily positioned by the physician. The orientation of the C-arm can be defined by the primary and the secondary angle. With respect to the patient, the primary angle corresponds to a rotation in the transversal plane and the secondary angle to a rotation in the sagittal plane. If the primary angle describes a tilt to the right then it is called the right anterior oblique (RAO) and if it describes a tilt to the left then it is called the left anterior oblique (LAO). If the secondary angle describes a tilt towards the head of the patient, then it is called cranial (CRAN), conversely it is called caudal (CAUD). During image acquisition the irradiation of the patient can be either continual or frame-rate-gated such as to reduce the time when he is exposed to X-rays. Usually, the images are discretized on eight to 16 bits on a 512×512 to 1024×1024 grid and with a frame rate of 10 to 60 fps.

Depending on the clinical application for which the images are acquired, the X-ray dosage can be increased, thus achieving images of a better quality, or decreased to reduce irradiation. To keep the image noise and the image brightness constant the X-ray dosage can be modified by a feed-back loop. Brightness constancy above the maximum radiation dosage is ensured in the signal acquisition chain by an automatic gain control unit (AGC). The automatic correction of image brightness, irrespective of how it is done, is considered by extension to be obtained by the AGC.

In the beginning, to store the interventional images for later analysis and postoperative studies, a camera with a rapid film changer was used. The images acquired had a very good spatial resolution, but since the digital technique has become ubiquitous, its obvious advantages had led to the disappearance of film-cameras from catheter laboratories.

The diagnostic-phase. In diagnostic imaging film-cameras have resisted a longer time. In such a case, the X-ray dosage is increased to obtain images of a better quality. At the same time, a contrast agent – usually an iodine-containing compound with maximum iodine concentration of $350\, mg/cm^3$ [27] – is injected in the coronary vessel and uniformly mixes with blood making thus the vessels visible under X-ray. Therefore, such a diagnostic-sequence contains mostly complete coronary angiograms showing the entire vessel tree.

These images are then recorded for some five to 10 seconds at a frame rate of 15 to 30 fps. Such sequences are usually called *cine-runs*. Sometimes the word "cine" is used to describe a coronary angiogram irrespective if it was acquired during the diagnostic phase or during the intervention proper and a distinction is made between cine images and *fluoroscopic* – or in short "fluo" – images, which are recorded with a lower X-ray dosage and mostly without contrast

agent, therefore only rarely showing vessels. Presently, film cameras are replaced by digital cameras and thus digital cine runs are obtained.

To set a diagnostic properly, several recordings from different angles – i.e. different geometric setups of the X-ray imaging system – are needed. However, the sequences recorded at this stage still cannot cover all position that the C-arm covers during an intervention. Each time a diagnostic sequence is recorded, contrast agent is injected into the vessels. For children in particular, the contrast agent can be poisonous in large quantities. To decrease the contrast load of the patient, biplane systems can be used. Such systems have two orthogonally positioned imaging systems – i.e. X-ray generation unit and X-ray detector – and acquire thus two sequences of different projections per contrast injection. They are however more difficult to position and sometimes the two imaging systems interfere generating images of lower quality.

The intervention. During an angiographic exam several images of the vessel tree are acquired with the imaging system set for optimal image quality, i.e., in the so-called *cine-mode*. If a narrowing or blockage appears during the exam, this is usually transformed on the spot into a PTCA, after making some small equipment changes.

During PTCA, the X-ray dosage is reduced to protect the patient and the physician, resulting in a lower image quality. This image acquisition mode is sometimes called *fluo-mode*. The physician chooses one of the previously acquired complete coronary angiograms as roadmap. This is then displayed on a screen next to the one showing live images. Then the guidewire is advanced through the catheter and the vessels to the lesion site using its steering tip. Once in place, the balloon is brought there and inflated to free the artery. Afterwards, in most cases, the stent is placed. The tip of the guidewire is made of an X-ray absorbing material and can be seen in the interventional images. The balloon however is transparent under X-ray and therefore two X-ray absorbing markers – of a circular shape – are placed at each of its ends to define its position. Of all surgical tools needed during the intervention, the tip of the guidewire together with the markers are primarily relevant for navigation, therefore we call these "surgical tools" as long as there is no danger of confusion.

During the intervention, the physician navigates mostly blindly, orienting himself on the static roadmap. If this is not sufficient, he injects from time to time a small quantity of contrast agent to find the exact position of the guidewire tip and then proceeds further. Unessential injections of contrast agent are avoided as a large total-dose of contrast agent can be toxic for the patient.

Both during the angiographic exam and during the PTCA the imaging system is brought in different positions with the help of a gantry to better observe the ill vessel. Also the operating table can be moved such as to image the desired location. To limit the X-ray exposure, special blocking plates are used to constrain the field of view of the imaging system. Also the X-ray imaging system is not active during the entire length of the intervention. A PTCA consists of many single viewing sessions. The physician can start and stop a viewing session by pressing or releasing a command pedal.

In some cases – i.e. when PTCA is performed as treatment to a heart attack – at the end of the intervention one or two more injections of contrast agent are given to asses the success of the intervention and make longer time predictions on the health of the patient by observing how much blood reaches the myocardium through the capillaries. This investigation is currently done semi-quantitatively using the Myocardial Blush Grade (MBG) [190]

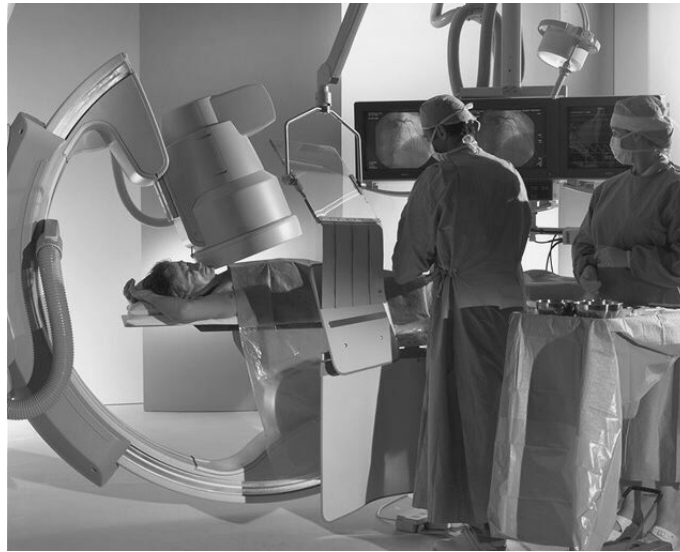The entire procedure takes usually between half an hour and an hour.

Figure 3.2: A modern catheter laboratory – courtesy of Philips.

Machine-vision support for PTCA

The computer has already found his place in the catheter laboratory. His purpose is not to replace the physician but to facilitate his work. One of the simplest applications is the digitization and recording of interventional image sequences. Such sequences can be later analyzed by several physicians to improve a diagnostic or they can be used for teaching purposes or as raw data for certain medical studies.

Once the computer technology was sufficiently developed to allow a time-effective processing of images acquired during a PTCA, the first efforts towards a machine vision support for PTCA were aimed at improving the diagnosis phase by precisely quantifying the projected stenosis area. The physician had to choose a coronary angiogram where the stenosis could be clearly seen. On such an image the vessel contours at the lesion site were found and the projected stenosis surface was measured [183], [122]. It has also been tried to analyze the blood flow through the coronary arteries by densitometric measures and to describe thus qualitatively and quantitatively the stenosis in an attempt to provide the physician with more information and enable him thus to decide on the right treatment method [168].

Recently the attention has shifted towards helping the physician during the intervention and by the assessment of the intervention's results. Several factors are responsible for these developments, the most important being:

1. The continuous advance of computer technology which enables the development of algorithms able to process more data faster.

2. The publication of different medical studies which show how the measuring of specific quantities, on the images acquired during the intervention can help improve the diagnostic and assessment phase.

3. The close collaboration between medical doctors and computer vision specialists.

3.1.3   Improved navigation in PTCA

The physician can see in the interventional images either the steerable tip of his guidewire or the vessels —provided he injects contrast agent into them. Consequently, during PTCA it takes a relatively long time until he reaches the lesion spot, because in order to navigate through the vessels he has to mentally register the diagnostic angiogram used as roadmap with the interventional images, which mostly show only the guidewire and its steerable tip.

To facilitate the navigation and thus shorten the duration of the intervention, and implicitly the period over which the patient and the physician are exposed to the noxious X-ray radiation, it has been proposed to show the physician an improved roadmap, which includes the precise position of the guidewire tip with respect to the vessels. This enhanced roadmap has to compensate also for the heart and respiration motions.

The enhanced static roadmap

Shechter et al. have proposed a method to correct the motion of a X-ray image of the heart using a patient-specific 3D+t coronary motion model [176], [174], [23], [175]. The model requires the cardiac and respiratory phases to correct the motion of a new image with respect to a certain initial complete coronary angiogram. It also needs a 3D static-model of the coronary arteries, which is constructed from one biplane angiogram acquired at the end of the diastole. Each and every image, together with the information it carries – e.g. the position of the guidewire tip – is motion-corrected and projected to the initial coronary angiogram which would take then also the role of a roadmap. For a sequence, with respect to the coronary vessel tree, a "freezing" effect may be achieved using this method. Therefore, the physician gets to see an enhanced static roadmap, where his tools no longer move abruptly in the image with respect to the contrasted vessels.

Schrijver [168] has described an elastic-model-based method to compensate the heart motion and hold the vessels still during angiographic exam. However, it was intended to be used only during the time the vessels are visible, and thus it cannot be applied directly for the computation of the roadmap.

The usefulness of such "freezing" methods is limited by the quality of the motion model. For a good 3D+t coronary motion model a correct segmentation of vessels is needed, as well as a robust method to estimate the cardiac and respiratory phase. The vessel segmentation is done off-line and it necessitates user interaction. The respiratory phase is computed using a region of interest manually positioned over the diaphragm at the beginning of the analyzed sequence. This makes the whole roadmap computation semi-automatic and it also implicitly requires a certain degree of image processing knowledge for the operator. Also, to achieve optimal results, the coronary vessels need to be observed for a period of time which is longer than usual. This implies a change in the intervention routine and rises an acceptance problem for the physician. The motion model is based on a static 3D reconstruction of (at least) the main arteries of the coronary vessels tree. The reconstruction uses two coronary angiograms (one in a LAO and the other in a RAO view) recorded at end-diastole presumably in a ECG-triggered way. To compute the model properly the vessels used should not overlap. This condition is enforced but not ensured by recording the images at end-diastole.

The dynamic roadmap

Alternative to "freezing", a dynamic roadmap can be used [6], [74], [48], [50]. A dynamic roadmap does not project interventional images – according to a certain motion-model – to a single complete coronary angiogram, but chooses for each interventional image the best suited angiogram from a set and projects the interventional tools – i.e. guidewire tip, balloon, etc. – on it under an affine transform. The transform should compensate for the patient movement between the moment when the set of coronary angiograms was acquired and the moment when the analyzed interventional image was acquired. For each interventional image, a corresponding coronary angiogram is chosen such that their cardiac and respiratory phases are similar. Only complete coronary angiograms are suited for roadmap computation and they should cover at least a heart beat. The complete coronary angiograms are selected from cine runs or from cine-run-like sequences. Such cine-run-like sequences include complete coronary angiograms, but are acquired in fluo-mode during PTCA.

Throughout the duration of the intervention, the imaging system is often repositioned to obtain a better view of the coronary vessel tree. Clearly the set of coronary angiograms from which the instantaneous roadmaps are chosen should correspond to the interventional images under this aspect. Small variations can be compensated by affine transformations. If the coronary angiograms acquired during the diagnosis phase are not sufficient to cover the entire intervention with regard to the set of positions taken by the imaging system, then during the intervention, the physician will have to inject contrast agent – a typical bolus of contrast agent is 10 cc over a period of 5-10 seconds – to generate the required set of angiograms (similar to a cine-run). These are than stored together with the information with respect to the position and reused if the imaging system is brought again in a similar position. In a worse case scenario, when none of the positions taken by the imaging system are similar to one another, the physician should have to inject a new bolus for each position. The necessary set of coronary angiograms thus obtained is then used to compute the dynamic roadmap.

In practice, after the physician has brought the imaging system in a new position, it is first investigated if there already exists set of coronary angiograms acquired with a similar geometry. If this is the case, the dynamic roadmap is displayed immediately, conversely no roadmap or a static one is shown, as an indication for the physician that he needs to inject contract bolus to see the roadmap. The moment when the contrast bolus becomes observable is detected and then images are recorded for a few heart-beats. Once a proper set of coronary angiograms is available, they are matched with respect to cardiac and respiratory phase to the live-images. The cardiac phase is computed from the ECG signal recorded simultaneously with the X-ray images. The respiration phase is determined automatically by tracking the diaphragm. Table and patient displacements are compensated by affine transformations using points from the tracked diaphragm as landmarks. If the diaphragm is not visible in the images which are analyzed, a set of X-ray-absorbing marks can be stitched to the operating table to provide the set of landmarks needed for table-motion correction. In such a case the respiration phase is deduced by investigating the similarity between different images in certain ROIs.

After choosing the corresponding instantaneous roadmap for an interventional image, the surgical tools are segmented and overlayed. In the case of the guidewire tip a final correction may be needed to compensate for potential morphological deformations which may be induced by the guidewire tip on vessels with a comparable diameter.

If the physician is not satisfied with the results presented by the roadmap, or if he wants

to asses the position of the guidewire on the interventional images directly, he simply gives a burst of contrast agent. Once the beginning of such contrast burst is detected, the computation of the roadmap is stopped. To show again the roadmap the physician has to start a new viewing session.

The computation of the dynamic roadmap is fully automatic, and it does neither require any change in the medical routine from the part of the physician nor additional knowledge from the part of the medical staff in the catheter laboratory.

3.1.4 Quantification of the Myocardial Blush

The PTCA is also used to treat a myocardial infarction. The purpose of the treatment is to reestablish a normal blood-supply of the heart muscle (or myocardium). Consequently, not only the coronary blood flow needs to be restored, but also the myocardial micro-circulation. Reopening of a blocked coronary artery does not necessarily imply integrity of the corresponding myocardial micro-circulation (or perfusion). Both for precise diagnostics and to assess the success of PTCA it has been proposed to investigate the myocardial perfusion by means of a dedicated measure: the MBG. The presence of blood with contrast agent in heart tissue can be observed during an angiographic exam as a darkening of the target region – i.e. angiographic myocardial blush (MB) – and the MBG can be established by an analysis of this darkening. Presently the MBG is assessed "semi-quantitatively" (grades 0-3) by the angiographer during a visual inspection and thus the procedure is afflicted by inter-observer variability.

Automated estimation of the MBG

To overcome these limitations it has been proposed to compute the MBG in an automated manner by a semi-interactive procedure [54]. The physician has to define by hand the region where he expects the MB to appear, this region is then automatically tracked, i.e., motion corrected with respect to the heart and respiration phases, to allow the observation of the same region during the entire analyzed sequence despite the heart motion. A subsequent robust analysis of the gray-level variations within the tracked region allows the characterization of the MB and establishes then the basis for the sought "quantitative" MBG.

In a clinical setup, the physician should have to record two sequences typically at the end of the intervention. These sequences are then analyzed after the procedure was concluded. One sequence is needed to estimate the MB for one of the patient's healthy arteries and the other sequence, should allow the estimation of the MB for an artery that has just been treated. The first sequence sets a ground truth and the MBG is then computed in a patient-specific way by reporting the results obtained for the second sequence to the ground truth.

3.2 A dynamic roadmap for enhanced navigation in PTCA

To build the dynamic roadmap, one needs to detect the vessel-frames and to differentiate between images showing only a part of the coronary vessel tree and those showing the entire vessel tree (Section 3.2.1). Then, after selecting those images showing the complete vessel-tree, they need to be matched to the live-images (Section 3.2.2) such that the surgical tools can be placed on the roadmap (Section 3.2.3).

Analysis of contrasted images.   The analysis of contrasted images – i.e. images acquired when contrast-agent was present in the blood and therefore parts or even the entire coronary vessel-tree is observable – should result in the detection of coronary angiograms and segmentation of images showing the complete vessel-tree. Such images can then be used as static individual roadmaps for each live-image.

Sequence matching.   To find a match between a live image and a static roadmap, their similarity needs to be measured by, e.g., correlation between the image-backgrounds and/or correlation between the corresponding heart and respiration phases.

Registration of surgical-tools.   To segment the surgical tools in the live images, first they are enhanced and then separated from the background by thresholding. Finally small discrepancies between their positions in the live images and in the dynamic roadmap are compensated.

3.2.1   Analysis of contrasted images

The analysis of contrasted images starts with an *enhancement* step – where the contrast of the vessels is increased – followed by *feature extraction*. For enhancement, the vessels are defined as dark, elongated structures of a certain size. For every image, we compute then one feature related to the vessel area, i.e., a vessel-feature. The curve of feature value over frame index, which we call the feature curve, is analyzed using statistical methods. The purpose of this analysis is to detect cine-run-like sequences during the intervention and to find complete coronary angiograms both there and in genuine cine-runs.

Clearly, the analysis of interventional-images with the purpose of *detecting coronary angiogrmas* requires strictly causal processing. To find out when vessels become observable in these images, we compare the vessel-feature with a threshold, which is set by a significance test [156].

Usually a sequence of coronary angiograms contains three phases:

- **Inflow**, when the contrast agent first enters the vessels.

- **Complete-state**, when the contrast agent can be found throughout the entire vessel-tree.

- **Washout**, when the contrast-agent leaves the coronary arteries.

Only complete-state coronary angiograms are useful for the computation of the dynamic roadmap. To the best of our knowledge, there is no prior art on how to *segment the the complete state* of a sequence of coronary angiograms. Thus, we present here two different approaches to accomplish this:

1. **The first** models the observations in each phase of a coronary angiograms sequence by a polynomial and seeks the segmentation which allows the best fit of three polynomials as measured by a Maximum Likelihood (ML) criterion [104], [156].

2. **The second** models the process which generates the corresponding feature curve by a two state (filled /non-filled) Hidden Markov model [162], [184], which we estimate using the Maximum a Posteriori (MAP) criterion.

These are then evaluated and compared on a number of sequences recorded in clinical routine

Vessel enhancement and feature extraction

To detect vessel images in sequences of X-ray projections of the heart, a feature related to the vessel area is extracted for each image. We start by computing a *vessel map* such that its histogram can be assumed to consist of two distinct distributions, one from background and one from potentially occurring contrasted vessels. The vessel map is the result of several *enhancement steps*. First we seek to equalize non-vessel background information, thus reducing its standard deviation. The additional absorption of contrasted vessels is then transformed such that its mean is considerably larger than the background mean. Since enhanced contrasted vessels show up with high intensities in the vessel maps (see Figure 3.5 (b)), we analyze their histograms and use a *high percentile* as a *measure of presence* of contrast agent and therefore also *of vessels*.

The *feature curve* obtained by observing the vessel feature over time is segmented by different statistical methods. In the case of fluoroscopic images the principle of strict causality is respected.

As we seek to detect the presence of vessels from the vessel map histograms rather than accurate segmentation or enhancement geared toward the human observer, issues like border accuracy and preservation of a certain "harmony" in the processed images are of less concern [9],[83].

Background attenuation by morphological processing.   The first step towards enhancing vessel contrast exploits the property that, being filled by contrast agent, vessels are locally darker than their immediate surroundings [9]. The varying background is equalized by a Bothat operator [67]. This operator is defined as the result of subtracting the original image from its closing. If the window size of the dilation filter is chosen slightly larger than the largest vessel diameter, the (dark) vessels will be suppressed after image opening, leaving only the background (see also Appendix A.1). The Bothat filter equalizes background information, while preserving the gray-level difference between background and vessels. A processing example is shown in Figure 3.3.

Motion-based vessel enhancement.   An additional clue to vessels is their mostly strong and jerky motion, which shifts them quickly over a distance of several vessel diameters and is caused by the beating heart [9]. When the vessel moves to a new position, absorption at this position will increase due to the contrast agent within the vessel. Therefore, when calculating the pixel-wise difference image between any given frame and its predecessor, pixels with the new vessel positions will tend to exhibit negative differences, while the vessel positions in the previous frame will tend to show positive differences. This effect is even more pronounced in Bothat-filtered angiograms, since the Bothat filter reduces background structures which would cause clutter in the difference images when moving themselves. We therefore clip positive values in the difference images to zero, and add the thus clipped difference image to the current Bothat-filtered frame. Consequently, the moved vessel tree regions will become darker than before. Local minima of small extent, which are unlikely to be caused by moving vessels, can optionally be removed by a morphological closing with a 3×3 structuring element. A block diagram of the processing chain and a processing result are shown in Figure 3.4.
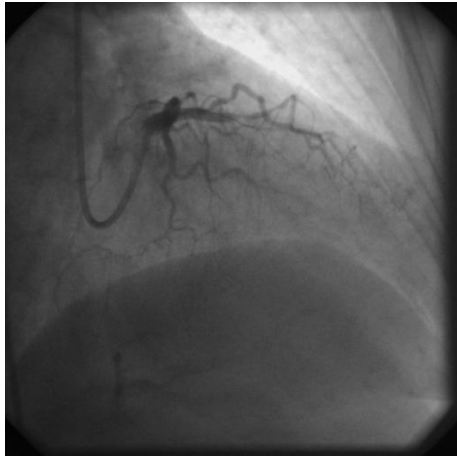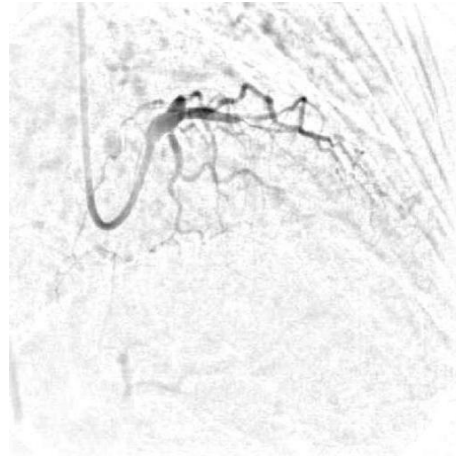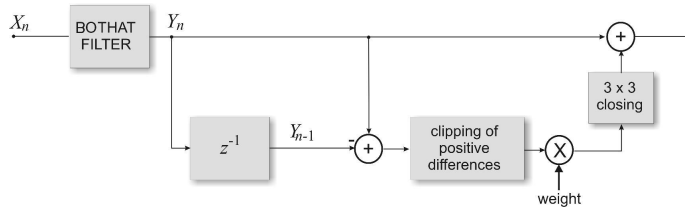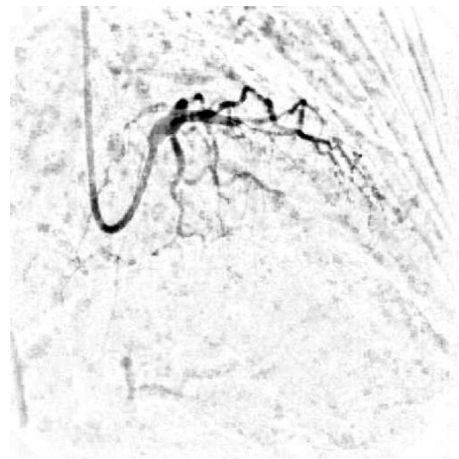
(a)   (b)

Figure 3.3: Coronary angiogram (a) and result of morphological processing (b)

(a)   (b)

Figure 3.4: Block diagram of spatiotemporal filtering (a) and filtration result (b).
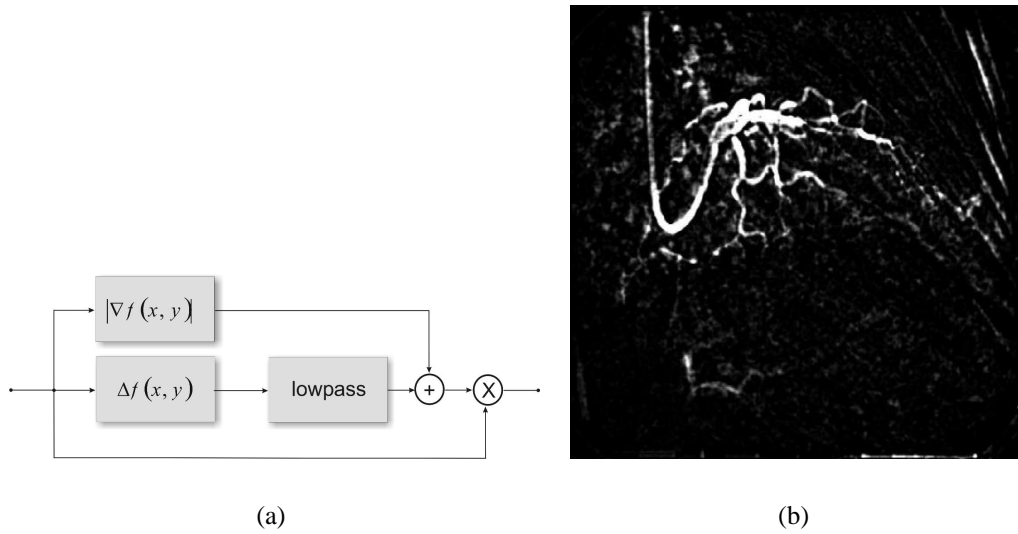
(a)   (b)

Figure 3.5: Block diagram of derivative-based enhancement (a) and final vessel map (b).

Computation of the vessel map.   The Bothat operator selects all dark background structures comparable in size with the vessels, including some which are not genuine vessels but vessel-like artefacts, i.e., vessel noise. As the motion based enhancement step is effective only for moved vessels, such noise still affects motionless vessels. These noise structures have a small surface, are dark but less contrasted than vessels, have diffuse edges and their shape is mostly patch-like, although some of them can be elongated and thin. The patch-like noise is due to the soft tissue which the X-rays encounter on their way through the body and is thus ubiquitous. The surface it covers varies from image to image as a consequence of the motion induced by the heart and by the respiration. The elongated noise is mainly due to bone tissue, it does not cover such a large area and its surface is approximately constant.

To attenuate the patch-like noise, we use a two steps approach. Starting from the result of the Bothat filtration, we compute first the gradient norm and add it to the second derivative [139, 186] using empirically determined weights (0.55 for the gradient norm and 0.45 for the second derivative). Then, we multiply this result pixelwise with the result of the motion-based enhancement to obtain the vessel map. The first step is equivalent to a logical OR between the results of the gradient norm which responds to contrasted structures with sharp edges but does not respond to vessel-like ridge profiles and the second derivative which does. The second step is the equivalent of a logical AND between the result of the derivatives-based enhancement and that of the motion-based enhancement as each procedure is based on different vessel properties.

To be able to correctly apply the AND, the result of the motion-based enhancement is rescaled before multiplication so that vessels are brighter than background. A block diagram of the processing chain and a vessel map result are shown in Figure 3.5.

A percentile-based feature of vessel surface.   As on the vessel map vessels appear brighter than background, the frequency of occurrence of bright pixels – and thus the bright tail of the histogram of the vessel map – varies with the vessel area as shown in Figure 3.6. We use this relationship to compute a feature related to the vessel surface – i.e. a vessel-feature – by means of a properly chosen percentile of the histogram of the vessel map:
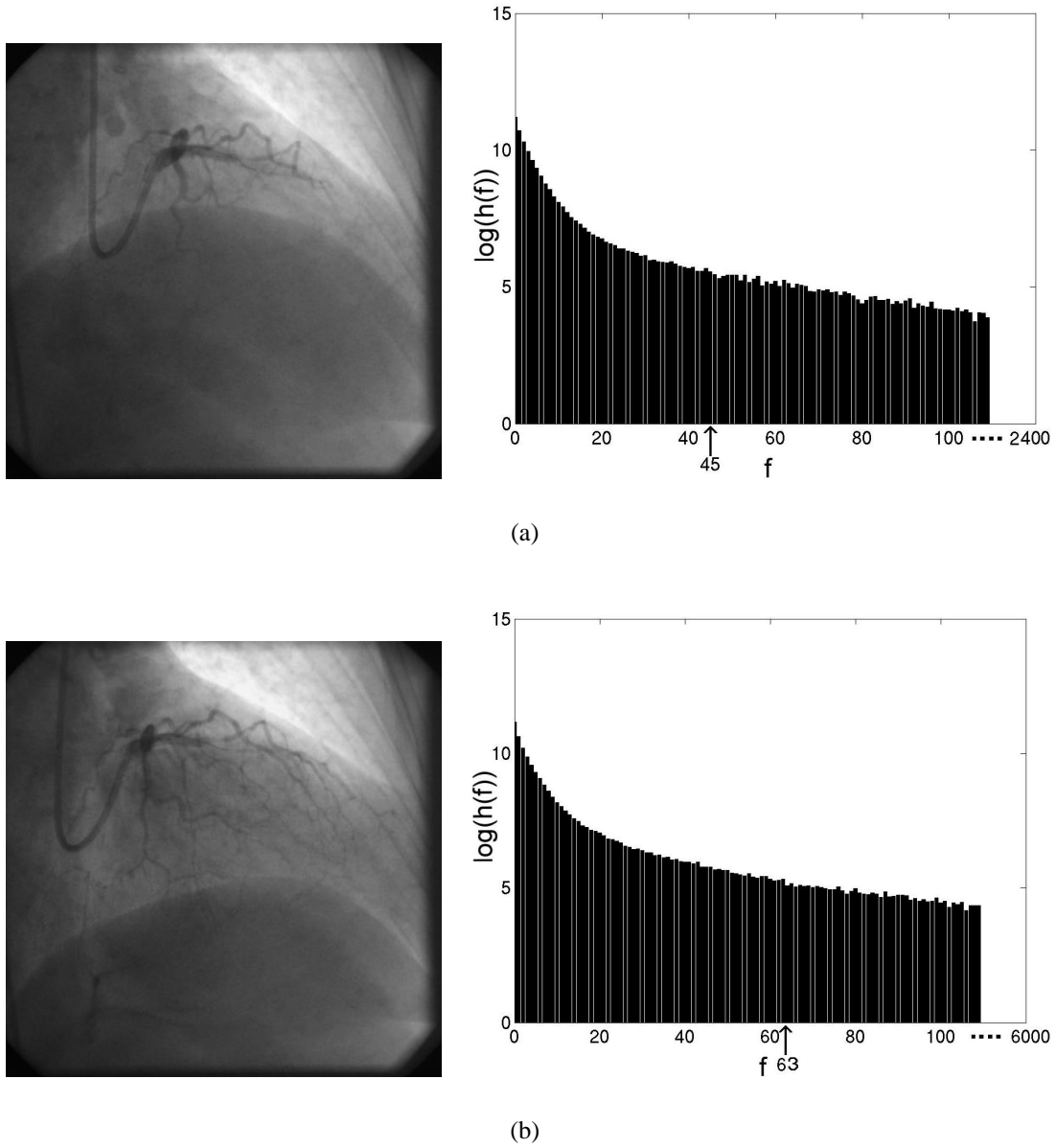
(a)

(b)

Figure 3.6: Angiogram and vessel map log-histogram for an inflow frame (a) and a complete-state frame (b). The 95-percentile is given by the arrows. Observe also the modification of the tail of the histogram.
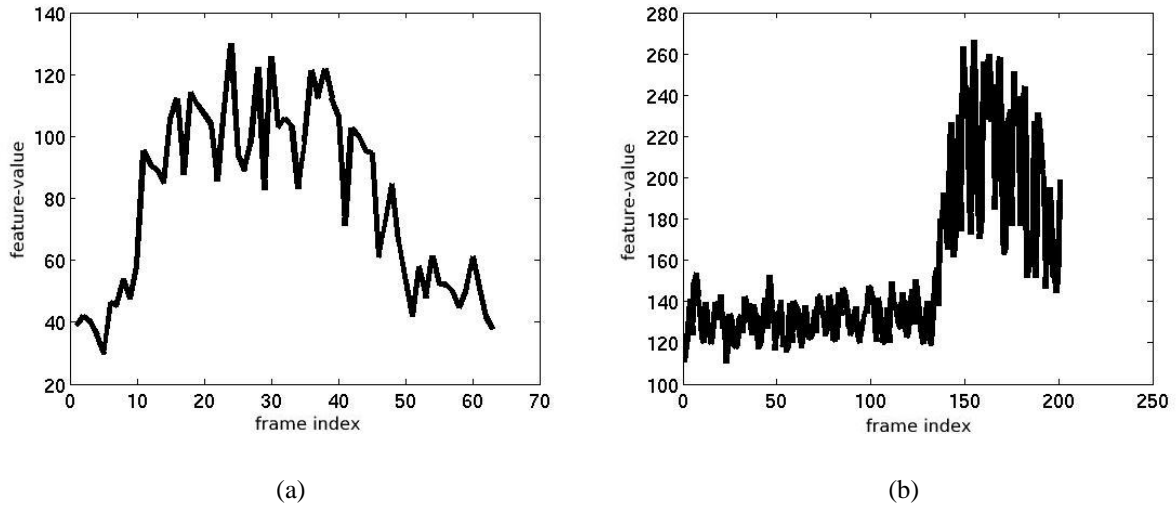
Figure 3.7: 95-percentile (a) and 98-percentile (b) feature over frame index for a pre-interventional and an interventional sequence respectively.

1. To detect the *moment when contrast agent is observable* in the interventional frames – i.e. the beginning of a contrast burst – we desire a feature which reaches its peak once the slightest trace of contrast agent appears. As the area covered by vessel-noise in an angiogram is around one percent, we choose the *98-percentile* as feature. Then, for images where contrast agent is present – i.e. more than two percents of all pixels in the analyzed image have high intensities – the feature will exhibit higher values than for images which show no vessels.

2. To *segment only the complete-state* from a sequence of coronary angiograms, we would like to have a feature which exhibits a maximum value over frames showing the complete vessel tree. As the vessel surface covers typically around six percents of the entire angiogram area, we choose the *95-percentile* as feature. For images showing the entire vessel tree, i.e., when more than five percents of all pixels in the analyzed image have high intensities, the feature will exhibit higher values than for images showing only a part of the vessel tree.

The feature curve. The evolution of such vessel-features over frame index generates a feature curve where the complete-state and the contrast burst frames respectively may already be identified as shown in Figure 3.7. Such feature curves have to be segmented in two classes, i.e., valid and invalid, where the valid class should contain complete-state and contrast burst frames respectively, depending on the target application.

A typical feature curve will inherently exhibit a certain variance related to the heart and respiration movement. The feature curve is also influenced by the AGC, which modifies the intensity of the X-ray radiation in each frame to ensure constant brightness [27].

Signaling of contrast-agent injections

The purpose of vessel enhancement was to obtain a vessel map such that its histogram can be assumed to consist of two distributions: that from the background pixels and that from potentially occurring vessel pixels. Thus, the vessel feature – in this case, the 98-percentile – ideally builds a separable *feature curve* with only *two classes*, because an interventional image either shows some vessel structures or shows only background. Considering also that the processing of interventional images is subject to *strict causality*, we propose to detect the beginning of a contrast burst by a *significance test*.

Particularities of the 98-percentile feature curve.   When contrast agent is injected into the vessels, it first travels trough the catheter. Once into the vessels, it uniformly mixes with blood and renders the vessels through which it travels visible under X-rays. These vessels remain visible as long as contrast agent is injected. After that, they slowly become invisible again as the blood with contrast agent is replaced with fresh blood. However, this washout does not affect the contrast agent within the catheter, and thus the catheter remains visible in the analyzed images.

A viewing session may begin right after a new catheter has been introduced, or after a previous injection of contrast agent. In the former case, depending on how much of the catheter is visible, it may happen that the 98-percentile responds to the catheter and not to the vessels. This makes difficult the segmentation of a sequence of interventional images showing contrast agent and justifies the heuristic rule proposed for this purpose, i.e., record images for eight heart beats after the contrast agent was detected.

The 98-percentile is linked to the vessel area. When contrast agent is injected into the vessels, the 98-percentile increases rapidly as it is computed from gray levels characteristic for pixels belonging to the large well contrasted vessels. From the large vessels the blood carrying the contrast agent travels to the mid and small vessels, than to the capillaries and finally reaches the myocardium. As the large vessels will be free of blood carrying contrast agent sooner than the small ones, the 98-percentile reaches its maximal value very fast when the vessels become visible but it decreases slower when the contrast agent leaves the vessel tree as then it is influenced mainly by gray levels characteristic for pixels belonging to the small less contrasted vessels.

Detection of the first frame of contrast burst by significance test.   To find the ideal percentile threshold T which separates the the two classes, we need to estimate the corresponding class conditional *pdfs*. However, in this case the thresholding procedure is constrained by the following observations:

1. The strong patient variability requires each case to be treated independently and thus it is pointless to gather data from several patients to estimate the *pdfs* and use the computed threshold on other patients.

2. The analysis of interventional images clearly requires a strictly causal processing and thus one cannot wait until the end of the sequence to analyze it.

We proceed using the observation that the images acquired during the first seconds of an intervention are most probably free of contrast agent. The *pdf* of the vessel-feature given class
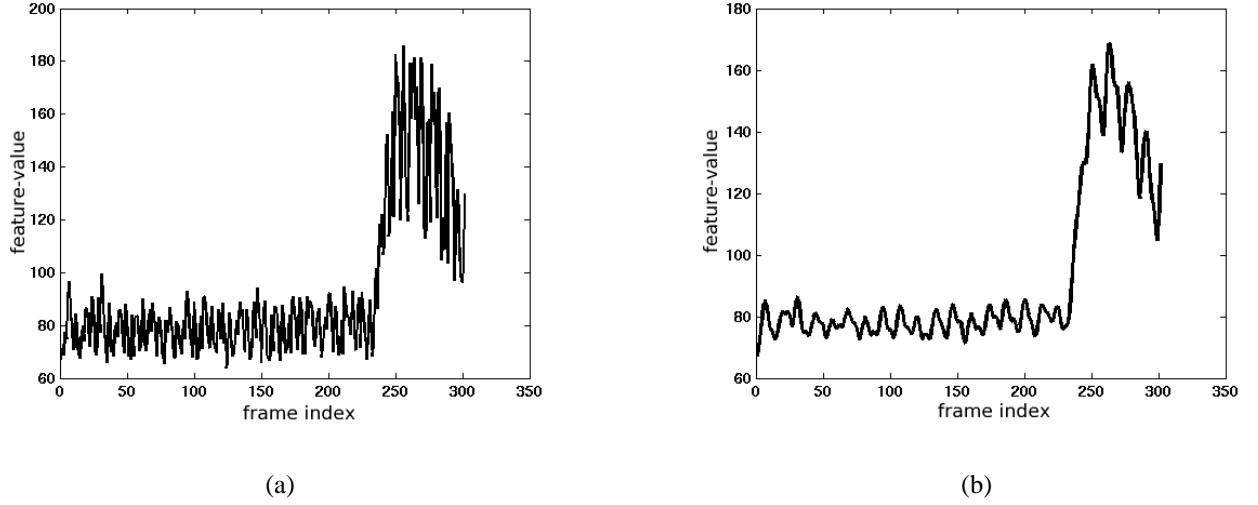
Figure 3.8: Feature curves of 98-percentile over frame-index computed before (a) and after (b) low-pass filtration for the same interventional sequence.

show-background-only may then be estimated using these images. As an interventional image can belong to only one of the two mutually excluding possible classes, we propose to detect the contrast burst frames by a significance test [156] whose null hypothesis H_0 is that the investigated frame shows no vessels.

From the learning phase, over the first seconds of the intervention, which are free of contrast agent, we can estimate non-parametrically [184] (see also Section 4.3.2 at page 140) the distribution $p(y(n)|H_0)$ of the feature $y(n)$ given the null hypothesis. The threshold T is then determined such that the probability of $y(n)$ exceeding T given H_0 is α, which is the so-called significance level, equivalent to the false positive rate (a typical value for α is $\alpha = 5 \cdot 10^{-4}$). T is thus given by inverting $\Pr(y(n) > T|H_0) = \alpha$ based on $p(y(n)|H_0)$. Whenever the 98-percentile $y(n)$ exceeds this threshold, the corresponding frame is classified as containing contrast agent.

To improve detection, before hypothesis testing we attenuate the inherent variation of the feature curve by filtering with a causal recursive first-order low-pass filter. A feature curve before and after filtration is shown in Figure 3.8. The difference equation characterizing this filter is:

$$y(n) = ax(n) + (1 - a)y(n - 1) \qquad (3.1)$$

$$\text{where:} \quad \begin{cases} 0 < a < 1 & \text{if } x(n) - x(n-1) \leq 4\sigma_0(x) \\ a = 1 & \text{else} \end{cases}$$

where $\sigma_0(x)$ is the standard deviation estimated for the null hypothesis from the first frames of the sequence. The filter thus smoothes within stationary time intervals, but preserves what it assumes to be a transition.

For the learning phase we take 72 images, which are acquired over a period of six seconds, with a frame rate of 12 fps. This is justified by the need to sample at least one complete respiration cycle, considering that under normal circumstances, a human breathes between 15 to 20 times per minute.

Signaling of contrast agent injections. Results.   We have analyzed a total of six sequences from different patients.

The signaling of arbitrary contrast injections is a binary classification problem, as for each sequence there can be just one contrast-bolus which is either correctly or falsely detected. We evaluate our methods by discrepancy methods, using a gold standard. The gold standard was achieved after detecting by expert visual-analysis the first frame which shows vessels. Taking into consideration the variability of such a visual segmentation, we consider that the automatic detection was successful if it falls in an interval of plus-minus two frames from the expert reference.

In all six cases the detection was successful.

Signaling of contrast agent injections. Discussion.   For the signaling of a contrast injection we have used the 98-percentile. A higher percentile will in most cases be insensitive to the presence of vessels as it will respond to noise. A lower percentile will detect vessels too late, i.e., when the image area which they cover is already large.

For the signaling of contrast agent injections a training set is needed. This training set is acquired during the first few seconds of the intervention which must be free of contrast agent. For these images the 98-percentile is computed from gray levels belonging to the background.

During respiration, for projection angles which permit also the visualization of the diaphragm, the 98-percentile follows the variation of the projected diaphragm area. We believe that this is a consequence of the AGC trying to compensate the variation of the image brightness caused by the moving diaphragm. This is why the training set is acquired over a period long enough to cover a respiration cycle.

Depending on the projection angle more or less of the catheter is visible. When filled by contrast agent, the catheter appears as an additional vessel. Because the washout does not affect the contrast agent from within the catheter, this "vessel" stays visible once contrast agent has been introduced. This makes the detection of all images acquired during a control injection difficult, as in such circumstances the training images will not show the catheter. Depending on the projection angle the image area occupied by the catheter varies.

The expected value of the 98'th percentile when no vessels are present is affected by noise, which is linked to artefacts in the feature map generated by respiration or by the movement of the heart. Consequently, the percentile value given the null hypothesis is considered to be Gaussian distributed and it is parametrically estimated using the percentiles recorded for the first frames of the intervention. For a new image, the null hypothesis is that its corresponding 98'th percentile lies somewhere in a region around the expected value whose borders are determined by the significance level. If the percentile varies from the expected value more than it is allowed through the significance level, then this can be caused only by the fact that vessels are present. To improve the segmentation the feature curve is filtered by a causal recursive filter. We have tested our algorithm on several appropriate sequences with very good results.

Selecting images showing the complete vessel tree

We describe two different methods to segment the complete state in a sequence of coronary angiograms. The first method uses a *ML* and the second a *MAP* approach.

The segmentation is based on the assumption that there are three and *only three states* observable on an analyzed feature curve: inflow, complete state and outflow. Clearly, these states

are *coherent* – i.e. there are no inflow/washout images in the complete state and no complete state images in the inflow/washout.

The result of segmentation can be judged based on several quality criteria. The first is that the complete state should be coherent. As this is a binary segmentation, other criteria are linked to the number of CC and FP. Specifically, when segmenting the complete state a FP is more critical than a FN. A FP is an image showing only parts of the vessel tree, and is thus obviously unsuited as roadmap. A FN means only that there are fewer frames available when building the dynamic roadmap. This has negative consequences only when there are too few images – i.e. not enough to cover a heart beat.

Consequently, a good segmentation method should yield a coherent complete-state with a very small fp rate – we mean here the FPs that appear before the beginning and after the end of a coherent complete-state – and a cc rate to cover at least one heart beat. We show that the fp rate can be reduced by using *ECG-based information* which is recorded synchronously with the interventional-images.

Particularities of the 95-percentile feature curve. Assuming that the contrast agent is uniformly mixed with the blood and all vessels are visible, the vessel contrast in an angiogram depends on the speed at which the blood flows through the arteries. This happens because the more contrast agent passes in front of a sensor element in the time interval between two frames, the less X-ray quanta reach the sensor as they are better absorbed. The contrast of individual vessels however, depends on their size, i.e., the large vessels have a good contrast and the small vessels a weaker contrast as they contain more respectively less contrast agent.

The vessel surface observable in X-ray projections, as well as the speed at which the blood flows, vary within a heart beat between two extreme values, as a consequence of the heart motion. The vessel surface reaches a minimum towards the end of the ventricular systole, i.e., when the ventricles contract to push the blood in the vascular system. The flow of blood reaches a peak at around the same time, more precisely during the ejection phase. The vessel surface reaches a maximum before the ventricular systole, i.e., when the ventricles are filled with blood. The flow of blood reaches a minimum at around the same time, more precisely at diastasis and during the atrial systole.

During inflow and washout, the 95-percentile varies mainly with the projected area of the vessels. However, during the complete state – i.e. when the contrast agent is uniformly mixed with blood and the area of the vessels is around six percents and therefore the 95-percentile is certainly computed from gray levels which belong to vessels – the 95 percentile of the vessel map follows mainly the variation of the vessel contrast and thus it varies with the blood flow, i.e., somewhat inversely to the vessel area. This actually means that the contrast of small vessels during rapid-ejection is larger than the contrast of slightly larger vessels at diastasis. We believe that this is due to the higher speed at which the blood flows during the rapid-ejection and which influences the contrast as described above[5]. When the flow is maximum, the contrast and the percentile reach also a maximum. Conversely, when the flow is minimal, the contrast and the percentile reach also a minimum. This is shown in Figure 3.9.

---

[5]One can argument that the AGC could be another reason for this. It increases the X-ray dosage to compensate for the darkening of the image when the vessel area is maximum and it therefore decreases the contrast of vessels. Conversely it decreases the X-ray dosage when the vessel area is minimal. However in this case the X-ray system should work near the saturation limit – such that an increase in intensity leads to a decrease in contrast – and the AGC should have an almost zero latency.
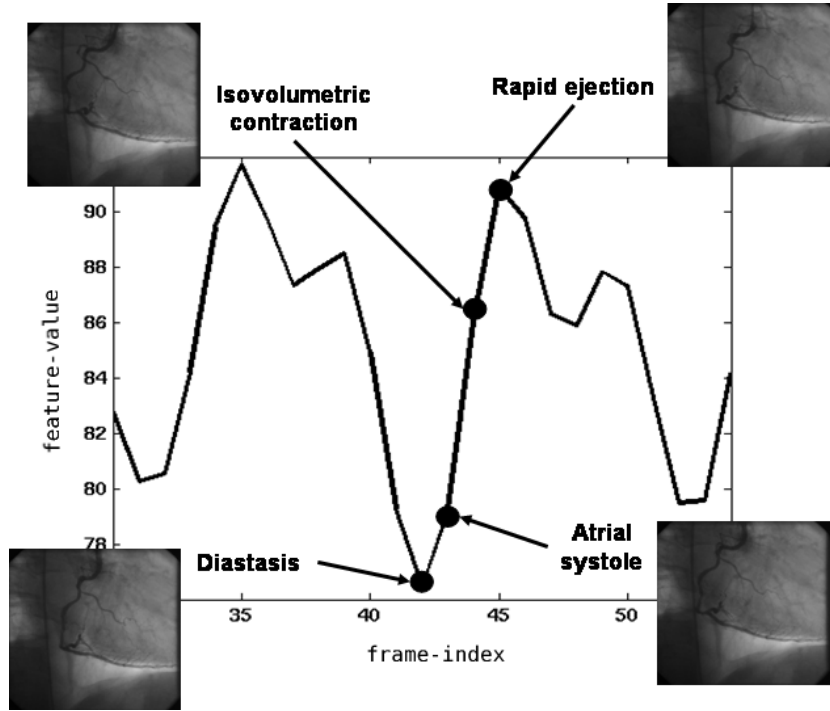
Figure 3.9: The curve of vessel features over frame index recorded during two heart beats of the complete state.

Consequently, the 95-percentile feature curve has two important characteristics:

1. The features extracted for images of the complete state have a rather large variance due to the modification of the contrast of the vessels. We call this the heart beat-induced variance to differentiate it from the inherent variance.

2. There is a certain amount of ambiguity with respect to the values of the feature curve recorded at the beginning and at the end of the complete state. It is uncertain if they are true complete-state images or are just inflow/washout images recorded at a maximum of the percentile feature, i.e., during the ventricular systole (see Figure 3.10).

Maximum Likelihood segmentation. To segment the complete state we propose to least squares fit three polynomials of a given degree to the feature curve. This approach makes implicit use of the a priori knowledge with respect to the coherency of the three states of the feature curve, which it also enforces. There are:

$$M = \binom{N}{2} \approx \frac{N^2}{2} \tag{3.2}$$

possibilities to segment a curve with N points into three coherent regions. As the typical standard length of an analyzed angiogram sequence is $N = 87$, this means that we have $M = 3741$ segmentation candidates. The correct segmentation is then found by using the Maximum Likelihood (ML) criterion as:

$$\hat{Q} = arg \max_Q p(S|Q) = arg \max_Q \prod_{j=1}^{3} p(\mathbf{s}_j|m_j, \sigma_j^2) \tag{3.3}$$

(a)   (b)
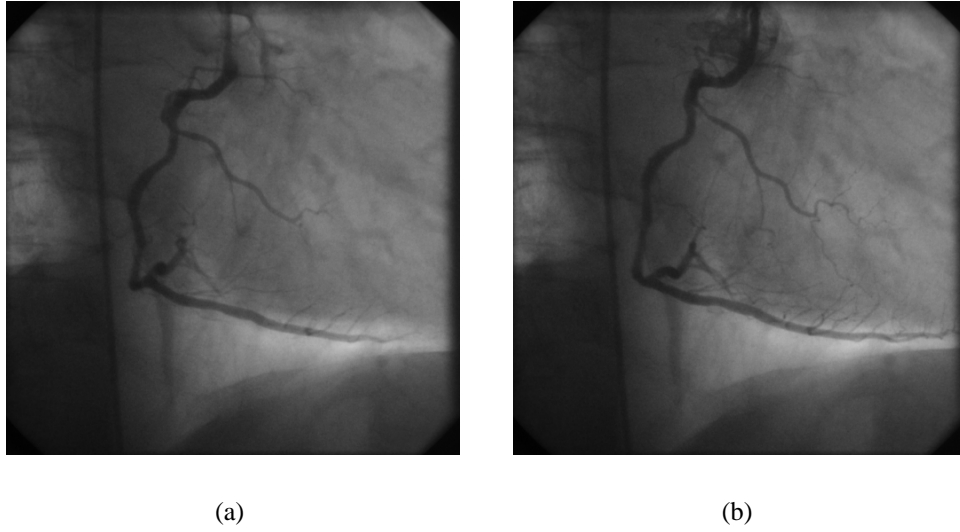
Figure 3.10: Images of a sequence of angiograms showing a partially filled coronary vessel tree (a) – frame 24: rapid ejection – and a completely filled coronary vessel tree (b) – frame 32: atrial systole. The value of the 95'th percentile feature is similar in both cases.

with $Q = \{q_1, \ldots, q_N\}$ one of the M successions of states where q_i is the state label – i.e. complete state or not-complete-state – for the ith frame of the sequence of angiograms and $S = \{s_1, \ldots, s_N\}$, the sequence of percentiles with s_i being the percentile observed for the ith angiogram frame, $j = 1, 2, 3$ is the index of region R_j, and $\mathbf{s}_j$ is a vector with ordered observations in each region such that $\bigcup_{j=1}^{3} \mathbf{s}_j = S$. Modeling the observations as Gaussian distributed, m_j and σ_j^2 are the region-specific parameters of the distribution. Also, we have made the assumption that the observations are region-wise independent [184]. We furthermore assume that the observations within each region are conditionally independent, i.e., that their dependencies are fully captured by the coherency of the region [184], [7], [199]. We then obtain for the region likelihood:

$$p(\mathbf{s}_j | m_j, \sigma_j^2) = \prod_{n \in R_j} p(s_n | m_j, \sigma_j^2) = \left(\frac{1}{\sqrt{2\pi\sigma_j^2}} \right)^{N_j} \exp\{-\frac{1}{2\sigma_j^2} \sum_{n \in R_j} (s_n - m_j)^2\} \qquad (3.4)$$

with N_j the number of frames in region R_j and m_j and σ_j^2 estimated by:

$$\hat{m}_j = \frac{1}{N_j} \sum_{n \in R_j} s_n \, , \quad \hat{\sigma}_j^2 = \frac{1}{N_j} \sum_{n \in R_j} (s_n - \hat{m}_j)^2 \qquad (3.5)$$

Then, replacing the true parameters in equation (3.4) with their estimates from (3.5) and introducing (3.4) in (3.3) we seek:

$$\hat{Q} = \arg\max_Q \left[\prod_{j=1}^{3} \sqrt{\hat{\sigma}_j^2}^{-N_j} \right] = \arg\min_Q \sum_{j=1}^{3} N_j \cdot \log\left(\hat{\sigma}_j^2\right) \qquad (3.6)$$

where we have dropped all factors and constants not influencing the extreme problem and have taken into consideration that: $N_1 + N_2 + N_3 = N$. Thus, for all practical purposes, the sought

succession of states $\hat{Q}$ should be chosen such that $\hat{\sigma}_j^2$ is minimized for each region R_j with $j = 1, 2, 3$. This is equivalent to seeking the segmentation which permits fitting a polynomial of degree zero – i.e. $l_j(s_{n_j}) = \hat{m}_j$ – optimally to each region. The quality of the fit is measured by the minimum mean square error as defined by equation (3.5).

The specificity and sensitivity of this segmentation method clearly depend on the quality of the polynomial model, i.e., how well can the three polynomials fit to the feature-curve. To allow a better fit we have used polynomials of higher degrees. In such a case we need to solve a typical ordinary least squares problem [105].

For region R_j, the fitting error term obtained for a polynomial of degree k and expressed in matrix-vector notation is:

$$\mathbf{e}_j = \mathbf{s}_j - \mathbf{A}_j \mathbf{b}_j \qquad (3.7)$$

where

$$\mathbf{e}_j = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_{N_j} \end{bmatrix}, \quad \mathbf{A}_j = \begin{bmatrix} n_1^k & n_1^{k-1} & \ldots & 1 \\ n_2^k & n_2^{k-1} & \ldots & 1 \\ \vdots & \vdots & \vdots & \vdots \\ n_{N_j}^k & n_{N_j}^{k-1} & \ldots & 1 \end{bmatrix}, \quad \mathbf{b}_j = \begin{bmatrix} b_k^j \\ b_{k-1}^j \\ \vdots \\ b_0^j \end{bmatrix} \qquad (3.8)$$

The variance estimate is then

$$\hat{\sigma}_j^2 = \frac{1}{N_j} \mathbf{e}_j^T \mathbf{e}_j \qquad (3.9)$$

and, minimizing $\hat{\sigma}_j^2$ by the estimate $\hat{\mathbf{b}}_j$ of the parameter vector, we find

$$\frac{\partial \hat{\sigma}_j^2}{\partial \hat{\mathbf{b}}_j} = \mathbf{0} \Rightarrow \hat{\mathbf{b}}_j = (\mathbf{A}_j^T \mathbf{A}_j)^{(-1)} \mathbf{A}_j^T \mathbf{s}_j = \begin{bmatrix} \hat{b}_k^j \\ \hat{b}_{k-1}^j \\ \vdots \\ \hat{b}_0^j \end{bmatrix} \qquad (3.10)$$

Introducing (3.10) in (3.9) we can compute $\hat{\sigma}_j^2$ for each region R_j and the sought segmentation is given by:

$$\hat{Q} = \arg \min_Q \sum_{j=1}^3 N_j \cdot \log\left(\hat{\sigma}_j^2\right) \qquad (3.11)$$

For each segmentation candidate, the optimal polynomial-fit can thus be determined. The sought ML-segmentation, is the one with minimum optimal fitting error. Since the number of segmentation candidates is limited by the coherency constraint (equation (3.2)), full search is practically feasible. The degree of the polynomials can be varied depending on the state which they should model. For optimal results, the degrees of the polynomials should be readjusted for each analyzed sequence, making this method unsuited for automatic usage. However good results are obtained using polynomials of degree two for the inflow and washout and one for the complete state.

Maximum a Posteriori segmentation.   A straightforward modality to automatically segment the complete state is by a fixed threshold. However, such a simple approach is generally ill suited, particularly due to the large heart-beat-induced variance of the complete state. A fixed threshold will either break the coherency constraint or return too many false positives.

To improve the segmentation we propose a variable threshold, which adapts by taking into account prior information with respect to the coherency within the three states on the feature curve.

For this purpose we use a Hidden Markov Model (HMM), which is a combination between a number of stochastic processes and a Markov chain [162], [184]. A first-order Markov chain is a random process of states in which the transition probability to the next state depends only on the previous one. In a HMM the states sequence can be observed only through the stochastic processes which produce a sequence of observations. A HMM is parameterized by:

  i The state transition probability matrix.

  ii The observations' *pdf* conditioned on the states.

  iii The initial states probability vector.

As a coronary angiogram can either belong to the complete state or not, we use a two states (complete/non-complete) HMM for our segmentation purpose. The observations sequence S is then the 95-percentile feature curve and it was obtained by combining the observations from two stochastic processes: one corresponding to the complete state and one to the non-complete-state (i.e. inflow and washout). We seek then the optimal sequence of states Q such that $P(Q|S)$ is maximum (i.e. subject to the MAP criterion).

The state transition probability matrix is chosen as:

$$A = \left\{ \begin{array}{cc} 0.9 & 0.1 \\ 0.1 & 0.9 \end{array} \right\} \tag{3.12}$$

thus enforcing the coherency constraint as the probability of keeping the current state is always larger than that of leaving the state.

We approximate the two state conditional observations' *pdfs* with normal distributions and use the segmentation results obtained by a fixed threshold (see Appendix A.6.2) to estimate their respective means and variances. We impose by means of the initial probability vector that the initial state of our model is not a complete state.

We obtain thus an *adaptation* of the initial fixed threshold by means of prior information, as it is the fixed threshold which initially separates the observations of the two states thus allowing the computation of the state conditional *pdfs*.

An adaptive threshold copes better with the large complete state variance than the fixed threshold. To further improve the segmentation we propose to explicitly enforce the coherency constraint by comparing the result achieved with the adaptive threshold with all possible coherent segmentations (Equation (3.2)) and choose the most similar. As distance measure we use the Hamming distance.

ECG enhanced complete state segmentation.   If only image information is used to segment the complete state, then the result will probably contain many false positives, e.g., images showing only a partially filled coronary vessel tree, acquired during the ejection phase of the ventricular systole. For such images, the corresponding percentile feature is similar with that of a genuine complete state frame acquired during the atrial systole. To reduce the false positives rate we use ECG information.
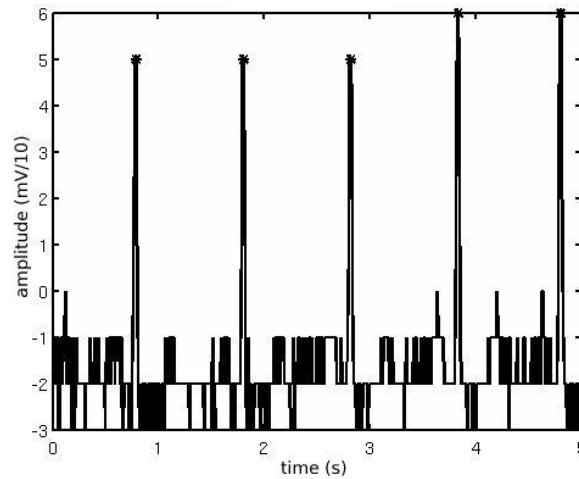
Figure 3.11: ECG signal with detected peaks marked by stars.

To make sure that neither inflow nor washout image are included in the complete state, we propose to constrain the complete state so that in may begin and end only at minima of the feature curve – i.e. before the ventricular systole. To detect such moments we use the ECG signal which is typically recorded synchronous with the analyzed sequence. The ECG records the electrical activity of the heart. The actual heart phase lags behind its electrical counterpart, as there is a small delay between the moment a command is issued and the moment when the system responds. Thus, the R peaks which represent the peak of the electrical ventricular systole are recorded right after the actual atrial systole – i.e. before the ventricular systole – and mark thus the moment when the projected surface of the vessels is maximal and the flow of blood and the contrast of vessels is minimal – i.e. minima of the feature curve.

To detect the R peaks we use a fixed percentile threshold. We choose the 90 percentile of the ECG values. With this threshold we select the ECG samples recorded right before and immediately after the peak. To precisely find the peak, we analyze the zero crossings of the derivative of the signal composed from the samples selected at the previous step. Results obtained for a sequence from our data base are shown in Figure 3.11. The final segmentation of the complete state is obtained by choosing from the initial result only images acquired between two R-peaks of the ECG signal. If only one R-peak was detected, then the final segmentation begins there and ends together with the initial result.

Selecting images showing the complete vessel tree. Results.   We have tested the algorithms on 31 sequences from our data-base. These sequences have been segmented by hand two times, obtaining thus two ground truths. The first one (A) was computed using the following rule: the complete state begins with the first frame when all vessels are visible – including the vessels with the smallest diameter – and ends before the first image where at least one vessel is poorly visible. The second one (B) was computed using a more permissive rule: the complete state begins when no new vessels are observed for several frames in a row and ends before the first image where at least one vessel is invisible. Using each rule, segmentations were conducted independently by the author and by a student help, images were selected into the ground truth only if both persons agreed. The mean number of frames in a sequences was 87. The mean

	$cc(\%)$	$fp(\%)$	R
SetB	99.5	14.1	0.28
ML	76.1	16.2	0.37
ML+ECG	59.3	8.3	0.28
MAP	93.1	25.1	0.43
MAP+C	99.5	26.2	0.42
MAP+C+ECG	87.4	13.5	0.30
T	96.5	28.6	0.45

Table 3.1: Results for the complete state, using set A as ground truth.

	$cc(\%)$	$fp(\%)$	R
ML	73.3	9.3	0.19
ML+ECG	57.1	2.9	0.08
MAP	91.9	14.8	0.23
MAP+C	98.9	15.1	0.22
MAP+C+ECG	81.6	4.3	0.08
T	95.5	18.3	0.26

Table 3.2: Results for the complete state using set B as ground truth.

length of the complete state by the first ground truth was $\overline{ncc} = 23$ and by the second ground truth was $\overline{ncc} = 31$ and the mean length of the inflow/washout was $\overline{nfp} = 64$ and $\overline{nfp} = 56$ respectively.

In Table 3.1 are shown the results achieved on the more restrictive ground truth (A) by the maximum likelihood approach without the ECG-based constraint (ML) and with the constraint (ML+ECG), by the maximum a posteriori approach (MAP), with the coherency constraint (MAP+C) and additional with the ECG-based constraint (MAP+C+ECG) as well as the results achieved when segmenting according to the second ground truth (SetB). In Table 3.2 the results achieved on the more permissive ground truth (B) are shown. As reference we give also the results achieved by a fixed threshold (T). The results are mean percentages of correct classifications (cc) and false positives (fp) over all sequences. Using these values and the mean length of the complete state and of the inflow/washout, we have computed also the relative fraction of incomplete roadmaps in a sequence as $R = \frac{nfp}{nfp+ncc}$ with $nfp = \overline{nfp} \cdot fp$ and $ncc = \overline{ncc} \cdot cc$.

For the ML method the degrees of the polynomials used to approximate the three phases were: two for the inflow and washout and one for the complete state. This combination was chosen after testing all possible combinations up to polynomials of degree five with respect to the fraction between the rate of correct classifications and that of false positives.

Selecting images showing the complete vessel tree. Discussion. It is considered that the feature curve has only three states who are practically always present. Consequently, cases when the analyzed sequence of coronary angiograms shows more than one injection, are not supported and the results obtained may be erroneous. If the entire washout or only a part of it was not recorded, correct results may still be achieved provided the complete state is long

enough.

The performance of the methods we have described is measured by computing correct classifications and false positives rates for each method, using a manually labeled test set. To evaluate each method – i.e. find out how appropriate it is for the automatic segmentation of the complete state – we have imposed empirical-bounds on these rates. These bounds are computed from the minimal number of true roadmaps and the maximal number of fake roadmaps which are acceptable. As already pointed out before, a successful segmentation of the complete state should yield very few if none at all false positives and enough correct classifications to cover a heart beat. We consider that the maximum allowed number of incomplete (fake) roadmaps – i.e. false positives – is three and the minimal number of complete (true) roadmaps – i.e. correct classifications – is 14. The minimal number of true roadmaps was computed such that it fully covers a heart beat: as the frame rate is 12 fps and a **normal** heart beats around 70 times a minute there are approximately ten images per heart beat. Considering the coherency constraint – which is embedded in the ML approach and can be enforced in the MAP approach – we can be sure that we have selected a succession of images and not only images showing the heart at the same position, e.g., right before the ventricular systole. At the same time this implies that in a worst case scenario, all three allowed false positives will be displayed one after the other, but no more than once every heart beat, meaning that they will cover less than one third of a heart beat which is considered acceptable.

The mean length of a sequence of angiograms is 87 frames and the mean length of the complete state according to the restrictive ground truth (A) is 23 frames, thus there are 64 frames in the inflow and washout states – i.e. potential false positives. A maximum of three false positives implies a false positives rate of 5% and 14 correct classifications implies a minimal correct classifications rate of 60%.Together they make a relative fraction of false positives of 0.18. As shown in Table 3.1, no classification method verifies these bounds when (A) is used as ground truth. However, taking the results achieved by (B) as reference, then the ML approach with the ECG-based constraint and the MAP approach with the coherency and ECG-based constraints are close if one considers the relative fraction of false positives as measure.

If (B) is used as ground truth, than there are in mean 31 images in the complete state and 56 images in the inflow and washout. This means that the maximum rate of false positives is also 5% and the minimal rate of correct classifications is 45%. Clearly, the fraction of false positives remains 0.18. As shown in Table 3.2 both the ML approach with the ECG-based constraint and the MAP approach with the coherency and ECG-based constraints are successful.

The results show that the ML approach yields less correct classifications and less false positives as well. The reason for this is that the ML approach is less adaptive than the MAP approach. This is due to the rigid way the assumption about the typical form of a feature curve is implemented. The ML approach provides a good solution for the segmentation of the complete state only if the analyzed sequence fulfills all implicit assumptions on which the method is based. If this is not the case, than it will fail and if it fails, then there is a high chance that it will fail completely, meaning that the correct classifications rate will be zero, for a false positives rate higher than zero. This is shown in Figure 3.12 using (A) as ground truth. From eight cases when it was considered that the ML segmentation failed (cc <60%), in four cases it failed completely (cc =0%). Most of the feature curves of the sequences where the ML approach failed were somewhat "unusual" (see Figure 3.13). The reasons for this varied from particular anatomies of the vessel tree to the way the physician injected the contrast-agent – i.e. the instantaneous flow of contrast agent – and length of the sequence – i.e. how much of the washout

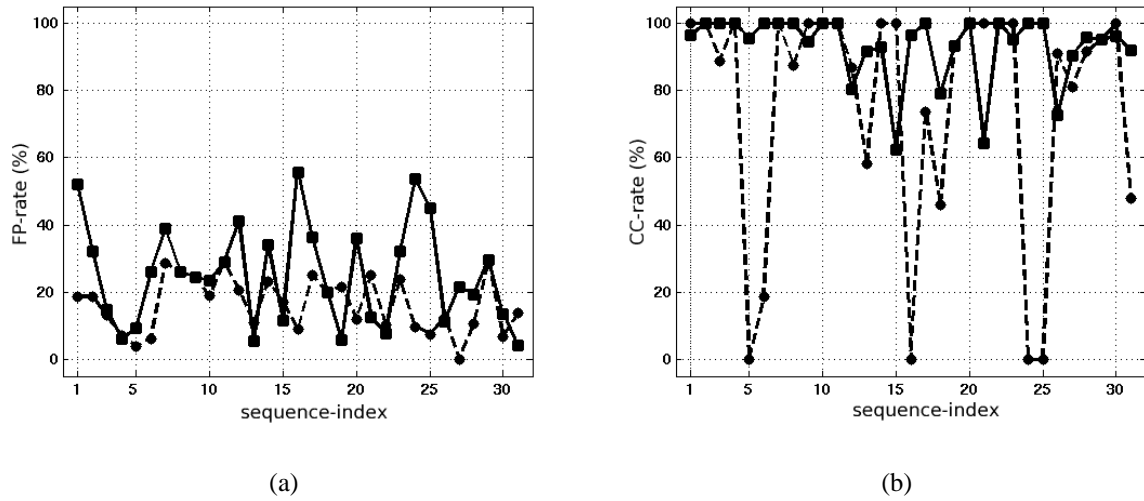(a)   (b)

Figure 3.12: False positives (a) and correct classification rates (b) for each of the tested sequences achieved by the ML approach (dotted line + bullets) and by the MAP approach (continuous line + squares).

was recorded. The ML approach takes longer to compute than the MAP.

We use the ECG to select into the complete state only images acquired between local minima of the feature curve. However there are case when the excluded images do belong to the complete state. Thus the mean correct classifications rate will decrease when the ECG-based constraint is applied, irrespective of what method was used for the initial segmentation.

One could enforce a stricter ECG-based constraint by selecting only those images recorded between local maxima of the feature curve. For this purpose, we consider that the complete state begins not before a ventricular systole , but during the ejection phase of the next one. Conversely it ends during the ejection phase of the last but one complete-state-ventricular systole. In such a case, using set A as ground truth, the ML approach yields a correct classification rate of 37.2% and a false positives rate of 3%, while the MAP approach yields a correct classification rate of 63.4% and a false positives rate of 5%. Thus the MAP approach verifies both bounds and the ML approach does not verify the bound on the correct classifications rate. The relative fraction of false positives in both cases 0.18.

We consider that the MAP approach, using the coherency and the ECG-based constraints, is the recommended method for the segmentation of the complete state. In comparison to the ML approach, the MAP approach never fails completely and if properly constrained it can verify both quality bounds on the most exigent ground truth. For the ML-based segmentation of the complete state, improved results can be achieved if the degrees of the three polynomials are determined for each analyzed sequence individually. However, in such a case, the method would be unsuited for automatic processing of sequences of coronary angiograms.

3.2.2 Sequence matching

Once a set of static roadmaps – i.e. complete-state coronary angiograms – is available, we need to choose from here the best match for each interventional image to compute the dynamic

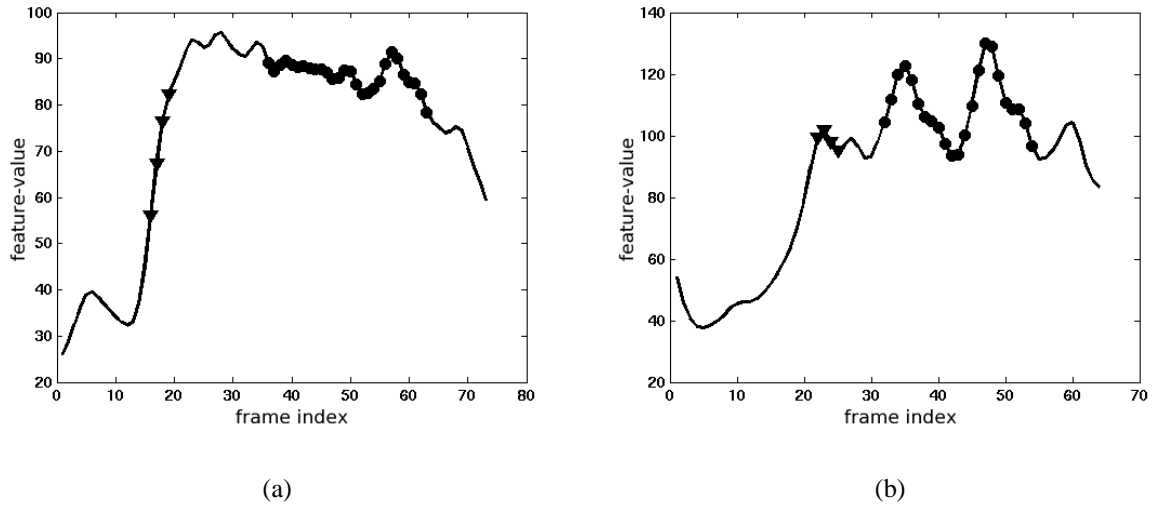|	
(a)	(b)

Figure 3.13: Feature curves where the ML approach fails. The complete state as detected by the ML approach is marked with triangles. The manually segmented complete state is marked by bullets.

roadmap.

To find the best match between an interventional image and a roadmap, one needs to measure the similarity between the two of them. We describe *two methods* to accomplish this. The first method *compares directly the two images*. The second method implies the computation of a *feature vector* which includes information on the corresponding heart and respiration phases. The similarity is then measured in the feature space.

Image-based matching

We would like to compare two images in spite of the fact that one shows vessel and the other does not. The straightforward solution is to proceed with the comparison after *eliminating the vessels* in the static roadmaps. Alternatively, one can design comparison methods which *ignore the vessels*.

Vessel-sensitive methods.   Assuming that only image information is available, the most simple way is to computing the sum of squared differences – i.e. the mean square error. Alternatively, as the two images are 2D signals, one can also compute their cross-correlation [143]. Such methods however are vessel-sensitive, thus first the vessels need to be eliminated from the roadmaps such that only backgrounds are compared.

A closing operation – i.e. maximum filtration to eliminate the vessels followed by minimum filtration to repair damaged background structures – will suppress the vessels leaving only the background. The size of the filtration windows need to be chosen such that is at least comparable with the size of the largest vessels – which are 19–21 pixels at a resolution of 512×512 pixels. Due to the relatively large size of the filtration windows, the closing will significantly alter the background eliminating potentially useful information and thus negatively influencing the computation of the similarity.

<div align="center">(a)   (b)   (c)</div>
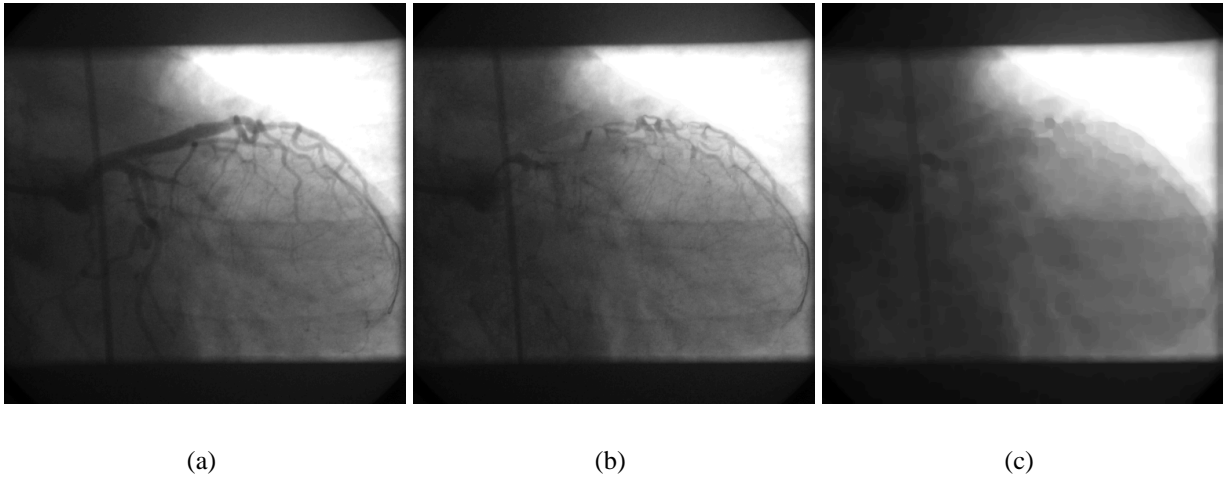
Figure 3.14: Original coronary angiogram (a), result of motion-based vessel suppression (b) and final result, after closing (c).

The strong and jerky motion of the coronary vessels can be used to enhance vessels but also to suppress them. If instead of adding to the current frame the clipped difference image between two consecutive roadmaps we subtract it, then the effect will be that the moved vessels will disappear, leaving an intact background. Although still and overlapping vessels will remain unaffected, most of the largest vessels – which usually exhibit the strong motion – will be suppressed. This effect is even more pronounced when not consecutive frames but frames further apart on the time axis are used. Thus, if in a second processing step one applies the closing operation, smaller filtration windows are now necessary and better results can be achieved. A result of such vessel suppression is shown in Figure 3.14. For optimal results, the angiograms are processed in a similar way – i.e. by applying a closing operation – before the two backgrounds now are compared. The most similar angiogram-roadmap pair is then chosen.

Vessel-insensitive methods.   Such measures can be computed based on the histogram of the image obtained by subtracting the roadmap from the interventional image [31], [32]. If the two images are similar, then a "peaky" histogram is expected. By peaky, we mean a histogram with two prominent peaks, one corresponding to the vessels and another one centered at zero and corresponding to the background. Conversely, a "flat" histogram is expected. Peaky and flat histograms can be differentiated by measures such as entropy [180] or *histogram energy* [32]:

$$HE = \sum_{k=1}^{N} p_k^2 \tag{3.13}$$

which is simply the sum of the squares of the histogram values p_k.

The plot of histogram-energy criteria over frame index for an interventional image and a sequence of coronary angiograms is shown in Figure 3.15, together with the corresponding best match.

Image-based matching. Discussion.   The quality of such vessel-sensitive criteria is afflicted by the loss of information associated with the closing operation. Thus, it is expected that a
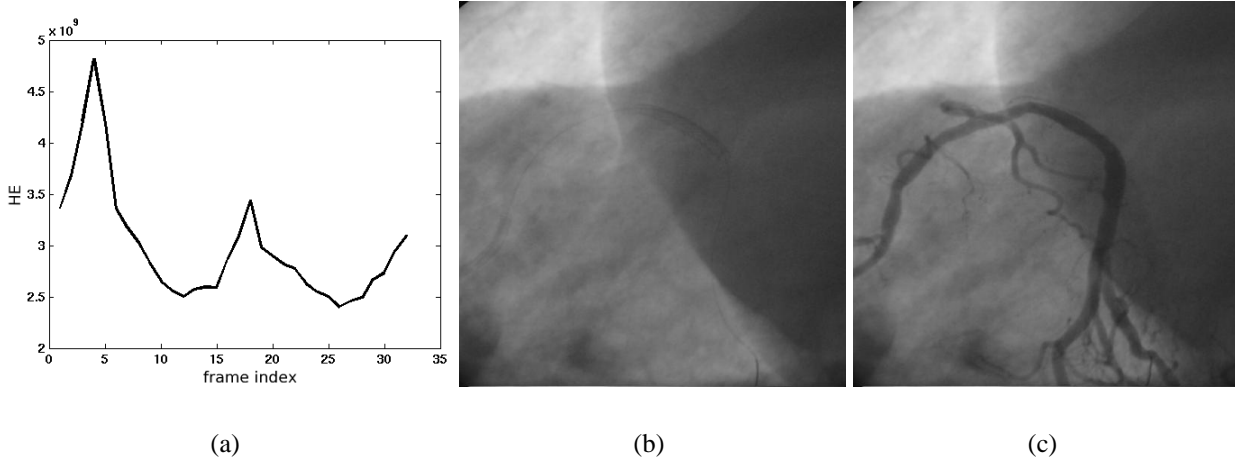
| (a) | (b) | (c) |

Figure 3.15: Curve of histogram energy criteria (HE) over frame index (a) for an interventional image (b) and the corresponding most similar complete angiogram (c).

vessel-insensitive criteria, which process unaltered image-information will yield better results.

Vector-based matching

An interventional image and its corresponding static roadmap should correspond also with respect to the position of the heart at the respective time instances. The position of the heart is influenced by the heart and the respiration phases. Therefore, two images correspond if they show the same heart and respiration phase.

Considering that the *ECG* of the patient is usually recorded synchronous with the images, it can be used to provide information with respect to the instantaneous *heart phase* at each image.

The *respiration phase* can be computed with the help of image-based similarity measures, by analyzing the correlation between the backgrounds of interventional images and roadmaps. Alternatively, it can be computed by analyzing – with similar means – the correlation between the current image and an image of known respiration state [74], [140]. However, if the diaphragm is visible, and provided that some assumptions are met – e.g. table and patient movements are compensated beforehand – the respiration phase can be more robustly estimated by *tracking the diaphragm* whose motion is directly linked to respiration [50]. Alternatively to such image-processing-based methods, a sensor can be placed on the patient, but again similar to the case of contrast agent injections such hardware solutions are unwanted.

Then, for each image a heart-position feature-vector is computed describing the heart and the respiration phases and the similarity between an interventional image and a roadmap can be computed as, e.g., the Euclidean distance between the corresponding feature vectors [74]. If the diaphragm is not visible, and e.g., the histogram energy is used as image-based similarity measure to provide respiration information, then its normalized inverse enters the computation of the distance function directly. Conversely, if a reference image is used and also when the diaphragm is visible, the information about the respiration phase is included along the heart phase in the heart-position vector.
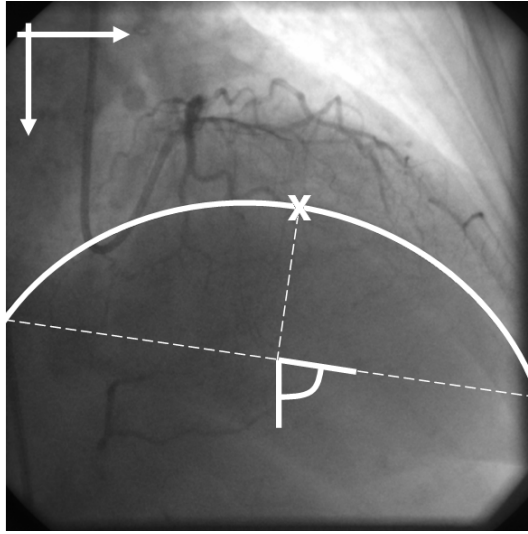
Figure 3.16: Diaphragm-position encoded in a 2D-vector: angle and Y-axis-position of the middle point. The system of reference is placed in the upper left corner.

Computation of the heart phase.   The heart phase can be easily computed from the ECG signal which is recorded synchronous with the analyzed sequences. If one cardiac cycle is defined as the time interval between two R-peaks of the ECG signal, then the heart phase can be computed as the portion of this interval which has passed before the analyzed image was acquired. Thus, the cardiac phase is $\Phi = 0$ at the first R-peak and $\Phi = 1$ at the last one. According to [176] absolute variations of this interval can be normalized as described in [195].

Computation of the respiration phase by diaphragm tracking.   We describe the position of the diaphragm in an image starting from the boundary of its projection. The position is encoded in a 2D vector which can then be used to describe the respiration phase. When computing this vector, we assume that the diaphragm boundary meets the image border always in two points only and we establish an image coordinate system whose origin lies in the upper left corner. One component of this vector is the angle which the line underlining the diaphragm (from one point where it meets the image border to the other one) makes with the Y axis and the second component is the y coordinate of the middle point of the diaphragm border. This is shown in Figure 3.16.

To find the boundary of the diaphragm – which we assume to be of a circular shape – we use the Hough transform for circles (see Appendix A.4), applied on the edges of the image [34]. To avoid that the detection algorithm is mislead by spurious edges we remove structures smaller than a certain size (i.e. potential vessels) by morphological closing. At the same time we also eliminate X-ray shutter edges. To restrict the search in the Hough parameter space and thus provide real time capabilities to our algorithm, prior anatomical knowledge about position and size of the diaphragm is considered. In subsequent frames, approximative diaphragm position and size are predicted from previous results.

There are also cases when the computed Hough circle does not correctly approximate the boundary, e.g., when the circularity assumption fails. To deal with such difficulties, for each detection result a confidence measure is computed. If the confidence measure indicates a poor

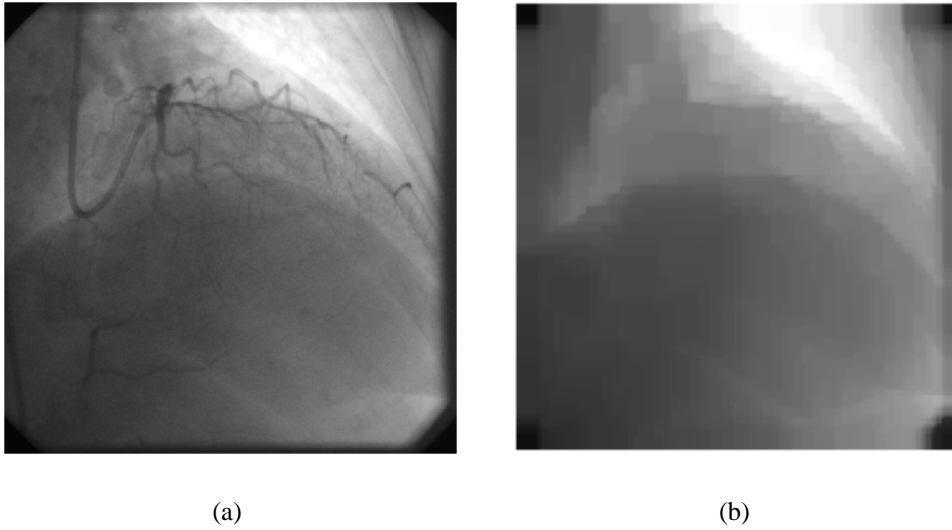<p style="text-align:center">(a)   (b)</p>

Figure 3.17: Original cardiac X-ray image (a) and result of morphological closing (b).

fit, the result is refined by an active contour algorithm (see Appendix A.3).

Preprocessing. To improve the performance with respect to speed, the input images are first down-sampled by a factor of 16, thus we work on 128x128 pixels large images down from the original resolution of 512x512 pixels. Before applying the Hough-transform every image is preprocessed to eliminate spurious edges given by, e.g., vessels or the X-ray shutter.

A gray level morphological closing operator is applied to the analyzed image prior to edge points selection, to eliminate vessels while retaining the diaphragm. The results are shown in Figure 3.17.

To eliminate the influence of shutter edge points we have applied two methods. The first most straightforward method is to simply ignore all the edges at a certain distance away from the image borders. However, it is difficult to properly establish this distance as there is no universally valid shutter setup. Alternatively, by analyzing the edge maps obtained for a number of images one can identify the shutter edge points as those belonging to objects which remain static. The edges of static objects are obtained in the the minimum-intensity projection of the edge maps of a number of N previously acquired images. The value of a pixel in the projection-frame is computed as the minimum over all values recorded at that site in the N images. To attenuate still edges, the minimum-intensity projection is subtracted from the edge maps of each analyzed image before computing the circular diaphragm fit (see Figure 3.18). Due to image noise and other external influences, as e.g., the AGC, the edge strength of the static edges varies lightly from image to image. Therefore, after subtracting the projection-frame the static edges are not completely eliminated but rather strongly attenuated (see Figure 3.18 (c)).

Tracking. Tracking is done with the help of the Hough transform for circles. To speed up the computation in the case of diaphragm tracking, we make use of prior knowledge on where the diaphragm can be expected and how large it is and search only in a certain ROI in the Hough-space. Typically, we start with a large ROI to detect the diaphragm in the first frame. The ROI is then reduced and kept constant as the diaphragm cannot travel far between consecutive frames.
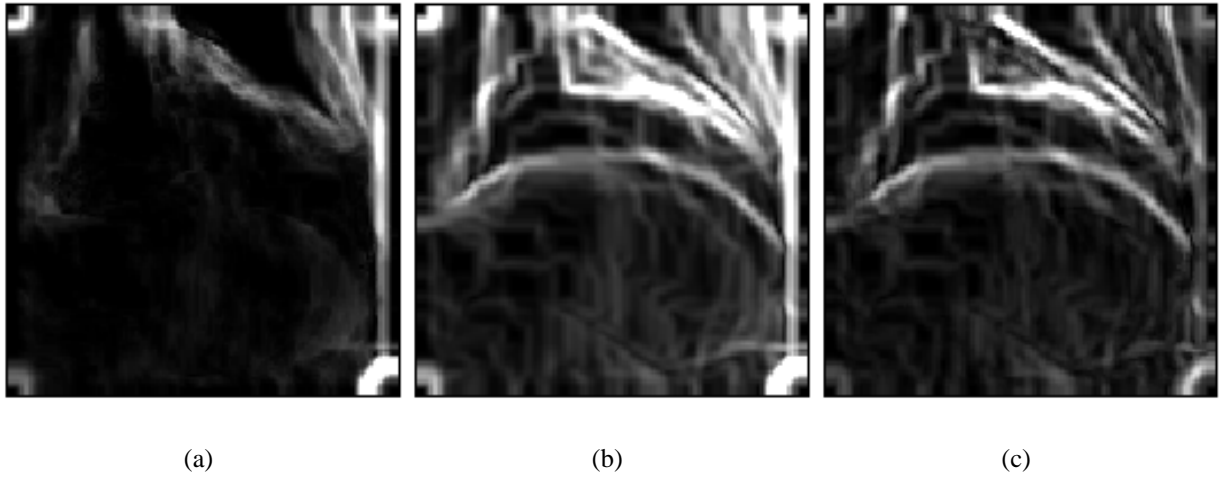
(a) (b) (c)

Figure 3.18: Projection-image (a), edge map for one image (b) and result of still-edges-attenuation (c).

In the first image the ROI is defined expecting the diaphragm to be situated in the lower part of the image.

During tracking, the ROI is centered always at Hough-point corresponding to the best fit found in the previous image. Figure 3.19 shows the results of Hough based diaphragm detection for a high-dose sequence.

Confidence measure. The Hough transform may fail to correctly capture the diaphragm in two cases:

- **ROI too small.** This case is encountered when the diaphragm does have a circular form, but this cannot be found with the Hough transform because the ROI is too small.

- **Diaphragm not-circular.** This case appears if the diaphragm does not have a circular form.

To detect such cases we define a measure of confidence in the Hough result, which is evaluated for each frame.

We propose two confidence measures: one is based on the Hough accumulator, while the other one is image-based. The accumulator-based confidence measure is simply the number of votes obtained in the ROI by the best circular fit (i.e. the ROI maximum). Assuming that the observable diaphragm border-length remains approximately constant during the entire analyzed sequence, the accumulator maximum should vary only within a small interval over the frame index. A large decrease indicates then that the corresponding Hough circle does not correctly approximate the diaphragm. Under such circumstances, the quality of the peak may be additionally characterized by relative confidence measures which show how well the peak is expressed in the accumulator, like e.g., the entropy[6]. In Figure 3.20 (a) the accumulator-based confidence measure is plotted against the frame index for a sequence where the observable diaphragm border length remains approximately constant.

---

[6]It is assumed that in this case that the Hough space is appropriately sampled. We observe that a very dense sampling of the accumulator will always produced well expressed peaks.
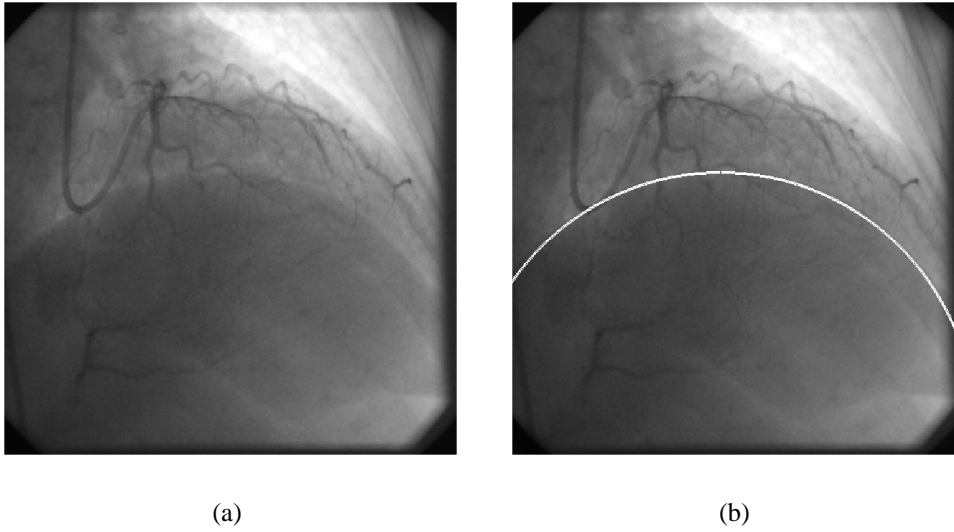
(a)   (b)

Figure 3.19: Original image (a) and result of diaphragm detection by Hough-based circular approximation (b).

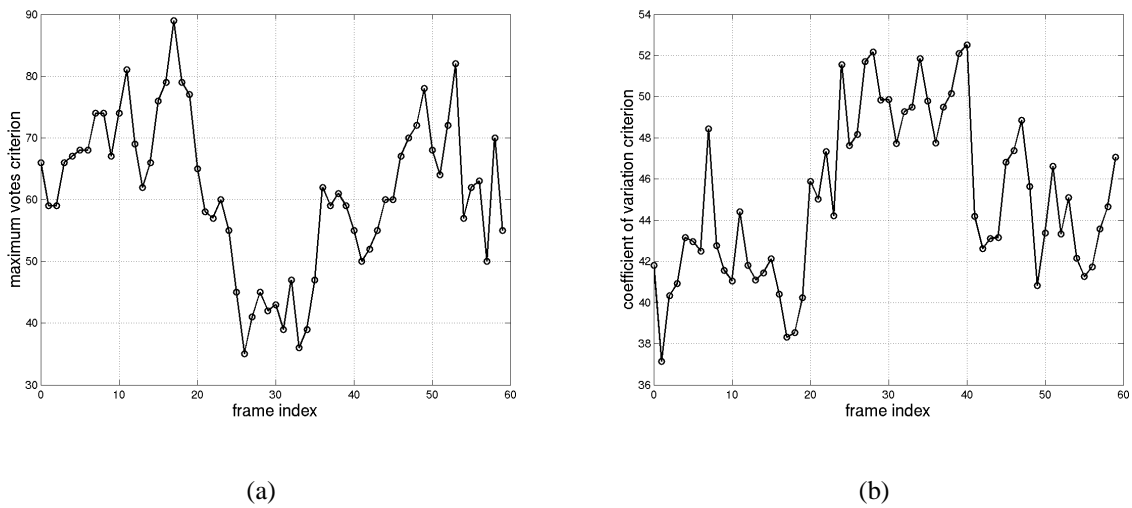(a)   (b)

Figure 3.20: Accumulator based confidence measure (a) and image-based confidence measure (b) for a sequence where the visible diaphragm border remains constant in length.

We propose also another confidence measure, which is not based on the Hough accumulator but on the image itself. This confidence measure is computed as the coefficient of variation (CV) of the pixels along the detected circular approximation of the diaphragm in the corresponding edge map. A high CV signals a poor fit irrespective of the visible diaphragm perimeter. The CV is defined as:

$$CV = \frac{s_x}{\overline{x}} \tag{3.14}$$

where s_x is the standard deviation of the samples x_i and $\overline{x}$ is their mean. An example is shown in Figure 3.20 (b).

Evaluating the confidence measure.   The confidence measure is compared to a certain threshold to decide if the corresponding result is correct. Conversely, an active contour is used to refine it. Since the image-based confidence measure is a decreasing function of the quality of the fit, a value above the threshold indicates a poor result. The accumulator-based confidence measure being an increasing function of the quality of the fit behaves in the opposite way. A threshold is computed for each available sequence of roadmaps during an intervention and then used also for all interventional images of similar projection.

The algorithm is based on the assumption that for a large majority of images the Hough circle is sufficient to correctly approximate the diaphragm. Thus, we cannot say beforehand that in the sequences of roadmaps available before the intervention, there are enough (if any) poor Hough fit results to allow setting the threshold using standard unsupervised classification algorithms. On the other hand we cannot be sure that no poor Hough fit appears. At the same time, we have no access to a labeled training set of confidence-measures. Therefore, the threshold is established in a heuristic manner.

We propose two methods for establishing the threshold. In the first case, the threshold is based on the mean confidence measure over the entire analyzed sequence. For the accumulator-based confidence measure we consider (on empirical basis) a value which is 15% less than the mean value while for the image-based confidence measure we consider a value which is 10% above the mean value. Alternatively, instead of basing the threshold on the mean, we consider the extremum of the confidence measure corresponding to the best possible fit – i.e. the maximum for the accumulator-based measure and the minimum for the image-based measure. In this case, for the accumulator-based confidence measure we consider (also on empirical basis) a value which is 40% less than the maximum while in the case of the image-based confidence measure we consider a value which is 25% above the minimum. In our experiments both thresholds proved successful, however the one based on the extremum is more robust.

Active contours-based refinement.   If the confidence measure indicates a poor fit, an active contour is initialized using the Hough circular-approximation. The snake evolves then to the correct diaphragm boundary. The result obtained with this approach for an image where the Hough-based tracking has failed is shown in Figure 3.21. To speedup the procedure, the outline of the diaphragm, as found by the Hough transform, is first sampled obtaining thus a reduced number of snakels. The positions of the converged snakels are linearly interpolated to obtain the snake-refined diaphragm outline.
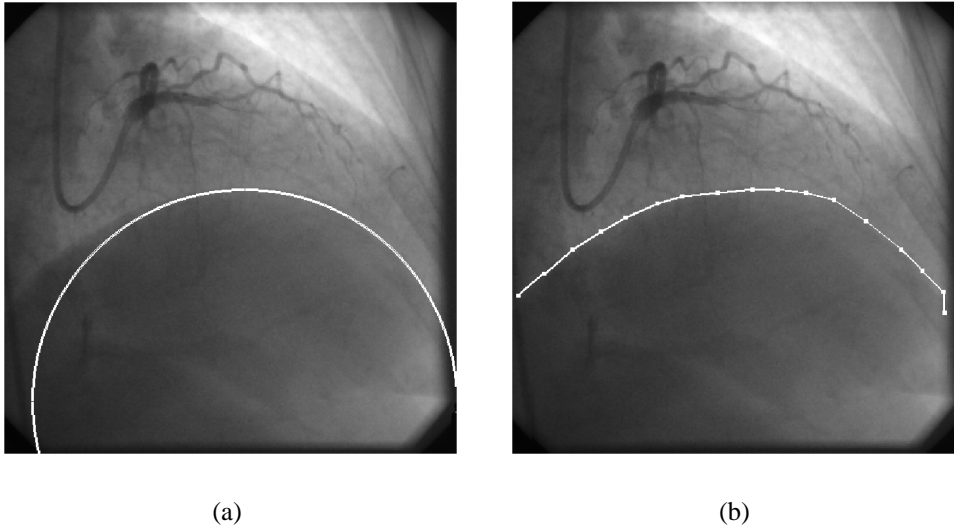
(a)   (b)

Figure 3.21: X-ray image with erroneous circular approximation of the diaphragm (a) and the result of the active contours based refinement (b).

Diaphragm tracking. Experiments and discussion.   For testing we have used six sequence-pairs from four different interventions. A sequence pair consists of a set of coronary angiograms and a set of fluo images showing no vessels.

– **Circular approximation.**   We assume that in general a circular approximation is well suited for the diaphragm border. To verify this assumption, we compared the hand segmented diaphragm border to an optimal circular approximation[7]. We have observed the approximation error (i.e. minimum mean distance) at a resolution of 512×512 pixels. The mean error was 0.24 pixels (standard deviation 8.05).

To show that the Hough transform is able to correctly find the diaphragm, using the same experimental setup as above we have compared the hand segmented diaphragm border to the circular fit found by the Hough transform. The mean error was 2.37 pixels (standard deviation 9.35) in this case.

To show that the snake-based refinement improves the Hough-based diaphragm-detection, we have compared the manual ground-truth also to the detection-results obtained using snakes to refine the Hough-fit when the confidence in that result was poor. In this case the mean error was 1.20 pixels (standard deviation 8.43).

– **ROI-size.**   A parameter of major importance for the speed of the tracking is the size of the accumulator-ROI. A large accumulator ensures that the optimal Hough-fit is always found and thus the snake is used only when the diaphragm does not have a circular outline. However, the larger the ROI, the slower the tracking and therefore, a balance has to be found between the size of the ROI and the tracking-time.

In this case of major concern for the optimization of the ROI-size are the intervals where

---

[7]The optimal circular approximation was taken as the one having the minimum mean distance to the hand segmented diaphragm border.
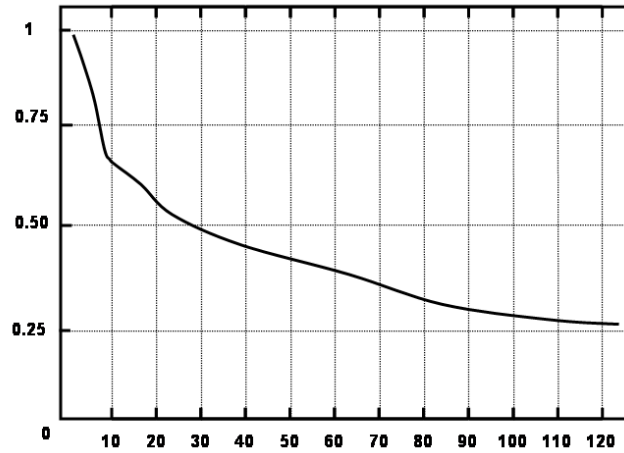
Figure 3.22: Normalized mean CV for ROIs of different sizes (in pixels).

the coordinates of the center of the circular-fit should be sought. The radius should be searched for in a very small interval – i.e. a few (e.g. seven) pixels – or even be kept constant after being found in the first tracked image. The reason for this is that in the analyzed images the variation of the size/radius of the diaphragm is very small in comparison to the variation of its position. In the following, the optimization of the ROI-size is done mainly with respect to the center-coordinates. The projection of the ROI on the center-coordinates plane of the Hough-space is a square and therefore when we talk about size we refer to the side of this square. The ROI-size on the radius-axis of the Hough space is in the following calculations constant and equal to seven.

To optimize speed, one has to consider both the largest ROI-size for which the search for the Hough-optimum still takes less time than the snake needs to converge and the smallest ROI-size for which the diaphragm is practically always correctly found, if it is circular. Thus, the snake is used only when the diaphragm does not have a circular form.

Empirically we found that the largest ROI-size, such that a Hough-search is faster than the snake, is larger than 90 pixels. To find the smallest ROI-size such that the snake is used for a minimal number of times, we have computed the mean CV along the Hough-diaphragm-outline over all available sequences for ROIs of different sizes. The smallest size that we seek will be found at the beginning of the interval of values for which the CV is constant. This curve of mean CV values over the ROI-size is shown in Figure 3.22 and it was computed for images with a reduced resolution of 128×128 pixels. Starting from a ROI with a side of 80 pixels, the CV remains largely constant and therefore staring there, the discrepancies between the Hough-fit and the true outline of the diaphragm appear only as a consequence of the latter not being a true circle. Thus, to use the snake only when the diaphragm is not circular and thus for a minimal number of times, the size of the ROI should be larger than 80.

Therefore, the optimal size of the center-coordinates ROI should be between 80 and 90 pixels. We have used in our experiments a ROI of size 81 for the center coordinates and seven for the radius.

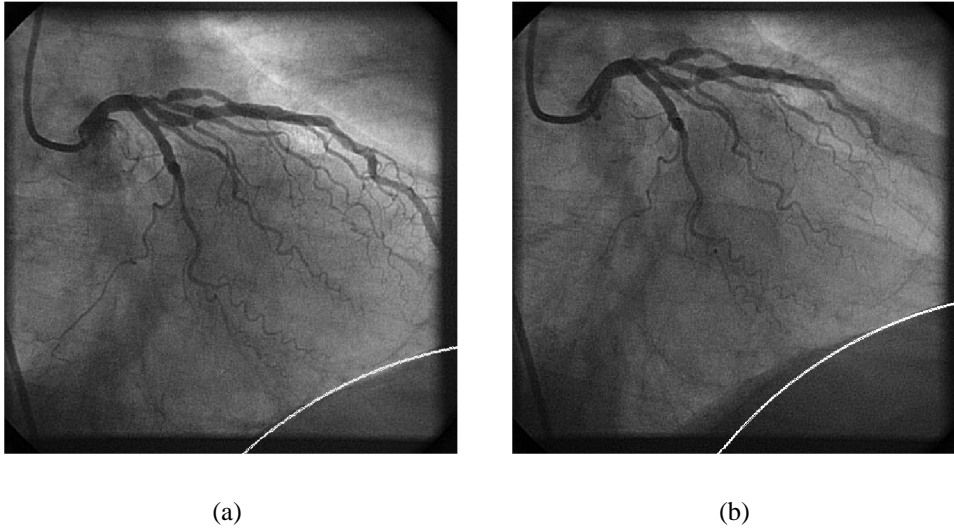(a)   (b)

Figure 3.23: Two images where the accumulator-based confidence measure has signaled a poor circular approximation of the diaphragm. Such a conclusion is false for image (a) and correct for image (b).

– Confidence measures.   The accumulator-based confidence measure will fail, if the assumption with respect to the constancy of the observable diaphragm border length is not fulfilled. Such a case occurs for certain X-ray projections. Then, the accumulator-based measure will exhibit low values both in the case of a poor fit in an image where the visible diaphragm border length is large and in the case of a good fit in an image where the visible perimeter is small (see Figure 3.23).

In Figure 3.24 the image-based confidence measure is plotted against the frame index for a sequence where the visible diaphragm border length varies. For comparison, the accumulator-based confidence measure is also plotted. Unlike in the previous case (See Figure 3.20), the two measures are now less correlated – correlation coefficient $q = -0.3140$ compared to $q = -0.6907$. The image-based measure reflects better the ground-truth.

In our experiments we have primarily used the image-based confidence measure as it has a better generalization performance, although it has a smaller dynamic range.

We illustrate the respiration information extraction by diaphragm tracking on two representative cases. In the first case, the length of the visible diaphragm border remains constant in the analyzed sequences. Figure 3.25 (a) and (b) shows the image-based confidence measure for both high-dose and low-dose images plotted against the frame index. The threshold for the image-based confidence measure was computed using the extreme value over a high-dose sequence. The value of the threshold was 46.42. Using this threshold, for a number of 24 high-dose images out of a total of 60 and for a number of five low-dose images out of a total of 126 it was necessary to improve the Hough circle fit by active contours.

In the second case, the length of the visible diaphragm border varies in the analyzed sequences. The confidence measure plots for both high-dose and low-dose sequences are shown in Figure 3.25 (c) and (d). In such a case, the accumulator-based confidence measure is not able to provide reasonable results (see also Figure 3.23). The threshold for the image-based confidence measure was in this case 54.30. With this threshold, for a number of 18 high-dose

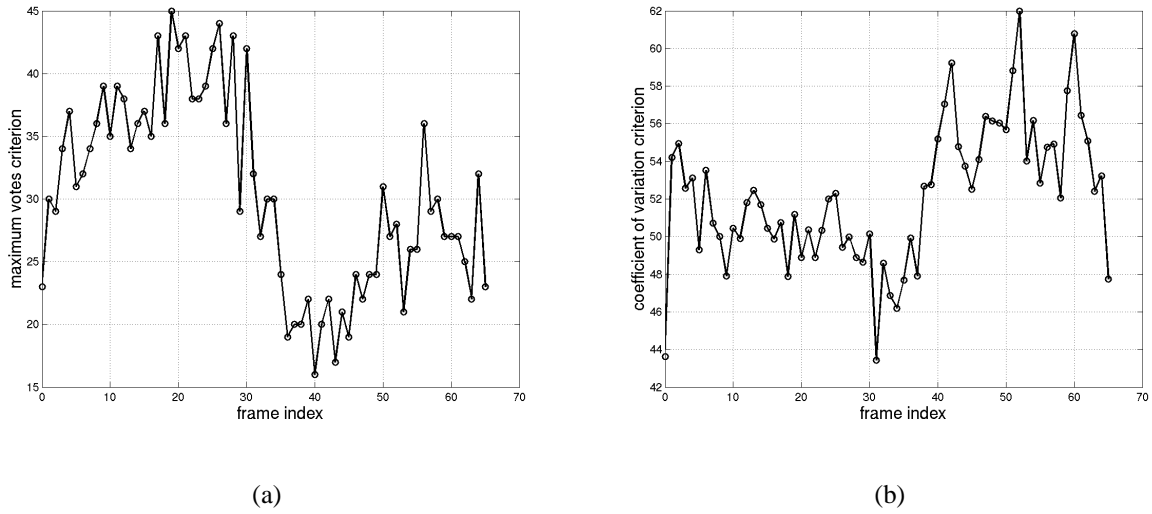(a)   (b)

Figure 3.24: Accumulator-based confidence measure (a) and image-based confidence measure (b) for a sequence where the visible diaphragm border is no longer constant in length. The results obtained with the image-based confidence measure correspond better to reality.

images out of a total of 66 and for a number of 18 low-dose images out of a total of 111 it was necessary to improve the Hough circle fit by active contours.

– Diaphragm description vector.   In Figure 3.26 the similarity – measured by the Euclidean distance between the corresponding 2D description-vectors and therefore the brighter the gray level the lower the similarity – between the tracked diaphragm positions in a low-dose and a high-dose sequence of the same patient for both example-cases is shown. The repetitive pattern which may be observed, is linked to the respiration frequency and to how many respiration cycles were recorded in the analyzed sequences. In the first case (Figure 3.26 (a)), the high-dose sequence showed one and a half respiration cycles while the low-dose sequence showed four and a half respiration cycles. In this case, the patient was allowed to breath freely. In the second case (Figure 3.26 (b)), the patient was allowed to breath freely during the acquisition of a high-dose sequence and he was told to hold breath for a while during the acquisition of a low-dose sequence.

Figure 3.26 illustrates that by tracking the diaphragm and specifying its position, as described in our algorithm, pairs of images with similar respiration status can be found. Thus, our approach represents a viable method of extracting image based respiration information.

The evolution of the two components of the 2D-description vector for both high and low-dosage sequences for the case when the patient was allowed to breath freely are shown in Figure 3.27.

– Abrupt and sudden motion of the diaphragm.   To speed up the algorithm we use a small, constant Hough accumulator ROI. This strategy however may lead to poor fit results also in the case of sudden and large diaphragm motion when the center of circular fit jumps out of the ROI. Such cases are detected using the confidence measure and corrected by the snake which in turn takes extra time. For further implementations, a variable size ROI could be used. In such a case
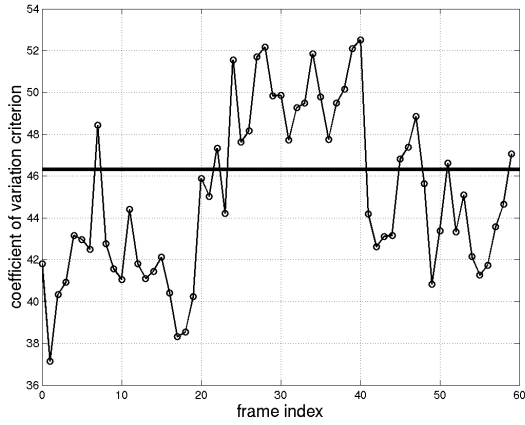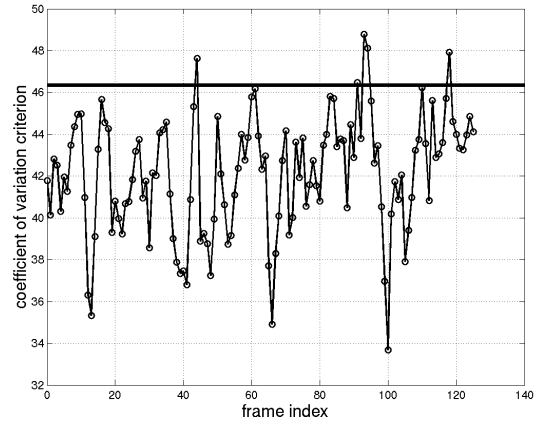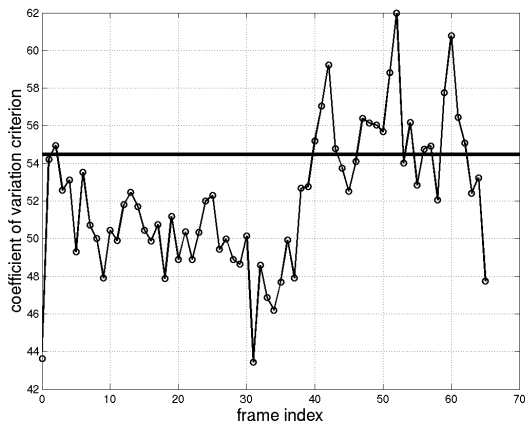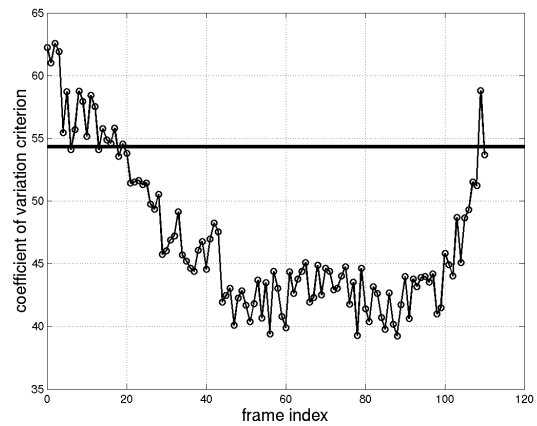
Figure 3.25: The image-based confidence measure for the high-dose sequence (a) and for the low-dose sequence (b) when the length of the visible diaphragm border was constant and the image-based confidence measure for the high-dose sequence (c) and for the low-dose sequence (d) when the length of the visible diaphragm border was not constant. The black horizontal bar marks the value of the threshold for each case.
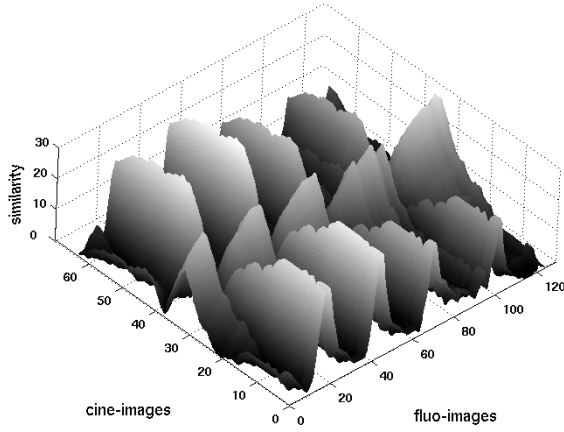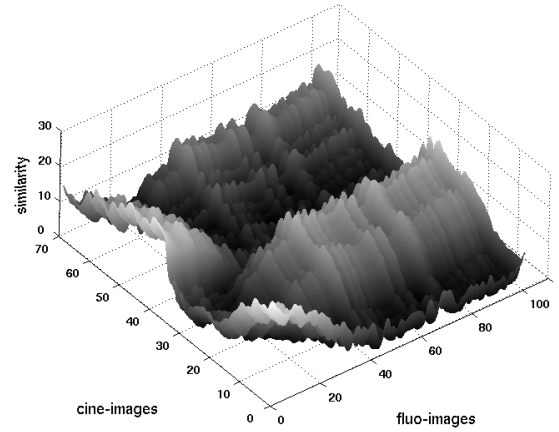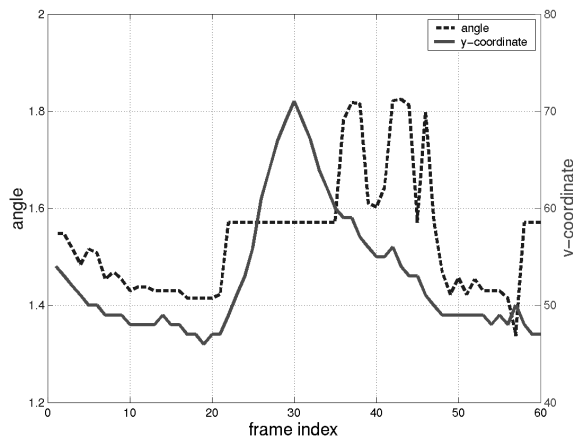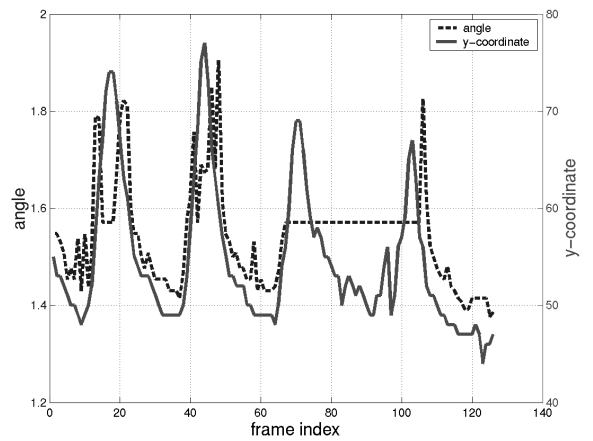
Figure 3.26: Respiration-phase information using diaphragm tracking when the visible diaphragm length is constant (a) and when the visible diaphragm length is not constant (b).

Figure 3.27: The evolution over frame index of the components of the 2D description-vector for the contrasted sequence (a) and for the interventional sequence (b).

the ROI size would decrease constantly as long as the confidence measure is 'clearly' above the threshold. If this is no longer the case, the ROI is reinitialized to a large size and the tracking continues. The decision interval within which the confidence measure is no longer assumed to be 'clearly' above the threshold, as well as the ROI reinitialization size, may be set empirically and will directly influence the algorithm's performance with respect to speed.

3.2.3 Segmentation and registration of surgical tools

After a correspondence has been established between a live-image and a static roadmap, the guidewire tip – which is the main navigational instrument – and eventually the balloon markers need to be superimposed.

Introduction

There are two strategies which can be followed to superimpose the surgical tools after a roadmap has been found for the current interventional image. The first possibility is to segment them and then to display them in a *highly contrasted manner* on the roadmap. The second possibility is to *combine* the two images such that the roadmap is displayed in a weaker contrast over the interventional image and a well-perceived overlay is achieved [26]. The latter method is impracticable if the contrast in the interventional images is very low. Also, it can encounter an acceptance problem by the physicians, as they can be confused by such overlay where vessels can become similar to other background structures. Therefore, we propose to use the first method.

To realize the overlay, first we need to *segment the guidewire-tip and the balloon markers*. Second, the roadmap and the interventional image need to be *geometrically correlated* such that the guidewire tip will be displayed into the vessel and thus the perceived quality of the overlay is good.

Compensate table and patient motion.   Potential mismatches, between an interventional image and its corresponding roadmap, can appear as a consequence of voluntary or involuntary patient and table movement. The importance of such mismatches depends on the amplitude of the motion. Typically, table motion has a large amplitude, while patient motion has a smaller amplitude. We propose to compensate them by geometric transformations, whose parameters are computed using landmarks, which are applied on the operating table and on the skin of the patient respectively.

However, only rigid transformations [189] – i.e. rotation and translations – are allowed, for two reasons:

- **First.** The overlay should function in frame-rate and thus the processing should be kept as simple as possible.

- **Second.** Affine, projective and elastic models can modify the input image to such an extent that in some cases the perceived quality of the overlay decreases beneath the acceptance threshold.

Register vessels and surgical tools. Usually, a rigid model will do the job, but sometimes, e.g., in the case of small rotations of the patient, the projective geometry of the acquired image changes beyond the reach of such model. Then, the guidewire tip typically lies close to the vessel but nevertheless out of it. In such a case, an additional correction is needed to bring the guidewire tip into the vessel. This correction needs to be done using only image cues.

Again two strategies can be followed:

- **Bring the guidewire tip to the vessel.** This strategy implies modifying the interventional images and displaying the roadmap with the registered guidewire tip into the corresponding vessel.

- **Bring the vessel to the guidewire tip.** This strategy implies the computation of a set of parameters of an again linear transform which is supposed to modify the roadmap.

Although tampering with the interventional images is potentially more critical than a modification of the roadmap, the first method has the advantage of avoiding the additional computation of a transformation. Thus, we propose to segment the guidewire tip and register it to the closest vessel.

Assuming a successful sequence-matching, the balloon markers are practically always correctly positioned – due to their small size and circular form – and therefore need not be registered to the vessel.

Segmentation of the guidewire tip

The guidewire tip appears as a thin dark oriented structure in the interventional images. To achieve a segmentation, a two step approach is proposed. In the first step, guidewire-like structures are enhanced obtaining thus a guidewire map where the object pixels are lighter than the rest. The guidewire is then segmented in the guidewire map by a percentile threshold.

The guidewire map. We have conducted experiments with two different methods to obtain a guidewire map. In the first method, high-frequency structures are enhanced by means of a high-pass filter. In the second method, oriented structures are enhanced by analyzing the first (i.e. largest) eigenvalue of the structure tensor which describes local orientation [108], [118] (see also Appendix A.5).

The high-pass filtration result is obtained by subtracting from the original image the low-pass filtered image. The low-pass filtration kernel was Gaussian. The optimal cut-off frequency of the low-pass filter – i.e. the size of the σ parameter of the Gaussian function – was empirically determined (see Table 3.4). An enhancement result is shown in Figure 3.28 (b). During processing shutter artefacts are ignored.

The structure tensor can be used to describe only one orientation. As the guidewire tip is an oriented structure, we propose to use as guidewire map the first eigenvalue image of the structure tensor. Provided the size of the neighborhood is chosen in relation to the diameter of the guidewire, then the first eigenvalue will exhibit large values mostly over the guidewire. An enhancement result is shown in Figure 3.28 (c). Prior to computing the structure tensor, a Bothat-filter is applied such that only guidewire-like structures are select and the rest are suppressed.
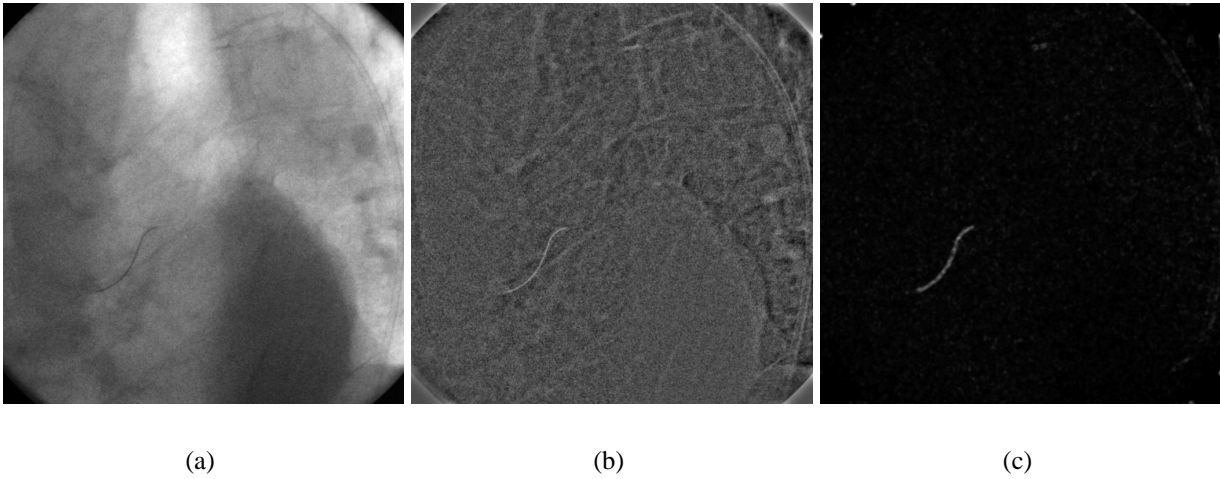
| (a) | (b) | (c) |

Figure 3.28: Original image (a), result of high-pass filtration (b) and result of tensor-based enhancement (c).

The segmentation.   One can assume that after enhancement, in the guidewire map there are only two well separated pixel classes viz. background and guidewire. Then, to segment the guidewire a fixed threshold should suffice. As the image area which is covered by the guidewire tip can be computed in advance, based on the size of the guidewire tip and the resolution of the image acquisition system, we propose to use a high-percentile threshold. Precisely which percentile should be used can be established using the available ground truth by simply trying different values. The best results were achieved by the 99'th percentile of the histogram of the guidewire map. However, this threshold does not return perfect results, there are still many false positives (see Figure 3.29 (b)), which we consider to be a consequence of the poor SNR of interventional images. To improve the segmentation, we use the observation that the guidewire tip, being the "most oriented" structure in the image, will contain also the overall maximal value in the guidewire map. Thus, we select into the segmentation all object pixels connected to the position of this maximum. A final segmentation result is shown in Figure 3.29 (d). To make sure that the entire guidewire is selected, we apply a binary closing operator to the initial segmentation results, before the maximum-based selection. (see Figure 3.29 (c)).

When the orientation-based guidewire map is used, due to the localization-detection trade-off, the segmentation result contains rather a region where the guidewire can be found with a high probability. To obtain the final guidewire, the result of segmentation is morphologically thinned to a line.

Segmentation of the guidewire tip. Experiments and results.   To find out which enhancement method is best suited for the computation of the guidewire map, we have computed the AROC of a percentile threshold used to segment the guidewire tip on each of the enhancement results obtained by every proposed method. For this purpose, we have used a manually labeled ground truth, which includes 120 images from four different interventions (30 images per sequence). Each sequence has different SNR from high to low. One sequence shows also balloon markers. The human expert was asked to draw a line (one pixel thick) along the guidewire tip, at the position where he thinks this is found. To compensate for the poor quality of the
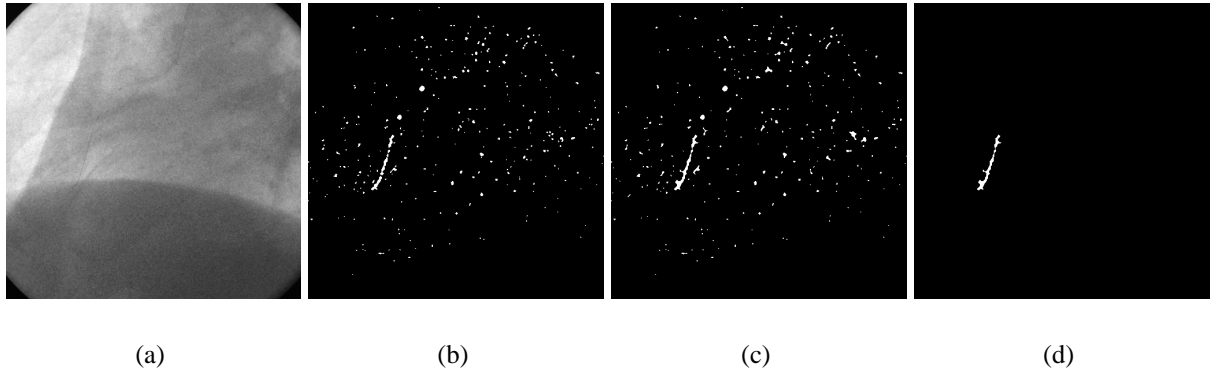
<div align="center">(a)   (b)   (c)   (d)</div>

Figure 3.29: Original image (a), initial segmentation result (b), result after closing (c) and after maximum selection and elimination of balloon markers (d).

manual ground-truth [116], these results were dilated and when compared with the automatic result a successful segmentation was considered only when the mutual overlap between the manual result and the automatic result was above a certain value [75]. Such a procedure is justified also by the classical localization-detection tradeoff problem [34], which is encountered mainly when using the structure tensor-based enhancement method. Therefore, the automatic results does not contain a precise localization of the guidewire tip but rather a region where the guidewire is located with a high probability.

The main parameter of the high-pass filter is σ the standard deviation of the Gaussian filtration kernel. The larger σ is the smaller the pass-frequency of the high-pass. The size of the filtration kernel is chosen in agreement with σ such that $Size = 5 \cdot \sigma$ [188]. The main parameter of the structure tensor-based enhancement is the size of the neighborhood, which is clearly identical with the size of the Bothat-window. The size of the derivative kernels is chosen in agreement with the size of the neighborhood, i.e., slightly smaller [108]. The results are shown in Table 3.4 and Table 3.3 for the high-pass and tensor-based enhancement respectively. Each time for different σ and neighborhood sizes.

As guidewire map we have used the tensor-based enhancement. The segmentation results obtained on the four test-sequences for different percents of mutual overlap (O) and for neighborhood sizes $N = 5, 7, 9$ are given in Table 3.5.

Segmentation of the guidewire tip. Discussion. Analyzing this data it appears clear that the larger the standard deviation σ and the neighborhood size N, the better the results achieved. However, the rate of improvement decreases significantly starting from $\sigma = 7$ and $N = 9$, which we consider to be the optimal parameters for each method respectively. It can be also observed that the tensor-based enhancement returns better results than the high-pass filter. Thus, we choose the tensor-based enhancement method to compute the guidewire map.

As shown in Table 3.3, starting from $N = 9$, the quality of the vessel map improves only slowly, and thus it seems that for the computation of the guidewire map $N = 9$ should be chosen. However, the results in Table 3.5 show that the $N = 7$ is actually the best choice. We consider that although at the scale of the entire image the contrast of the guidewire-tip to the background improves with N, there are also more small dark background-structures (i.e. background-noise) from the direct vicinity of the guidewire which are also enhanced. Such

N	3	5	7	9	11	13
S_1	87.92	96.86	98.47	99.05	99.23	99.39
S_2	71	79.90	87.06	94.03	95.68	96.47
S_3	76.22	85.47	90.69	94.03	95.23	96.27
S_4	70.37	78.27	84.68	88.55	90.18	91.57

Table 3.3: AROC for the tensor-based guidewire map for different neighborhoods N.

σ	3	4	5	6	7	8
S_1	73.90	81.99	87.65	90.83	93.02	94.18
S_2	56.94	63.31	69.01	72.99	76.41	78.82
S_3	54.04	57.95	61.64	64.46	67.38	69.57
S_4	50.78	54.00	56.92	59.22	61.18	62.66

Table 3.4: AROC for the filter-based guidewire map for different σ.

background-noise is then segmented in the initial segmentation where it appears linked to the guidewire and is thus selected with it in the final segmentation.

From $O = 70$ the relatively bad precision of the manual segmentation starts having an influence. Clearly this is also the mutual overlap where the most inferences with respect to the quality of the automatic segmentation can be made.

The segmentation of the guidewire-tip is based on the assumption that it is the sole and best oriented X-ray absorbent in the image. This assumption is no longer valid when sewing wires or balloon markers are present in the image. Such structures need to be segmented in advance and ignored during the maximum-based selection of the guidewire tip.

Registration of the guidewire tip

After segmenting the guidewire tip in the interventional image, it is projected on the corresponding roadmap. At this stage it is assumed that the guidewire tip is either within the vessel, or very close to it. To precisely place the guidewire into the vessel, we use an active contour-based approach. The thinned guidewire (see Figure 3.30 (b)) – i.e. the guidewire centerline – is seen as an elastic curve – i.e. a snake – which advances on the roadmap from its initial position to the vessel situated next to it under the influence of internal and external energies. The internal energies define elasticity and stiffness and constrain the curve such that "unnatural" fitting results

N	5	7	9	5	7	9	5	7	9
S_1	100	100	100	100	100	100	85.71	85.71	66.67
S_2	100	100	100	100	100	100	66.67	76.19	57.14
S_3	100	100	100	100	100	100	80.95	80.95	57.14
S_4	95.24	95.24	85.71	76.19	76.19	66.67	33.33	47.62	23.81
O	50			60			70		

Table 3.5: Percents of correct classifications for different neighborhood size N and overlap O.
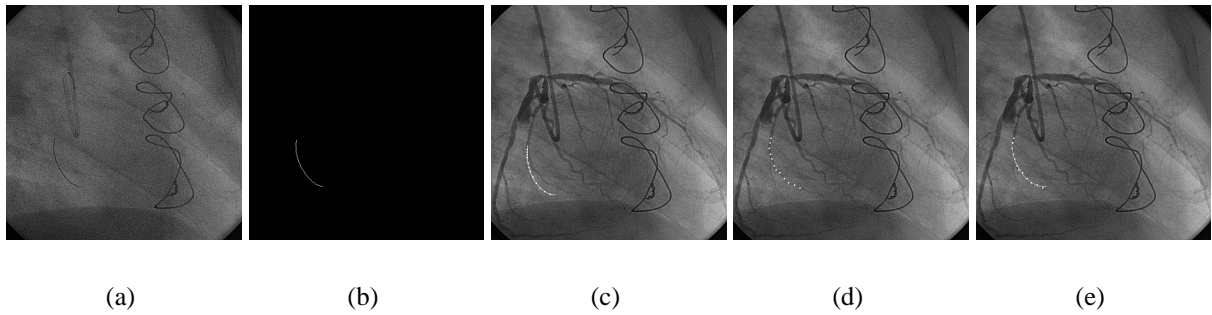
(a) (b) (c) (d) (e)

Figure 3.30: Interventional image (a), result of guidewire-tip segmentation (b), guidewire-tip projection on the roadmap (c), snakels positions after registration (d) and final result (e).

are excluded. The external energies are defined to attract the curve into the vessel.

To decrease the time needed to find a good fit to a minimum, the guidewire is sampled at equidistant positions – obtaining thus so-called snakels – and only these points are fitted (see Figure 3.30 (c)). The continuous guidewire is then interpolated among snakels. If the guidewire enters relatively small vessels, then it changes their appearance. Therefore, small vessels with a rather curved appearance will be straighten as soon as the guidewire reaches them. Clearly this phenomenon cannot be observed in the roadmaps, thus the guidewire tip will be fitted to the curved vessel and the final result will show a guidewire with an "unnatural" form. Due to the internal energy terms, the final fitting result shows a "graceful degradation" in such cases, as it will not follow each bending of the target-vessel. However, the form of the guidewire will be changed. As it is generally considered bad practice to modify live results, we would like to make sure that the original form of the guidewire – i.e. the guidewire as segmented from the interventional images – remains the same. For this purpose, either the internal energies are defined accordingly, or the guidewire – as originally segmented – is fitted among the final snakels positions according to the minimum mean square error criteria (see Figure 3.30 (e)). The latter solution has the advantage that it practically guarantees that the guidewire-pattern remains the same.

As already pointed out, we propose to use markers to geometrically correct ECG and respiration-matched roadmaps before overlay and registration of the guidewire tip. As we had no such sequences available, the proposed algorithm was tested on sequences recorded from patients which have underwent an open-chest intervention, in which case the sewing wires are observable and can be used to achieve the geometric correction. Some results are shown in Figure 3.30.

The segmentation of balloon markers and sewing wires.

The balloon markers are highly absorbing, thin but circular in shape. They appear in the initial segmentation because in such cases both eigenvalues of the structure tensor exhibit large values. The can be separated from the guidewires mainly by their circular form. The sewing wires are also highly absorbing thin oriented structures. They can be separated from guidewires by their form and by the fact that they are static in the analyzed sequences.
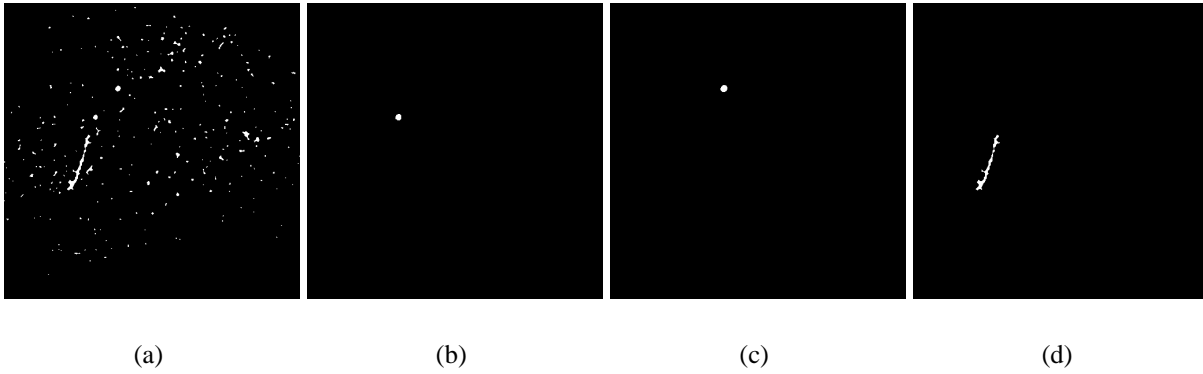
(a)   (b)   (c)   (d)

Figure 3.31: Intermediary segmentation result (a), maximum-selection result at first iteration C=0.79 (b), at second iteration C=0.74 (c) and at the third and final iteration C=0.10 (d).

Segmentation of balloon markers. The balloon markers are segmented together with the guidewire tip by the 99'th percentile of the histogram of the guidewire map. As they absorb X-rays slightly better than the guidewire tip, often the maximum value in the guidewire map lies on such markers. To detect if a marker and not the guidewire tip was segmented, we compute the compactness of the segmentation result. For an object of area A and perimeter P, the compactness is defined as:

$$C = \frac{4\pi A}{P^2} \tag{3.15}$$

The compactness of a circle is one and the lengthier the analyzed structure, the smaller the compactness. Empirically, we consider that if the compactness is above 0.5, the result shows a marker (see Figure 3.31). Then, the corresponding structure is eliminated from the initial guidewire-segmentation result and the procedure starts again. This repeats until the value of the compactness falls bellow 0.5. As sometimes, e.g., for sequences with a particularly bad SNR, despite the closing the guidewire tip is still fragmented, it may happen that we reject the two balloon markers and then the guidewire chunks and end up segmenting background structures. Such an error is clearly more critical than segmenting only a part of the guidewire, thus if after four iterations the compactness does not fall bellow 0.5, the current result is accepted as guidewire. The structures which have been eliminated from the guidewire segmentation due to their high compactness are considered the balloon markers.

Segmentation of sewing wires. To segment the sewing wires we record 24 frames – which correspond approximately to two heart beats – and each time compute the guidewire map and segment it using the 90'th percentile of its histogram as threshold, such as to make sure that all sewing-wire pixels are selected. We than add all these binary results – zero for background and 1 for guidewire/sewing wires – and compare their sum with a threshold. This threshold has the value 23 such that only the pixels corresponding to the static sewing wires will have values larger than the threshold. A processing result obtained after eliminating shutter-artefacts is shown in Figure 3.32. We wait for two heart beats, because we want to make sure that only the static sewing wires will be segmented and there are no artefacts from guidewire and/or balloon-marker pixels. 24 frames are equivalent to two seconds of recording. During this period, no overlap of surgical tools on the roadmap is possible.
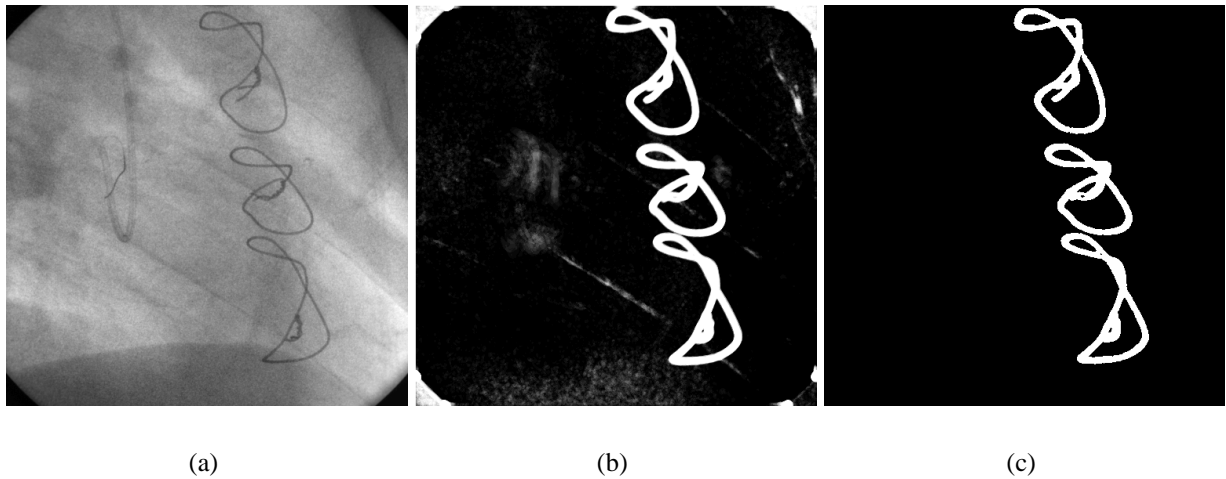
Figure 3.32: Interventional image showing wires (a), sum of 24 segmentation results (b) and segmentation result after eliminating the shutter artefacts (c).

3.3 Quantitative analysis of the myocardial-blush

The MBG is an important measure for cardiac revascularization and it is used to make long-time predictions on the health of patients which have suffered a heart attack and have been treated by PTCA. Presently, the MBG is assessed "semi-quantitatively" (grades 0-3) by the angiographer and thus the procedure is afflicted by inter- and intra- observer variability [190]. We describe methods which permit a "quantitative" assessment of the MBG mainly by tracking a predefined region of interest (ROI) of the moving heart over a sequence of X-ray projections (Section 3.3.1). Tracking is needed to allow the observation of the same region during the entire analyzed sequence, despite the heart motion. The MBG is then obtained by analyzing the gray-level variations within this ROI during the investigated sequence (Section 3.3.2).

The MBG. Myocardial infarction (MI) is the most common cause of morbidity and mortality in the industrialized world. In most cases, an abrupt occlusion of a coronary artery leads to MI. Diagnosis and treatment usually takes place under X-ray supervision, the treatment being performed by PTCA. The PTCA is instrumental in reopening the clotted coronary artery, which does not necessarily imply integrity of myocardial micro-circulation. The presence of blood with contrast agent in heart tissue can be observed as a darkening of the target region – i.e. myocardial blush (MB). For precise diagnostics, to assess the success of such intervention, as well as for making long-time predictions on the health of the patient, it has been proposed to investigate the MB by means of a dedicated measure: the MBG [98].

A heart-motion-compensated region of interest. The MB is an artery-related phenomenon. The position of the region where it can be observed – i.e. the ROI – depends on the position of artery which is investigated. Therefore, the position of the MB-feature changes during a heart-beat due to the change in position of the corresponding artery. To measure the MBG we need thus to track the MB-region. This region is initially defined by hand by the physician in the first image of the analyzed sequence.

Robust estimation of the MBG. The MB is a time-dependent phenomenon. It becomes observable when blood with contrast agent reaches the target region. It then becomes more salient the more blood is pumped-in and it finally vanishes, when the blood with contrast agent leaves the target region. After finding the MB-region in the entire analyzed sequence, the MBG is estimated by a robust analysis of the gray level variations within the tracked-region.

3.3.1   A heart-motion compensated ROI

Typically, the ROI has no dominant cues to support the tracking. As the vessels exhibit bendings and bifurcations – i.e. *junctions* – while other structures do not, and the vessel motion is directly linked to the *heart motion*, we estimate the latter by *tracking junctions* between consecutive images.

To project then the motion of the vessels on the ROI, its *border* is modeled as a *closed elastic string*, which is deformed under the action of certain external forces, defined in relation to the tracked junctions by attaching *virtual springs* between each junction and each border point. The behavior of the springs is regularized by internal forces [119].

Clearly, the ROI can only be tracked in images acquired during the complete state (Figure 3.33 (a)). The MB appears when the blood reaches the myocardium. This moment varies from sequence to sequence, depending on the way the contrast agent is injected. In some sequences it may appear as early as in the middle of the complete state, and continue until close to the end of the washout. For images acquired during the washout, the vessels are poorly or not at all visible, thus the ROI cannot be tracked there. To analyze the MB during its entire duration, the images showing insufficient contrast are matched, with respect to the heart and respiration phase, to the contrasted images, which are acquired during the complete state. The ROI is then projected from the contrasted to the non-contrasted images.

Detection and tracking of junctions

By junctions we mean the bendings and bifurcations which – in this context – can be observed only on vessels. The junctions are detected by a tensor approach (see Appendix A.5).

We see two possible approaches to junction tracking:

- **The rigid approach.** In this case the same set of junctions are tracked over the entire sequence. Clearly the appearance of junctions should not change during tracking.

- **The adaptive approach.** In this case a set of junctions are tracked only between consecutive images. Therefore in an entire sequence the set of junctions can change. This type of tracking adapts to potential changes in appearance of junctions.

The vessels and together with them the junctions change their appearance when the heart moves. Therefore, we choose the adaptive approach for tracking as it is practically impossible to track the same junctions over an entire heart beat. Also, as long as junctions can be found only on vessels, the motion of vessels – and thus also the motion of the heart – can be obtained irrespective of what junctions are tracked.

Junction detection.   Junctions are points in whose vicinity at least two prominent local orientations can be found, where by local orientation we understand the direction along which the

gray level profile shows the least variations in a certain neighborhood. Junctions are characterized by relatively large values of the smallest eigenvalue (λ_2) of a 2×2 tensor used to fully describe a single orientation [12].

To detect junctions, we set a threshold over the λ_2-image (see Appendix A.6.2). Thresholding returns junction surfaces and background. The junction surfaces consist of several points likely to be junctions. We consider as true junctions that points exhibiting the largest λ_2 value in a neighborhood (Figure 3.33 (c)).

(a)   (b)

(c)   (d)
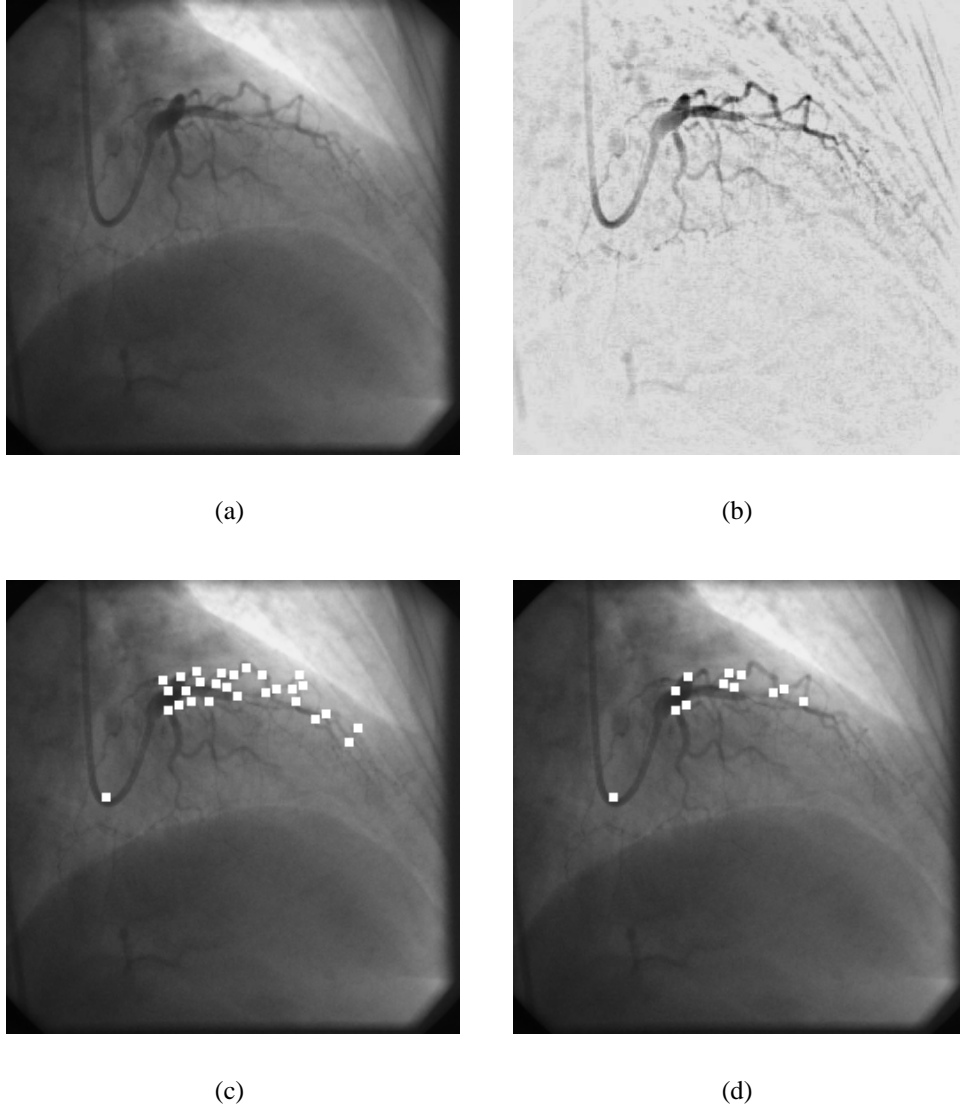
Figure 3.33: Detection and tracking of junctions: original image (a), Bothat result (b), set of detected junctions (c) and tracked junctions (d).

The assumption on which the ROI tracking is based is that in a coronary angiogram vessels and only vessels do show junctions. With the purpose of eliminating potentially spurious junctions, the background of the coronary angiogram is equalized – while preserving the gray-level difference between background and vessels – by applying a Bothat-operator [50] (Figure 3.33 (b)). On the result obtained after applying the Bothat-operator we compute then the second

eigenvalue of the structure tensor.

Junction tracking.   Once the junctions have been detected in two consecutive angiograms, they are tracked with the help of a similarity function consisting of three terms: the first term is based on the Euclidean distance between junctions, the second term uses the correlation coefficient, and the third term is based on a mixed orientation vector [12].

A junction travels from a past frame to a current frame over a certain distance which – considering the speed at which a heart moves during a beat, together with the frame rate at which the images are acquired ($12\ fps$) – is usually far smaller than the distance between two neighboring junctions in the angiogram. Thus, to find the same junction again in the current frame, one should take the closest one. The similarity between two junctions can be expressed using the Euclidean distance between their position vectors as:

$$s_p = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{d^2}{2\sigma^2}}$$

(3.16)

where σ weights the relevance of the distance d. In our experiments $\sigma = 1$. Clearly, the smaller the distance, the larger the position-similarity term and $s_p \in \left(0, \frac{1}{\sqrt{2\pi}}\right]$

However, there are cases when distance alone is not enough, e.g., for images acquired during a ventricular systole, when the distance the vessels travel between consecutive frames is large. For a successful tracking we need also junction-specific information which we extract in two ways: by investigating the correlation between junction neighborhoods and by the mixed orientation vector. The correlation between junction neighborhoods is given by the correlation coefficient:

$$r = \frac{K(\overline{X_p}, \overline{X_c})}{\sigma(\overline{X_p})\sigma(\overline{X_c})}$$

(3.17)

with K the covariance and σ the standard deviation for the neighborhoods $\overline{X_p}$ and $\overline{X_c}$ for the past and current frames respectively. Theoretically $r \in [-1, 1]$, but in our case $r \in [0, 1]$. To fully describe two orientations, an extended 3×3 tensor is needed which involves second order derivatives. This tensor's eigenvector corresponding to the smallest eigenvalue is called *the mixed orientation vector* and can be used as a juciton descriptor, as it contains a full but implicit description of the sought orientations [12]. We propose to use the Euclidean distance between the corresponding mixed orientation vectors d_o to build the orientation-similarity term:

$$s_o = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{d_o^2}{2\sigma^2}}$$

(3.18)

again with $\sigma = 1$.

The similarity function used to track junctions between consecutive frames will then be:

$$s = w_1 s_p + w_2 r + w_3 s_o$$

(3.19)

The weights $w_i, i = 1, 2, 3$ can be used to force the tracking to rely more heavily on a certain term. In our experiments however, they were all set to one.

Outlier detection.   The tracking may fail, e.g., when false junction correspondences are used. Such outliers are detected end eliminated by comparing their similarity against a certain threshold. This threshold is set using a significance test. For the first two frames at the beginning of the analyzed sequence, the similarities of all possible junction pairs are computed. For each past junction, the pair made with a current junction for which the similarity is maximal, is considered as a true pair only if it is also the pair with the largest similarity from among all pairs formed between the corresponding current junction and past junctions. The similarities of such pairs are eliminated from the set. What remains are only similarities between outliers which are used to estimate the junction pairs' similarity distribution under the outlier assumption H_0. The outlier-detection threshold is then computed such that it corresponds to a significance level $\alpha = 10^{-4}$. Junction pairs with a similarity below the threshold are eliminated (see Figure 3.33 (d)).

Modeling the ROI

The tracked junctions can be used as landmarks to compute the parameters of an elastic transform [24] which models the heart motion. The ROI image – as observed in the past frame – should then be modified according to this model and applied to the current frame to achieve the heart-motion-compensated ROI-tracking. Such a strategy is highly sensitive to the number of landmarks and the presence of outliers and thus an outlier robust implementation of the elastic transform is needed [165]. The results obtained with an elastic transform are usually negatively influenced by the fact that the landmarks are not dense around the ROI, due to its typical positioning (Figure 3.35). Such difficulties can be partially solved if the solution is further constrained, e.g., by introducing artificial landmarks.

In a more robust approach to tracking, one could use a physically motivated ROI model which is deformed under the action of forces generated by the junctions' displacements between consecutive frames. Such a model can be constructed by sampling the ROI's surface – or only its border – by control points whose displacements are internally constrained by mass, damping, elastic and stiffness terms defined in agreement with the ROI's physical properties, i.e., the corresponding properties of the heart-tissue. Such models can be generated, e.g., based on the finite element method (FEM) [158], or on energy minimization methods [119].

We propose here a physically-based ROI model whose border is properly sampled by control points. The displacement of these control points is constrained by imposing an energy minimization condition. As there are typically less tracked junctions than control points, the junctions are used to build an external attraction energy field where the model evolves to a minimum energy position, under the additional influence of internal energy terms, much in the way active contours are defined. The total energy to be minimized for a model m is:

$$E(m) = E_{ext} + E_{int} = E_{ext} + E_{el} + E_s \qquad (3.20)$$

where E_{ext} is the external attraction energy, and E_{int} is the internal energy consisting of two terms which control the elasticity E_{el} and the stiffness E_s of the model.

External attraction energy.   Provided the number of tracked junctions equals the number of control points and the same set of junctions is tracked over the entire analyzed sequence, to find the position of the ROI in a current frame one could establish junction-control-point pairs and link them by a virtual string so that the motion of the junctions is transmitted to the control
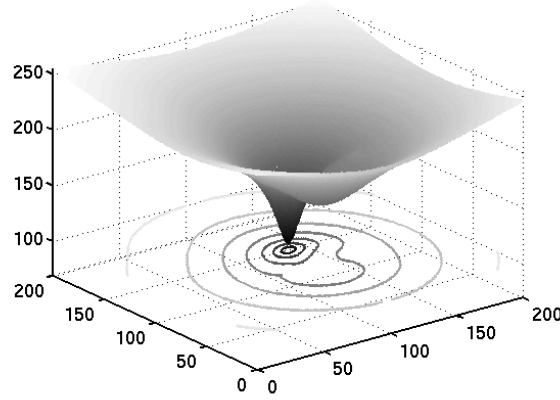
Figure 3.34: External energy field.

points. The potential elastic energy which appears when a string anchored at a certain point is stretched over a distance x is:

$$E_p = \frac{1}{2}kx^2 \tag{3.21}$$

where k is a constant describing the elastic properties of the string. The internal energy defined between control points would then condition the behavior of the ROI as it travels to a new position seeking to reach a minimum energy state.

However, the number of junctions is typically less than the number of control points and the set of tracked junctions can vary from frame to frame. Thus, to track the ROI between two consecutive frames, we attach a string between each control point and each tracked junction. The energy of this entire system – which is zero in the past frame – increases in the current frame due to the tensing of the springs as one of their anchor points (i.e. the junctions) move over a certain distance. For all possible positions in a certain vicinity of a control point – as it is positioned in the past frame – the elastic energy obtained by spanning the springs between each junction in the current frame and each position are computed. The external energy which would influence that particular control point is then determined by considering at each position the mean contribution over all junctions. The external attraction energy will then be:

$$E_{ext}(m, n) = \gamma \frac{1}{N} \sum_{i=1}^{N} E_p^i(m, n) \tag{3.22}$$

where γ is a weight factor, N is the total number of tracked junctions and (m, n) are the Cartesian coordinates of a point in the vicinity. An example is shown in Figure 3.34. Clearly, the vicinity should be large enough to include all possible end-positions of the control point.

Internal energy.   To define the internal energy we place rods between each pair of control points. The elasticity and stiffness of these rods are controlled by special energy terms. The elasticity is modeled by:

$$E_{el} = \alpha \left[\sum_{i=1}^{M} \|\vec{p}_i - \vec{p}_{i-1}\| - L_0 \right]^2 \tag{3.23}$$

with α a weight factor, $\vec{p}_i = [x_i, y_i]$ the vector containing the Cartesian coordinates for each of the $M + 1$ control points and L_0 the average distance between them. This energy term increases

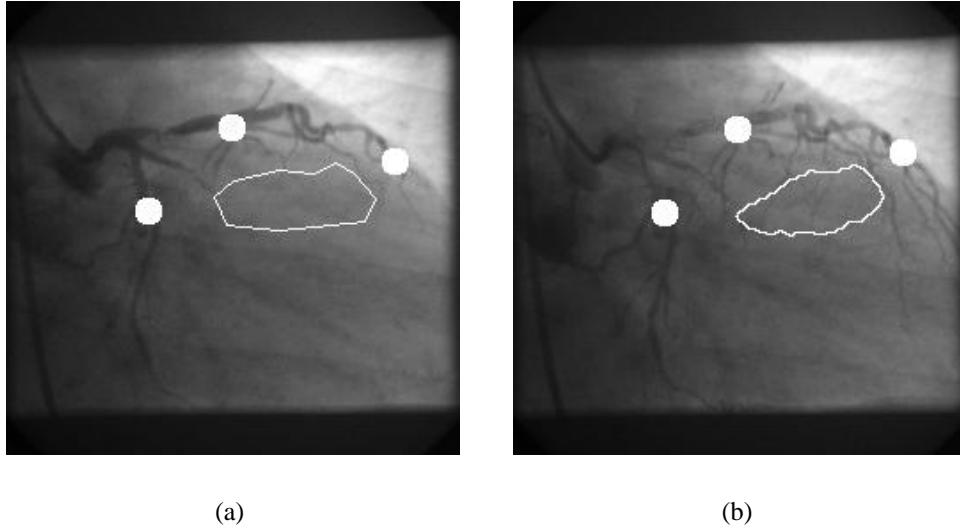(a)   (b)

Figure 3.35: First (a) and last (b) image of an analyzed sequence with markers and ROI.

if the model is stretched or compressed. The stiffness is modeled as:

$$E_s = \beta \sum_{i=1}^{M-1} \|\vec{p}_{i-1} - 2\vec{p}_i + \vec{p}_{i+1}\|^2 \qquad (3.24)$$

with β a weight factor. This energy term increases if the model is bended. As we model a closed curve, $\vec{p}_{-1} = \vec{p}_M$ and $\vec{p}_{M+1} = \vec{p}_0$, thus the sum in equations (3.23) and (3.24) runs from 0 to M. Adjusting the parameters α, β and γ controls the relative importance of the respective energy terms inducing thus a specific model behavior.

A heart-motion compensated ROI. Results

We have tested our algorithms on a number of seven sequences of 15 to 28 images each, showing only the complete state. The angiograms had a resolution of 512×512 pixels and were acquired at a frame rate of 15 fps. To increase the processing speed during experiments, the resolution was reduced to 256×256. The size of the Bothat-window was chosen to be nine pixels. The size of the derivative kernels used to compute the orientation tensor was chosen to be seven pixels and that of the pixel-neighborhood where the orientation tensor is computed was again nine pixels. The energy weights were chosen: $\alpha = 15$, $\beta = 3$, $\gamma = 5$. The sequences had neither table nor patient movement and were acquired in clinical routine with different projection angles. We have computed several measures to show that the result is conform with an expert opinion, and that it follows the variation of the heart surface. As reference we have artificially built a failed tracking result for each sequence by defining a static rectangular ROI in the lower left corner of each image. We have then computed all measures again for this tracking result. The results are given in Table 3.6 and an example is shown in Figure 3.35.

To verify if a tracking result is in agreement with an expert opinion, we have defined by hand the ROI in each frame of the analyzed sequences and defined correct and incorrect tracking results in relation to this ground truth. A result was considered correct if the mutual overlap

	cc	std	Av	Pv
Results	94.19	3.02	7.87	4.06
Reference	0	5.56	7.23	3.28

Table 3.6: The results obtained on the test data.

between the tracked ROI and the manual ground-truth was larger than 50%. We have computed the mean percentage of correct classifications over all sequences (cc).

To verify whether the tracked ROI keeps its position relative to the projected heart surface, we have marked three vessel-points distributed over the entire vessel tree such that they are situated around the tracked ROI. We have typically chosen bifurcations and bendings of the main vessels (see Figure 3.35). Then, we have determined the center of mass of these three points and the center of mass of the ROI and computed the Euclidean distance between them in each frame of the sequence. We have then computed the standard deviation in each sequence and the mean of standard deviations over all sequences (std).

To verify whether the tracked ROI changes its surface in agreement with the changes of surface of the heart, we have computed the area and perimeter of the triangle defined by the three marked points and then we have computed their variation in percents between consecutive frames. We have compared this with the same variation computed for the ROI's area and perimeter by taking the difference between the ROI and the triangle variation and then computing the mean of absolute differences over the entire sequence. We have then taken the mean over all sequences: Av and Pv for area and perimeter respectively.

A heart-motion compensated ROI. Discussion

The results show that the ROI tracking method is able to follow the movement of the heart and that the tracking results are in agreement with human intuition. However, although the ROI as a whole follows the heart, it seems that its area and perimeter do not follow the local variations of the heart tissue but they remain approximatively constant. One possible reason is that a majority of the tracked junctions are typically situated on the same vessel or on vessels that were very close and thus they are representative for the motion of the heart, but they are not representative for the local variations of the heart tissue in and around the ROI. We believe that a 50% overlap between the tracked ROI and the real MB region suffices for an automatic estimation of the MBG by a robust feature. Experience shows that it is practically impossible to define the ROI to include all and only the blush region. Thus, the ROI will normally include also regions where no blush has been observed.

The surface of the heart which is observable in an angiogram varies as a consequence of the heart changing its volume. The coronary vessels follow this variation by bending and stretching, thus many of the junctions change their appearance sometimes drastically over a heart beat. Instead of tracking a set of particular junctions over an entire sequence, we track different sets of junctions only between consecutive images, as only in such a case the change in appearance of most junctions is small enough – considering the speed of the heart beat in relation with the frame rate – to permit tracking. It is very important for the tracking that the junctions do not change their appearance drastically between consecutive frames. This may happen, e.g., during the ventricular systole when the speed at which the heart moves reaches a maximum. Thus it is

expected that the tracking improves with the frame rate.

By its energetic formulation and by its adaptive approach to junction tracking, the ROI-tracking is also marginally robust against slow, small-amplitude table and patient movement. However, to analyze the MB in those cases when it extents over the washout, and generally for a precise tracking, motion-compensation clearly is needed. Ideally the table should not be moved during image acquisition and the patient should remain still. Depending on the angle under which the the X-ray imaging system looks at the patient, the coronary vessels can be imaged under different projections. Only those projections where the vessels show a minimal overlap – ideally none whatsoever – are suited for processing. Otherwise, e.g., as the heart has also a rotation motion this may result in some vessels traveling upwards and some downwards in the projection images, then in a worst case scenario, the algorithm will track alternatively junctions on vessels traveling in opposite directions, resulting in a stationary ROI.

3.3.2   Robust analysis of the myocardial blush

After the physician has defined the ROI by hand in a first image, this is tracked throughout the sequence. The MB appears as a persistent dark staining of the target region. Usually, the target region and the ROI do not correspond 100% – because practically, the ROI cannot be defined to include all and only the MB region – but they do overlap strongly. Thus, there will be also other items present in the ROI besides the blush and to measure it, we need a robust *blush feature* which responds to blush only. From the *blush curve*, computed using the blush feature, the *MBG is estimated*.

The blush feature, the blush feature-curve and the blush curve

The simplest way to obtain information related to the blush is to compute the mean of the gray levels from the ROI. However, the mutual overlap between the tracked and the target region – although is usually above 50% from the ROI – varies from frame to frame. Consequently, the mean is a rather unreliable blush feature. A feature is needed which is insensible to the variations of the overlap. Such a feature must be computed only from gray levels which correspond always to the blush region, irrespective of the variations of the overlap. We propose to use as *blush-feature* a *low percentile* of the histogram of the gray levels found in the ROI.

The blush feature measures blush in each frame of the analyzed sequence. Similar to the analysis of contrasted images (Section 3.2.1), we call the curve of MB feature over frame index the *blush feature-curve*. From the blush feature curve, the *blush curve* is computed, which is then used to estimate the MBG.

A percentile-based MB-feature.   The percentile selects a certain gray level and at the same time defines a certain area within the ROI which we call percentile-area. The percentile-area is represented by the pixels with gray levels below that corresponding to the percentile and is not necessarily connected. The value of the percentile defines also how much of the area of the ROI – expressed in percentages – is the percentile-area. Then, e.g., if the dark target region and the ROI overlap such that the former occupies some 50% of the latter, a percentile lower than 50 will select gray levels corresponding in a large majority to pixels over which the blush appears. Clearly this observation holds only as long as the overlap between the target and the tracked regions is larger than the value of the percentile.

A good blush-feature should respond to blush alone. Thus, the choosing of the specific percentile of the ROI histogram to be used as blush-feature involves several aspects:

- **The first** is that of the minimal mutual overlap between the target region and the ROI, achieved during tracking. The value of the percentile should be smaller than the percentage of the ROI-area corresponding to this minimum overlap.

- **The second** is related to the type of structures – other than the blush region – which can be found in the ROI. Assuming that such structures are likely to be comparable or lighter than the blush, the value of the percentile should be chosen such that its corresponding gray level is smaller than the gray levels of these structures, i.e., they are excluded from the percentile area. Assuming that the structures are darker than the blush, the value of the percentile should be chosen such that its corresponding gray level is larger than the gray levels of these structures, i.e., they are included in the percentile area. These conditions form a balance which should be kept even, when choosing the percentile.

Thus, assuming that the mutual overlap between the target and the tracked regions is around 50%, we choose the 25'th percentile as blush feature.

Particularities of the blush feature-curve. The values of gray levels corresponding to the target region will vary with the heart beat by a mechanism similar to that described at the segmentation of the complete state (see "Particularities of the 95-percentile feature curve" at page 65). It will reach a minimum during the ventricular systole, i.e., when blood with contrast agent flows at relatively high speeds and thus the absorption and the contrast of tissue containing blood mixed with contrast agent is maximum. Because a darkening of the tissue means increased contrast, minima of the blush-feature observed during the complete-sate can be correlated with maxima of the complete-state feature. Such a behavior is shown in Figure 3.36.

The blush feature-curve will also exhibit some variance due to the AGC, or due to other dark structures which may appear in the ROI at a certain moment and whose gray levels are bellow the percentile gray level. It is important to observe that such dark structures will influence the MB feature only if it appears in regions which previously had gray levels above the percentile gray level. Examples of dark structures afflicting the curve are the patch-like noise (see "Computation of the vessel map" at page 57, in Section 3.2.1) and "passenger" vessel chunks. Passenger vessel chunks appear shortly in the ROI and then disappear again. Such a situation occurs, e.g., if the ROI-tracking fails.

Consequently, the blush-information exists already in the blush feature-curve but it is affected by noise.

The blush curve. In contrast to a blush feature-curve, a blush curve should reflect only the influences of the MB on the ROI. We describe two methods to compute a blush-curve: a "direct" and a "relative" method. The MBG can then be estimated from blush curves directly.

The "direct" blush curve. The variations of the blush-feature induced by the heart are short in comparison with the time-length of the blush, thus a blush curve can be interpolated from the values of the blush feature. This is equivalent to a low-pass filtration of the blush feature-curve to eliminate the noise. An example is shown in Figure 3.37.
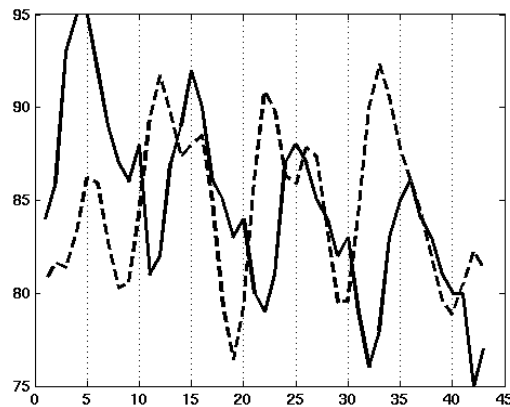
Figure 3.36: The curves of blush feature (continuous-line) and complete state feature (dashed-line) over frame index for the complete state of a sequence showing a MBG=3.

The "relative" blush curve. One can imagine an alternative "relative" blush curve. In this case in each image of the sequence where the MBG should be estimated, a reference gray level is chosen such that this reference reflects all angiogram influences other than the MB. Such a reference can be, e.g., a percentile from the gray levels of the vessels, assuming that the vessels have been segmented (see Chapter 4) and are visible throughout the entire duration of the blush. Then the relative variation of the blush feature between consecutive images is compared to the relative variation of the reference gray level by subtracting them.

If the reference gray level increases by, e.g., seven percents from one image to the other and the blush-feature increases only by three it is clear that the blush region becomes darker, i.e., blush appears. By plotting such differences against the frame index another type of blush-curve is obtained – eventually after interpolation/filtration. This is somewhat equivalent to observing the first derivative of the blush curve from the "direct" method.

Influences of respiration. Due to the influence of the AGC, the blush feature will also vary with the observable surface of the diaphragm[8] and therefore with the respiration cycle. Depending on how much of the diaphragm is visible, this variation can be more or less important. As the period of respiration is comparable with the time-length of the blush, the estimation of the MBG will be disturbed by the moving diaphragm, thus we recommend that the patient halts his breath during the time the sequence of angiograms containing the blush is recorded.

Estimating the MBG

To estimate the MBG, two blush curves are needed from the same patient: one from a healthy coronary artery and one from an ill coronary artery that has just been treated. The blush curves need to be then synchronized (see Section 3.2.2) and compared. We propose to compare them by evaluating the difference area between the two blush curves (see Figure 3.38).

---

[8]For those projections where the diaphragm is observable.
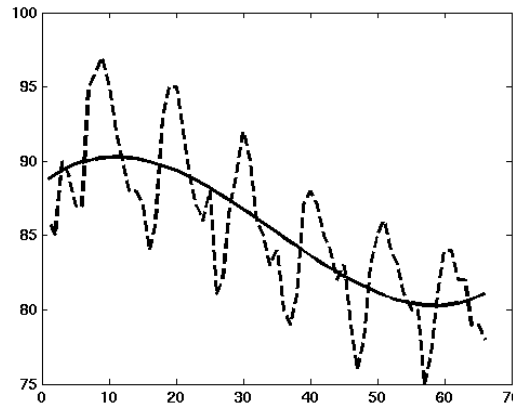
Figure 3.37: Blush feature-curve (dashed-line) and "direct" blush curve (continuous-line).

Particularities of the MBG estimation.   First one has to eliminate from each curve the influences foreign from the myocardial blush. Assuming that such influences can be modeled by a shift of the blush curve, we propose to subtract for each of the blush curves the the initial values so that both start with the value zero and then to compute the area between the two of them.

In the example shown in Figure 3.38 (a), the two curves were obtained for the same sequence of coronary angiograms by linear interpolation from the values of the blush-feature and are therefore "direct" blush curves. The sequence shows only the beginning of the blush, i.e., it contains only the complete state of a larger sequence. The sequence exhibits a strong blush, which receive a MBG of three in an expert evaluation. One curve was obtained by choosing the ROI over the target blush region and the other one was obtained by choosing the ROI in the vicinity of a vessel where no blush could be observed, i.e., MBG = 0.

To validate this results, we have computed a third blush curve, also from the same sequence and by choosing the ROI in another region with no blush. The result is shown in Figure 3.38 (b). The area between the two curves is here approximately four times less than the area computed in the first case.

3.4   Conclusions and discussion

PTCA is an intervention commonly used in the treatment of the CAD and of MI. We have described machine vision-based methods to support the physician during this important intervention.

We show how to construct a coronary angiograms-based moving – and therefore dynamic – roadmap with superimposed surgical tools for improved navigation during the intervention. Currently, navigation is done with the help of one, inherently static, complete coronary angiogram, the physician needing to mentally register the position of his surgical tools into the vessels.

We also show how to measure the MBG in an automatic manner. Currently, the MBG is assessed visually by the physician. An automatic measurement offers the basis for an improved analysis as it is observer-independent.

(a)   (b)
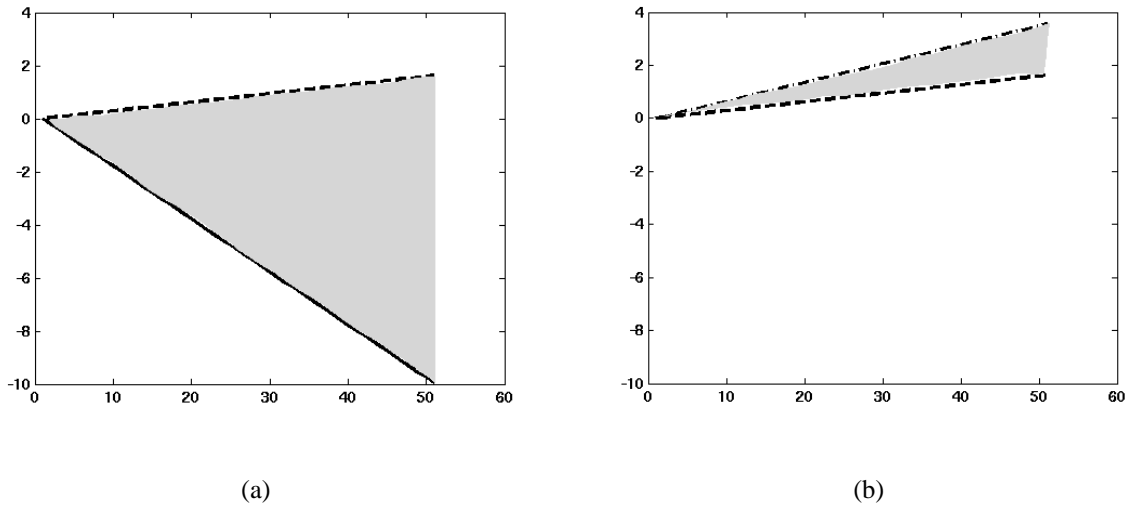
Figure 3.38: Blush curves for a region where the MB can be observed (continuous line) and for a region where no blush is present (dashed-line) (a) and blush curves for two regions where no blush is present (dashed and dash-dotted line respectively) (b). The area between the curves is shown in light gray.

3.4.1   The dynamic roadmap

The computation of the dynamic roadmap has to be fully automatic, so that is does not need additional training or work from the part of the medical staff. The computation of the dynamic roadmap should only-marginally/if-at-all modify the interventional routine. No changes in the interventional hardware are permitted.

In praxis the computation of the dynamic roadmap would proceed as follows: assuming that no roadmaps are available for the current projection, the physician would inject a standard contrast bolus. The moment when the contrast bolus appears in the interventional images is detected. Starting then, images are recorded for a few heart beats. These images are analyzed to find complete-state frames. If roadmaps are available, the complete-state frames are extracted from a data base. Then, they are matched to the live images using the heart and respiration phases. Potential sewing wires are first detected and eliminated, then for each image the balloon markers and the guidewire tip are segmented in the live images. The guidewire-tip is then registered with the closest vessel from the corresponding roadmap and is displayed in a contrasted manner together with the balloon markers.

Due to patient variability, the presented methods are unsupervised. To segment the complete-state and detect the contrast-burst we use a feature related to the vessel-area shown in the analyzed images. This feature is computed as a high-percentile of the histogram of a vessel map obtained after enhancing vessel-like structures in the images. As the detection of the contrast burst has to be done during the intervention, the proposed processing is subject to strict causality.

Vessel enhancement and feature extraction

To directly estimate the vessel area in angiograms we need vessel segmentation, which is difficult especially on images with a small SNR – as it typically happens in the case of angiograms and particularly in the case of fluoroscopic images. Thus, instead of computing the vessel area, we use a *vessel-area-related feature*, i.e., a percentile of the histogram of a *vessel map*. A vessel map is obtained after *enhancing the contrast of vessels to the background*. As in the vessel map the vessels appear with high intensities, higher percentiles are used. Depending on the application, different percentiles are chosen in relation to the total possible vessel-covered area in the images – for the segmentation of the complete state – or in relation to the are covered by vessel-like structures (i.e. noise and catheter) – for the signaling of arbitrary contrast agent injections.

Signaling of arbitrary contrast injections

For signaling of arbitrary contrast-agent injections a threshold is set on the vessel area-related feature by means of a *significance test* whose null hypothesis is that the frame under consideration shows no vessels. To estimate the null hypothesis pdf, the frames recorded in the first few seconds of the intervention are used, under the assumption that they show no vessels. The results show that the method is successful in detecting the first frame of a burst of contrast agent.

Segmentation of the complete-state

A successful segmentation of the complete state has to return sufficient images to cover a heart beat and very few false segmentations. To evaluate our algorithms we have accordingly imposed bounds on the rates of correct classifications and false positives. Only the *ECG-enhanced coherent MAP-based segmentation* method fulfills these constraints. The other methods fail to achieve the desired false positives rate.

The direct way to segment the complete state on a feature curve is by a fixed threshold. However, such a threshold cannot compensate the variance of the feature curve – especially the heart beat-induced variance – and does not achieve satisfactory segmentation results having a too large false positives rate. The MAP-based adaptive threshold can lower the false positives rate, but not enough to verify the corresponding bound. By explicitly enforcing the coherency constraint, we can improve the correct classifications rate, while keeping approximately the same false positives rate and make sure that we select a succession of images and not only images showing the heart at the same position. To lower the false positives rate, we choose from among the already selected images which show the heart at different positions during one or several beats, only those acquired between ventricular systole. We are thus able to achieve a segmentation result verifying both the bound on the correct classification rate and that on the false positives rate. Our experiments show that the bound on the false positives rate is the most difficult to reach, and this is due to the large heart-beat-induced variance of the feature curve.

For a successful analysis we propose an *image acquisition protocol*. For filled-state segmentation, a sequence of angiograms should be recorded such that there are sufficient complete-state images, i.e., at least three hart beats, such that a 60% correct classifications rate returns enough true roadmaps. In the case of contrast burst detection, the first seven to ten seconds should be burst-free. In both cases it is desirable that the X-ray dosage remains constant and there are no table nor patient movements – eventually these should be compensated.

Sequence matching

Image-based methods have the disadvantage of taking a rather long time to compute even if special search rules are used, like e.g.: after a match is established, the roadmap for the next interventional image should be searched only among complete-state angiograms acquired after the one selected for the initial match. In the context of the dynamic roadmap, which should function in frame rate, faster methods are required. Speed can be increased at the cost of decreasing robustness by using downsampled images, or considering only portions of the images.

Vector-based methods are fast and robust, because they include also ECG-information and the respiration phase is usually computed by detection and tracking of an anatomical organ: the diaphragm. Image-based similarity measures are used within the vector-based method, to return respiration-related information, when the diaphragm is not observable.

Therefore, we propose to achieve the matching of contrasted and non-contrasted images by using a *similarity feature vector*, which contains information related to the heart and respiration phases and fully describes the position of the heart in the analyzed images.

Diaphragm-based respiration phase. We have described how to compute the respiration phase from images acquired under projections which permit the visualization of the diaphragm. For this purpose, we need to track the diaphragm in the analyzed sequences. In doing so we assume that:

1. The diaphragm is present in all analyzed images.

2. Both contrasted and non-contrasted images are acquired with the same projection geometry.

3. Potential table and patient movements are compensated beforehand.

4. The roadmaps sequence contains at least one respiration cycle.

The diaphragm appears in the observed images as a dark large circular pattern, whose position we describe by a two dimensional vector starting from its edges. To avoid misleading effects from vessel and shutter edges, we eliminate them by morphological filtering the image and minimum intensity projection in the edge map respectively. We then find the diaphragm by a Hough transform for circles. To allow on-line usage of this algorithm, we use prior knowledge, both anatomical and about the image acquisition system, to speed up the diaphragm tracking process. Knowing the typical respiration frequency and the frame rate at which images are acquired, we assume that the diaphragm cannot travel far between consecutive frames. Thus, we do not analyze the whole Hough accumulator to find the best fit but only a small region of interest (ROI) around the previous best fit. The whole procedure is initialized with a large ROI and continues then with a smaller one. In some cases the Hough based tracking may fail. There are mainly two reasons for this:

1. A sudden move of the diaphragm which causes an accumulator maximum which is out of the investigated small ROI.

2. The diaphragm edge has no longer a circular pattern.

To detect such cases we propose to measure the confidence in the Hough result by a dedicated measure. This measure is computed as the variation coefficient of the pixel gray-level values along the detected circular approximation of the diaphragm in the corresponding edge map. The confidence measure is compared to a threshold to decide on the quality of the investigated Hough-result. This threshold is established for each available sequence of roadmaps and used for all interventional images of similar projection. If the confidence measure shows a poor fit, this is refined using active contours.

We cannot make any prior assumptions on the success of the tracking and thus on the evolution confidence measure in a roadmaps sequence and we have no access to a labeled training set, therefore the threshold is established heuristically.

Due to the Hough-ROI-based diaphragm tracking, our algorithm is fast enough to allow on-line implementations and the confidence measure ensures the robustness needed in medical applications. We have tested our algorithm on several appropriate sequences from our data base and an initial visual inspection has shown good results.

Segmentation of surgical tools

To segment the guidewire tip, we proceed in two steps: enhancement followed by segmentation. The result of enhancement is a *guidewire map*. This is computed by the first eigenvalue of the structure tensor, using an appropriate – i.e. small – neighborhood. Segmentation is done by a *percentile threshold*. This threshold however returns not only the *guidewire tip* but also – if visible – the *balloon markers* and the *sewing wires*.

Selection of the sewing wires.   The sewing wires are selected and then eliminated from the segmentation of the guidewire tip by using their property of being static in the analyzed sequences. They can then be used as source of additional markers for the correction of table and patient motion.

Selection of the balloon markers.   The balloon markers are considered surgical tools. Due to their small size they will be practically always found in the corresponding vessels and thus they need no registration. They are segmented by the percentile threshold together with potential sewing wires and the guidewire tip. They are then selected using their high compactness.

3.4.2   Automatic estimation of the MBG

As in the case of the dynamic roadmap, the analysis of the MB should not modify the interventional routine and no changes in the interventional hardware are allowed. It requires only minimal interaction from the part of the physician that has to define the ROI only once for every analyzed image sequence. Initial reactions from the part of the medical staff indicate that this does not constitute an acceptance problem [54].

The quantitative analysis of the MB results in the awarding of the MBG. During this analysis, the blush region has to be identified in each image of the input sequence after being defined by hand by the physician in the first one. We showed how to accomplish this by tracking the initial ROI. The tracking is robust to the influences of the motion of the heart on the position of the MB region. We have described also methods to measure the MB in the tracked region by means of a blush-curve and use these results to compute the MBG.

A heart-motion compensated ROI

The behavior of the ROI model is controlled by the energy weights, which gives our approach adaptability and robustness. Heart-motion information is extracted from complete-state images by tracking the vessel junctions between consecutive images.

Clearly the ROI can be tracked only as long as the coronary vessels are visible. To measure the MB also during and after washout, the tracking results obtained for contrasted images are extended to non-contrasted images after matching the corresponding images (see Section 3.2.2).

Junction tracking. Junction tracking is done using a similarity function which takes into account: the distance between junctions, correlation between their neighborhoods and similarities in their orientations. Falsely tracked junctions are detected by comparing the corresponding values of the similarity function with a threshold and accepting only those above it.

ROI modeling. The tracked junctions are then used to fit a physically-motivated model of the ROI, which evolves under internal and external constrains – expressed as energy terms – to a position with minimal energy. Currently, we model only the border of the ROI. In a similar manner it is possible to model the whole ROI surface by sampling it with control points and linking the them by elastic rods. Also, an additional energy term can be introduced to model the mass of the ROI. A more adaptable model can be obtained like this.

Analysis of the MB

To measure the blush in each frame, we propose to use the 25'th percentile of the histogram of the gray-levels from within the ROI as blush feature. The MBG can then be estimated from blush curves, which are obtained from the blush-feature curves, i.e., the curves of blush feature over frame index.

The blush curve. A blush curve shows the blush for the investigated artery. We have described two types of blush curves, a "direct" and a "relative" one.

At a first glance, the "relative" blush curve is more robust as it is designed to be independent to external influences to the analyzed angiogram. Such influences would affect both the blush-feature and the reference and are eliminated by subtraction. Further experiments are needed to see which modality is better suited for the estimation of the MBG and to validate these initial results.

We believe that the "relative" curve shows a better potential.

The MBG. The MB is artery dependent. Therefore, to assess the MBG for a patient, one has to compare the blush measured for the investigated ill artery with the blush computed for a reference healthy artery of the same patient. The MBG can be estimated as the difference area between two blush curves, one for the ill and one for the healthy artery respectively.

Chapter 4

Vessel segmentation in 2D-projection images

Vessel segmentation is a key component of many modern, computer-based cardiovascular image analysis tools. Such tools are used in applications, like e.g., diagnosis and intervention planning for various vascular diseases [121], diagnosis of the diabetic retinopathy [100], [43] [44], measuring of angiogenesis [52], [51], [151], registration of images [121], [176], densitometric measurements of the blood-flow [168] or blood-flow analysis. Yet again, in many such tools quantification of the target vasculature plays a pivotal role.

We describe a framework for segmenting vessels in 2D-projection images (Section 4.1). This framework includes an enhancement step (Section 4.2), followed by classification (Section 4.3). We also discuss two applications to support our approach (Sections 4.4 and 4.5).

4.1   Introduction

Vessel images are acquired by a multitude of imaging techniques, we concentrate here mainly on imaging methods which yield 2D-projection images (Section 4.1.1) and approach the problem of segmentation from a pattern recognition perspective (Section 4.1.2) to obtain a vessel segmentation framework (Section 4.1.3).

4.1.1   Vascular imaging methods

Currently there are several imaging methods available for vascular investigation. 3D imaging techniques like computed tomography (CT) and magnetic resonance imaging (MRI) return high-quality data. However, they are rather costly with respect to both money and time and the necessary hardware is bulky. They are also unaccessible to some categories of patients, e.g., patients with implants, like mechanical artificial heart valves, can not go through an MRI exam. 2D-projection imaging techniques, like X-ray angiography and retinal photography, are in many applications the methods of choice as they return results of sufficient quality without the overhead required by 3D imaging [44], [82]. Currently, they also have a better performance with respect to the rate of data acquisition, although this may change in the near future [103]. In 2D-projection imaging, only one image showing the target vasculature is available for analysis. There are also 2D+t imaging techniques, where a set of consecutive 2D-projection images is

available. For example, several X-ray projection images are acquired to record a short injection of contrast agent. This can be then used, among other, for blood-flow measurements, which are needed for improved diagnosis of the vascular disease. Currently, some research is done to apply multispectral imaging methods for vascular observation [41]. Finally, there is also ultrasound imaging, where a probe is introduced through the vessels to the investigation site similar to the way a catheter is introduced in PTCA. The focus of this chapter is on vessel segmentation from 2D-projection and 2D+t imaging modalities.

4.1.2 Vessel segmentation as a pattern recognition problem

The main vessel characteristic, which is instrumental in separating the vessels from their surroundings, is the contrast. The methods discussed here use primarily the contrast of vessels to achieve a segmentation.

We differentiate between pattern recognition-based vessel segmentation methods and other methods. When applying a pattern recognition-based method, each pixel is represented by a point into a feature space of one or several dimensions. This feature space is divided by a decision surface into two regions, one corresponding to the vessels and one to the background. Based on the positioning of the pixel feature vector with respect to the decision surface, it is assigned to one of the two classes, i.e., a binary classification is performed. Sometimes, neighborhood-based information is also used, e.g., by including it in the pixel feature set. In the other methods, certain connections among vessel pixels are implicitly used to decide for one pixel, or for a collection of pixels if they belong to the vessel class. These include tracking methods, where vessel centerlines are first segmented, e.g., by following valley courses [77], front-propagation methods where, e.g., the vessel region is segmented from inside-out by evolving a delineating curve to the vessel-borders [136], [171], as well as methods where explicit vessel models are used [122], [119].

As opposed to pattern recognition-based methods, the other methods have usually the disadvantages of very long computation times and comparatively poor results. The tracking methods need elaborated initialization steps and they can not handle branchings properly. The front-propagation methods need good vessel borders to avoid leakage, a condition which is very hard to fulfill in particular for the small weakly contrasted vessels. Vessel-model methods are inherently limited to specific vasculature or to the main vessels. Consequently, we concentrate here on pattern recognition-based methods.

Pattern recognition-based methods can be *supervised* or *unsupervised*. Most unsupervised methods are data-clustering algorithms [112]. In this case the input consists of a single image where each pixel represents a data-point. The output is a partition of the image-pixels into two classes, i.e., vessels and background. Conversely, a supervised method needs a training set of vessel images where in each image both vessel and background pixels have been labeled previously.

Unsupervised segmentation

Here we differentiate between an *unsupervised a priori* classification method and an *unsupervised learning* classification method. In neither case is there any labeled set of examples available to determine the parameters of the respective method. In the former case, the parameters

are imposed a priori by the user and in the latter case, the parameters are learned from unlabeled data.

Usually, for vessel segmentation an unsupervised approach is used, as there is no proper set of labeled examples available. There are *several reasons* for this, the major one being that in most applications a *different training set* would be needed not only for each type of vasculature, but indeed for each patient.

Another reason is the typical *bad quality* of such a training set. Many errors within the training set are due to the tiredness of the human specialist labeling the data, which at the fourth angiogram, after at least two hours of work, does not maintain the same level of attention as at the beginning and omits some vessel pixels or includes in the segmentation some background pixels. The false omissions affect usually the weakly contrasted small vessels. The false inclusions happen usually at the borders of vessels. Errors can also appear as a consequence of the unfounded expectations of the human operator with respect to the position and form of the vessel tree.

Yet another reason is the *insufficient number of components* of the training set. Due to the complications related to the labeling process, one can realistically expect training sets of 20 to 40 labeled images for supervised algorithms. Although this seems like a lot of pixel-examples, – assuming a resolution of, e.g., 512×512 pixels for an analyzed image – considering alone the variation induced by different types of pathologies, it appears clear that 40 images are not enough to build a training set to cover the entire variability of the vessel data. Also, it is possible that the parameters of the image acquisition system change between two recording-sessions. This represents an additional source of variability which needs to be covered in the training set.

Supervised segmentation

There are some applications, like e.g., segmentation of retinal images, where – under consideration of the points already mentioned – supervised segmentation methods can be successfully used. Usually, supervised methods are preferred over unsupervised ones if: (i) the number of images which need to be segmented grows large, (ii) all of them show the vasculature of the same organ and (iii) they are acquired with a similar setup of the imaging system. Under such circumstances, we believe that it makes sense to build a training set.

Segmentation of the retinal vasculature in images of the ocular fundus – i.e. retina photographies – helps physicians in the field to infer on retinal, ophthalmic and even systemic diseases like diabetes, hypertension and atherosclerosis. A central feature in such diagnosis is angiogenesis, therefore one needs to measure the area of the retinal vasculature and compare it against measurements taken at a previous time. Assuming that the area of the vasculature is measured from a segmentation result, then a fast, accurate and automated vessel-segmentation algorithm is needed. Such an approach will allow then ophthalmologists to screen larger populations for abnormalities and achieve thus an improved health care [181], [101], [129].

Segmentation of the retinal vasculature can be also instrumental in biometric recognition [113]. According to reference [99], the pattern of the retinal vessel-tree is stable and unique in each individual. A vessel segmentation result can be used to add the vessel length and surface to the biometric feature-set but also to capture the constellation of branching points which can be used to compute additional biometric features. It remains to be investigated how does angiogenesis in the case of people suffering from diabetic retinopathy influence the permanence

characteristic[1] of such a biometric system.

4.1.3   A framework for vessel segmentation

To this date there is no universal solution to the problem of vessel segmentation in 2D-projection images – apart from manual segmentation which is a tedious, prone to errors endeavor and constitutes a motivation for the development of computer-based supportive methods[2]. Being convinced that an algorithmic universal solution for this problem does not exist, we would like to introduce a set of rules – i.e. a framework – to guide a practitioner while devising the best algorithm for a special vessel-segmentation problem. We describe then practical implementations of these rules. Our approach is supported by examples spanning different projection imaging techniques and types of vasculature.

We believe that vessel segmentation should proceed in two steps: first vessel enhancement and then segmentation. Enhancement should modify and improve the description of vessels such that they are afterwards easier to segment. The enhancement of vessel structures in 2D-projection images is discussed in Section 4.2. During segmentation, one should make use of all available information about vessels, to decide for each pixel of the analyzed image if it is a vessel pixel or not.

Segmentation methods which return a result without any user help are called automatic [53], [44], [33], [101], [39], [117]. Automatic methods are discussed in Section 4.3. For some applications, better results can be achieved by semi-automatic methods such that the user is allowed to interact with the algorithm [170], [52], [51]. Such an application is described in Section 4.4. Finally vessel segmentation for 2D+t imaging methods is discussed in Section 4.5.

During enhancement, we eliminate or reduce also the influences of the imaging method, which permits in turn to approach segmentation in a more general manner. We introduce hysteresis segmentation as such a general approach (Section 4.3.2).

The segmentation methods presented here have as ultimate goal the analysis of vasculature based on measurements of, e.g., vessel length and surface, or on the description of the vasculature including, e.g., the number of vessel branches and branching points.

Usually, gray-level images are sufficient to conduct such an analysis and additional color information is useless. Most of the vascular 2D-projection and 2D+t imaging techniques return gray-level images and the algorithms described here use only such images.

4.2   Vessel enhancement

Vessel enhancement should accomplish two tasks. First, it needs to act as a buffer between the raw image data and the segmentation algorithm, such that similar segmentation algorithms can be applied on different data. Second, it has to improve the separability [87] between vessel and background pixels by decreasing the intra-class variance – i.e. a more homogeneous gray-level representation for background and vessels – and increasing the inter-class variance – i.e. better contrast.

To achieve such purposes, we proceed in two steps: (i) first we make use of prior information about the imaging system and about vessels to attenuate the background and decouple the

---

[1]Permanence is related to the invariance of the biometric feature-set over time [113].

[2]Manual segmentation of an angiogram takes at least half an hour.

(a) (b) (c)
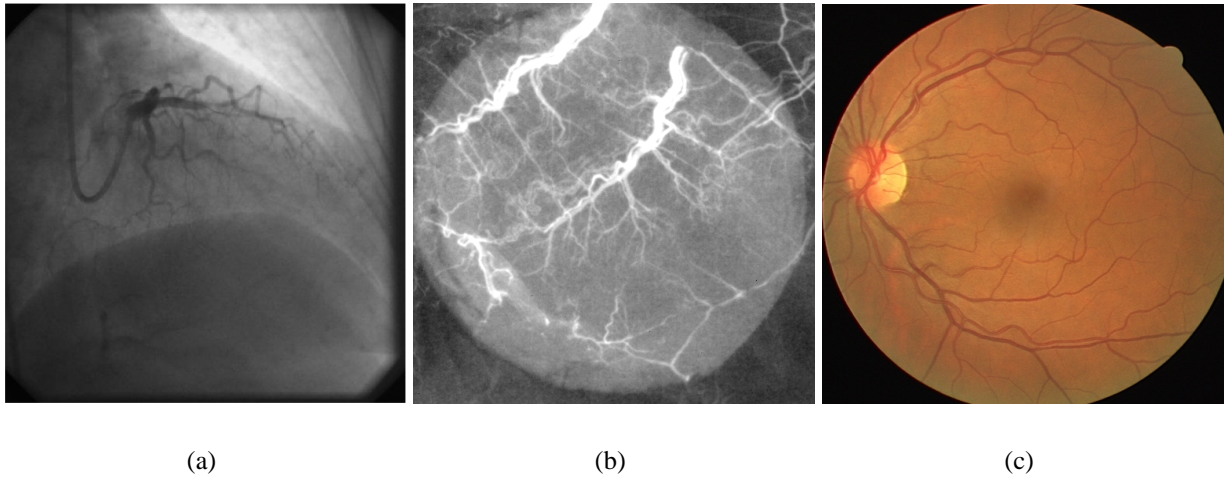
Figure 4.1: Examples of vessel images: a coronary angiogram (a), the transplant region of a microangiogram (b) and a retina fundus photography (c).

segmentation problem from the imaging system (see Section 4.2.1) and (ii) we use mainly prior information about vessels to improve their contrast and the homogeneity of their gray-level representation (see Section 4.2.2). The result obtained after any of these two enhancement steps is called here a *vessel map*.

For an optimal description of vessels and background, several vessel maps can be combined to obtain multidimensional pixel feature-vectors, as described in Section 4.2.3. In Section 4.2.4 we discuss the results of several experiments conducted with the purpose of validating the vessel enhancement methods described here. Section 4.2.5 contains concluding remarks.

4.2.1 Background attenuation

For vessel segmentation, an ideal background should be homogeneous, such that its corresponding pixel-class has a small variance. Therefore, all non-vessel image items should be attenuated or eliminated.

Depending on the application, vessels of different diameters are imaged against different backgrounds, with different resolutions and they can be lighter or darker than the background. This is shown in Figure 4.1. The coronary angiogram in Figure 4.1 (a), has a rather heterogeneous background showing besides vessels different items like the dark circular diaphragm in the lower part, or the bright lung-tissue in the upper right part. The background of the skin-transplant microangiogram in Figure 4.1 (b) is also heterogeneous. One can clearly see the circular transplanted skin-flap and its irregular borders. There are also variations within the flap as some regions are better irrigated by contrast-agent-carrying blood than others. The retina photography in Figure 4.1 (c) shows besides vessels, the fovea with its high concentration of cones and the blind spot where the optical nerve leaves the retina.

By convention we impose that *vessels are darker than the background* [9], thus images where the vessels appear lighter should first be rescaled. Color images, like e.g., fundus photography are first transformed to gray-level images, usually by ignoring the blue and red channels

<table>
<tr><td>(a)</td><td>(b)</td></tr>
</table>
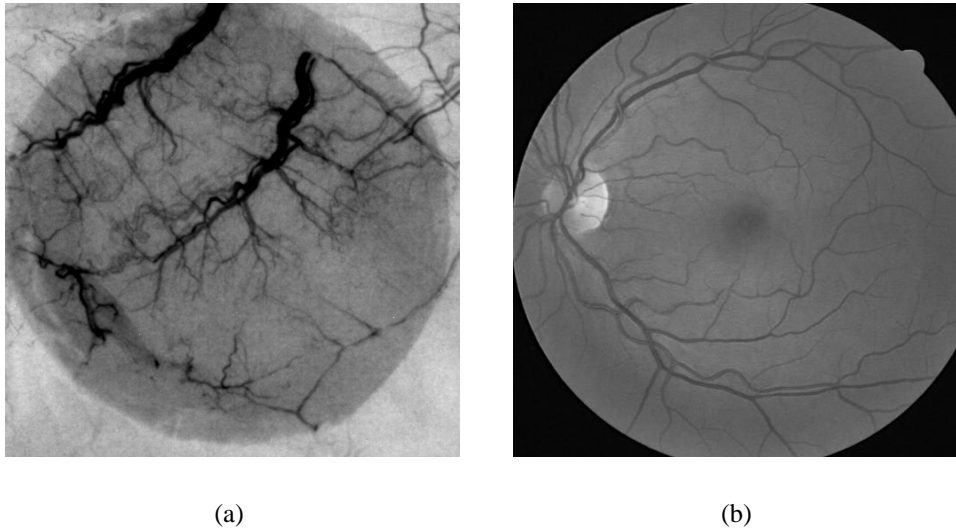
Figure 4.2: Examples of convention-conform vessel images, with dark vessels on a lighter background.

and using only the green one, which contains the most contrast[3]. Suited input images are shown in Figure 4.2.

For 2D+t imaging, background attenuation is simply the subtraction of the mask image from the contrast images. If necessary, the two images need to be first registered [31] to avoid the appearance of artefacts.

Usually, background structures are larger than vessels, thus the background will be attenuated by a *high-pass filtering* approach. We describe two methods to implement such an approach: (i) by *linear shift-invariant processing* [90] and (ii) by nonlinear *morphological processing* [67], [90].

LSI processing

The *high-pass filter* is a linear shift-invariant system, which is designed in the frequency domain. Special care has to be taken to adapt the filter to the *imaging model of the input image*, which in this case can be *additive or multiplicative*.

LSI high-pass filter for background attenuation.   A LSI high-pass filter can be implemented by a kernel, or by the difference between the original image and a low-pass filtered version [90]. For vessel enhancement, the pass frequency of the filter has to be chosen in relation to the diameter of the largest vessel. Theoretically, this frequency can be precisely computed, if one knows the vessel diameter. Practically – assuming an appropriate ground truth is available – it is better to choose it by analyzing some quality measure of the filtering result for an interval of frequencies. As a heuristic rule, if the high-pass is implemented as the difference between the original image and a low pass filtering result obtained with a Gaussian-kernel, then the size S of this kernel should be at least two times larger than the diameter of the largest vessel. Its standard deviation should be $\frac{S}{5}$ [188]. The result of a high-pass filter is shown in Figure 4.3.

---

[3]Assuming images acquired with visible light.

(a)   (b)
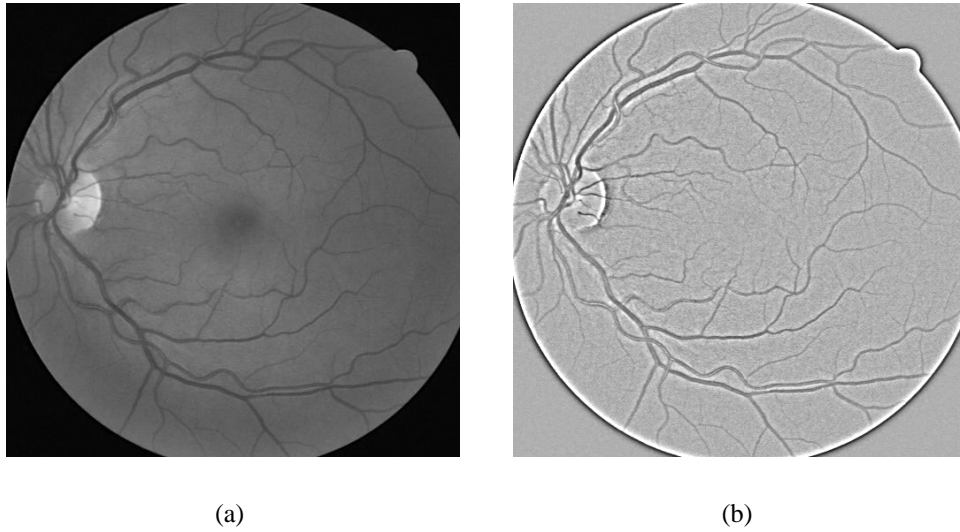
Figure 4.3: Vessel image (a), result of background attenuation by high-pass filtering (b).

The homomorphic filter.   Suppressing the background and selecting the vessels by high-pass filtering takes into consideration an additive composition of objects and background. Such a composition should be expected in the case of an additive imaging model, which is based only on the reflective properties of the imaged tissue. This approximation is sufficient for vessel images acquired with light, under uniform illumination conditions, but is not correct for X-ray vessel images. By the absorption integral (see Equation 2.9) such images are better described by a multiplicative imaging model. Under such circumstances it is to be expected that the composition of objects and background is also multiplicative. Thus, for X-ray images, the background should be attenuated by a homomorphic filter [153], [90]. In a fist step, the logarithm of the image is computed, thus obtaining again an additive combination of background and object. Then, the background is eliminated by high-pass filtering and finally the contrast in the result image is restored by exponentiation. The result of such a homomorphic filter is shown in Figure 4.4.

Morphological processing

The background can be suppressed also by morphological processing. We propose to use the *Bothat transform* for this purpose. The Bothat transform assumes an additive imaging model and therefore *special care* has to be taken in *the case of X-ray angiograms*.

The Bothat transform for background attenuation.   Taking into consideration the size of the vessels and the fact that they are darker than the background, dilating the original image with a structuring element slightly larger than the largest vessel diameter will suppress the dark vessels but it will also affect some dark structures larger than the vessels. Such structures are then "repaired" by a subsequent erosion step. The final result of the closing operation will show only the background. A vessel-only image can be then obtained by subtracting the original image from the image obtained after closing. This chain of operations is called a Bothat transform or an inverted Tophat transform [90] (see also Appendix A.1).

The main parameter of the Bothat transform is the size and form of the structuring element.

(a)   (b)
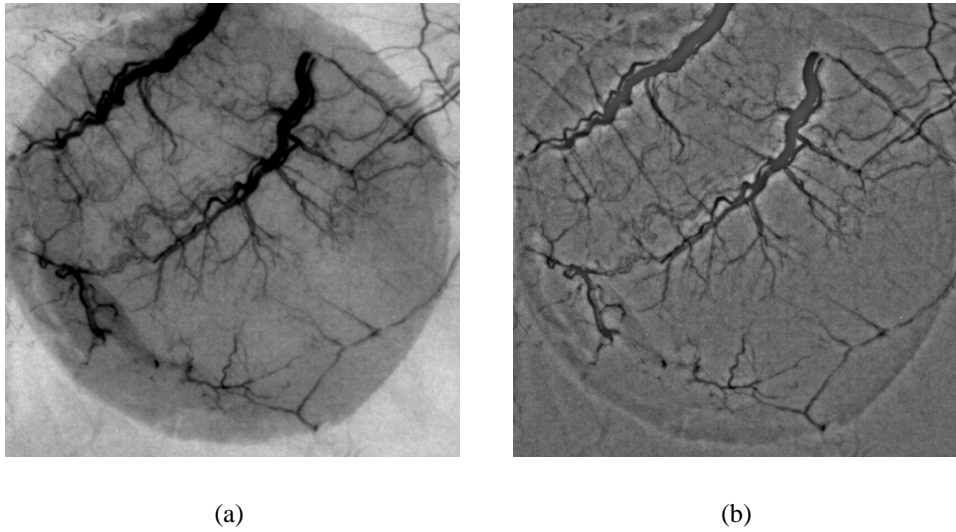
Figure 4.4: Vessel image (a), result of background attenuation by homomorphic filtering (b).

The size of the structuring element has to be chosen such that is slightly larger than the diameter of the largest vessel. Again, assuming an appropriate ground truth is available, it can be chosen as the value optimizing a measure of the quality of the result. The form of the structuring element has to be isotropic such that vessels are selected irrespective of their orientation. Thus the optimal structuring element is a disk. Sometimes, a square approximation of the disk is used for efficient implementations. A result of the Bothat transform is shown in Figure 4.5.

Adaptation to a multiplicative imaging model.   Similar to the LSI high-pass filtering, the Bothat transform assumes an additive imaging model. To accommodate a multiplicative model, a logarithmation should be effectuated before the transformation[4].

Such an approach is also instrumental in reducing the intra-class variance. Before logarithmation, the intensity representation of a vessel of a given size depends not only on its own absorption, but varies with the absorption of different background structures situated beneath the vessel. Background equalization after logarithmation removes this influence on the vessel representation so that only the variability of the vessel thickness remains.

In a last processing step, the logarithmation is reversed by exponentiation. Such a processing result is shown in Figure 4.6.

4.2.2   Vessel augmentation

There are two objectives to be followed while increasing the separability of the vessel pixel-class, i.e., improve the contrast of vessels to non-vessel structures and reduce the vessel variance.

Improving the vessel contrast.   Clearly, the attenuation of background, which is primarily based on the vessel size, returns a result including many non-vessel structures of similar size

---

[4]Alternatively, one could replace subtraction by division.

(a) (b)
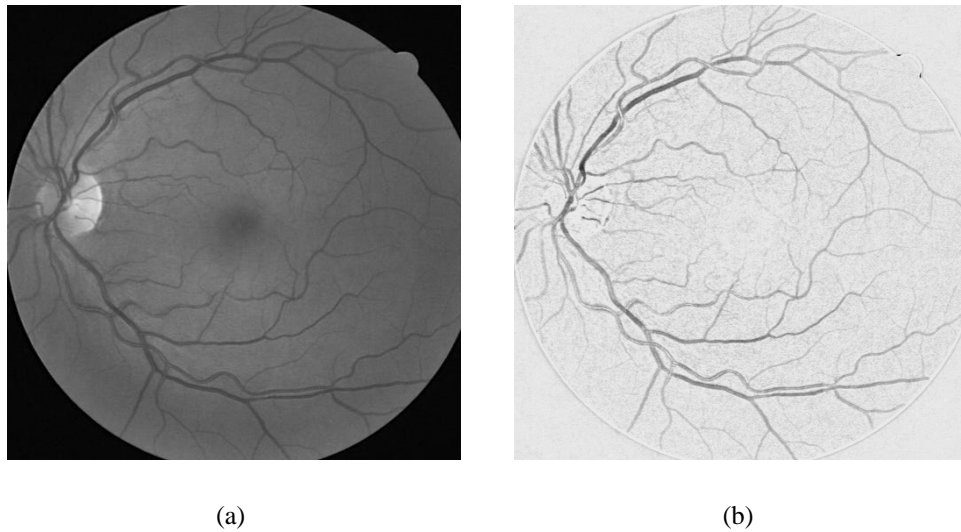
Figure 4.5: Vessel image (a), result of background attenuation by morphological processing (b).

(a) (b)
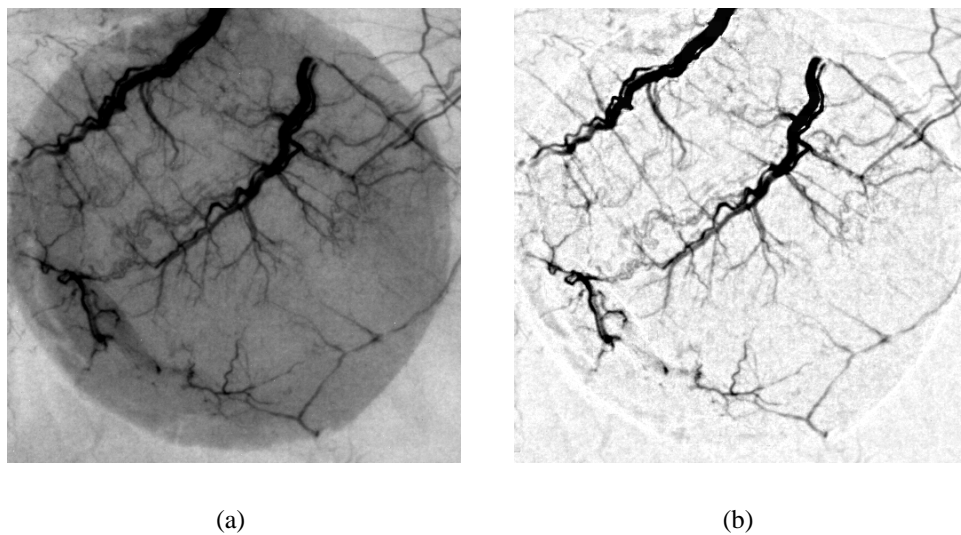
Figure 4.6: Vessel image (a), result of background attenuation by morphological processing adapted to a multiplicative imaging model (b).

and sometimes even artefacts from larger structures. Also, the background attenuation does not affect the usual high-pass image noise, like e.g., the sensor noise (see Section 2.2.4). Thus, vessels need first to be *selected* from among other items in the background-attenuated images, to increase then the inter-class variance in a targeted manner.

Reducing the vessel variance. The relation between contrast and vessel size.   Vessel contrast in 2D-projection images is a consequence of the interactions between the primary information carrier – be it light or X-ray – and compounds found in the blood, either natural ones or ones that have been mixed with it on purpose. Assuming that such compounds are uniformly distributed in blood, the larger the vessel size – i.e. the more blood passes through – the better the contrast. Thus, smaller vessels have a poorer contrast than larger ones and hence the variance of the vessel pixel-class is rather large. We propose to *reduce this vessel intra-class variance* in a *multiscale approach* to vessel augmentation.

Vessel selection

Assuming that after attenuating the background, artefacts *larger* than vessels and *high-frequency* noise are still present, one can select the vessels by their size, e.g., by a LSI *band-pass filter*.

Until now we have used extensively the vessel size both to attenuate the background and select the vessels. However, the vessels have also other properties, like e.g., their tubular form, which causes them to appear as ridges in angiograms. Therefore, the vessels can be also obtained at the output of a ridge-detector. Such a detector can be obtained using the *Hessian matrix*.

Band-pass filter.   In the case of a real band-pass filter, the transition between the stop and the pass-band in the range of lower frequencies is not infinitely abrupt and thus some low frequencies are rather attenuated than eliminated. This characteristic is instrumental in reducing the variance of the vessel pixel-class, as low frequency and high contrast large vessels are more attenuated, while higher frequency and small contrast small vessels are less or even not attenuated.

A band-pass filter can be defined as a combination of two low-pass filters whose transfer functions are subtracted one from the other. These define the low and high cutoff bands respectively. The impulse response can then be computed by inverse Fourier transform.

Practical implementations of a band-pass filter include, e.g., the Difference of Gaussians (DoG) and the Laplacian of Gaussian (LoG) filters. Such filters approximate also the second derivative of the image function and have therefore a strong response over ridge-like structures such as the vessels [108].

To obtain the band-pass vessel map we use the DoG, because it is easy to parameterize and it is more isotropic – i.e. direction insensible – than the LoG [108]. To implement the DoG, one needs to define only the standard deviations of the two Gaussian filtering-kernels. To achieve the desired improvement of the inter-class variance the result is multiplied by a weight. The parameters of the optimal filter can be chosen in a target-oriented manner, with the help of some training data for which the ground truth is known. Usually, they correlate with the vessel-size, i.e., the size of one filter should be larger than the largest vessel-diameter, while the size of the other filter should be smaller than the smallest vessel-diameter. A result obtained when applying the DoG on a vessel image is shown in Figure 4.7 (b).

λ_1	λ_2	
L	L	noise, no ridge
H(+)	L	dark ridge
H(-)	L	bright ridge
H(+)	H(+)	dark blob
H(-)	H(-)	bright blob

Table 4.1: Classification of eigenvalues. H/L high/low modulus and (+)/(-) positive/negative

Hessian-based analysis. Ridges are important image features and ridge enhancement and detection algorithms have been extensively studied in low-level image processing [107].

A ridge contains points of maximal curvature and thus they can be detected at extrema of the second-order derivative. The second order structure of an image can be investigated at each pixel by means of the Hessian matrix:

$$H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \tag{4.1}$$

This is a symmetric matrix which has thus orthogonal eigenvectors and real eigenvalues. The size of the eigenvalues is related to the amplitude of the variation along the eigenvectors. The eigenvector corresponding to the largest eigenvalue is oriented along the direction of maximal variation in a neighborhood of an investigated pixel. The sign of the eigenvalues differentiates between positive and negative second-order derivatives. An analysis of the eigenvalues of the Hessian matrix gives a rich insight in the second-order structure of the image [17] as shown in Table 4.1.

Consequently, the first eigenvalue λ_1 of the Hessian matrix can be used as ridge – i.e. vessel – detector, while the second eigenvalue λ_2 can be used to differentiate between blobs and ridges. Alternative methods for vessel enhancement include steerable filters [85], and other equivalent matched filter approaches [84], [159], but it can be demonstrated that ridge-estimation by eigen-decomposition of the Hessian matrix can be interpreted in terms of steerable filters [107].

In the background-suppressed images vessels are dark ridges. Therefore, we propose to construct a vessel map by subtracting at each pixel the second eigenvalue of the Hessian from the first eigenvalue. Clearly vessels will appear brighter than other structures in this vessel map. To be conform with our convention the vessel map is rescaled such that vessels are darker than background (see Figure 4.7 (c)).

It can be shown that an improved ridge detector can be obtained at only minimal additional computational complexity from the eigen-analysis of a modified Hessian matrix [107]:

$$H_m = H\frac{1}{3}(P^T H P) \tag{4.2}$$

with

$$P = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \tag{4.3}$$

In the practical implementation of a Hessian-based vessel map we use H_m. The derivative kernels are optimized, they are computed as the derivative of a Gaussian low-pass filtering kernel to avoid noise-boosting. The size of the derivatives and that of the Gaussian kernel

<div align="center">

(a)   (b)   (c)

</div>
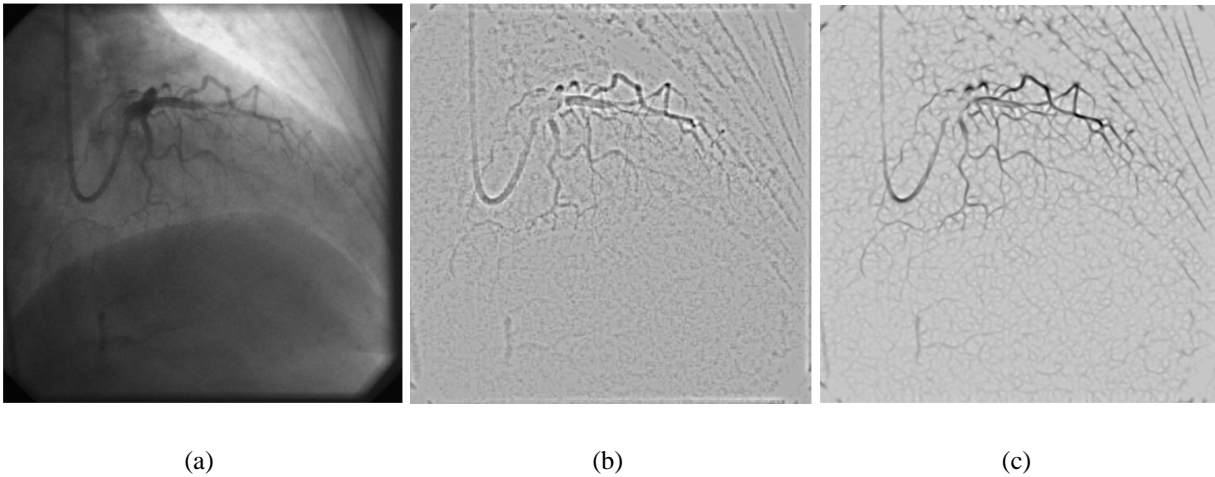
Figure 4.7: Original vessel image (a), result of band-pass-based enhancement (b) and of Hessian-based-enhancement (c). The Bothat eliminated during background attenuation that part of the vessel where contrast-spill was observable.

represent the parameters of this vessel map. Similar to the band-pass vessel map, they can be adapted to the analyzed data in a calibration step and they correlate with the vessel size. An example is shown in Figure 4.7 (c).

Reduction of the variance of the vessel-class. Extension to multiple scales

A multiscale approach to vessel enhancement is motivated by the variance of the contrast with the vessel-size. It offers individual access to each category of vessels – i.e. small-, mid- and large-size – such that they can be then enhanced independently by multiplication with different weights, reducing thus their original contrast-difference and increasing the overall vessel contrast.

We show how to compute a vessel map using the *Laplacian pyramid*, which can be seen as a multiscale extension of the band-pass filter vessel map. We also show how to extend the *Hessian*-based analysis at *multiple scales* and how to compute then a vessel map.

Laplacian pyramid. A band-pass decomposition of the image can be obtained by representing the image in a differential scale-space [108], where the change of the image with scale is emphasized in the difference between consecutive scales. In practice one encounters mainly discrete implementations of a differential scale-space, which use, e.g., a DoG filter bank. A wavelet transform could be also used [39] but it can be shown that the results achieved are comparable or even worse [3], [4], [66].

An image representation based on DoG has the disadvantage of large data-size and computational cost, considering that the latter increases with the size of the filtering kernels. Therefore, it is more efficient to use a Laplacian pyramid [30]. The Laplacian pyramid however has two small disadvantages: (i) it decomposes the image in octaves – as it is a discrete version of a differential scale-space – and it may happen that such a decomposition is not optimal to select each category of vessels, although usually this is a very good approximation and (ii) downsampling

usually introduces alias, which can not be ignored if the subbands are processed to increase the contrast of vessels. Alias needs thus to be kept very low by an appropriate design of the decomposition and (implicitly) reconstruction filters.

For vessel enhancement we use a Laplacian pyramid with five levels, to model: noise, small, mid and large vessels and the rest-background. For enhancement, we multiply each subband by a weight, before reconstructing the image. To both increase contrast and achieve a reduction of the variance of the vessel class, the weights should be chosen according to a positive Gauss-like function of the subband index (see "The relation between contrast and vessel size", at page 122). Therefore, the noise subband will have a small, subunitary weight, the small vessels a large supraunitary weight, the mid vessels a smaller supraunitary weight, the large vessels a weight only slightly larger than the unit and the rest-background a weight very close to zero. The rest-background receives a weight larger than zero as this enhancement step is usually applied on the background-suppressed images and thus the rest-background may still contain some vessel information. The parameters of the pyramid, i.e., the frequency characteristics of the decomposition/reconstruction filters[5] and the enhancement weights can again be chosen with the help of some hand-labeled test data or using some labeled training data, if this is available. A processing result is shown in Figure 4.8 (b).

Hessian multiscale decomposition. The computation of the Hessian matrix includes several filtering operations, like e.g., high-pass filtering to compute the derivatives and low-pass filtering to avoid noise boosting. Clearly the scale of these filters defines the size of the ridges–vessels to which the Hessian responds. Scale mismatch results in information loss in the enhancement result.

Working at a single scale, the size of the filters should be chosen in agreement with the diameter of the largest vessel. However, considering the typical size-induced variation of the vessels' contrast, a multiscale approach to Hessian eigen-analysis is better suited for vessel enhancement [83].

To obtain a scale-space representation of the image, a Gaussian pyramid [108] can be used with three levels to model each of the three main vessel categories, i.e., large, mid and small. Now – due to downsampling – one can select vessels at each level by Hessian eigen-analysis using filters of similar size, which is good practice, considering the goal of reducing the internal variance of the vessel pixel-class. At the same time, the successive low-pass filters, used to compute the pyramid-levels, have as secondary effect a reduction of the contrast of larger vessels. This contributes additionally to the reduction of the variance of the vessel class.

To obtain the vessel map we proceed the following way: first the results of the eigen-analysis at different pyramid levels are interpolated back to the original resolution and weighted such that this compensates for the size-dependent contrast-loss and therefore reduce the variance of vessels. Finally the results are combined across scales by always choosing the maximum gray-level value at each pixel position:

$$V(i,j) = \max_k (w_k \cdot v_k(i,j)) \qquad (4.4)$$

with V the vessel map at pixel (i,j), k the index of the levels of the pyramid and w_k the weight of the pyramid level v_k. The weights w_k should be chosen in agreement with the vessel-size

---

[5]The 'a' parameter [30] controls the amount of low-pass filtering in the subbands, the higher 'a' the closer to a band-pass rather than a low-pass are the filters.

| (a) | (b) | (c) |
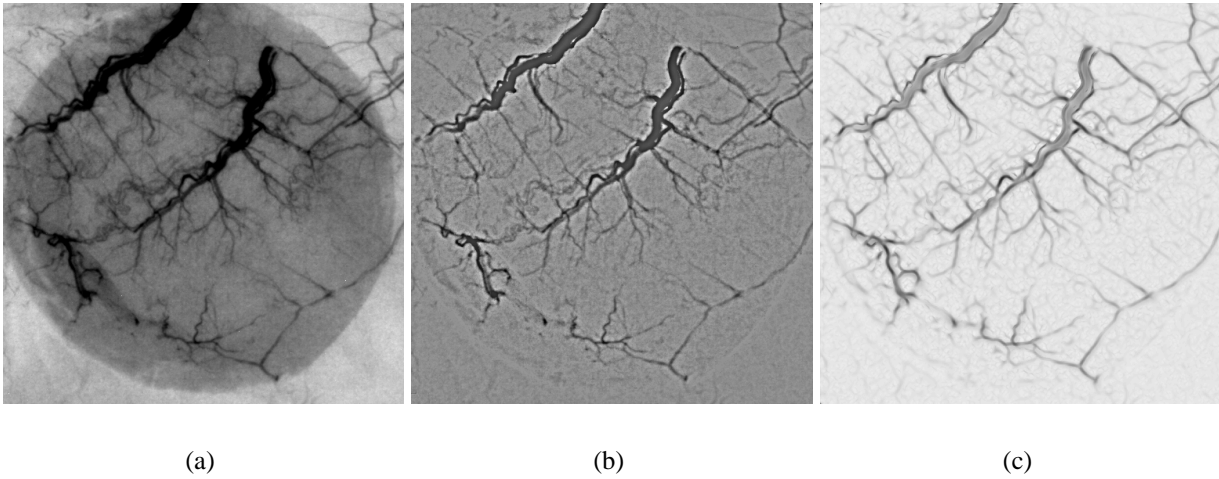
Figure 4.8: Original vessel image (a), enhancement using the Laplacian-pyramid (b) and enhancement using the Hessian multiscale decomposition (c).

such that small vessels are stronger enhanced than larger ones.

The optimal parameters for this enhancement method can also be computed in a calibration-step using, e.g., hand-labeled data. Such an enhancement result is shown in Figure 4.8 (c).

4.2.3 Pixel-based multidimensional description of vessel and background

A multidimensional description of vessels and background is obtained by combining the results of several different vessel enhancement methods. For each pixel we build a feature vector by ordering its scalar features in each vessel map into a vector [52], [51].

Within each vessel map we seek to increase the separability between the background and vessel pixel-classes. Our strategy is in this case to combine the results of several different enhancement methods, in the hope that together they constitute a more separable representation of vessels and background than any of them taken alone. We believe that a multidimensional pixel feature-space is better than a single vessel map, as it includes more information about vessel, acquired from different perspectives. A schematic representation of the way a pixel-based feature vector space is computed is shown in Figure 4.9.

Which vessel maps should be used to compute the multidimensional pixel feature-vector, can be decided with the help of a feature selection procedure [160], [80]. In Section 4.3.2 we describe also a feature selection procedures dedicated to computing the optimal set of such pixel-features for a special type of classifier: the hysteresis classifier.

4.2.4 Vessel enhancement: Experiments and discussion

To assess the value of the vessel enhancement methods introduced here, hand-labeled coronary angiograms, skin-transplant microangiograms and retina photographies were used to compute several measures of separability on each vessel map. These measures are: the J_1^a, J_1^b and J_1^c criteria (see Appendix A.6) as well as the area under the ROC of a percentile threshold (AROC).

vessel-image   vessel-maps   pixel feature-vectors
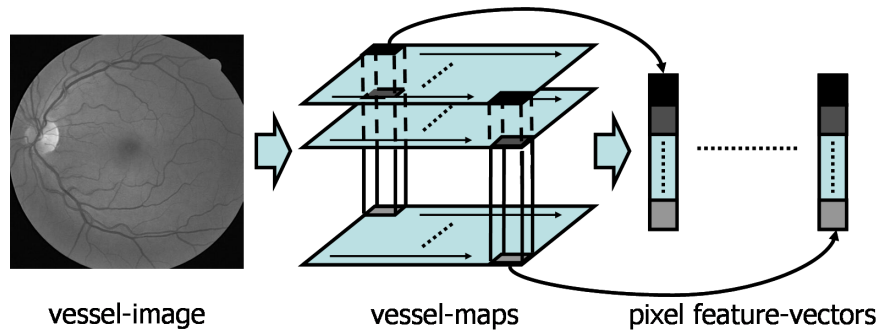
Figure 4.9: Schematic representation of the method to achieve a pixel-based multidimensional description of vessel and background.

It is considered that the AROC is a better indicator of the quality of a vessel map, with respect to the segmentation results that can be achieved there. This is mainly due to the target-oriented manner by which this criterion is computed, i.e., by observing how the most simple segmentation procedure would behave. The J_1 criterion correlates with the AROC only under the assumption of Gaussian distributed classes. For class-conditional distributions of different skew and kurtosis the results obtained by the J_1 criterion, analyzed together with that of the AROC, can be seen as an indication of how close are they to the Gaussian hypothesis. Such information can be useful in choosing a specific classification algorithm. The J_1^a criterion considers both the inter- and the intra-class variances, i.e., both contrast and homogeneity within each class. The J_1^b criterion is more related to the contrast and the J_1^c criterion to the homogeneity of the classes.

For experimenting we have used 11 coronary angiograms and 11 skin-transplant microangiograms, from our own data-base [43], [52] as well as 20 retina images from a publicly available data-base [181]. The results we have computed are mean values over all images of the corresponding data-set.

In this section we discuss only the background attenuation and the vessel augmentation methods, the pixel-based multidimensional description of vessel and background has been tested in Section 4.3.3, together with the hysteresis classifier.

Background attenuation

Results.   In Table 4.2 we show the separability measured on a vessel map obtained after background attenuation with: the High-Pass filter (HP), the homomorphic filter (log-HP), the Bothat transform (B) and the logarithmic Bothat transform (log-B). For comparison we have measured also the separability in the original images for each analyzed category. The results are mean values over all components of each image set.

Discussion.   Suppressing the background by the Bothat transform returns better results than if a high-pass is used. The main reason for this is that the practical implementation of the high-pass can eliminate only the zero-frequency component, while all other frequencies in the stop band are more or less attenuated and produce disturbing artefacts. In other words, the Bothat transform – seen as a high-pass filtering operation – can achieve a steeper transition between the cut and the pass bands. Another disadvantage of the high-pass filtering is that the relation

	Skin				Coronary				Retina			
	J_1^a	J_1^b	J_1^c	AROC	J_1^a	J_1^b	J_1^c	AROC	J_1^a	J_1^b	J_1^c	AROC
HP	0.18	0.15	0.85	76.86	0.10	0.09	0.91	82.34	0.40	0.28	0.72	87.29
log-HP	0.17	0.14	0.86	77.21	0.09	0.09	0.91	82.33	0.39	0.28	0.72	87.08
B	0.33	0.23	0.76	81.75	0.27	0.20	0.80	87.06	1.03	0.50	0.50	94.88
log-B	0.39	0.28	0.72	83.96	0.15	0.13	0.87	87.77	1.15	0.53	0.47	88.98
Org.	0.15	0.12	0.88	80.21	0.002	0.002	0.10	66.55	0.05	0.05	0.95	59.98

Table 4.2: Separability after background attenuation, measured as mean value on different image sets by different methods.

	Skin				Coronary				Retina			
	J_1^a	J_1^b	J_1^c	AROC	J_1^a	J_1^b	J_1^c	AROC	J_1^a	J_1^b	J_1^c	AROC
BP	0.28	0.22	0.78	84.23	0.16	0.13	0.87	84.79	0.51	0.34	0.66	90.31
LapP	0.51	0.33	0.67	91.32	0.33	0.24	0.76	92.25	0.97	0.49	0.51	95.60
Hss	0.32	0.23	0.77	84.71	0.04	0.04	0.96	88.23	0.88	0.47	0.54	94.07
Hms	0.46	0.31	0.69	90.11	0.44	0.30	0.70	90.46	0.86	0.46	0.54	95.01
(log-)B	0.39	0.28	0.72	83.96	0.27	0.20	0.80	87.77	1.03	0.50	0.50	94.88

Table 4.3: Separability after vessel augmentation, measured as mean value on different image sets by different methods.

between the vessel size and the pass-frequency is rather difficult to grasp.

Taking into consideration a multiplicative imaging model and working with the logarithm, returns better results for the skin-transplant microangiograms. However, only limited improvements can be seen for the coronary angiograms. For retina images, which are acquired with light, a logarithm operation is not necessary.

As the J_1 criteria and the AROC correlate for skin transplant microangiograms, it can be concluded that the two class-conditional *pdfs* obtained after background suppression are Gaussian-like. This is also the case for coronary angiograms, but not for retina images.

Vessel augmentation

Results. Table 4.3 contains the separability measured on a vessel map obtained after vessel augmentation by the methods that have been proposed: band-pass filtering (BP), Laplacian pyramid (LapP), Hessian single-scale (Hss) and Hessian multi-scale (Hms). The parameters for each method and each data set were optimized using hand-labeled examples. For comparison we have computed also the separability on the background attenuation result that was used as input ((log-)B). For the microangiograms this was the log-Bothat and for the coronary angiograms and for the retina images it was the Bothat. The results are mean values over all components of each image set. We judge the results mainly by the AROC criterion and give also the J_1 criteria, to investigate the Gaussianity of class conditional *pdfs* in the computed vessel maps.

(a)   (b)
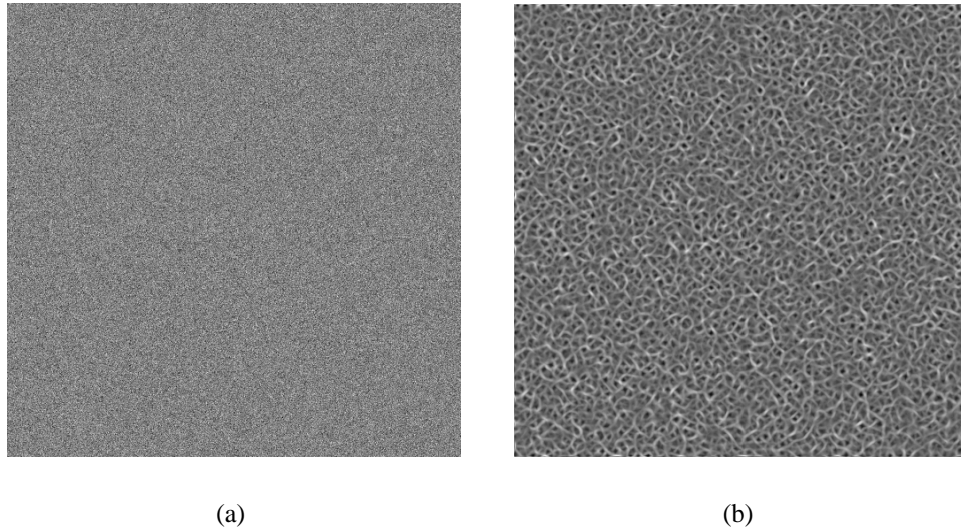
Figure 4.10: Input noise image (a) and result of Heesian single-scale vessel enhancement (b). Observe the thin vessel-like artefacts in the vessel map.

Discussion.   The results clearly show that the multiscale methods are better than the single-scale ones. They achieve both a contrast enhancement – i.e. increasing the inter-class variance – and an increase in the homogeneity of the vessel representation – i.e. decreasing the intra-class variance.

It appears also that the vessel map obtained by decomposition and enhancement on a Laplacian pyramid, is better than that obtained by the multiscale analysis of the eigenvalues of the Hessian matrix. However the differences are quite minor and could be also due to the imperfect ground truth.

The band-pass vessel map achieves an improvement over the result of the background attenuation step only for skin microangiograms. The Hessian single-scale does not further improve the separability measured after background attenuation for retina images. This shows again that single-scale enhancement methods are only of limited use. However, such enhancement methods may prove useful if used in multidimensional pixel feature-spaces (see Section 4.4).

One can dichotomize the methods proposed here in LSI-methods, including the band-pass filtering and the decomposition on the Laplacian-pyramid and Hessian methods, including the single-scale and the multiscale approach. Generally the Hessian approach returns smoother more connected vessels than the LSI approach, but it also has the potential of returning many false positives, as it has the characteristic of "gathering" vessels even where none are present. This is illustrated in Figure 4.10. The image in Figure 4.10 (b) is the Hessian single-scale vessel map obtained for the input shown in Figure 4.10 (a), containing only uniformly distributed noise. We believe that choosing between LSI- and Hessian-methods should be done in an application-dependent way.

4.2.5   Vessel enhancement: Conclusions

The purpose of vessel enhancement is to obtain a vessel map by transforming the original vessel-image such that better segmentation results are achieved. This implies improving vessel contrast

and the homogeneity of the gray-level representation of vessels. To compute the vessel map, we make use of prior knowledge about the form, contrast and – in connection with the parameters of the imaging system – size of the vessels, to improve the separability of vessels and background. During enhancement, the vessels are brought to a "standardized form" – i.e. they are dark, surrounded by a brighter and homogeneous background – thus allowing the same segmentation algorithm to be also used for images acquired by different imaging modalities.

The vessel map. To obtain a separable representation of vessels and background, we seek to increase the inter-class variance and to decrease the intra-class variance, i.e., increase the contrast between vessels and background and improve the homogeneity of the gray-level representation of vessels and background. Enhancement proceeds in two steps:

- **In the first step**, the background is attenuated thus improving the homogeneity of its gray-level representation. The best way to attenuate the background is by using the Bothat transform. Depending mainly on the imaging system, a logarithmation may be necessary to ensure an additive imaging model.

- **In the second step**, the contrast of vessels relative to the background is improved together with the homogeneity of the gray-level representation of the vessels. For this purpose, we found that a multiscale approach yields the best results, because it provides individual access to each category of vessels, i.e., small-sized, mid-sized and large. They can be then enhanced, such that both contrast and homogeneity are improved. With respect to the last point, one should remember the fact that vessel contrast usually varies with size in 2D-projection images. We have described two novel multiscale vessel-augmentation methods. Which method should be chosen in practice, depends on the application.

The result of an enhancement step is called a vessel map.

The set of parameters. The optimal set of parameters for a vessel map can be established in a calibration step using images with known ground truth. If such a calibration step is not feasible, than:

- **For background attenuation**, the size of the Bothat structuring element should be correlated with the diameter of the largest vessels. The structuring element should be a disk.

- **For vessel augmentation**, the weights of the Laplacian pyramid should be chosen such that the noise and background are suppressed and the smaller vessels are better enhanced than the larger ones, i.e., as if they were a chosen from a Gauss-like function of the subband index. The same observation is valid for the Hessian multiscale vessel map. However, as noise is not modeled separately in this case, the processing at the first level of the Gaussian pyramid should obtain a weight larger than one, comparable with that of the third level, but smaller than that of the second level.

We have assessed the quality of the methods proposed here by measuring the separability of the obtained vessel maps by general separability criteria – i.e. the J_1 criteria – and by the AROC of a percentile threshold. Together they can give a deeper insight into the data structure within a vessel map. If the two criteria correlate, than this usually means that the class-conditional probability densities are Gaussian-like. Such information can then be used to choose a specific classification algorithm.

Multidimensional description of vessels and background. A pixel-based multidimensional description of vessels and background can be achieved by combining several different vessel maps. In this case, a pixel is characterized by a vector composed from all scalars representing the values assigned to the pixel in each vessel map. We believe that such a feature space is more separable than any of the vessel maps used to compute it taken alone (see also in Section 4.3.4, the "Comparison between supervised and unsupervised hysteresis methods" at page 160).

4.3 Automatic vessel segmentation in 2D-projection images

Semi-automatic vessel segmentation algorithms rely on additional information provided by the user to achieve a result. Such a strategy is clearly optimized for quality, as the human operator can guide the algorithm to extract the vessels he is interested in. However, this strategy fails if the number of images which need to be segmented is very large, e.g., for the analysis of vessel images acquired during screening of a larger population. Also, automatic algorithms are usually used if a segmentation is needed only as an intermediary step in a more elaborated analysis tool.

The following presentation of vessel segmentation starts with the introduction of some novel threshold-based methods in Section 4.3.1 and culminates with the hysteresis classification paradigm and its practical implementation as a hysteresis threshold and a hysteresis classifier, described in Section 4.3.2. All segmentation methods are tested as described in Section 4.3.3 and their results are discussed in Section 4.3.4. The conclusions are presented in Section 4.3.5.

Thresholds and classifiers. Assuming the segmentation result is computed on a single vessel map, the simplest way to do this is by using a threshold. Thresholding is a scalar (1D) procedure, meaning that each pixel is described only by one value, which is used to separate only two classes. A classifier on the other hand can handle also vectorial (multidimensional) inputs and indeed more than two classes. We consider the threshold to be the most simple classifier. However, on some occasions within this section a distinction is made between thresholds, which work only on scalar inputs and classifiers, which work on vectorial inputs. Where this is not clear from the context we will highlight this difference when it is made.

4.3.1 Segmentation by thresholding

The result of vessel enhancement is a vessel map with improved separability between vessels and background. The simplest modality to segment the vessels is by setting a threshold. This threshold is said to be fixed – i.e. a *fixed threshold* – if it is compared against the gray-level values of all pixels within the image. A fixed threshold can change from image to image but not within the same image. If different thresholds are used for different pixels, e.g., by refining the result of the fixed threshold in an additional step, then this is called *adaptive thresholding*.

The results obtained by thresholding – be it fixed or adaptive – can be further improved by making use of the connectivity of vessels. Vessels are connected structures which appear always in a tree-like pattern, because the blood flows from larger vessels into smaller-vessels up to the level of capillaries such that cells from different tissues can exchange food and waste products with the blood-cells to complete their life-cycle. Therefore, if one can find some high-confidence vessel pixels, then such *vessel markers* can be used to *select the remaining vessel pixels* from the thresholding result.

Fixed thresholds

A fixed threshold is obtained here either as a percentile on the histogram of the input image, or as the scalar expression of a general purpose unsupervised classification algorithm. We discuss several such thresholding algorithms.

The percentile threshold.   In the case of vessels segmentation, the histogram of the input image can be successfully used to establish a fixed threshold. The simplest such threshold is a percentile threshold, which is usually chosen such that it reflects some prior expectations on the number of vessel pixels, i.e., the surface of the vessel map covered by vessels.

Thresholds obtained from statistical unsupervised classification algorithms.   A percentile threshold is rather rigid and in some way it imposes compliance of the input data with the prior expectations, which motivated the threshold. However, such prior expectations express general rules and the data complies with them only within certain limits.

The Otsu threshold.   A more flexible approach is obtained if instead of keeping the surface of segmented vessel pixels constant, one seeks a threshold such that the separability measured between the two class conditional *pdfs* obtained after thresholding is optimal. Such a method was proposed by Otsu [155] (see also Appendix A.6.2). It measures the separability between two class conditional *pdfs* as computed from the results of several potential thresholds and chooses finally the threshold corresponding to optimal separability.

Other thresholds.   The disadvantage of the Otsu-method is that the separability measures which it uses are appropriate only for Gaussian-like *pdfs*. If the Gaussianity assumption does not hold, then the threshold which is returned is not the optimum. Assuming the parametric form of the class-conditional *pdfs* is known a priori, one could use the Expectation-Maximization (EM) algorithm [65], [21], [20] for a mixture with two components to estimate the *pdfs* and then find the threshold by a likelihood-ratio test [184]. However, usually one can only guess the parametric form of the *pdfs*. If the distributions are assumed Gaussian – which is often the default guess – one obtains results similar to the Otsu-method.

Thresholds obtained from non-statistical unsupervised classification algorithms.   Many unsupervised classification algorithms return a fixed threshold, when applied to scalar data. If the analyzed vessel map is seen as the training set of such an algorithm, then a segmentation can be achieved by reading the output of the algorithm when the training phase ends. Such a process takes often a long time to return a result. A typical vessel image has 512×512 or more pixels which means that the training space has some 262144 components.

One could also train for several images and then compute the result directly for later images. However, such an approach is likely to decrease the quality of the segmentation considering the typical variability of vessel images (see the discussion about "Unsupervised segmentation" at page 114).

Yet another difficulty is represented by the large class-skew – i.e. strong unbalance between the number of components in each class – typical for pixel-based vessel segmentation and which represents a challenge for most unsupervised learning algorithms.

The Kohonen-map. In the case of an unsupervised neural network, like e.g., a Kohonen-map (or the Self Organizing Map - SOM) [123], [149], the architecture of the network has to be chosen such that a binary segmentation is achieved, i.e., there are only two output neurons. This is rather counterintuitive, particularly in the case of the SOM. Similar to other unsupervised algorithms, the training time is in this case long, because it depends both on the number of components of the training-space and on the number of training-epochs. After the weights of the two output neurons are found, a pixel is assigned to a class depending on the distance between its corresponding feature vector and the nearest neuron/class-prototype. Assuming that the feature vector has just one component, i.e., the gray-level in the vessel map and that the distance function is the usual Euclidean distance, this is equivalent to comparing the gray-level of each pixel to a threshold situated at the middle of the distance between the two class prototypes.

Adaptive thresholds

Adaptive thresholds can achieve better results than fixed thresholds, mainly by making use of prior knowledge about the vessels. An example in this sense is given by the *sliding window-based adaptive threshold*. We introduce also a *novel adaptive threshold* based on a *Markov Random Field* (MRF) model.

The sliding window approach. The large overlap of the two pixel gray-level classes suggests that the vessels are only *locally* darker than the background. An adaptive threshold, which makes use of such knowledge is obtained by deciding for the class of each pixel individually in a sliding window, which is centered at the investigated position [173], [100]. The size of the window defines the pixel-neighborhood which is considered when taking the decision. Within each window where the variation of contrast is large-enough, a fixed threshold is computed – e.g. the Otsu threshold – otherwise the default label (background) is applied. Clearly the window should be larger than the largest vessel.

Such an approach has the disadvantage of being extremely time-intensive and sensible to parameterization. We will introduce next an improved adaptive threshold which is: fast, less sensible to initialization and more accurate, as it is able to make use of a larger amount of prior information

A MRF-based adaptive threshold for vessel segmentation. An adaptive threshold does not need to decide on the class of a pixel solely on the basis of the gray-levels and the local distribution of contrast. For improved results, it should take into consideration also other vessel characteristics. Such vessel characteristics can be, e.g., context-related, like the connectivity of vessels and their compactness.

An angiogram contains both contrast information and context information. A gray-level threshold – be it fixed or adaptive – uses usually the contrast information to return a segmentation. Context information can be used to further improve such an initial segmentation.

From a statistical perspective a decision can be achieved for each pixel by a likelihood ratio test. For this purpose, one needs to know the class-conditional *pdfs* for background and vessels. Context information can then be introduced through the class priors. We describe here such a statistically motivated adaptive threshold which we have also published in [53].

The algorithm.   Starting from an initial over-segmentation result several refinement steps are required to obtain a set of vessel candidate pixels. During these refinement steps every pixel is reclassified according to a decision rule which takes into consideration: (i) the class conditional pixel gray level probability and (ii) the influence of the classification results obtained for a certain set of neighboring pixels. Thus, if for a certain pixel a majority of his neighbors have been labeled as vessels, the chances are high that the pixel is itself a vessel, because vessels are compact. Accordingly, the prior-probability for class vessel should be in this case large. Such a behavior is modeled by a first-order MRF [184], [88].

A first-order MRF is the extension to images (2D) of a first-order Markov random chain (see also the "MAP segmentation" at page 68). For a MRF, the dependency condition appears now in the form of a certain neighborhood which influences the state of the current pixel. The states are assimilated to classes and the over-segmentation result serves to initialize the MRF. Let $Y = \{y_j\}$ denote the gray level image with a gray level y_j at each pixel j, and $S = \{s_j\}$ the segmentation result, with a binary label s_j indicating whether pixel j belongs to background or vessels. Starting from the maximum a posteriori criterion and invoking the Bayes theorem, we seek S such that $p(Y|S)P(S)$ is maximized. Here, $p(Y|S)$ is the likelihood linking the data Y and the segmentation result S, while $P(S)$ denotes the a priori probability of S. Since the gray level distribution of a pixel j depends only on its class s_j, the y_j are *conditionally* independent, implying $p(Y|S) = \prod_j p(y_j|s_j)$ [184]. The conditional marginal probability $p(y_j|s_j)$ is for each class estimated within the Otsu thresholding method described above.

The Markovian property implies that the conditional probability of s_j given all other labels depends only on the labels in a small neighborhood N_j around the investigated pixel j, i.e., $P(s_j|S\backslash s_j) = P(s_j|s_i, i \in N_j)$. For a first-order cross-shaped neighborhood, considering all cliques, we compute $P(s_j|s_i, i \in N_j)$ as in [88] by:

$$P(s_j|s_i, i \in N_j) = \frac{e^{-s_j(\alpha+\beta \sum_{i\in N_j} s_i)}}{1 + e^{-(\alpha+\beta \sum_{i\in N_j} s_i)}} \tag{4.5}$$

where α and β are parameters to be determined experimentally.

The decision rule is then: chose the class with the largest posterior probability which we express using the Bayes rule in terms of prior probabilities and class conditional probability density functions. This decision rule is used to iteratively reclassify each pixel until the classification result remains unchanged between two consecutive iteration steps or until a certain number of iterations have passed.

The over-segmentation result is obtained by thresholding the vessel map with a percentile-based threshold. After establishing empirically that the vessels always cover less than 50 % of the image area, we chose the 50'th percentile as threshold. The class conditional pixel probabilities are presumed Gaussian and parametrically estimated using the Otsu threshold.

Junction-based selection of true vessels. Vessel markers

In the case of vessel segmentation, both fixed and adaptive thresholds select besides vessels, more or less background structures. To eliminate such false positives, we use again prior knowledge about the vessels. We use the fact that vessels are the only items in the analyzed images that usually show bendings and bifurcations, i.e., points with more than one orientation: junctions. Thus, if such junction seed-points are detected, than there is a good chance that they belong to the vessels. We use also the fact that vessels are connected structures.
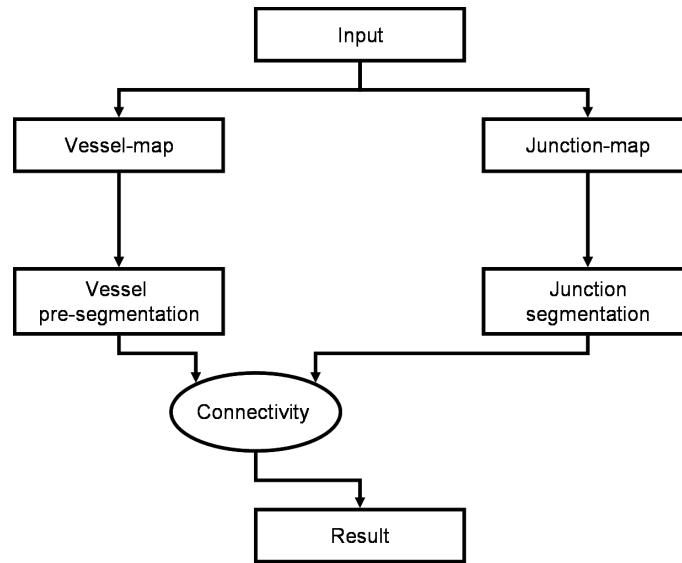
Figure 4.11: Flow chart of a vessel segmentation algorithm using thresholding and junction-based vessel selection.

We therefore segment vessels the following way: first we use any segmentation algorithm – e.g. an adaptive threshold – to select vessel-candidates, then we detect junctions and use these junctions to choose only vessel candidates connected to such a vessel marker, because vessels are connected structures. The connectivity can be both direct and it can extend also over other vessel-candidates. In the case of direct connectivity, the junction and the vessel-candidate pixel are neighbors in, e.g., a standard square-neighborhood. In the case of extended connectivity the junction and the vessel-candidate pixel are related over a chain of direct neighbors. We then consider the selection result to be the true vessels. To detect junctions, a junction map is first built by computing the second eigenvalue of the structure tensor on the input vessel map (see "Detection and tracking of junctions" at page 96). Junctions are then found by segmentation on the junction map, using, e.g., the Otsu threshold. A flow-chart of the algorithm is shown in Figure 4.11. We have published this algorithm, together with the MRF-based adaptive threshold, in [53].

The strategy used to achieve the final segmentation can be summarized like this:

1 Find a first set of points such that at least a majority of them are true vessels, even if there is a large number of false vessels.

2 Find a second set of points such that all of them are true vessels even if not all vessel points are selected.

3 Choose – based on the connectivity of vessels – from among the points in the first set only those connected to a point in the second set.

4 Return the result of this selection as the final segmentation result.

For such a procedure to work, one has to take extra care that true vessels are not connected to false vessels. To achieve a perfect segmentation one has to ensure that all vessels are selected into the first set and that at least one point on each vessel chunk is selected into the second set.

Using the connectivity, one can decide correctly even in the case of a vessel-pixel for which the probability of its gray-level given class background is larger than the probability given class vessel.

This type of reasoning constitutes the foundation on which the hysteresis classification paradigm is introduced in the next section.

4.3.2   Hysteresis segmentation

A hysteresis classifier represents a pair of classifiers which are *coupled* by a relation which permits to use the results of one classifier to condition the results of the other, such that a better classification performance is obtained. Each of the classifiers of the pair is called a base-classifier. If the two base-classifiers are linear then a linear hysteresis classifier is obtained.

The simplest form of a hysteresis classifier working in a scalar feature space is a bi-threshold procedure called hysteresis thresholding [34], [151], [43]. The hysteresis threshold gained recognition through the seminal work of Canny [34] on edge detection. The problematic of segmenting edges and vessels is relatively similar: both weak edges and small, weak-contrast-vessels are difficult to segment, both edges and vessels are coupled through their connectivity. Therefore, the hysteresis threshold has been used also for vessel segmentation, as proposed almost simultaneously in [43] and also in [151].

Currently, irrespective of the application, the two thresholds are established rather heuristically. Canny [34] recommends that their ratio should be between two and three. In [151] they are chosen in an application-dependent way. We introduce here a set of rules which, in our opinion, permit the computation of the *hysteresis threshold* on any data from any application – as long as the *hysteresis paradigm* can be applied – both in an unsupervised [43] and in a supervised manner. We show also how to apply the hysteresis paradigm to multidimensional inputs in a supervised manner, obtaining thus the *supervised hysteresis classifier* [44].

Within the frame of this work we give only examples from one application: vessel segmentation. For such an application we show also how to choose the parameters of a *vessel map* such that it is optimally *suited for hysteresis classification*.

From vessel markers to hysteresis classification

The hysteresis segmentation of vessels and then the concept of hysteresis classification which we will introduce here, evolved from our research on automatic vessel segmentation, in particular from the junction-based selection of vessels. For clarity of presentation we summarize this path here.

In the case of vessel segmentation the two class-conditional gray-level *pdfs* corresponding to vessels and background, usually exhibit a large overlap and consequently correct results are difficult to achieve. Prior knowledge can be successfully used to improve the classification performance. For example, it is known that vessels are connected structures, then, if one finds some points which belong to vessels with a high probability – i.e. *vessel markers* – one can safely assume that all other potential vessel-points in their vicinity – as selected by an imperfect segmentation algorithm – are also true vessels and so on until the vicinity is empty and all vessel points chained to the vessel marker have been selected. Clearly this selection makes sense only if the imperfect segmentation is actually an over-segmentation, with all vessels but also many false-positives which are not connected to vessels.

Therefore, in hysteresis segmentation we use *two classifiers* to correctly segment vessels: (i) one to find vessel markers and (ii) one to find all vessels. They are then *coupled over the vessel connectivity* such that together they return an improved vessel segmentation.

The search for vessel markers is based again on prior knowledge. In Section 4.3.1 we have used the fact that *vessels show junctions*, but one can also use the fact that *the largest connected structures* in an imperfect segmentation result are vessels or that the *points of highest contrast* in the analyzed images belong to vessels.

Junction-based selection of vessels. It is known that vessels build junctions, while other structures present in the analyzed images usually do not. Hence, one can use junctions as vessel markers. However, junction-detection in vessel images is a rather complicated procedure due to the particular multiscale structure of vessels.

Junction detection by segmentation on the tensor-eigenvalue-based junction map does not function correctly for dark blob-like structures appearing in the input image. The value of the second eigenvalue of the structure tensor for blobs is similar to that computed for junctions. However, an enhanced orientation analysis [10] can potentially solve such problems at the expense of a longer computation time. We have achieved also good results by dividing the second eigenvalue of the structure tensor by the the first one to compute the junction map.

Size-based selection of vessels. In the vessel map all other structures, with the exception of vessels, are attenuated. Therefore, one could assume that an over-segmentation will show only vessels and noise and therefore the largest connected structures will be vessels. Then, one can use as vessel markers some points belonging to the largest vessels and find the true vessels by morphologically processing the over-segmentation result. In this case we apply opening by reconstruction [179] with an isotropic structuring element.

During opening by reconstruction, the binary result is eroded until all structures other than vessels are eliminated and then the remaining points – i.e. vessel markers – are used to choose from the initial result, again over the vessel connectivity, the true vessels. Such an approach functions only as long as after erosion only vessels remain, which means – assuming an isotropic structuring element – that the non-vessel structures should all be smaller than the largest vessel.

Because the over-segmentation should select all vessels – i.e. including the small ones with gray-levels similar to that of the background – there are usually also falsely segmented background structures present comparable in size or even larger than the vessels and thus the morphological processing may fail.

Contrast-based selection of vessels. The contrast of vessels depends on their size (see Section 4.2.2, page 122), and the vessel map is computed such that only two homogeneous pixel-classes are present: the dark vessels and the brighter background. Therefore, one can use as vessel-markers the points with the smallest gray-levels. In comparison to opening by reconstruction, it is the vessel contrast – enhanced in the vessel map – which is used to segment vessels and not the vessel size. Thus we eliminate the influence of large falsely segmented background structures on the computation of vessel markers.

Consequently, one needs two thresholds together with the connectivity constraint and the requirement that falsely-segmented background and vessels are not connected to one another to

achieve a successful segmentation of vessels. Also, part of the vessel tree should appear with a higher contrast, such that vessel markers can be segmented with a virtually zero fp rate.

We consider that the contrast-based selection of vessels is the optimal selection modality. The hysteresis threshold used later in this chapters is of this type.

The hysteresis classification paradigm.

We can now state the hysteresis classification paradigm. In a binary classification problem, if the two classes are not linearly separable in the feature space and under the assumption that no new features can be added, nonlinear techniques need to be used to obtain satisfactory results. Such nonlinear techniques are typically expensive in training time and can easily overfit. However, if the components of one of the classes (object class) exhibit certain connectivity in some additional space different from the feature space, while at the same time vectors of the two classes which are close in this space are different enough in the feature space, then two linear classifiers working in the original feature space suffice to achieve a good classification. One of these classifiers, the "pessimist" should select only object points, having thus a zero false positives rate and the other one, the "optimist" should select all object points irrespective of the number of false positives. We make here the realistic assumption that neither class has an infinite support. Then, based on the connectivity property from the additional space, the points which with a high confidence belong to the object class and which were selected by the "pessimist" can be used to choose true object points from among those selected by the "optimist" as long as they are linked with one another over a vicinity [34], [131]. A flow chart of a hysteresis classification algorithm is shown in Figure 4.12.

The pair of linear classifiers coupled over the connectivity constraint builds then a linear hysteresis classifier. Here we concentrate on linear classifiers, but other classifiers can be used also as long as they are trained to fit the role of a "pessimist" and an "optimist", for example by training on specially chosen sets. Of course if the connectivity can be described by a feature or set of features, these can then be added to the original feature space and a standard classifier can be used.

If the parameters of the hysteresis classifier are computed from a labeled training set, then the classifier is called supervised, otherwise it is called unsupervised. In the former case, it is different from a classifier committee [72], [71], [86], [133] as the two classifiers are coupled over the connectivity of the analyzed data rather than over a certain combining rule and the training takes place over the entire data-set for both classifiers. In the form in which it is introduced here it can be used mainly to segmentation or to two-class classification. One simple modality to extend it to multiple-class classification is to train disjunct classifiers for each class. Application of the hysteresis classification paradigm to scalar inputs yields a hysteresis threshold, while for vectorial inputs a bi-class hysteresis classifier is obtained.

Hysteresis thresholding

The practical implementation of the hysteresis paradigm for scalar inputs yields a set of parametric rules used to compute two thresholds, which taken together constitute the hysteresis threshold. As vessels are darker than background, one threshold with a small value – the "pessimistic" – will select vessel markers and the second, with a larger value – the "optimistic" – will return an over-segmentation containing all vessels. Thus, pixels whose feature value is
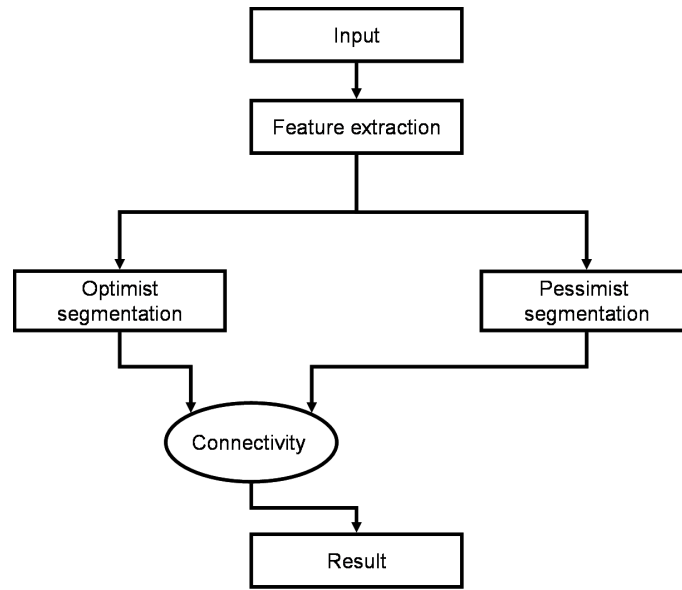
Figure 4.12: Flow chart of a hysteresis classification algorithm.

below the low threshold are most probably vessels and are kept and pixels with a feature value above the high threshold are most probably non-vessels and are eliminated. Those in between the thresholds are vessels only if they are connected to a certain vessel point.

The parameters of the hysteresis threshold are either specified by the user, based on the prior information usually available about the analyzed data, or learned from examples. In the case of the *unsupervised hysteresis threshold*, the parameters are specified by the user. For the *supervised hysteresis threshold*, we show how to compute these parameters by learning from a set of labeled examples.

Unsupervised hysteresis threshold. We propose *hypothesis testing* as a general procedure to compute the two thresholds and show how this can be translated to a simple and *fast percentile-based threshold-selection rule* [43]. This is then an unsupervised a priori procedure as the parameters of the hysteresis threshold are not learned from a set of examples but are specified a priori by the user.

Hypothesis testing. The "pessimistic" threshold can be set using the background's class conditional *pdf* by means of a significance test. The null hypothesis is that the pixel under investigation belongs to the background class. To obtain the threshold we choose a small significance level. We have then:

$$\int_{-\infty}^{t_p} p(x|\omega_b) = \alpha \tag{4.6}$$

with t_p the threshold, ω_b the background pixel-class, x the pixel-gray-level and α the significance.

If the gray level of an investigated pixel is below this threshold, then the probability of it belonging to the class background is very small (i.e. less than the significance), which means that the available data does not sustain the null hypothesis in that case. As on the vessel map we have either background or vessel pixels, then the pixel under consideration must belong to the

vessel class. The error probability of this decision (i.e. the type I error probability) equals the significance level. The "pessimist" threshold separates thus high confidence vessel pixels from the rest.

The "optimistic" threshold can be set using the vessel class conditional *pdf*. This time we hypothesize that the pixel under consideration is a vessel pixel. To compute the threshold, we impose that the error probability for a decision in agreement with our hypothesis – i.e. the probability for a decision against our hypothesis, allthough it is true and we should agree – is very small:

$$\int_{t_o}^{\infty} p(x|\omega_v) = \beta \tag{4.7}$$

with t_o the threshold, ω_v the vessel pixel-class, x the pixel-gray-level and β the probability of error.

If the gray level of the investigated pixel is above this threshold, this means that the probability of it being a vessel is very small and again, as there are only two classes on the vessel map, it must belong to the background class. Thus, the "optimistic" threshold separates high confidence background pixels from the rest.

The two parameters of the unsupervised hysteresis threshold are then the two error probabilities α and β. They should be chosen empirically and have a small value – e.g. in the interval $(10^{-4}, 10^{-2})$.

– **Estimation of the class conditional *pdfs*.**   If some samples from each density are known, then the class conditional *pdfs*, needed for the thresholds' computation, can be estimated. We find these samples by means of the Otsu threshold. However, any other similar algorithm or even a percentile threshold could be used. Of concern here is that only a majority of samples do belong to the respective class and not all of them.

If we assume some generic form of the distribution, its parameters can be determined using the ML or the MAP estimation. Such a parametric estimation has some disadvantages: (i) it is considered that the random variable can take values in an infinite interval and then the computed thresholds may fall outside the vessel map's range, (ii) assuming a generic distribution form limits the flexibility of our approach providing thus an additional source for errors if the assumption is incorrect or if the samples used at estimation were not uniformly extracted. If no assumptions are made with respect to the distribution's generic form, then it can be found, e.g., using Parzen estimation [184]. The *pdf* at position $\hat{y}$ is then computed as:

$$\tilde{p}(\hat{y}|\omega) = \frac{1}{N} \sum_{i=1}^{N_\omega} \gamma(\hat{y} - y_i) \tag{4.8}$$

with N_ω the number of available samples y_i for class ω and γ the kernel potential function or Parzen window. A function should satisfy certain conditions to be a kernel. Two kernels are usually used for estimation: (i) the Gaussian kernel:

$$\gamma_g(z) = \frac{1}{\sqrt{2\pi}h} e^{\left(-\frac{z^2}{2h^2}\right)} \tag{4.9}$$

or (ii) the exponential kernel:

$$\gamma_e(z) = \frac{1}{2h} e^{\left(-\frac{|z|}{h}\right)} \tag{4.10}$$

where h is the size of an interval with $\hat{y}$ as midpoint.

This nonparametric estimation procedure is more flexible than the parametric one and the thresholds remain practically always in the vessel map range, if the parameter h is properly chosen (empirically $h < 5$). However, the better performance is balanced by a severe increase in the time needed to compute the thresholds.

Percentile-based selection of thresholds. When the thresholds are set by hypothesis testing, they are chosen such that the probability of a certain event is very small, i.e., at most equal to the significance.

For the "pessimistic" threshold we have : $P(x < t_p) = \alpha$. Previously we have considered x to be a continuous variable, but actually with respect to a vessel map x is a discrete variable and then Equation 4.6 can be approximated as:

$$\sum_{i=v_{b\,min}}^{t_p} \frac{n_{bi}}{N_b} = \alpha \tag{4.11}$$

with v_{bmin} the minimum gray-level on the histogram of the background gray-levels, n_{bi} the number of background pixels with gray-level i and N_b the total number of background pixels in the image. t_p is then the α'th percentile of the histogram of the background gray-levels.

The histogram of the vessel map is the approximation of the mixture of vessel and background class conditional *pdfs*. Therefore, t_p is also a percentile of the histogram of the vessel map and can be found using:

$$\sum_{i=v_{min}}^{t_p} \frac{n_i}{N} = \alpha_{vm} \tag{4.12}$$

with v_{min} the minimum gray-level on the histogram of the vessel map, n_i the number of pixels with gray-level i and N the total number of pixels in the image. t_p is then the α_{vm}'th percentile of the histogram of the vessel map and it should be chosen such that it selects *only* vessel pixels.

Proceeding similarly, one can show that the probability of error β from Equation 4.7 is related to the β_{vm}'th percentile of the histogram of the vessel map – i.e. t_o – which selects *all* vessel pixels.

Consequently, one can use prior information about the approximate surface covered by vessels in the analyzed angiograms to define the two thresholds by means of percentiles of the vessel-map histogram. This procedure is illustrated in Figure 4.13.

As vessels are always darker than background, the percentile defining the "pessimistic" threshold t_p is rather small and is related to the minimal image area which is *certainly* occupied by vessels – this is the black area under the curve in Figure 4.13. Theoretically this should be the largest percentile selecting *only* vessels.

The percentile defining the "optimistic" threshold t_o should be choose as $\beta_{vm} = 100 - V$ where V defines the largest percentage of the vessel-map surface which can be covered by vessels and is thus related to the minimal image area which is *certainly* occupied by background – this is the dashed area under the curve in Figure 4.13. Theoretically this should be the smallest percentile selecting *all* vessels.

The two parameters of the unsupervised hysteresis threshold are then the minimal image area certainly occupied by vessels, which yields α_{vm} and the minimal image area certainly occupied by background, which yields β_{vm}. Practically, it is sufficient that α_{vm} is chosen such
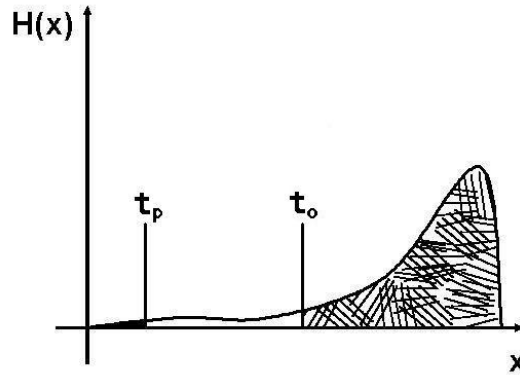
Figure 4.13: Schematic representation of the vessel-map histogram with the "pessimistic" and "optimistic" thresholds, t_p and t_o respectively.

that *only* vessels are selected and β_{vm} such that *all* vessels – but without including more than half of the background – are selected.

Supervised hysteresis threshold.   The parameters of the practical percentile-based hysteresis threshold are the percentages of images surface covered only by vessels and background respectively, which can then be directly converted to percentiles of the histogram of each image. If a set of labeled vessel-images is available, one can then determine these parameters rather than impose them – i.e. train the hysteresis threshold. We introduce two main training modalities: *sequential training*, where each image in the training-set is introduced separately and *batch training*, where all images in the training-set are introduced together. Both use ROC-based analysis [25], [76].

   Sequential training.   We describe two types of sequential training which we call: *indirect* and *direct*. In the case of *indirect sequential training* we proceed the following way:

- For each possible "pessimist" try all possible "optimists" and each time build the ROC. Each point on the curve represents the mean percentage of true and false positives over all images in the training set. Proceeding e.g. in unit steps, this yields 100 classifiers and accordingly 100 ROCs.

- Choose as "pessimist" the one which yields the largest AROC.

- Choose the "optimist" on the ROC from the previous step as the percentile which gives the point on the ROC closest to 0% fp and 100% cc.

   Choosing the "optimist" as the one which gives the closest point on the ROC to 0% fp and 100% cc may return an unacceptably high false positives rate, thus it is better to define it over a maximum false positives rate. Good results have been achieved with $T_o = arg \max(cc|fp < \alpha\%)$ with α as close as possible to 0%.

   As the two classifiers are actually percentiles, they can also be computed from a percentile-ROC. We call this *direct sequential training*. In this case we proceed the following way:

- For each percentile threshold, record in unit steps the mean rates of cc and fp over all images of the data set.

- Determine the two components of a supervised hysteresis threshold by choosing the values for which $\max(cc|fp = \alpha_p)$ and $\min(fp|cc = \alpha_o)$ for the "pessimist" and "optimist" respectively.

These could then be used to define the image areas covered only by vessels and background. α_p has to be chosen as close as reasonably possible to 0%, and α_o as close as reasonably possible to 100%.

The training methods described until now can be applied to compute the parameters of an unsupervised hysteresis threshold. Such a training procedure will clearly take a long time to compute, as each image is introduced separately and for many times, first to compute the "pessimist" and then to compute the "optimist".

Batch training. One can use also faster training procedures. In this case, first a training matrix has to be computed. The first row of this matrix contains the class labels for each pixel from each image in the training set and the second row contains the corresponding gray-levels from the vessel maps. Using these gray levels, a global histogram for the entire training set can be computed. The training algorithm will have the following steps:

- Use the training matrix to build the ROC of a global percentile-threshold for the entire training space.

- Choose the "pessimist" threshold T_p such that $T_p = arg \max(cc|fp = \alpha_p)$ and the optimist threshold T_o such that $T_o = arg \min(fp|cc = \alpha_o)$ with α_p as close as reasonably possible to 0%, and α_o as close as reasonably possible to 100%.

Such a segmentation strategy has to be carefully used. It implicitly assumes that the gray-levels characteristic to vessels are very similar in each image of the data-set – i.e. vessels are homogeneously represented not only within each image but also between images – but it does nothing to enforce this assumption. Thus, it may happen, e.g., that the same gray-level that in a certain well-contrasted image selects some vessel-points, selects none in another image. Conversely, to obtain also vessel markers in the latter image, one should choose a value which will include also background among vessels in the former one. This represents a justification for choosing α_p as close as reasonably possible to 0%. Such a training procedure is very fast.

The supervised hysteresis classifier

In the case of vessel segmentation, a multidimensional feature space can be obtained by describing each pixel with a vector whose components are the corresponding gray-levels in several different vessel maps (see Section 4.2.3).

To build a hysteresis classifier [44] we propose to use as "optimist" and "pessimist" two *Fisher linear classifiers* [184] respectively. In terms of the parameters of each Fisher base-classifier – i.e. the *parameters of the hysteresis classifier* – the set of weights is the same for both, but the two thresholds differ and they are established by training. We show also how to *select the optimal set of features* for such a hysteresis classifier.

The Fisher classifier.   The parameters of the Fisher classifier are a set of weights $\vec{w}$, defining a transformation from the original input feature space to a 1D space where the two classes are optimally separated and a threshold T, to discriminate between the two classes. In the transformed space the homogeneity of the representation of the two classes is improved over all images in the training-set. The separability criterion is:

$$F = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} \tag{4.13}$$

where $\mu_{1,2}$ and $\sigma_{1,2}$ are the means and variances in the transformed space respectively. Then the weights vector is:

$$\vec{w} = (\vec{m}_1 - \vec{m}_2)^T \Big[\frac{1}{2}(\mathbf{K_1} + \mathbf{K_2})\Big]^{-1} \tag{4.14}$$

where $\vec{m}_{1,2}$ and $\mathbf{K_{1,2}}$ are the means and covariance matrices in the input feature space respectively. Thus the classifier has the form:

$$\vec{w}^T \vec{y} \underset{\omega_2}{\overset{\omega_1}{\underset{<}{\overset{>}{}}}} T \tag{4.15}$$

with $\vec{y}$ the input feature vector. T can then be determined in terms of the statistical decision theory. If the left side is larger than the threshold we decide for the class ω_1, conversely for ω_2.

The parameters of the hysteresis classifier.   For supervised classification with hysteresis, the weights $\vec{w}$ are the same for both "pessimist" and "optimist" but the thresholds are different. They are computed by training, using the ROC of the decision rule (Equation 4.15).

Establishing the thresholds by batch-training.   In this case the training matrix is built having at the first row the class labels and on the next rows the gray-level values in different vessel maps for each pixel from each image in the training-set.

The thresholds represent values in the transformed space and not percentages of image area covered only by vessels or background. The training algorithm will have the following steps:

- Use the training matrix to build the ROC of the decision rule for the entire training space.

- Choose the "pessimist" threshold T_p such that $T_p = arg\max(cc|fp = \alpha_p)$ and the optimist threshold T_o such that $T_o = arg\min(fp|cc = \alpha_o)$ with α_p as close as reasonably possible to 0%, and α_o as close as reasonably possible to 100%.

Sometimes in practice, e.g., for a sub-optimal feature space, the rule by which T_o is selected leads to over-segmentation due to the large fp rate associated with high cc percentages. In such cases it is better to define the "optimistic" classifier over a maximum fp rate. Good experimental results were obtained by choosing $T_o = arg\max(cc|fp < 15\%)$. For a higher fp rate, misclassified points connected to the object points appear and the overall results worsen.

Establishing the thresholds by iterative training   The parameters α_p and α_o and implicitly the thresholds can be imposed by the user, based on some prior-knowledge, or they can be *trained* the following way: starting from a standard Fisher classifier – with T corresponding to the ROC-point most distant to the base-line linking the ROC-points with $fp = 0\%$ and $fp = 100\%$ – used as "optimist", try all "pessimists" – building thus a hysteresis-ROC – and choose again the one corresponding to the point most distant to the base-line. Repeat, this time keeping the "pessimist" constant and continue then for a predetermined number of steps or until the thresholds do not change anymore.

Feature selection.   We use the ROC to characterize the feature space. Clearly, the larger AROC, the more separable the feature space. To the limit, the two classes are linearly separable when AROC is one, i.e., 100% cc for 0% fp. If several features have been computed, then those which build the best feature space will also yield the largest AROC. In the case of hysteresis classification we are interested also in the derivative of the ROC curve, especially in the region where the fp rate is small. A hysteresis classifier trained on a ROC curve with a large integral but a comparatively mild increase over the region with small fp rate, will yield rather poor results because the "pessimist" classifier will select too few true vessels. Thus, it is better to consider only a partial AROC, i.e., the pAROC. In our experiments we have computed only the AROC bounded by a 30% fp rate.

Two strategies can then be followed: a full search strategy when all possible combinations of features are investigated and a sequential search strategy [160] when first the single best feature is selected then the best combination between that feature and another one and so on until the optimal feature set is found.

Practical estimation of the parameters of the vessel map for hysteresis segmentation

In Section 4.2 it has been pointed out that the optimal parameters of each enhancement method can be established in a target-oriented manner during a calibration step. We have already given practical rules to set the parameters without any calibration, however, if some hand-labeled examples are available, the optimal parameters can be established with their help.

We have used such data to compute the AROC of a percentile threshold, to *show the quality of a vessel map* (see Section 4.2.4). At a first glance, the same measure could be used to establish the parameters of a vessel map as those for which the mean AROC[6] over all test images is the largest. However, this will yield optimal vessel map parameters with respect to a linear separation between classes and may not be the best solution when hysteresis classification methods are used.

At the heart of the hysteresis classifier is the connectivity which object-points exhibit in some additional space. This information is used to achieve a correct segmentation in the feature space despite the fact that the data is not linearly separable. At the same time, it is important that in the feature space, object points situated on the border and non-object-points are well separable, otherwise the connectivity constraint will bring false object-points in the segmentation. This should be enforced particularly during feature extraction – i.e. the computation of the vessel map – but also during feature selection in the case of vectorial inputs. Here we de-

---

[6]Most of these quality criteria are based on the AROC of a percentile threshold, which we call simply AROC. When the AROC of another segmentation method is discussed we will make this clear in the text.

scribe *quality criteria* which permit the computation of a *vessel map well suited for hysteresis segmentation* of the vessels.

The AROC as quality criterion.   The AROC will favor blurry vessel maps, where only the large vessels are well observable and the background is well suppressed. There are two reasons for this: (i) the background suppression does not eliminate properly small vessel-like artefacts and (ii) the ground truth is usually flawed. The main reason why the ground truth is flawed, is that the human expert usually goes a little over the actual vessel-border during the segmentation. He either sees the vessel thicker than it really is or he deviates slightly to the left or right from the path of the vessel (observe Figure 4.14 (b) and (d)).

The vessel-like artefacts will count as false positives and therefore, to obtain a small fp rate, one would favor vessel maps where they are suppressed. At the same time, if only the true vessel is segmented, the pixels of the difference-area between the segmented vessel and the flawed ground truth will count as misclassified true vessels and therefore decrease the measurable cc rate. To improve this rate, one would favor vessel maps yielding results where these pixels are also vessels. In both cases, i.e., small vessel-like artefacts and flawed ground truth, such vessel maps as those favored, are achieved after stronger low-pass filtering, and they are therefore "blurry". Strong low-pass filtering eliminates background artefacts but at the same time eliminates also weak-contrasted vessels and "spreads" vessels of larger diameter and higher contrast.

The AROC can be successfully used to show the enhancement potential of a vessel map. Because it favors an enhancement result where background is strongly suppressed, the AROC can be also used to compute the parameters of a vessel map where a linear segmentation algorithm, like e.g., a fixed threshold, will be applied. However, the AROC is unsuited to choose the optimal parameter-set for a vessel map used with a hysteresis classification algorithm. This happens because a hysteresis algorithm can segment weak vessels despite background artefacts and in a "blurry" vessel map the weak vessels are attenuated.

This is shown by the example in Figure 4.14. The vessel maps are computed using the Laplacian pyramid. For the "blurry" vessel map in Figure 4.14 (c) the AROC = 93.90 and for the "sharp" vessel map in Figure 4.14 (d), the AROC = 81.65. The two segmentations in Figure 4.14 (e) and (f) respectively are computed using the 12-percentile of the corresponding vessel map, which empirically proved to be the best choice of a fixed threshold in this case. The correct classification and false-positives rates are [68.57, 3.71] and [59.9, 4.98] respectively. Comparing (e) with the ground truth (b) it is clear that many weak vessels have not been segmented and larger vessels are "spread". Conversely, comparing (e) with (f) there is less falsely segmented background.

To show the link between the imperfect ground truth and the "blurry" vessel map obtained with the AROC, we added the result achieved on this vessel map with the ground truth. This is shown in Figure 4.14 (g). For comparison, we show also the sum between the ground truth and the segmentation result achieved on the "sharp" vessel map in Figure 4.14 (h). One can clearly see how the white areas, which signify, according to the ground truth, correct decisions, increase for the "blurry" result, particularly in the vicinity of large vessels.

Quality measures to compute vessel maps for hysteresis segmentation.   For hysteresis classification, a "good" vessel map means a vessel map where all vessels can be segmented by a fixed threshold – irrespective of how many vessel-artefacts are selected together with the vessels

(a) Original image

(b) Ground truth

(c) 'Blurry' vessel map

(d) Vessel map

(e) 'Blurry' vessel map segmentation

(f) Vessel map segmentation

(g) 'Blurry' segmentation & ground-truth

(h) Segmentation & ground-truth
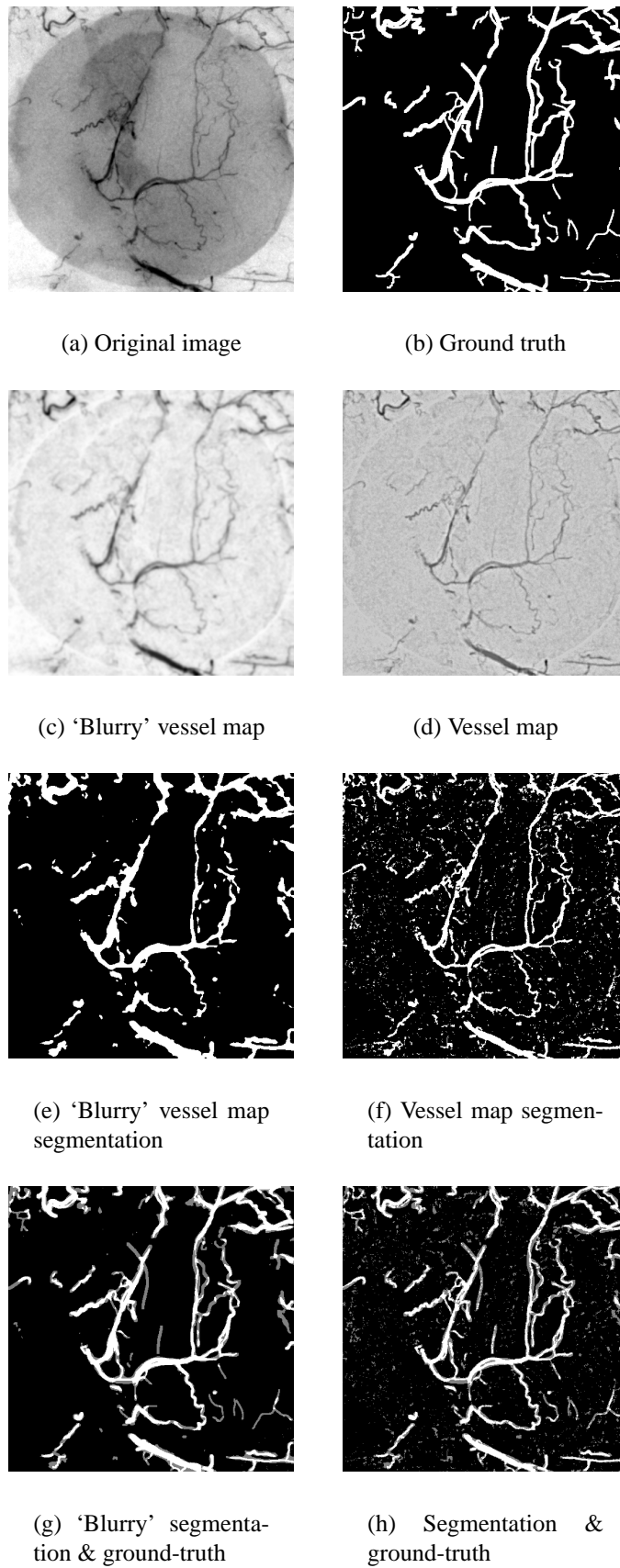
Figure 4.14: Comparison between a 'blurry' vessel map computed using the AROC quality criteria and a "sharp" vessel map, computed with quality criteria designed specifically for hysteresis vessel segmentation.

(a)   (b)   (c)
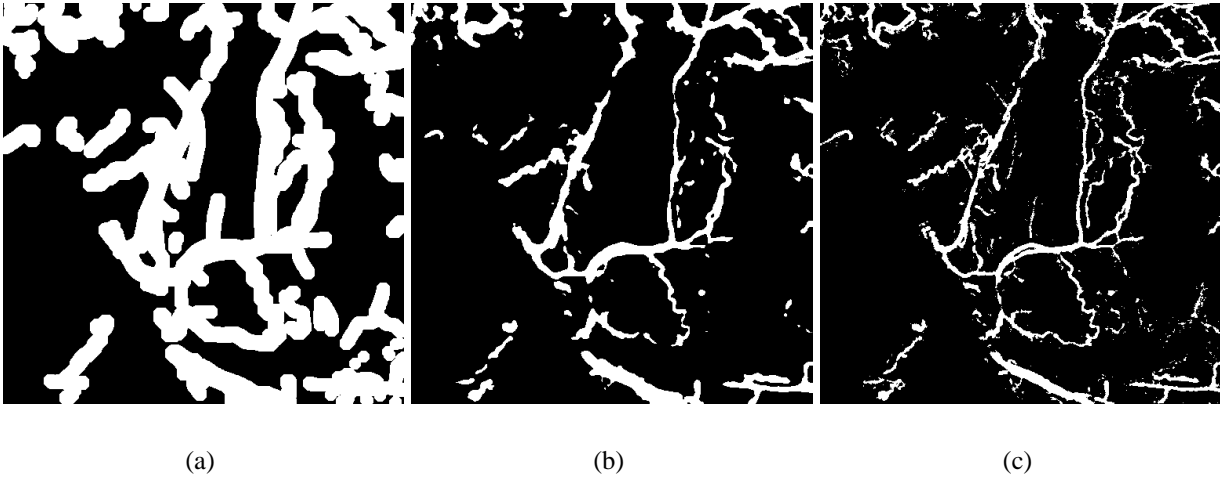
Figure 4.15: Computation of the background-less AROC by the mask method: dilated manual ground truth mask (a) background-less segmentation on the "blurry" vessel map (b) and on the "sharp" vessel map (c) after pixelwise multiplication with the mask.

– and vessels are well separated at the margins from non-vessels. Thus, a measure is needed that "ignores" the background and weights more the errors in the vicinity of vessels.

The background-less AROC.   One possible solution is to compute another type of AROC – we call this the background-less AROC. This time the correct classifications and false positives are no longer estimated over the entire vessel map but only in the direct vicinity of the vessels. We propose two methods to achieve this: the mask method and the vessel marker method. The first one selects the vessels and a small vicinity around them by multiplying pixelwise each percentile result by a dilated ground-truth mask. Then, estimates the AROC using as input this modified segmentation result. This is shown in Figure 4.15. The second one uses the ground truth vessel pixels as vessel markers to select from the percentile result only vessel-class pixels. Then, estimates the AROC using as input this modified segmentation result. This is shown in Figure 4.16.

Thus, most background structures are ignored when computing the AROC. Background structures linked to the vessel pixels, which may appear in the segmentation, have a stronger influence on the vessel marker method than on the mask method. Theoretically, such a measure – particularly when computed by the vessel marker method – should favor that parameter set which defines a vessel map such that as many vessels as possible are selected, irrespective of the background, as long as miss-segmented background is not linked to the vessels. Therefore, it returns vessel maps well suited for hysteresis segmentation.

The background-less pAROC.   Still we have the problem of the flawed ground truth. Therefore we propose to use as measure the area under the background-less percentile-AROC bounded by a reasonably small false-positives rate, assuming that for "good" vessel maps the corresponding ROC has a sharper increase in the region of small false-positive rates. We have obtained good results by setting this bound at 2%. One could also use the mean false-positives rate over all percentile results used to compute the ROC, but this measure is less sensible. Us-

(a) (b)
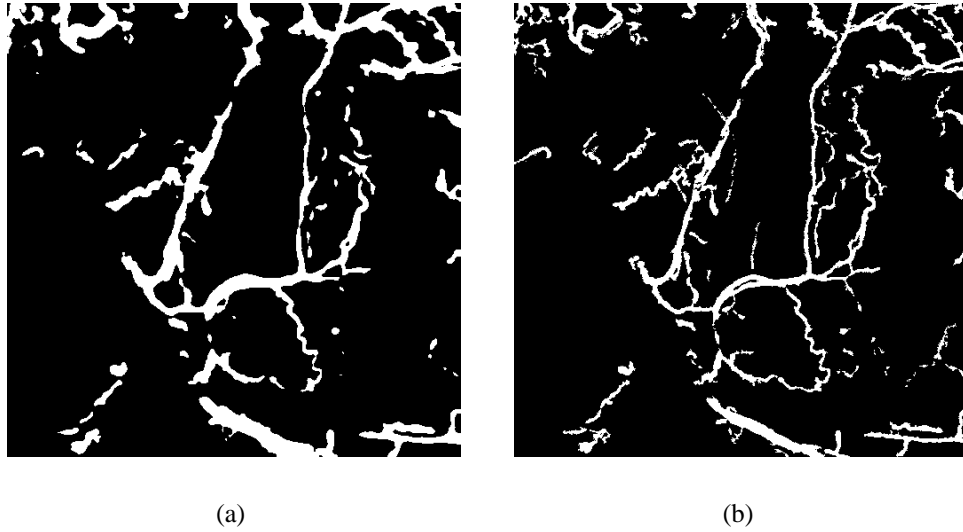
Figure 4.16: Computation of the background-less AROC by the vessel marker method: Background-less segmentation on the "blurry" vessel map (a) and on the "sharp" vessel map (b) after selection using the manual ground truth.

ing the 2% bound, the value of the measure is pAROC = 1.02 for the "blurry" vessel map and pAROC = 1.11 for the "sharp" vessel map in Figure 4.14 (c) and (d) respectively.

Vesselness measures. Alternatively, one can use a *combined vesselness measure* to find the optimal parameter-set.

The vesselness is measured by the length of the segmented vessels in the background-less percentile segmentation. Such a measure will favor vessel maps where the small vessels are well segmented. To ensure that the large vessels – with a large surface and comparably small length – are also well segmented, we combine vesselness with cc and fp rates, to obtain a combined vesselness measure.

We propose to compute the combined vesselness measure as the difference between the correct classification rate and the false positives rate for that percentile result where the vesselness is similar to that of the ground truth.

Establishing the set of parameters. Using such measures, the parameter-set can be sought: exhaustively, sequentially or iteratively.

An exhaustive search would try all combinations of parameters between certain limits and with appropriate increment-steps and then choose the one for which the quality measure reaches an optimum.

A sequential search will start from a certain initial parameter set and improve it. This is done by searching the best value – as indicated by the quality measure – of the first parameter, while keeping the others constant. This is then repeated for the next parameter starting from the same initial set and so on until the set is exhausted.

In the case of the iterative search, the parameters are established similar to the sequential search, but this time, instead of starting the search for a new parameter from the same initial set, one starts from the set including the already found optima. The procedure is then repeated with

the intermediary optimal set as initialization, until the parameter set remains unchanged for two iterations.

We have obtain good results using the pAROC of the background-less percentile-ROC and a sequential search. The optimization of the parameter-set using the combined vesselness measure is time intensive.

Outlook

Starting from the hysteresis paradigm, we have introduced several classification algorithms. We have described an unsupervised a priori hysteresis threshold, a supervised hysteresis threshold and a supervised hysteresis classifier.

All these algorithms have been introduced for binary classification problems. We believe however that in particular the hysteresis classifier can be used to separate *multiple classes* as well.

The MRF-based adaptive threshold together with the junction-based selection of vessels represents actually a type of unsupervised learning hysteresis threshold. However, as opposed to all other hysteresis classification methods introduced until now, this one is heterogeneous, as the two classifiers differ with respect to the input feature space: one segments vessels and the other one junctions. We believe that a *homogeneous unsupervised learning hysteresis threshold* where both thresholds have the same application domain can be developed as well.

Extension to multiple classes.   The hysteresis classifier was introduced here for binary classification. A direct extension to multiple classes is rather difficult. Although it is tempting to train several disjunct hysteresis classifiers – i.e. one for each class – such a scheme would then reach its limitation when more than one of the classifiers responds to an input. In this case, one needs to develop modalities to establish which classifier is more correct.

The main idea of the hysteresis classifier is to use some sort of connectivity between members of a certain class to extract little by little the class-components – even if they are rather different from the class-prototype – as long as they are connected to one-another and connected also to at least one component of a sub-set of the class, which is this time very similar to the class-prototype or which is known a priori to belong to the class.

An analogy can be made with the k-NN classifier. There, labeled-samples belonging to the training set are used to decide on a new sample. The connectivity is defined by the number of class-components found in hypersphere of radius R centered at the new sample. Usually R is defined such that it includes an odd number of labeled-neighbors. In this case it can be considered that the training labeled samples are the results returned by the "pessimist" and the new sample is one of the results returned by the "optimist". If the new sample is connected to a class ω_i – i.e. a majority of labeled-neighbors belongs to ω_i– then the decision on the new sample is taken in favor of class ω_i even if the prototype of class ω_i is more distant than that of say class ω_j. At this point the analogy ceases as for a hysteresis classifier, the now-labeled sample should be received into the training-set to help decide for other samples falling in that region of the feature-space. Different pruning schemes can however by devised to eliminate, e.g., training-set samples which had no contribution for the classification of the last k new samples, such that the number of components of the training set does not grow indefinitely. The eliminated samples of the original training-set can be reclassified at a later moment if one does not want the classes to start moving around the feature-space.

A homogeneous unsupervised learning hysteresis classifier. To develop a homogeneous[7] unsupervised learning hysteresis classifier for vessel segmentation, we propose to use the EM algorithm to find the parameters of a mixture with four modes on the histogram of a vessel map. We use one mixture component per vessel category and one component for the background. Then, the "pessimist" threshold can be chosen to minimize the probability of error between the mixture component with the smallest mean – corresponding to the class of the largest vessels – and the rest of the mixture and the "optimist" threshold can be chosen to minimize the probability of error between mixture component with the largest mean – corresponding to the background – and the rest of the mixture.

4.3.3 Automatic vessel segmentation: Experiments and results

The automatic segmentation methods proposed here have been tested using three different types of vessel images for which the hand-labeled ground truth was available: 12 skin-transplant microangiograms, 11 coronary angiograms recorded from different patients and under different projections and 40 retina fundus photographies. To establish the ground truth, the microangiograms and the coronary angiograms have been labeled manually by the author with expert help. The retina images belong to a publicly available data-base [181]. In this case, according to the authors, the ground truth has been established by experts. If needed, these data sets have been divided into a train and a test-set respectively.

We have dichotomized the methods proposed in *supervised* and *unsupervised* and experimented with them separately.

Unsupervised segmentation

The unsupervised methods proposed here for vessel-segmentation are: the MRF-based adaptive threshold with junction-based selection of true vessels and the unsupervised hysteresis threshold.

Input. The algorithms have been all tested on the vessel map obtained by multiscale Hessian analysis, after attenuating the background by the Bothat transform.

Usually, if an unsupervised method is used, there is no prelabeled data set available to permit the optimal choosing of the parameters set of the vessel map. However, here we have used the available hand-labeled images to compute the parameters (see "The practical estimation of the parameters of a vessel map", at page 145 in Section 4.3.2) as a confirmation of the empirical rules proposed in Section 4.2.

For background attenuation, for the microangiograms a multiplicative imaging model was considered. The size of the transformation window was: 13 for microangiograms, 21 for coronary angiograms and nine for retina images.

For vessel augmentation by multiscale Hessian analysis, for the coronary angiograms the weights are: 0.7 for the highest level of the pyramid – containing small vessels and noise, 2.1 for the mid-level – containing mid and small vessels and 0.8 for the lowest level containing large and mid-vessels. For the skin microangiograms the weights are: 1.3 for the highest level of the pyramid, 1.7 for the mid-level and 0.8 for the lowest level. For the retina images the weights are:

---

[7]See also "Connection to the hysteresis threshold" at page 162

(a)   (b)   (c)
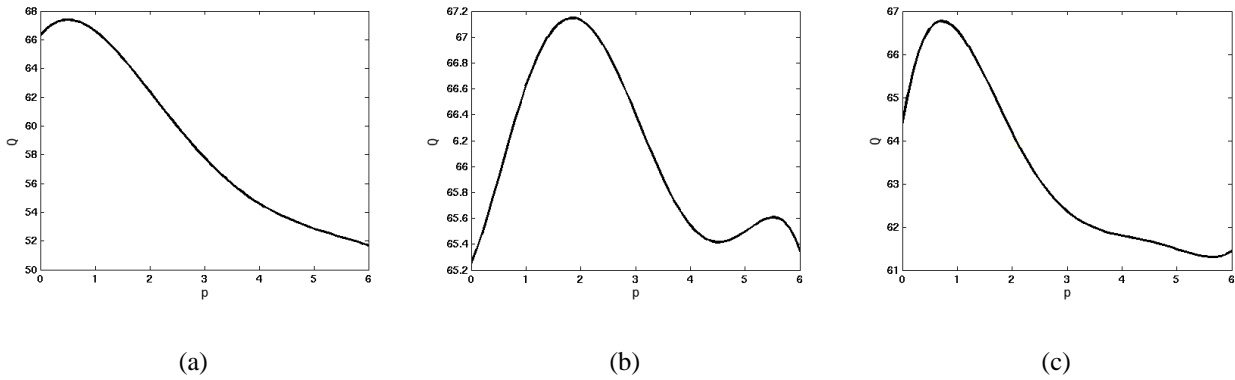
Figure 4.17: Estimation of the vessel map parameters using hand-labeled images. Quality criteria (Q) – the background-less pAROC – for different parameter values (p) for the Hessian-multiscale vessel map. We show results for the first pyramid level (a), second pyramid level (b) and third pyramid level (c) – all normalized to 100.

	Skin		Coronary		Retina	
	cc	fp	cc	fp	cc	fp
otsu	48.26	2.35	44.56	1.86	52.49	0.65
adaM	63.21	3.92	60.32	3.30	69.91	2.25
hyst	77.66	5.10	68.52	4.90	83.42	5.50

Table 4.4: Results obtained with unsupervised segmentation methods on different image sets.

2.7, 2.1 and 0.1. The parameters have been found during a sequential search. The parameter range was always between zero and six. The quality criterion was the background-less pAROC. The corresponding curves for coronary angiograms are shown in Figure 4.17.

Results.   We have computed the mean percentages of *cc* and *fp* on all images from the coronary and skin data-set and on the test images from the retina data-set for the MRF-based adaptive threshold with junction-based selection (adaM) and the hysteresis threshold (hyst). These results are shown in Table 4.4. A flow chart for each of the two algorithms used to achieve the results is shown in Figure 4.18. As reference we show also the results obtained by the Otsu-threshold (otsu).

Some results are shown in Figure 4.19. The first line shows the original angiogram, the second line the ground truth and the third line the segmentation result achieved by the MRF-based adaptive threshold with junction-based vessel selection and the fourth line the result achieved by the hysteresis threshold for a coronary angiogram, a microangiogram and a retina photography.

Methodology.   The parameters used for the MRF-based adaptive threshold were: $\alpha = 2$ and $\beta = -0.65$, mildly encouraging thus a decision in favor of the vessel pixel-class. These parameters remained unchanged irrespective of the type of analyzed angiogram. The algorithm was allowed to iterate until the classification result remained unchanged between two consecu-
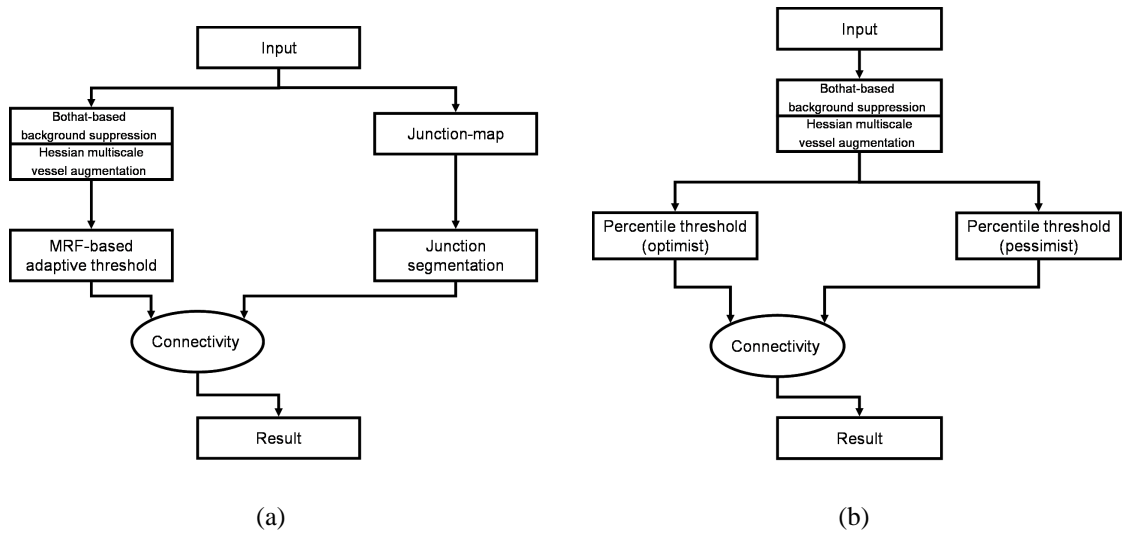
Figure 4.18: Flow charts for: the MRF-based adaptive threshold with junction-based selection (a) and the unsupervised hysteresis threshold (b). Compare these with those in Figure 4.11 and 4.12 respectively.

tive iterations. The junctions were detected by thresholding with the Otsu threshold the second eigenvalue of the structure tensor computed on the vessel map.

The parameters for the hysteresis threshold have been (in percentages of image surface certainly covered by vessels and background respectively): three and 85 for the skin-microangiograms, 0.1 and 84 for the coronary angiograms and six and 82 for the retina images.

Supervised segmentation

The supervised segmentation methods proposed here are: the supervised hysteresis threshold and the supervised hysteresis classifier.

Input. The available vessel images were divided into two sets for training and testing respectively. The training set used for the segmentation of coronary angiograms contained four images, that used for the segmentation of microangiograms contained five images and that for retina photographies contained 20 images. The remaining images in each set have been used for testing.

With respect to the size of the input feature space there are both scalar and vectorial segmentation methods. The scalar hysteresis threshold used again the Hessian multiscale vessel map. For the multidimensional hysteresis classifier and Fisher classifier we have used vectorial inputs. Multidimensional pixel-descriptions were obtained by combining in a vector the gray levels that a certain pixel receives in several different vessel maps (see Section 4.2.3). The initial set has been obtained by combining all vessel enhancement methods (see Section 4.2). The optimal features can be selected from an initial set by feature-selection as discussed in Section 4.3.2 page 145.

For the retina images, after feature selection, the optimal set was based on the following vessel maps: the Laplacian-multiscale, the Hessian-multiscale and the Hessian single-scale. The
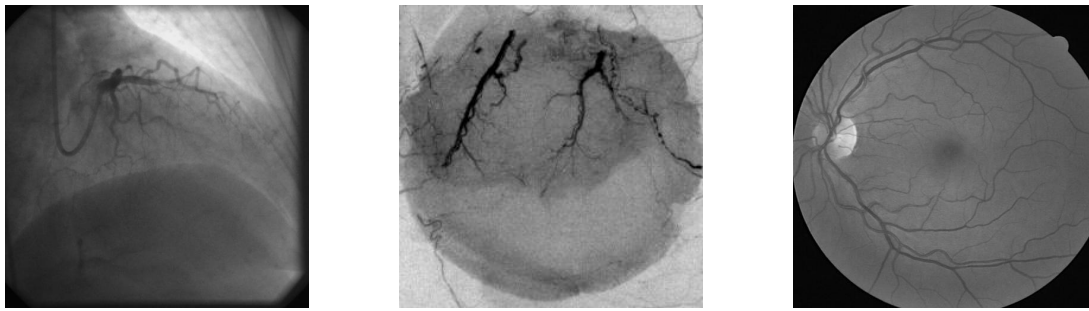
(a) original

(b) ground truth
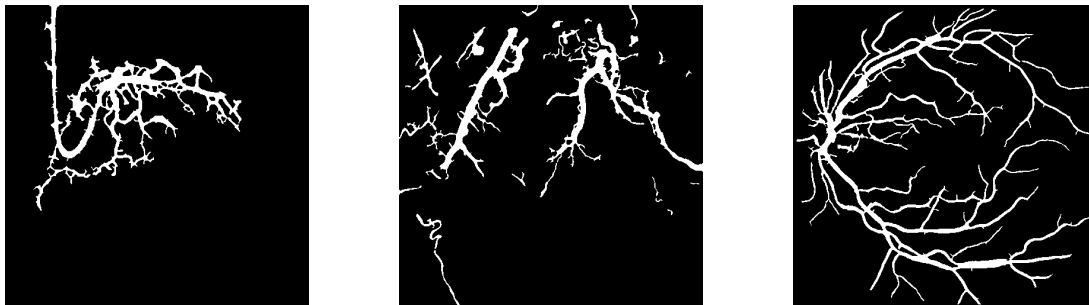
(c) adaM
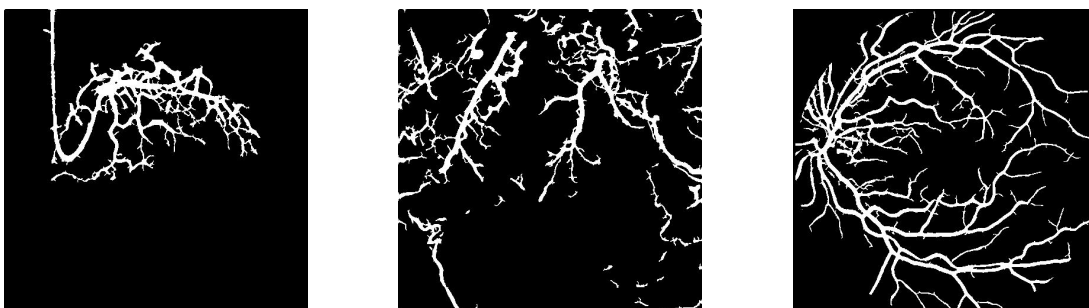
(d) hyst

Figure 4.19: Examples of unsupervised segmentation: Original image (a), manual ground truth (b), segmentation results of the MRF-based adaptive threshold with junction-based vessel selection – adaM (c) and for the hysteresis threshold – hyst (d) for a coronary angiogram a skin-transplant microangiogram and a retina-image.

	Skin		Coronary		Retina	
	cc	fp	cc	fp	cc	fp
histT	75.79	6.73	69.45	4.25	88.80	9.54
hystC	72.83	5.59	68.92	3.89	84.87	4
FisherC	74.93	10.35	70.1	5.4	87.61	10.94

Table 4.5: Results obtained with supervised segmentation methods on different image sets.
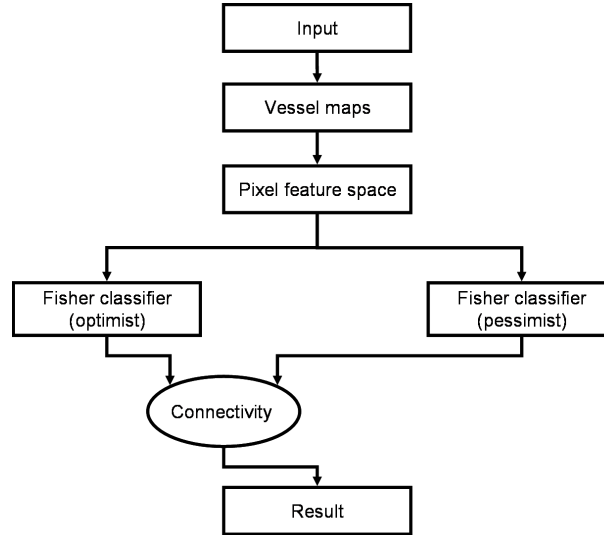
Figure 4.20: Flow chart of the hysteresis classifier. Compare it with Figure 4.12.

parameters for each map were: 0.8, 3, 2.3, 1.2 and 0.7 for the levels and 0.6 for the parameter a of the Laplacian-multiscale, 2.7, 2.1 and 0.1 for the Hessian multiscale and 15 and three for the Hessian single-scale.

For the skin-transplant microangiograms and for the coronary angiograms the optimal feature set contained: the Laplacian-multiscale and the Hessian-multiscale. For microangiograms, the parameters for each map were: 0.5, 2, 2.5, 1.1 and 0.7 for the levels and 0.5 for the parameter a of the Laplacian-multiscale and 1.3, 1.7 and 0.8 for the Hessian-multiscale.

For the coronary angiograms the parameters for each map were: 0.1, 2.5, 2, 1 and 0.3 for the levels and 0.3 for the parameter a of the Laplacian-multiscale and 0.7, 2.1 and 0.8 for the Hessian-multiscale.

Results. The results for supervised segmentation are shown in Table 4.5. The flow chart of the hysteresis classifier used to compute the results is shown in Figure 4.20. They represent the mean percentages of cc and fp over all images in each test data set for the supervised hysteresis threshold (histT) and the supervised hysteresis classifier (hystC). As reference we computed also the segmentation results achieved by a linear Fisher classifier (FisherC). The threshold of the Fisher linear-classifier was determined such that it corresponds to the ROC point situated at the largest distance from the line linking the ROC-points with $fp = 0\%$ and $fp = 100\%$.

Some segmentation results are shown in Figure 4.21. The corresponding original images and ground truth are shown in Figure 4.19 (a) and (b) respectively.

(a)   (b)   (c)
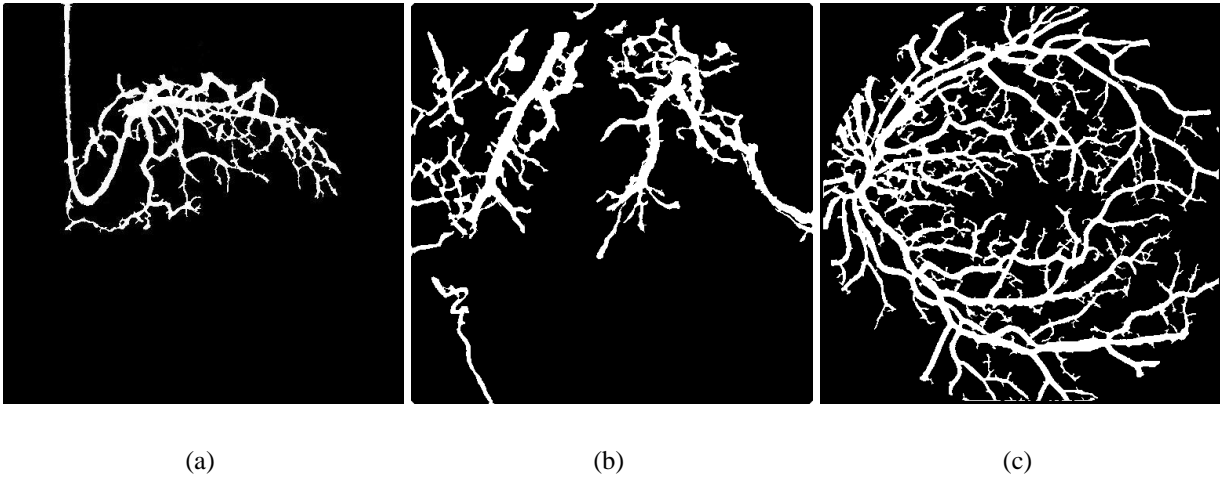
Figure 4.21: Examples of supervised segmentation using the hysteresis classifier. We show segmentation results obtained for a coronary angiogram (a), a microangiogram (b) and a retina photography (c). The original images and the corresponding ground truth are shown in Figure 4.19 (a) and (b) respectively

Methodology.   We show only the segmentation results achieved by the supervised hysteresis threshold trained by indirect sequential training. To compute the "optimist" a maximum of 7% fp was allowed. Then the parameters are (in percents of image surface certainly covered by vessels and background respectively): four and 86 for the skin-microangiograms, 0.1 and 82 for the coronary angiograms and nine and 76 for the retina images. In the case of direct sequential training one can achieve similar results by imposing a maximum fp rate of 0.6% to compute the "pessimist" and 12% to compute the "optimist".

In the case of the multidimensional hysteresis classifier, the two threshold for each data-set were obtained as: $T_p = arg\max(cc|fp < 0.5\%)$ and $T_o = arg\max(cc|fp < 15\%)$. We have obtained similar results when establishing α_o and α_p by iterative training rather than imposing them.

Experiments.   We have trained the supervised hysteresis threshold by batch training also. Then, we obtained worse results mainly due to the lack of homogeneity of the gray-level representation of the vessels among different images from a data-set. For example in the case of microangiograms the results obtained after sequential training had 75.79% cc for 6.73% fp and for the batch training 64.29% cc for 7.85% fp.

We have conducted several experiments to obtain a better comparison between a linear classifier and the hysteresis classifier. As input data we have used the retina images. First we have computed the static functioning point of the hysteresis classifier. Then we have computed the ROC curve of a Fisher classifier using the same weights-vector $\vec{w}$ on the test-set.

The static functioning point.   The static functioning point of a hysteresis classifier represents the performance of the classifier on the training-set. For the retina-images, the training-ROC curve used to find the thresholds of the "pessimistic" and "optimistic" classifiers is shown in Figure 4.22 (a). The points corresponding to the two thresholds are marked by squares. Using

(a) (b)
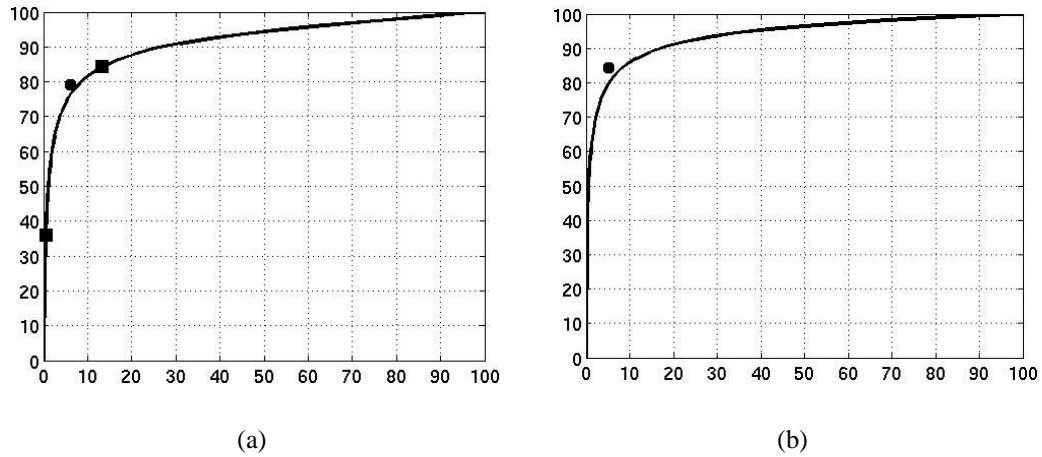
Figure 4.22: ROC curve used for training the hysteresis classifier (a). The squares mark the points corresponding to the pessimist and optimist respectively. The bullet marks the static functioning point. ROC curve obtained while testing a Fisher classifier with the same parameter-set $\vec{w}$ as the trained hysteresis classifier (b). The bullet marks the performance of the hysteresis classifier.

these thresholds, the static functioning point of the hysteresis classifier is 79.98% cc and 6% fp. This point is marked by a bullet in the graphic.

4.3.4   Automatic vessel segmentation: Discussion

We have described and investigated the performance of several methods designed to achieve a fast and reliable automatic segmentation of vascular structures in 2D-projection images. Among these methods we differentiate between standard methods, like statistically motivated *thresholding* and a new class of classification methods, with application to vessel segmentation: the *hysteresis classification*.

Thresholding

To segment vessels we have used off the shelf *fixed thresholds* and designed a new *MRF-based adaptive threshold*. All thresholding methods used here are statistically motivated. We believe that there is no difference in the quality of the results returned by non-statistical (e.g. the SOM) and statistical (e.g. Otsu) fixed thresholds.

To improve thresholding results, we have developed a junction-based selection mechanism and used it together with the MRF-based adaptive threshold. Together they represent a heterogeneous type of hysteresis threshold (see the "Outlook" at page 150, in Section 4.3.2).

Fixed thresholds.   Fixed thresholds like the percentile threshold and the Otsu threshold, are in general unsuited for vessel segmentation. The main reason for this is the strong overlap between vessels and background. This is shown for one typical example in Figure 4.23. In Figure 4.23 (a) it is shown the histogram of a skin-transplant microangiogram and in Figure 4.23 (b) the histogram of the corresponding vessel map computed by multiscale Hessian analysis. The
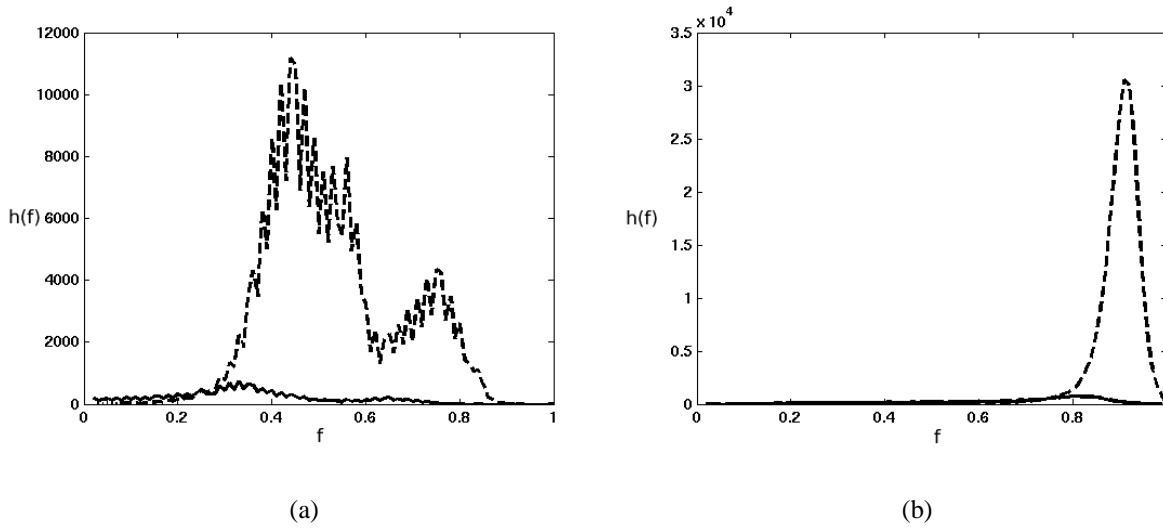
Figure 4.23: Vessel (continuous line) and background (dashed line) class conditional *pdfs* – approximated by histograms – for a skin transplant microangiogram (a) and for its corresponding vessel map (b). The interval of gray-level values has been normalized to the unit.

vessel enhancement improves the separability of the two classes, but the overlap between them is still strong. A fixed threshold can not successfully separate such an input space.

Yet another difficulty for fixed threshold is the dependency between vessel size and contrast, which is usually attenuated but not eliminated during vessel enhancement (see also Figure 4.23). Consequently, the vessel class includes actually several sub-classes corresponding to the different vessel categories. Also, the background has a lot more pixel than the vessels, which can constitute a problem for some unsupervised classifiers [14].

An unsupervised classification algorithm uses certain separability criteria, like the J_1, or the distance to a class prototype, to decide where to classify a certain vector. Yet again, many statistical unsupervised classification algorithms implicitly assume a Gaussian model of the analyzed data. If such methods are used to obtain a fixed threshold, than one must keep in mind that the Gaussian-assumption is not always correct for vessel segmentation and therefore the results achieved may be suboptimal. Usually the Gaussian-guess can be assumed to be correct for the background of the vessel map but it is less appropriate for the entire vessel class.

The MRF-based adaptive threshold with junction-based vessel selection.   The MRF-based adaptive threshold can be also seen as an improvement over the classical fixed Otsu threshold. Vessel-like artefacts, like e.g., ribs in the case of coronary angiograms – will appear in the final segmentation because they do look like vessels – i.e. they have a contrast mostly similar to that of small vessels and are also compact connected, elongated structures. Such artefacts are eliminated by the junction selection mechanism. On the other hand, the junction selection mechanism is likely to miss single straight vessels.

The results in Table 4.4 show that the proposed algorithm represents a genuine improvement over simple general-procedures as the Otsu threshold. This is mainly due to the introduction of prior-knowledge – specific to the problem of segmenting vessels – while taking the decision at pixel-level.

Hysteresis classification.

The hysteresis classification paradigm can be successfully applied to vessel segmentation because they are spatially connected structures. A hysteresis classifier for vessel segmentation uses the vessel contrast to separate potential vessel pixels from the background pixels and the *vessel connectivity* to select only the true vessels.

We have introduced and experimented with several practical implementations of the hysteresis classification paradigm which include *unsupervised* and *supervised methods*. To demonstrate the potential of the hysteresis classification for vessel segmentation, we have *compared* the unsupervised *hysteresis* threshold and the hysteresis classifier *with other state of the art algorithms* on the same data set.

Vessel connectivity. The hysteresis classifier can be used only if object points exhibit a certain connectivity in some additional space. In the case of vessel-segmentation, vessel object-points are connected spatially over a vicinity in the analyzed images. This connectivity needs to be conserved in the result returned by the "optimistic" classifier. Therefore, it should be enforced particularly during the computation of the vessel map. For this purpose we use specific quality criteria to establish the parameters of the vessel map, the best being the background-less percentile pAROC. However, for this purpose one needs to have access to labeled data.

Unsupervised methods. The hysteresis threshold is based on principles that have been implicitly used for the MRF-based adaptive threshold with junction-based selection: use high-confidence vessel points to select low-confidence vessel points over the connectivity of vessels. However in this case the high-confidence vessel points are chosen based on their gray level – as they are the image items with the largest interaction with the primary information carrier – and not on the local orientation. This makes possible the selection of straight vessels but at the same time tends to ignore single weak-contrasted vessels. Such aspects have to be considered while establishing the "pessimistic" and "optimistic" thresholds.

Supervised methods. We have compared the hysteresis classifier with state of the art general-purpose classification algorithms in the case of vessel segmentation. By comparison to a simple linear classifier, the linear hysteresis classifier can achieve results of a higher quality virtually as fast in training and computation time. By comparison to a SVM, the hysteresis classifier is faster and manages to segment a lot more vessel area with similar precision.

Comparison between a supervised hysteresis classifier and a linear classifier. The results on the test set (see Table 4.5) show that the supervised hysteresis threshold returns results comparable with the Fisher's linear classifier with regard to the fp rate for slightly improved cc rates. This also indicates that a multidimensional description of pixels leads to a more separable feature-space. In comparison with Fisher's linear classifier, the hysteresis classifier can significantly reduce the fp rate while keeping a similar cc rate.

As shown in Figure 4.22 (a) the static functioning point of the hysteresis classifier lies above the ROC-curve of the Fisher classifier which suggests a better performance with respect to the rate of cc for the same rate of fp.

Using the weights vector $\vec{w}$, a ROC curve for the test set can be computed, describing the performance of a linear classifier. In the case of retina images, for the same fp rate as the

hysteresis classifier – i.e. 4% – such a procedure will yield only 81.31% cc. This suggests an improvement of 3.56 percentage points for the hysteresis classifier over a linear classification. This ROC curve is shown in Figure 4.22 (b). The performance point of the hysteresis classifier is marked by a bullet.

Comparison between a supervised hysteresis classifier and a SVM.   We have also experimented with a SVM for supervised pixel-based vessel segmentation. The results achieved on the retina test-images were: 64.02% cc and 1.24 % fp. The poor results of the SVM can be explained by the large overlap and unbalance between the vessels and background classes and also by the poor choice of parameters. The parameters of the SVM, which should be set prior to the actual training, are the length C of the bounding box, the kernel type, and for Gaussian kernels the size of the dome. Training a SVM on a space with 4,296,431 components takes more than 48 hours under Matlab on a Pentium IV 2.4 GHz machine with 1024 MB of RAM, consequently finding the best C and the best dome-size is a tedious work. Training the linear hysteresis classifier on the same machine takes only 20 minutes. Segmenting an image means in this case classifying 262,144 vectors, which for the SVM takes around half an hour due to the large number of support vectors (a consequence of the strong overlap between the two classes). For the hysteresis classifier the segmentation of an image takes around two seconds.

Comparison between supervised and unsupervised hysteresis methods.   The supervised and the unsupervised hysteresis thresholds reach the same results for the same parameters set – the difference being that in one case they are imposed and in the other case they are trained. In practice, the results are comparable (see Table 4.4 and Table 4.5), which shows the quality of the proposed training method, as the parameters of the a priori unsupervised hysteresis threshold have been optimized manually.

Comparing the results obtained by the unsupervised hysteresis threshold and the hysteresis classifier, it can be concluded that the hysteresis classifier is better as it shows more cc for less fp. This is due mainly to the additional information contained in the 3D pixel-feature space.

Comparison with other dedicated segmentation methods on the retina data set.   Segmentation methods for retina fundus photographies have been also proposed in [181], [117] and [101].

In [181] it is reported on comparative results on the same data set. The quality criteria chosen there is the AROC of the decision algorithm. To allow a comparison, we have computed also the AROC for the both the unsupervised hysteresis threshold and the hysteresis classifier.

In the case of the threshold, we have hold the "pessimist" constant at six percents of the images surface covered only by vessels and varied the "optimist" between one and 100 percents of the image covered only by background. In the case of the hysteresis classifier, we have hold the "pessimist" constant at a fp rate of 0.5 and varied the "optimist" over the entire range of fp rates from 0 to 100%. Each time we have used as input data the test-set from the data-base from [181].

The AROC values obtained were: 0.9681 for the hysteresis threshold and 0.9723 for the hysteresis classifier. For comparison the best value reported in [181] is: 0.9520. For both hysteresis methods, the segmentation results have been achieved as described in Section 4.3.3.

4.3.5 Automatic vessel segmentation: Conclusions

We have introduced several automatic vessel segmentation methods. Most of them are thresholds which accept as input only a scalar feature-space, while the supervised hysteresis classifier can handle also vectorial inputs. All segmentation methods described here start from the assumption that some vessels are always visible in the analyzed images. They are likely to find some even when there are none, thus the analyzed data sets should not include any blank vessel images – i.e. images showing no vessels at all. Also, most of them use the assumption of vessel connectivity, therefore connectivity of vessels and separation from background at the borders has to be enforced particularly during the computation of the vessel map. We have introduced supervised and unsupervised segmentation methods.

To achieve a correct supervised segmentation, a properly labeled training set is needed. For any binary classification there are two types of possible errors: omissions and false inclusions. For vessel segmentation in particular, with respect to the vessel class, omissions are typically encountered for some small vessels and false inclusions for background points at the border of vessels. The training set was affected by such errors, mainly due to the tedious prelabeling process as manually segmenting an angiogram can take up to two hours. Therefore the accuracy of any vessel segmentation algorithm using such a "bronze" ground truth is negatively affected.

Vessel segmentation: from fixed thresholds to hysteresis classification. A timeline.   We set forth to accomplish the task of vessel segmentation. For this purpose initially we have investigated several state of the art algorithms [121] and found out that pattern recognition approaches represent the best choice. From among the pattern recognition approaches we consider the statistically motivated ones, like the Otsu threshold, to be the most promising with respect to the quality of the achieved results and the speed.

However, the quality of these initial results was considered unsatisfactory. Therefore, we have improved the Otsu threshold to the point of devising a new algorithm: the *MRF-based adaptive threshold*. The improvements have been achieved mainly by incorporating prior information about vessels into the algorithm. For even better results, we have continued to incorporate prior knowledge in the decision-making process by introducing the *junction*-based mechanism for *selection of true vessels*.

The junction-based selection makes use of the vessel connectivity to improve an existing over-segmentation. This constitutes the basis on which we have introduced the *hysteresis classification* as an almost optimal solution to vessel segmentation.

The MRF-based adaptive threshold with junction-based selection

The MRF-based adaptive threshold represents an improvement over the classical Otsu threshold [53]. For each pixel it is decided by the MAP criterion if it is a vessel or a background. The decision is supported both by the class-conditional *pdfs* – computed with the Otsu threshold – and by the prior-probabilities for each class computed from the MRF. The final result is achieved after junction-based selection, where additional information with respect to vessel connectivity and their appearance is used to select from the segmentation result only points linked to one another and linked to a junction.

Connection to the hysteresis threshold.   We consider the MRF-based adaptive threshold, together with the junction-based selection of true vessels, to represent a heterogeneous type of unsupervised learning hysteresis threshold. The MRF-based adaptive threshold which segments vessels represent the "optimistic" classifier and the Otsu threshold which returns junctions represents the "pessimistic" classifier. We call this method heterogeneous, as the "optimist" and the "pessimist" have different input feature spaces. In the case of the homogeneous hysteresis threshold, both "optimist" and "pessimist" have the same input feature space.

Hysteresis classification

The hysteresis threshold.   The MRF-based adaptive threshold with junction-based selection prompted the development of the hysteresis threshold for vessel segmentation [43]. The hysteresis threshold has been used before for edge segmentation [34], however in the case of edge segmentation the selection of the two thresholds was rather ad hoc. We have introduced better threshold selection methods and adapted them to the problem of vessel segmentation.

The unsupervised a priori hysteresis threshold needs two parameters which define the "optimistic" and "pessimistic" thresholds. In machine-vision terminology these two parameters are percentiles which actually should define the limits within which a large majority of the vessel-pixels can be found. These limits translate into the percentages of image area certainly covered by vessels and background. They need to be established by the user in advance for the entire set of analyzed images. This can be done, e.g., by trying several combinations, until the results which are obtained fulfill the expectations of the human operator.

If a labeled set of vessel images is available, these parameters can also be trained, obtaining thus a supervised hysteresis threshold. Training is based on an extended ROC analysis.

The supervised hysteresis classifier.   We have shown how to extend the hysteresis classification paradigm to multidimensional feature-spaces in a supervised manner [44].

The segmentation result can be improved if a better pixel description is available. This can be achieved by characterizing the pixels not by a single value but by a vector. We have shown how to extend the supervised hysteresis threshold to handle also multidimensional inputs. For this purpose we have used the Fisher classifier as base-classifier. First the input feature space is transformed to improve the homogeneity of the vessel and background representation over the entire training set. Then the "optimistic" and the "pessimistic" classifiers are defined such that the first selects all vessel pixels and the last only vessel pixels. Then the feature space transformation ensures that the training is valid for all images in the analyzed data set.

A supervised hysteresis classifier can also be conceived starting from other base-classifiers, as long as they are trained to fit the roles of "pessimist" and "optimist" and a connectivity between feature-vectors does exist. For example, perceptrons could be used: the first one is forced to completely exclude one class, e.g., the background class, and the second the other one, i.e. then, the vessel class. This can be achieved by appropriately weighting the errors of each class during training. However, one has to pay attention that background and vessel feature vectors appear in comparable numbers in the training space. This can be achieved by appropriately pruning the background class. The two results are then combined over the connectivity of vessels as in the case of the Fisher classifier.

Feature extraction for hysteresis classification. At the heart of hysteresis classification stays the connectivity which object points exhibit in some space other than their feature space. In the case of vessel segmentation, it is their spatial connectivity which justifies the application of a hysteresis classifier. We have described here also quality criteria for estimating the parameters of a vessel map, such that it fulfills to the best possible extent the connectivity and contrast conditions needed for a successful hysteresis-based segmentation.

4.4 A semi-automatic system for the analysis of angiogenesis in skin-transplant microangiograms

Angiogenesis is the process by which new vessels are formed from older ones. Both the process of inducing and that of inhibiting vessel growth can be amplified by certain substances. Clearly the potential of such substances for the purpose of healing different disease is enormous and therefore subject to active research [132].

Here we describe a semi-automatic vessel segmentation tool aimed at quantifying angiogenesis in treated fasciocutaneous skin-flaps transplanted on the back of laboratory mice. For this purpose the input images are first processed such that vessels are enhanced (see Section 4.2), then they are segmented. We describe two segmentation methods (Section 4.4.2). The first uses a fuzzy clustering algorithm and the second, the hysteresis threshold. Both methods return automatically a result which can then be manually improved by the human operator. The segmentation result is then used to compute some measures which should describe vascularization in the target tissue (Section 4.4.3). The results obtained on our data base are shown in Section 4.4.4. Section 4.4.5 contains the discussion and conclusions.

Although it was first applied to microangiograms, the tool we describe here can be used for all types of 2D-projection vessel images. We believe that the same framework can be also used to segment as well other types of vessel-like objects in 2D-projection images, like e.g., nerve fibers.

4.4.1 A tool for the analysis of angiogenesis: Introduction

Angiogenesis plays a major role in a number of situations, for example if specific environmental conditions impose the acceleration of the metabolic reactions at the level of an organ, new vessels are formed over a relatively short period of time to permit the organ to fulfill its purpose. Also the process by which a wound heals involves the construction of new tissue which has to be supported for which purpose vessels need to grow into it. In the case of transplanted skin sustained angiogenesis is needed to avoid necrosis of the transplanted tissue and other complications [132]. However, there are also cases when angiogenesis is not desired, e.g., in the case of a tumor, when one would like to cut the nutrients-supply to kill thus the tumor cells [151], [164].

The search for substances to improve *angiogenesis* represent currently an *active research field* in the pharmaceutical domain [132]. For this purpose, often, a specific X-ray imaging technique is used which returns images of the micro-vessels build during angiogenesis, these images are called *microangiograms*. During experiments one seeks to *quantify angiogenesis* in microangiograms to investigate the effectiveness of different substances.

(a)   (b)   (c)
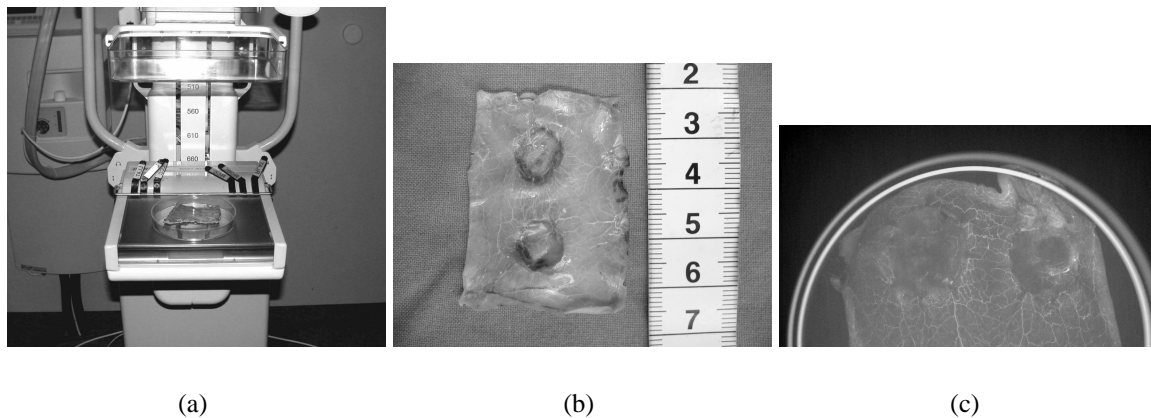
Figure 4.24: Imaging system (a) dermal matrices (b) and microangiogram (c).

Laboratory research for improved angiogenesis. The success of a skin transplantation depends on a proper revascularization of the transplanted tissue. To evaluate the effectiveness of different pharmacological substances designed for this purpose, experiments are conducted. During such an experimental run, drug-treated dermal matrices [94] are transplanted to cover two disk-shaped full-thickness skin defects (diameter: 15mm) on the backs of laboratory animals (nude mice, body weight about 30g) [93] (see Figure 4.24 (b)). The animals live then another three to 14 days with the transplants such that there is enough time to build new vessels. During this time it may happen that the transplants fall from their places, this represents one of the reason why there are two transplanted dermal matrices. A typical experimental run includes 12 animals. Two to three of them are used for control purposes and receive untreated transplants.

Microangiograms and other vessel images. For vessel imaging after a given time interval, blood is withdrawn via the left carotid artery using micro-surgical instruments, and slowly replaced by a contrast medium. The transplant samples (see Figure 4.24 (b)) are then harvested together with some surrounding tissue and imaged using an X-ray mammography system (typical settings: 9mAs at 24kV) shown in Figure 4.24 (a). As a result a microangiogram is obtained, shown in Figure 4.24 (c). This shows the vessels, potentially down to a size of about $20\mu m$ [93].

Currently, research is done to replace the microangiograms by by photographies of the tissue when this is placed on a transilluminator. The advantage is that the tissue can be reused after imaging, as it does not need chemical additives. This saves animals and allows to perform comparisons among a larger number of samples.

Quantification of angiogenesis. In such images we seek to quantify the angiogenesis in the dermal matrices, e.g., by measures such as the percentage of area covered by the blood vessels in the target sample (see Figure 4.28), the vessel length, or the micro-vascular index, which relates the total vessel length to the area involved (see [93] and the references in there). Yet another type of measurement returns the area and length of the newly built vessels in percentages with respect to a reference sample. This sample is selected on the microangiogram to cover an area comparable to that of the transplant, in a region where the natural vasculature can be observed.

4.4.2 A tool for the analysis of angiogenesis: The segmentation of vessels

For our quantification purposes, the vessels need to be segmented in the imaged transplant [52], [51]. Since the field of application of this system is an experimental laboratory setting for drug evaluation, rather than clinical routine, time constraints are of less concern and also a certain degree of interaction is feasible, with the purpose of improving the results.

Our system returns automatically a segmentation proposition which can be – if needed – improved in a semiautomatic manner. For automatic vessel segmentation, we need unsupervised methods as one would like to keep the generality of the method and does not have access to an expert-labeled training set. We propose two algorithms: (i) a *fuzzy clustering algorithm*, which can work also on vectorial inputs (see Section 4.2.3) and which has better classification performance in comparison to other standard techniques [14] and (ii) an unsupervised a priori *hysteresis threshold*. We show also how to improve the segmentation results by interacting manually with the two algorithms.

Vessel segmentation by fuzzy clustering

The clustering algorithm iteratively improves a performance measure computed on a fuzzy set decomposition, starting from an initial partition. To allow easy and comfortable interactive processing this partition is an over-segmentation. Furthermore, since thus practically every vessel is visible at a certain stage during the iterations, this permits to achieve better results at the end of the processing chain. After the iteration stops, we automatically choose the desired vessel segments by junction analysis (see "Junction-based selection of true vessels" at page 134, in Section 4.3.1). In case the user is not quite satisfied with the automatically provided result, he may then manually refine the segmentation by leafing through the different stages of the algorithm, and select the vessel segments.

The clustering is initialized with a vessel over-segmentation result computed by thresholding the Top-hat vessel map. Empirically, the vessel covered area is always less than 50% of the image area, thus we choose the 50th percentile as threshold. A result is shown in Figure 4.25 (b).

The fuzzy class memberships are computed by a function (i.e. affinity) which measures how closely related the investigated vector is to a certain class. Let $\vec{x}$ be one from a set of N feature vectors, and ω_i be one from a set of M classes. The affinity of $\vec{x}$ to ω_i is defined as:

$$r(\vec{x}, \omega_i) = 1 - \frac{1}{N} \sum_{\vec{y} \in \omega_i} h^\beta(\| \vec{x} - \vec{y} \|) \tag{4.16}$$

with $h^\beta : [0, \infty) \to [0, 1]$ and:

$$h^\beta(\nu) = \begin{cases} \frac{\nu^2}{\beta} & \text{if } \nu \leq \sqrt{\beta} \\ 1 & \text{if } \nu > \sqrt{\beta} \end{cases} \tag{4.17}$$

Then the class belonging coefficient for $\vec{x}$ and ω_i is: $u_i(\vec{x}) = P_i \frac{r(\vec{x}, \omega_i)}{r(\vec{x}, U)}$, with P_i the prior on ω_i. The parameter β is actually a bound on the cluster/class spread. Feature vectors further apart than $\sqrt{\beta}$ are ignored in the affinity computation. Choosing β such that $\max \| \vec{x} - \vec{y} \| = \sqrt{\beta_m}$ for $\vec{x}, \vec{y} \in U$ permits the consideration of all feature space points. Then, representing each class

(a)   (b)   (c)
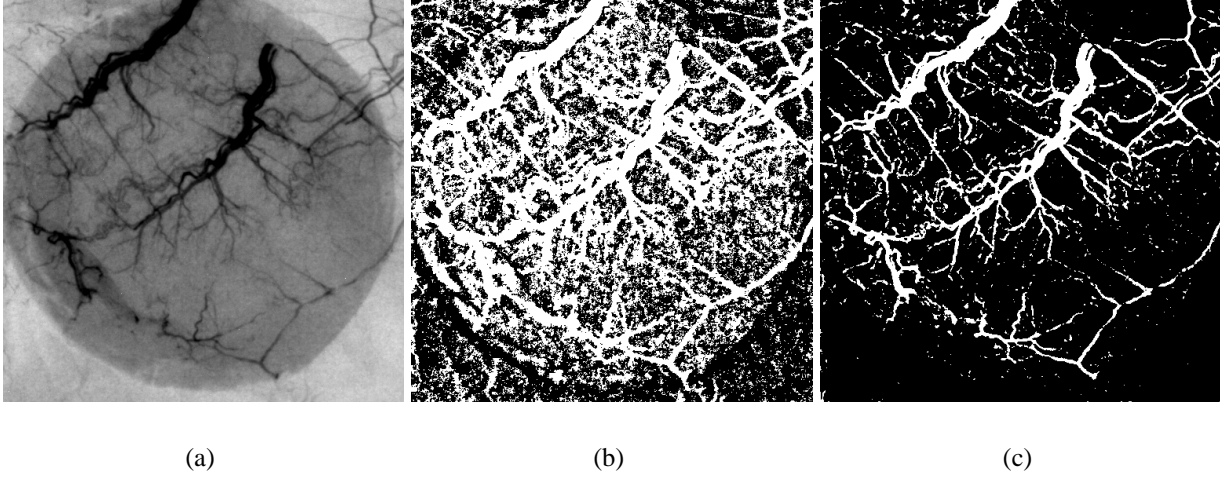
Figure 4.25: Original-image (a), clustering initialization (b) and clustering result (c).

by its mean vector alone, the fuzzy class membership coefficients will be:

$$u_i(\vec{x}) = \frac{\beta - \parallel \vec{x} - \mu_i \parallel^2}{\beta M - \sum_{j=1}^{M} \parallel \vec{x} - \mu_j \parallel^2} \tag{4.18}$$

and $\sum_{i}^{M} u_i(\vec{x}) = 1$. Although heuristic methods for determining β are available in [14], for vessel segmentation we propose on an empirical basis choosing β such that $\sqrt{\beta} = 0.8\sqrt{\beta_m}$. For vectors beyond the allowed spread $u_i = 0$. To measure the quality of a certain fuzzy partition – i.e. the amount of incertitude (fuzziness) present – the following function is used:

$$\Psi = \frac{1}{N(M-1)} \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} \sum_{\vec{x} \in U} (u_i(\vec{x}) - u_j(\vec{x}))^2 \tag{4.19}$$

with $\Psi \in [0, 1]$ and $\Psi = 0$ indicating the highest possible degree of fuzziness.

Then during an iteration a vector will change its class only if it leads to an increase in the quality of the fuzzy partition. The iterations stop when Ψ can not be increased anymore.

Vessel segmentation is characterized by strongly unbalanced class sizes – i.e. class skew – and inner-class variances. Under such conditions the fuzziness measure inherently favors good accuracy (see Section 2.5.1), implying high correct classification rates for the larger background-class, which in this case also has a smaller spread in the feature space. Clearly this leads to under-segmentation of the vessels and in extreme cases to results with a 0% cc rate.

To successfully meet the challenge of a correct vessel segmentation in microangiograms we have extended the fuzzy clustering algorithm to use an additional statistically motivated stopping criterion based on the same separability measure which guided also the feature extraction process. Then the algorithm will stop as soon as the separability measured on the current vessel class fuzzy coefficients set decreases. As ground truth the segmentation result obtained at the end of the previous iteration is used. A result is shown in Figure 4.25 (c).

Vessel segmentation by hysteresis thresholding

The fuzzy clustering-based segmentation of microangiograms is achieved by reclassifying some feature-vectors at each iteration. As it starts from an initial percentile-based over segmentation,

at each iteration the larger majority of reclassified pixels are miss-classified vessels. Thus it can be said that the background is "eroded" stepwise until only vessels remain. Finally, only true vessels are selected using the junction-mechanism. Experience shows that some weak contrasted vessels are also eliminated and straight vessels are not selected which makes then necessary the subsequent user-supported selection of true vessels, to achieve an optimal segmentation. This is then based on the assumption that the steps are fine enough such that at a certain step each vessel – irrespective of its contrast – will be visible and it can be selected into the final segmentation. However, sometimes the steps are larger than optimal. It happens that between two steps, a vessel, which at step t is still immersed in background noise, disappears completely at step $t + 1$. More often, a small vessel is visible at step t but is still linked to some background and at step $t + 1$ a large part of it has already been reclassified as background, thus the user has to additionally separate the vessel chunk from the background at step t before selecting it.

However, if the hysteresis threshold is used, the size of the "erosion-steps" and also the number of segmented vessel chunks can be controlled by the way the "pessimistic" and the "optimistic" classifiers are specified. One can use a default parameters-set – as the ones suggested in Section 4.3.3 – and achieve with them an automatic segmentation. Starting from this proposition the user can then either select or deselect some vessel-chunks. If he is not satisfied with the automatic proposition, than is offered the possibility to modify the two thresholds in steps as fine as he needs and until he can reach the desired segmentation. As the hysteresis threshold is very fast, the processing is time-efficient. Such a procedure provides an almost total control to the user.

4.4.3 A tool for the analysis of angiogenesis: Description of vessels

The area and length of the newly built vessels can be computed easily on the segmentation results. The area is related to the total number of object-pixels and the length can be computed as the total number of object-pixels left after selecting only the centerlines in the segmentation result, e.g., by morphological thinning (see Appendix A.1). Thinning directly the segmentation result may lead to poor results as the sometimes – particularly for small vessels – some artefacts hang on the vessel border giving it an irregular appearance. These artefacts influence the thinning-result and thus fake branchings appear on the analyzed vessel. This is shown in Figure 4.26 (b). A quick and efficient solution is to compute the thinning not on the original segmentation result but on a slightly dilated segmentation result. This is shown in Figure 4.26 (c). The structuring element used during dilation was a disk with a diameter of nine pixels. Such inexactitudes are less important if relative measurements of the target vasculature are effectuated, as they affect both the results for the transplant sample and for the reference.

Assuming knowledge on the parameters of the imaging system, precise measurements in the metric system can also be achieved. In some cases it is sufficient if a scale marking is imaged together with the transplant.

4.4.4 A tool for the analysis of angiogenesis: Results

For the purpose of building a feature vector space one can use any combination of vessel maps. If hand-labeled examples are available then an optimal combination of vessel maps can found during feature selection, based on the AROC of a Fisher classifier as described in Section 4.3.2

(a)   (b)   (c)
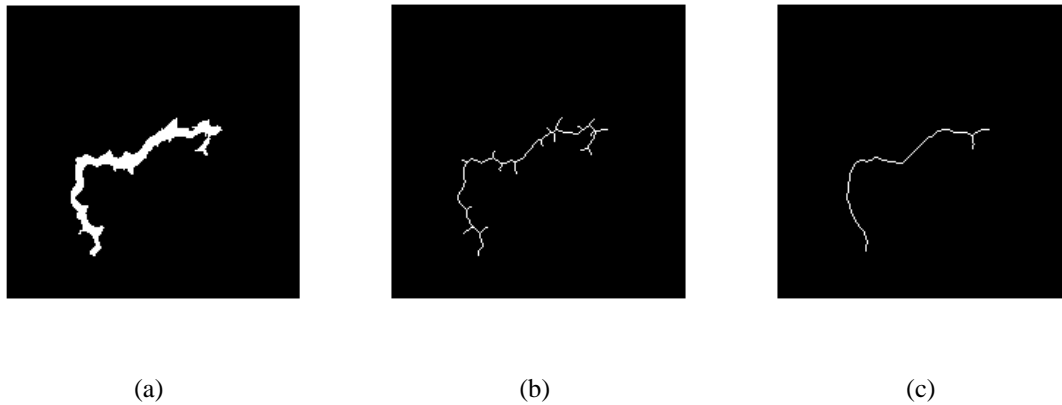
Figure 4.26: Original segmentation (a), vessel-centerline before (b) and after (c) dilation.

	automatic		semi-automatic	
	cc	fp	cc	fp
clustering	62.64	2.17	84.69	4.05
hysteresis	77.66	5.10	86.20	4.11

Table 4.6: Results obtained with and without user support for fuzzy-clustering and hysteresis-based segmentation.

or using other quality criteria, as e.g., the J_1 separability measure. Starting from the vessel maps described in Section 4.2 and based on 12 hand-labeled microangiograms, we found that the optimal set of vessel maps contained: the Laplacian-multiscale and the Hessian-multiscale vessel maps (see also Section 4.3.3).

The segmentation results achieved on the microangiograms from our data set, when applying each of the two methods described, i.e., the fuzzy clustering and the hysteresis threshold are shown in Table 4.6 both for automatic and semi-automatic usage. The flow chart of the algorithm is shown in Figure 4.27. A segmentation result achieved using the hysteresis threshold with user support, is shown in Figure 4.28.
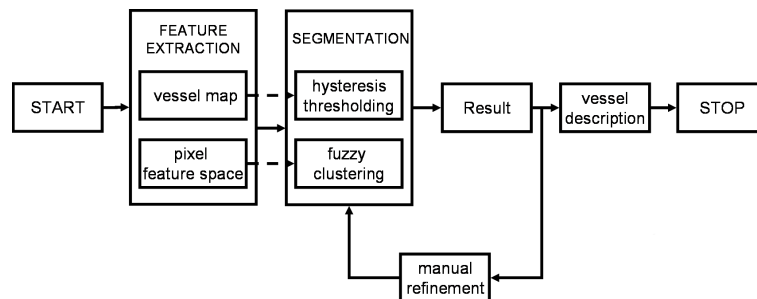
Figure 4.27: Flow chart of the semiautomatic segmentation of skin-transplant microangiograms.

(a) (b)
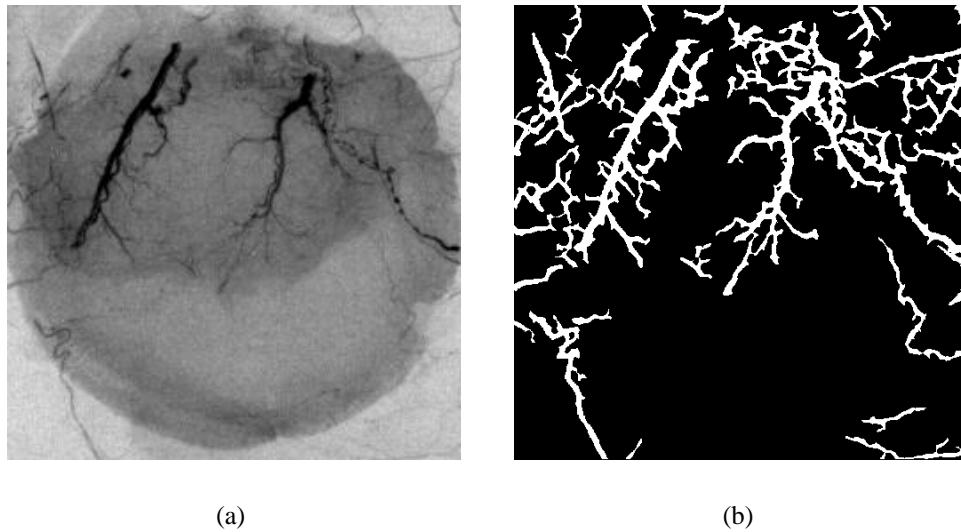
Figure 4.28: Original-image (a) and result of semiautomatic segmentation using the hysteresis threshold (b).

4.4.5 A tool for the analysis of angiogenesis: Discussion and conclusions

We have presented a novel framework for the analysis of micro-vessels in skin transplants in laboratory environments for drug testing. The analysis is supported by a micro-vessel segmentation algorithm. The segmentation-results are used for vessel area and vessel length measurements in skin transplant microangiograms. As we usually do not have access to a training set, but also to ensure the adaptability and thus the generality required for a successful segmentation of vessels, each microangiogram is individually analyzed.

Experience shows that the result may still contain some false positives and false negatives. Since our system's field of application is an experimental laboratory setting for drug evaluation rather than clinical routine, time-constraints are of less concern and a certain degree of interaction is feasible. Thus, the experimenter is allowed to manually refine the segmentation by browsing through the results of the iteration steps, and select or deselect vessel structures by placing "positive" and "negative" seed points respectively.

We discussed two segmentation algorithms: the *fuzzy-clustering* [14] and the *hysteresis threshold* [43] which return automatically a segmentation-proposition. *Comparing the two methods*, they both methods return good results but the one based on the hysteresis threshold is faster and allows a larger degree of interaction from the part of the user and therefore potentially better results.

Fuzzy clustering. The fuzzy-clustering, starts from an initial over-segmentation which it erodes step-wise. It has as input a multidimensional pixel feature space is build by gathering the results of several vessel enhancement methods so that together they constitute a separable description of vessels and background. On the final result true vessels are selected using seed-points computed automatically. As a key vessel feature, we use the fact that they typically exhibit branchings, i.e., local structures with junctions, which may be detected by a structure tensor approach. The detected branching points (junctions) are then used as vessel markers. A

structure which appears in the final segmentation result is automatically identified as vessel only if it is connected to such a seed point. If the user is not satisfied with the automatic segmentation, he can then edit the intermediary segmentation results obtained at each iteration-step and select the desired vessels from the results where they are optimally segmented.

Hysteresis threshold.   The hysteresis threshold requires from the user the specification of the image surface which in all analyzed angiograms will be covered only by vessels and background respectively. With these values it returns an automatic segmentation. If the user is not satisfied with this segmentation, he may then select or deselect vessel chunks in this result or modify the two parameters and compute a new segmentation until the desired vessel is good visible and can then be selected it in the final segmentation.

Comparison between the two segmentation methods.   The hysteresis threshold accepts only 1D inputs, however in semi-automatic usage, it returns results comparable or even better than the fuzzy clustering. This is due mainly to the fact that the step size – between intermediary results – can be also controlled and adapted to the targeted vessels.

The hysteresis threshold needs only a few seconds to achieve a segmentation and allows a larger amount of interaction from the part of the user than the fuzzy clustering which needs a few minutes minutes to segment an image. However, once the image is segmented, leafing through the intermediary results is than clearly very fast.

The results achieved by the hysteresis threshold are only slightly better than that of the fuzzy-clustering method, but they are computed faster – as now the iterative training step can be avoided – and with less intervention from the part of the human operator. Therefore, we believe that the hysteresis threshold based segmentation will prevail in practice.

Difficulties with the ground truth.   The hand-labeled ground truth usually contains errors – both ignored true vessel-pixels and additional fake vessel-pixels (see also Section 4.3.2). Most of these errors are rather uncritical being situated at the vessel-margins, e.g., when a thick vessel is under-segmented and a thin one over-segmented, but there are also critical errors, like e.g., thin vessels which are not captured. Such errors potentially fake the results obtained. Thus, results computed directly on the hand-labeled images should be rather seen as giving the upper bound on the rate of fp and the lower bound on the rate of cc.

4.5   Dynamic vessel segmentation

Atherosclerosis is a condition affecting arteries in general. In the case of the coronary arteries, this is called the coronary artery disease (see Section 2.1). If it affects a carotid artery is called the carotid artery disease and if it affects one of the arteries supplying blood to the arms, legs, stomach or kidneys, it is called the peripheral artery disease.

Here we describe vessel segmentation methods for improved X-ray-supported diagnosis and angioplasty-planning in the case of the peripheral and carotid artery disease. We start by describing specific X-ray imaging techniques like the Digital Subtraction Angiography (DSA) and how can vessel segmentation help during diagnosis and planning in Section 4.5.1. We then describe two segmentation methods in Section 4.5.2 which we compare in Section 4.5.3. Finally Section 4.5.4 contains the conclusions.

4.5.1 Dynamic vessel segmentation: Introduction

X-ray angiograms of the ill artery are used mainly for diagnostic purposes but also during treatment which may be done – similar to the case of the coronary artery disease – by an X-ray supported angioplasty intervention. Lately, ultrasound and MRA have been used as alternative imaging modalities. Ultrasound has the disadvantage of poor image quality and MRA can not be applied to all patients, e.g., patients with metallic implants like artificial heart valves, etc. Therefore, X-rays-based procedures are still the most often encountered in practice. Such *imaging* methods provide high-quality angiograms of still vessels. One of the most often encountered such procedures is DSA. It implies the recording of a sequence of angiograms showing the contrast-bolus traveling through the investigated vessel-tree.

We concentrate on vessel-segmentation in DSA sequences, which we call *dynamic vessel segmentation*. Such results offers the possibility not only to investigate the morphology of vessels and to measure their area and length, but also to analyze by *contrast-bolus tracking* the dynamic of blood through the investigated vasculature, offering thus new opportunities for enhanced data-assessment.

Imaging

Initially, diagnosis angiograms – of high image-quality – were acquired under constant flow of contrast agent such that the entire ill artery was visible. In some cases, the sheer quantity of contrast agent which had to be injected, especially when the arteries supplying blood to the legs were examined, could have risen a health problem. Although modern contrast agent is less toxic than before, it is good practice to reduce the inject quantity both for reasons of patient health and contrast-agent-cost. Advances in the computer technology have allowed the reduction of the quantity of contrast agent needed while keeping or even improving the quality of the images which were acquired. Currently, two imaging techniques are usually used in practice: *DSA* and *bolus-chase*.

Digital Subtraction Angiography. In DSA [31] several X-ray images of the vessels are acquired. One before the injection and the rest during the injection of short burst of contrast agent. The contrast-agent-less background is then subtracted from the contrast-agent-images dramatically improving contrast by elimination of still artefacts, like e.g., bones or other tissue. The DSA-images can then be used to reconstruct vessel-information. As only a short burst of contrast agent needs to be recorded, DSA is fast and therefore both patient and physician irradiation is minimal.

Bolus-chase. Bolus-chase is usually used to obtain angiograms of the arteries of the inferior limbs. In this case the bed with the patient moves synchronous with the bolus of contrast agent such that the C-arm containing the imaging system follows the contrast bolus on its way through the arteries. The obtained images are then registered to reconstruct the arteries. Usually, a set of non-contrasted images is also recorded to allow for DSA-style background attenuation. Similar to DSA, bolus-chase reduces both the quantity of contrast agent and the time required to obtain an angiogram.

Objectives of vessel segmentation

Vessel-segmentation can be used here to further enhance contrast, or to support the quantification of the stenosis. The images acquired for DSA can also be used to measure the blood-flow. For this purpose one can use densitometric measures [168] as well as the results of a vessel segmentation.

Vessel segmentation in DSA sequences can be also used to image the blood-flow – e.g. by a time-based display of newly segmented vessels – such imaging can potentially enhance the diagnostic capabilities as one obtains thus also a visual indicator of the speed at which the blood flows.

Blood-flow analysis by contrast bolus tracking. If the vessels are segmented in each angiogram of the exam-sequence, then one can show which vessels have appeared in each image, thus implicitly describing the flow of blood through each vessel. For example, time-coding the newly found vessels with the index of the analyzed frame, will have as effect that vessel which are reached by contrast agent later in the exam-sequence will appear brighter than others. Detecting bright vessels among darker ones is then a sign that some vessels in a certain region have been irrigated later than the majority of their neighbors, which may be a pathological sign. One obtains also an indication about the size of this delay.

4.5.2   Dynamic vessel segmentation: Methods

We describe two approaches to segment the vessels in each image of a DSA-sequences. The first one uses *change-detection*-related methods and the second one is based on *general-purpose vessel-segmentation* methods.

Using these segmentation results one can then *segment the entire target vasculature* and also *capture the blood dynamics* through the vessels, i.e., segment the blood-flow.

Segmentation by change detection

The straightforward solution to segment still-vessels in a DSA sequence is by change detection [8], [5]. As in this case special care is taken to properly register the mask image to the contrast images [31] and the imaging system has no eigen-movement, it can be safely assumed that most temporal variations in intensity are due to the appearance of contrast into the vessels. However, some can also be generated by sensor-noise (see Section 2.2.4). Consequently, vessel segmentation by change detection is still an ill-posed problem – as a pixel can vary its gray level both when is reached by contrast agent or due to noise – and it is thus afflicted by the tradeoff between true and false positives.

General-purpose change detection and its application to vessel segmentation   Change detection has been thoroughly studied in the context of motion estimation from gray-level changes in image-sequences [1]. In this case – under the assumption of a still camera – the temporal high-pass obtained by subtracting consecutive images from one-another ensures that slow changes in illumination, like e.g., the change from day to night do not influence the motion estimation problem and ultimately allows the analysis of motion from gray-level changes. At a

first glance, similar methods can be used for vessel segmentation by change detection in DSA-sequences.

Adaptive thresholds.   The simplest way to detect changes in a difference-image is by adaptive-thresholding. To improve the robustness of the threshold with respect to noise, the decision for a pixel should be taken while considering also the changes in its neighborhood. Thus, in a statistical approach one does not use the gray-level as statistic, but rather the variance-normalized sum of square differences – for Gaussian class-conditional probabilities – or the standard-deviation-normalized sum of absolute differences – for Laplacian class conditional probabilities – in a neighborhood [8].

The optimal threshold for each pixel can be determined in a MAP approach (Equation 2.12) by a likelihood ratio test (Equation 2.14). Such a threshold can adapt by means of the prior probabilities $Pr\left[\omega_c^i\right]$ and $Pr\left[\omega_u^i\right]$ of the decision changed c or unchanged u at pixel i.

Objects have usually a compact shape with smooth boundaries. Such prior knowledge can be included in the threshold selection by specifying the priors in terms of a Gibbs/Markov random field [88]. Again the previous change-mask and the previous detection results are used to adapt the fix threshold [8]. One obtains a set of thresholds thresholds which are stored in a table and used depending on the constellation of neighbors around the investigated pixel from the change-mask. Such an approach encourages compact and isotropic, patch-like objects, and is thus less suited to model vessels which are rather elongated.

Two-thresholds segmentation.   Assuming that compact connected objects move in front of a camera, one would expect that a pixel once it obtains the label changed will remain changed and thus it makes sense to adapt the threshold by lowering it for moved pixels and increasing it for unmoved pixels. Such considerations lead to a two-thresholds segmentation method [5] where the previous change-mask tells for each pixel which threshold should be used. Such modeling is more appropriate in some cases than in others. For example, if an object is constantly moving in front of a camera at a speed which is very small in comparison to the frame-rate, there is a high chance that most object-pixels, which have been marked as changed in the previous change mask, are still changed and then such modeling is appropriate. This is, e.g., the case of a person talking in front of a camera. However, if the object passes in front of the camera at a higher speed, as e.g., in the case of a car, then a a relatively high number of pixels that have been marked as changed in the previous change-mask are now unchanged – as the object has traveled away from this position since more than two frames ago – and are being compared against the smaller change-threshold, therefore potentially generating false-positives.

For bolus tracking, we are in such an unfavorable case, however as our ultimate goal is vessel segmentation and a vessel remains a vessel even after the contrast has left it, this is less critical. Also, we are interested in showing the flow of contrast agent, and then only newly segmented vessel-pixels – which are always selected by the higher threshold – need to be segmented. Therefore, for vessel segmentation, the two-thresholds algorithm results in a single, fixed-threshold approach.

Morphological processing.   Prior expectations in change-detection can be also considered by means of specific binary morphological image-processing. However, such processing is not

part from the classification method and it imposes compliance of the solution with the prior expectations rather than encouraging it [8].

General-purpose vessel segmentation

We propose here that each image in the DSA sequence is treated as an angiogram which needs to be segmented. Therefore, we follow the approach introduced in the beginning of this section and first *enhance the vessels* (see Section 4.2) and then *segment* them (see Section 4.3).

Vessel enhancement.   The temporal aspect, characteristic to DSA sequences, is used solely for high-quality background suppression. As usually, after background suppression follows vessel enhancement. The best enhancement methods use multiscale decomposition, which usually implies knowledge on the precise size of the analyzed vessels, or even better, a ground truth to allow the computation of the optimal set of parameters. All these can be achieved in a calibration step for different vasculature and setups of the imaging system. A good general-solution is to use the Hessian single-scale vessel map with an integration low-pass size of 15 and a derivative window-size of seven. (see Section 4.2.2)

Segmentation.   Segmentation is then achieved by unsupervised hysteresis-thresholding. The optimal parameters of the segmentation algorithm can also be determined during the calibration. Good results were obtained assuming 1% of the image surface definitely covered by vessels and 92% by background.

Here it is implicitly assumed that each analyzed image from the DSA-sequence shows vessels. To ensure compliance with this assumption, before segmenting an image, its variance is compared to a threshold. DSA-images showing vessels will exhibit a far larger variance than images showing no vessels. This threshold was empirically set to 0.8.

Segmentation of the target vasculature and of the blood flow

Until now we have shown how to segment vessels in each image of a DSA-sequence. Assuming that in each intermediary result vessel are represented by one and background by zero, to segment the entire target-vasculature one can simply add all intermediary results together and choose then only pixels with a gray-level above zero.

To segment the blood flow, one has to keep track of which new vessel are segmented in each image. Each time new vessel-pixels appear they are displayed alongside already-segmented vessels but with different gray-levels, e.g., the later the vessel pixel is segmented in the sequence the the brighter its gray-level in the result (see Figure 4.29). We obtain thus an image reflecting the evolution of the blood-flow during the sequence, i.e., we segment the blood flow.

4.5.3   Dynamic vessel segmentation: Results and discussion

We have segmented the blood flow both by change detection and by general-purpose vessel segmentation in three DSA-sequences. The results achieved by are shown in Figure 4.29.

In the case of change-detection, both the bi-threshold and the multi-threshold general change-detection methods were tested. However, in this case they returned similar results which – as already pointed out in Section 4.5.2 – makes them equivalent to a fixed-threshold approach.

(a)

(b)

(c)
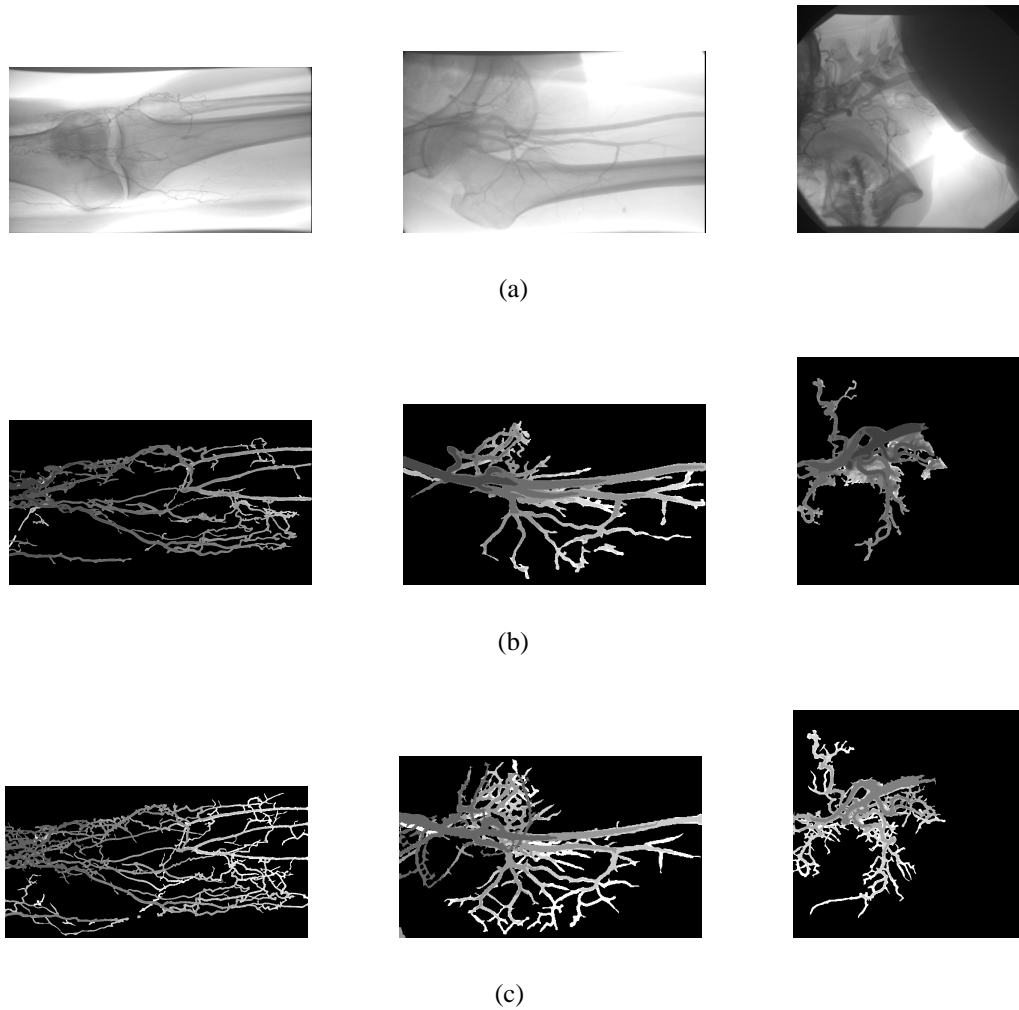
Figure 4.29: Dynamic vessel segmentation and blood-flow segmentation results. We show, one of the frames where vessels can be best observed in each of the analyzed sequences (a), the result obtained by change-detection (b) and the result achieved by general-purpose vessel-segmentation (c) for three different sequences.

In comparison to change detection, the general-purpose vessel segmentation approach is more stable with respect to parameterization. For change detection, no single parameter-set was found to return good results for all three sequences. The results obtained by general-purpose vessel segmentation usually yield more true vessels. The quality of the results was estimated by visual analysis.

The results show vessels crossing each other. This is a consequence of the imaging method, as vessels separated in depth are projected one on top of the other. A flow chart of the algorithm is shown in Figure 4.30.

4.5.4 Dynamic vessel segmentation: Conclusions

The best way to segment vessels in a DSA sequence is by general-purpose vessel segmentation. We believe that this is due to the use of vessel-specific prior knowledge – mainly in the vessel
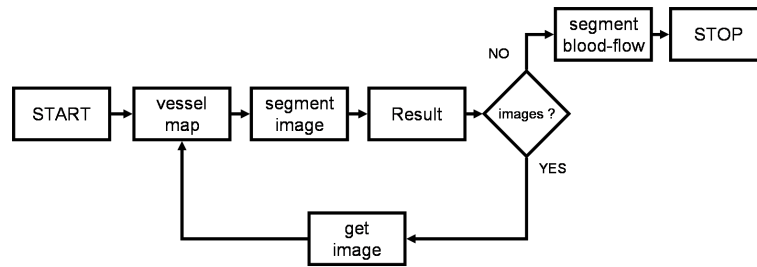
Figure 4.30: Flow chart of the dynamic vessel segmentation algorithm. Segmentation of the blood-flow includes segmentation of the target vasculature.

enhancement step but also during segmentation.

From each contrast image of the sequence the registered mask is subtracted. In the difference image, vessels are darker than their surroundings. A small median filter is applied to reduce the quantum noise before enhancement by single-scale Hessian analysis and segmentation by unsupervised hysteresis thresholding. To capture the flow of contrast agent/blood through the vessels, each time new vessels are segmented they are added to the segmentation result with different gray levels. The entire vessel tree can then be obtained from this results as all pixels with gray-levels above zero.

To achieve meaningful results the gray-level changes should be generated only by the contrast bolus passing through the vessels. Thus, before segmentation potential patient-motion or motion of the inner organs needs to be compensated. This is achieved during registration between the mask and the contrast images.

We have shown that the blood flow through still vessels can be also segmented from DSA-sequences. One obtains thus information with respect to the dynamics of blood through the investigated vasculature, a result which can be used for enhanced diagnosis of the artery disease.

Chapter 5

Quality inspection for xenograft valve-implants

The valves of the heart play a major role in the cardiovascular system. In many cases, an ill heart valve needs to be replaced during open-heart surgery by an implant. There are several types of implants, both mechanical and biological. From among the biological implants the xenograft ones – i.e., valves received from animals, e.g., pigs – are largely used. (see Section 5.1)

A good implant should exhibit some typical anatomical and functional characteristics, such that it can successfully replace the native tissue [145]. We describe here machine-vision-based, automatic methods for the measuring of quality parameters of heart-valve implants – mainly xenograft implants – including the area of the orifice (see Section 5.2) and the fluttering of the valve's leaflets (see Section 5.3). These methods provide a precise and reproducible way to infer the quality of an implant and can be therefore used for the analysis and automatic quality control of heart-valve implants. The fluttering analysis of the leaflets offers new opportunities for an enhanced quality inspection. We concentrate mainly on valves with three leaflets, i.e., the aortic, pulmonary and tricuspid valves.

5.1   Introduction

5.1.1   The heart valves and their role in the cardiovascular system

The rhythmic contractions of the heart move the blood through the vessels, such that it delivers nutrients and other essential materials to cells and removes waste products. In this respect, the heart is actually a pump containing four pressure cavities grouped by two into the left and right side of the heart. The two sides of the heart are separated by a septum, such that the blood cannot flow between them. The two cavities in each heart side are called atrium and ventricle respectively and are in turn separated by uni-directional valves.

The right side of the heart transports deoxigenated blood to the lungs and the left side transports oxygenated blood from the lungs. The deoxigenated blood coming through the veins is collected into the right atria and pushed through the tricuspid valve into the right ventricle and from there through the pulmonary valve into the pulmonary artery and then into the lungs. From the lungs, the oxygenated blood flows through the pulmonary veins into the left atria and from there it is pushed through the mitral valve into the left ventricle. The contraction of the left ventricle pushes the blood through the aortic valve into the aorta (see Figure 5.1).
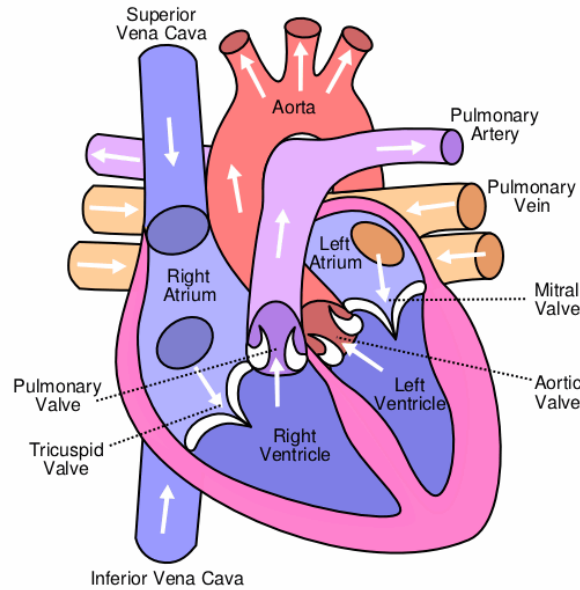
Figure 5.1: Representation of the heart showing chambers, valves and vessels.

The mitral and the tricuspid valves, which separate the atria from the ventricles, are called atrioventricular valves. They are anchored to the ventricular wall by cord-like tendons (chordae tendinae), which prevent them from opening in the reverse direction. The aortic and the pulmonary valves on the other hand have no such anchor mechanism.

The name tricuspid points out that this atrioventricular valve has three cusps (i.e. leaflets) and accordingly the mitral valve is sometimes also called the bicuspid valve. Both the aortic and the pulmonary valves have three leaflets and are thus tricuspid valves, but to avoid ambiguities only the atrioventricular valve separating the right atria from the right ventricle is called the tricuspid valve. However, in some individuals the tricuspid valve has two or even four leaflets, but this is not considered a valve condition, as long as the valve does not exhibit major leakage and the opening area is large enough.

5.1.2   The heart valve disease and its treatment

Pathology.   The pathology of a heart valve includes endocarditis, stenosis and insufficiency. In the case of stenosis, leaflets are fused together or they stiffen and do not open properly anymore. In the case of insufficiency, the valve does not close tightly anymore and relatively large amounts of blood – a small amount of leakage is normal particularly in the case of the atrioventricular valves – can flow in the reverse direction. Consequently, the heart stress increases with dire consequences, depending on the gravity of the condition. In the case of endocarditis the valve-tissue inflames.

Operative treatment.   If the condition is severe enough, the physician may decide to replace the diseased valve by an implant. This is then done during a major open-heart intervention, associated with a relatively high amount of risk. As of 1996, in the US there have been some 60,000 such interventions each year [192].

(a) Stented xenograft implant   (b) Stentless xenograft implant   (c) Mechanical implant
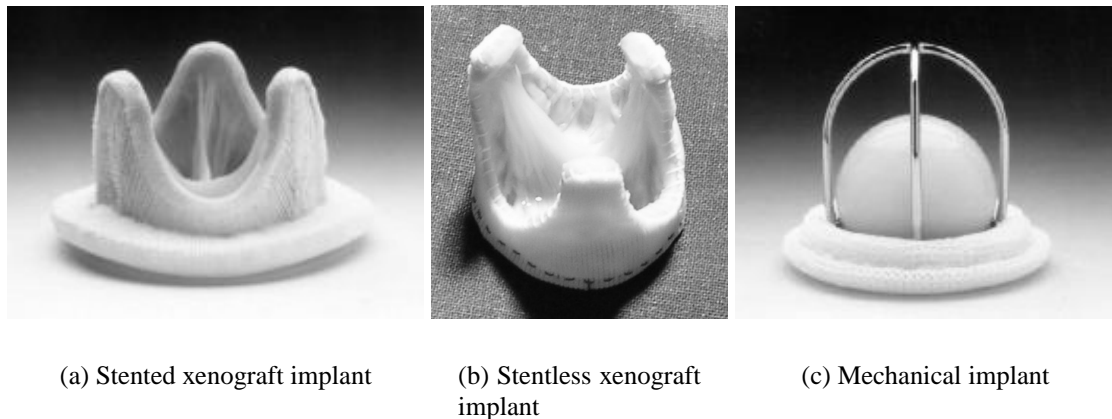
Figure 5.2: Examples of prosthetic heart valves.

Medical studies show that the implants do not reach by far the reliability of the native valves, mainly because they do not have the autorepair abilities of the native tissue. Thus in many cases, after a certain period of time complications reappear and a new intervention is eventually necessary.

Prosthetic heart valves. Since 1950 there have been developed some 80 types of implants. There are both mechanical and biological implants.

With respect to their origin, the biological implants are of three types: *xenograft implants*, which consist of similar tissue collected from specially raised animals [1] most often pigs, *homograft implants*, which consist of tissue received by donation from a deceased person and *autograft implants* where, e.g., the own pulmonary valve is used to replace the aortic valve and another – usually homograft – implant is used instead for the pulmonary valve. Such an exchange is advantageous as the aortic valve is subject to more stress than the pulmonary valve and at the same time, the native tissue is more robust.

With respect to their construction, the biological implants are of two types: *stented implants*, where the valve is placed on a support frame and *stentless implants*, without the support frame. Currently it is believed that stentless implants exhibit better characteristics than stented implants [64], [191], however this remains to be confirmed by long-term clinical studies.

Some implants are shown in Figure 5.2. The valve in Figure 5.2 (b) is one of those analyzed within the Results-section of this chapter.

5.1.3 Quality control for biological implants

Clearly, there is a major interest in the quality of these implants, as this is directly related to their life-span and the occurrence of complications for the patient.

From among the biological implants the xenograft implants are the most often used. To make sure that only optimally suited implants reach the patients, the characteristics of such potential implants are verified before them being accepted. An implant can be intrinsic-deficient,

---

[1]Clearly the food-processing industry can also offer such implants. But then the animals are not raised in proper conditions, e.g., they are administered different drugs, or take different substances through their food, which are deposited in the body and may be then harmful for the patient which would receive the implant.
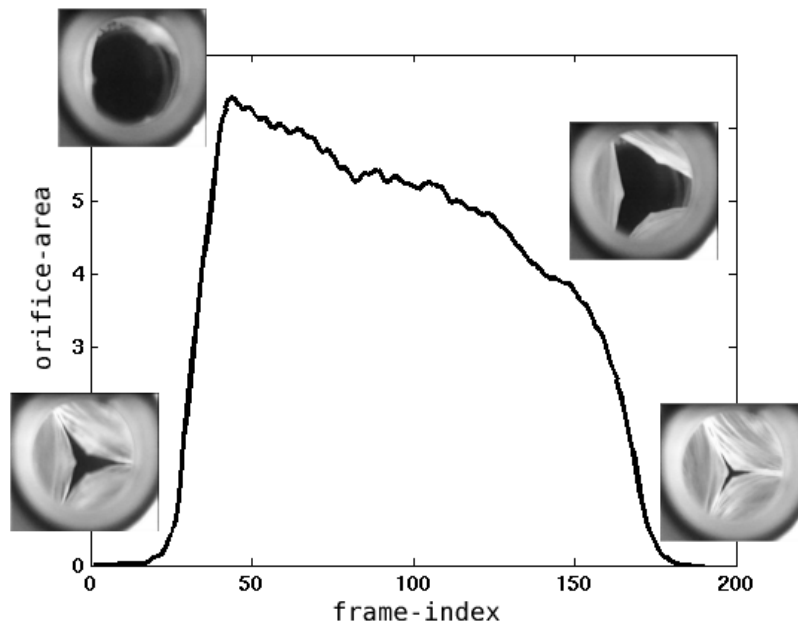
Figure 5.3: Evolution of the orifice area over the frame index in a test sequence.

i.e., if the donor was itself ill or extrinsic-deficient, e.g., if the heart valve was not placed properly.

Quality criteria

There are two major quality criteria which characterize a heart-valve implant [147], [148], [167], [78]:

1 **The area of the orifice** since the moment the valve opens and until it closes (see Figure 5.3).

2 **The fluttering** of the leaflets in the blood flow (see Figure 5.4).

Orifice area.   During the period when the valve is open, the area of the opening has to reach a certain maximal value. It has to evolve then in a certain predefined way in a certain interval of values. The typical evolution of the orifice area over a *valve cycle* – i.e. the time interval from the moment the valve starts opening until it is completely closed – may be observed in an *orifice-area curve* and can be divided into three parts. In the first part, starting from a closed position, the valve opens quickly, thus the area of the orifice has a steep increase. After reaching a maximal value, the area of the orifice decreases slowly over most of the remaining second-part. Finally, the valve closes during the third part and the orifice area decreases, but with a smaller slope than during the opening phase. Such a behavior can be observed on the orifice-curve shown in Figure 5.3. A valve which does not open properly will eventually force the heart to work harder, with negative consequences for the patient's health.

Fluttering of the leaflets.   Without further physical experiments, we hypothesize from our data that there are two types of fluttering: *eigen-fluttering* and *jerky-fluttering*. Valves may

(a)   (b)
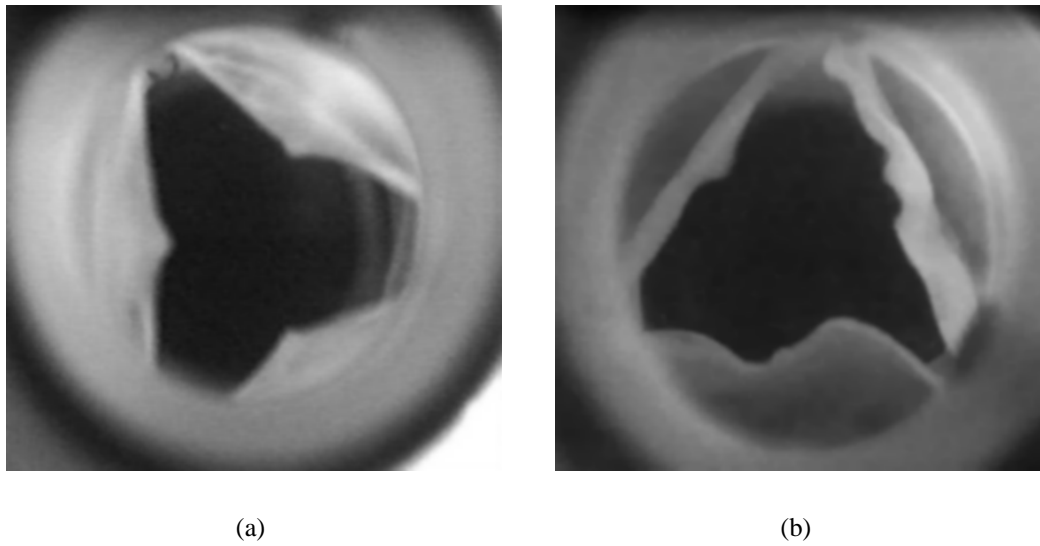
Figure 5.4: Examples of valves without fluttering (a) and with fluttering (b).

exhibit both types of fluttering or any of the two alone. A good implant should not exhibit any jerky-fluttering.

Assuming the leaflets are elastic, the flow of liquid through the valve is usually turbulent and can be thus modeled by a wide-band noise-signal which excites them. If the excitement signal contains also large-enough components of one of the eigen-vibration frequencies of the leaflets, these will start to vibrate as a whole, i.e., without major changes at the borders of the leaflets. Such *eigen-fluttering* has arguably a limited impact on the quality of a valve, as in this case the leaflets move over their virtual base-line, i.e., the line connecting the two end-points of a leaflet's border (see Figure 5.12 (a)). There is also *jerky-fluttering*, which is generated by the flow of liquid, the same way as the eigen-fluttering. This time though, the liquid flows through a valve with poor elastic properties, which has no eigen-vibration modes. In this case, the leaflets bend also along lines perpendicular to the base-line (see, e.g., the right and the lower leaflet of the valve-implant shown in Figure 5.4 (b) and the leaflets in Figure 5.11). This type of fluttering causes major stress on the leaflets – therefore reducing the longtime durability of the implant – and influences also the flow of blood – with negative consequences for the health of the patient receiving the transplant – being thus a major issue in heart valve quality control [191].

Fluttering affects valves with poor elastic properties [147], [148]. The elastic properties of a valve depend both on the valve-material and on the way the valve is anchored, i.e., if it is spanned. In the case of implants, usually it is not the valve-material which has poor elastic properties, but the valve is not fixed properly. For some valves, theoretically, the fluttering can be corrected by positioning better the valve on the implant. Thus, it is also interesting to know which leaflet flutters besides how much fluttering is there. For example, if two leaflets of a tricuspid heart valve do not flutter and the third flutters strongly, then it can be assumed that the valve was not placed properly.
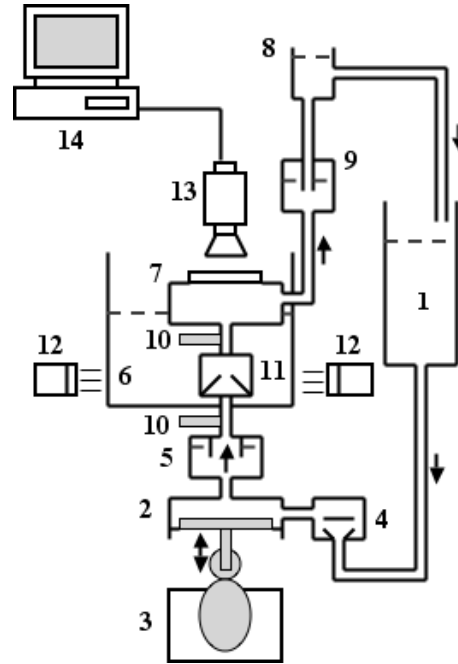
(a)   (b)

Figure 5.5: Experimental setup (a) and its schematics (b).

Imaging

Imaging is done with the help of a special test-setup [167], which is shown in Figure 5.5. From a reservoir (1) a transparent fluid – usually water – is transported through a disc valve (4) by a piston pump (2), which is driven by a waveform adapted cam plate (3). After passing an input compliance (5) the fluid is pressed through the inspected heart valve (11) into a visualization chamber (7) located in another fluid reservoir (6). Pressure sensors (10) are installed below and above the heart valve. Passing an aortic compliance (9) the fluid reaches a height variable column and flows back to the first reservoir. The heart valve is illuminated by light sources (12) outside the fluid tank (6) made from perspex. A high-speed video camera (13) takes images of the heart valve. The digitized images are stored on a PC (14).

Images of the heart valve are taken with 500 fps and an interlaced resolution of 480x420 gray pixels. A *test sequence* (see Figure 5.3), which shows one valve cycle, has a duration of some half a second leading, to approximatively 250 frames per sequence. Imaging by an interlaced camera introduces artefacts on moving objects. In this case, such artefacts can be observed on the margins of the leaflets. Illumination is provided by Neon-lamps. Such an illumination modality introduces also artefacts, as the 60 Hz flickering of the lamp is observable in the 500 Hz image sequence.

The sequences analyzed within this chapter have been acquired with the test-bed already described, therefore they show interlace- and illumination-related artefacts. Interlace-related artefacts are strongly attenuated by an opening operation with a disk-like structuring element of small size. To reduce the influence of the flickering, for each image a certain value is added or subtracted such that the mean gray-level value in a certain region – which under continu-

ous illumination should have constant gray-levels – remains constant over the entire sequence. Clearly, such problems can be avoided by improving the test environment, such that it uses a progressive-scan camera and a continuous light-source for illumination.

Usually, the orifice appears way darker than the rest of image items. However, there are also cases when it is strongly inhomogeneous, as some bright artefacts are visible (see Figure 5.11). These artefacts are generated by mechanical parts of the test environment which reflect some light. They can be easily eliminated, e.g., by painting them black, or by slightly modifying the test-setup. Therefore, they have been ignored while designing the algorithms.

Data analysis

To this date, usually only the area of the opening is measured manually [167]. To compute the opening-area, when evaluating, e.g., a tricuspid heart valve, the human operator has to mark the six points at the connections between leaflets and at the peak of each leaflet. Including also the midpoint of the opening, one can define like this six triangles, which approximate the area of the orifice. For each analyzed heart valve, this procedure, which neglects the rather curved boundary of the leaflets, has to be repeated for some images properly chosen over the test sequence. From such data, the behavior of the implant in the entire sequence is inferred. Even for a few images, this is a rather tedious job and it clearly returns imprecise results, because it does not consider the entire sequence.

During such a test-run, the fluttering is usually analyzed only visually by the operator if at all. In some special cases – related to medical research rather than medical routine – the fluttering is also analyzed with the help of a Bending Deformation Index (BDI) [78], which is again manually computed only for a few images.

We introduce here novel methods for machine-vision-based support of the quality inspection of xenograft heart-valve implants. We show how to measure the orifice area and the fluttering of the leaflets automatically. Such measurement will allow for an improved human-based quality control, as they are precise and offer reproducible results. Also, an analysis of fluttering becomes now available in medical routine. They may represent also the basis for fully automated analysis which can be conducted fast and for a large number of valves. This is clearly interesting for companies producing such implants.

5.2 Analysis of the orifice area of a heart valve

To obtain orifice-area curves, one needs to segment the orifice in each image of a test sequence. The inspected valve, placed on its supporting frame, is positioned inside a tube through which a transparent liquid flows. The portion of the tube downwards from the valve with respect to the camera is protected from light, such that when the valve opens the orifice appears dark, while the valve itself is light as it receives illumination (see Figure 5.5).

We describe two classes of methods for the segmentation of the orifice: (i) *thresholding-based* methods (see Section 5.2.1) and (ii) *active contour-based* methods (see Section 5.2.2). From each analyzed image-sequence we obtain one orifice curve (see Section 5.2.3), which can then be used to judge the quality of the inspected valve.

<div align="center">(a)   (b)   (c)</div>
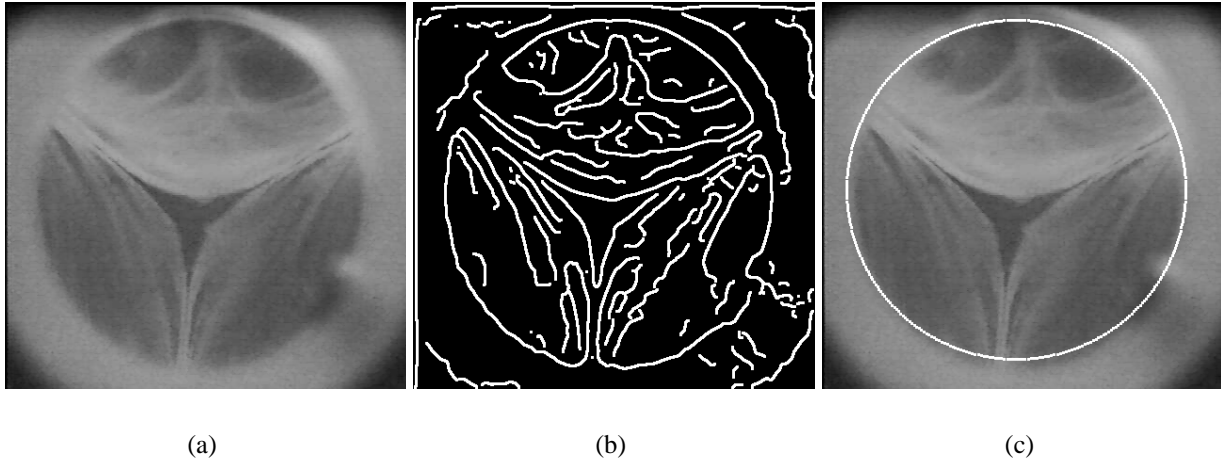
Figure 5.6: Original image (a), edge detection result (b) and Hough-segmented outline (c)

5.2.1 Segmentation of the orifice area by thresholding

An image of a test sequence shows besides the valve, also parts of the pipe through which the water flows and some of the area outside this pipe. The pipe appears similar to the valve and the area outside the pipe similar to the orifice (see Figure 5.6 (a)). To focus our analysis on the valve alone, we establish a *valve-ROI* and ignore the rest of the image. Since on each analyzed image we expect to see only two pixel classes, i.e., orifice and non-orifice/valve, the straightforward method to segment the orifice is by gray-level thresholding [95]. Thresholding can still be used if the underlying model is no longer *bi-class* but *multi-class*. In this case one has to use the prior information that the orifice gray-levels are the darkest ones in the valve-ROI.

One can either threshold each image individually, or establish one global threshold for an entire sequence. As we do not expect to see an orifice in each of the analyzed images, a global threshold is better suited.

Establishing a valve-ROI

The camera observes the valve through a cylindrical tube, using a normal-imaging optical system (see Chapter 2). Thus, the inner wall of the tube is also observable and delineates the contour of the valve. We use this circular outline as a mask and ignore everything outside it.

We segment the circular outline by the Hough transform for circles [102], [106] (see also Appendix A.4). As the approximate position of the center of this circle and size of its radius are known a priori, the search in the Hough-space is very fast. We search the coordinates of the center in a ROI of size 15 pixels centered at the image center and the radius on an interval of the same size centered at the value 150. The center and radius of the outline circle is then estimated by taking the mean over several images. This defines a circular ROI where only the valve is visible. An example is shown in Figure 5.6 .

Thresholding with a bi-class model

Using the global histogram of a sequence, which is computed from the gray-levels of all images, the threshold can be established either *manually* by the user or *automatically*. In some

(a)   (b)   (c)
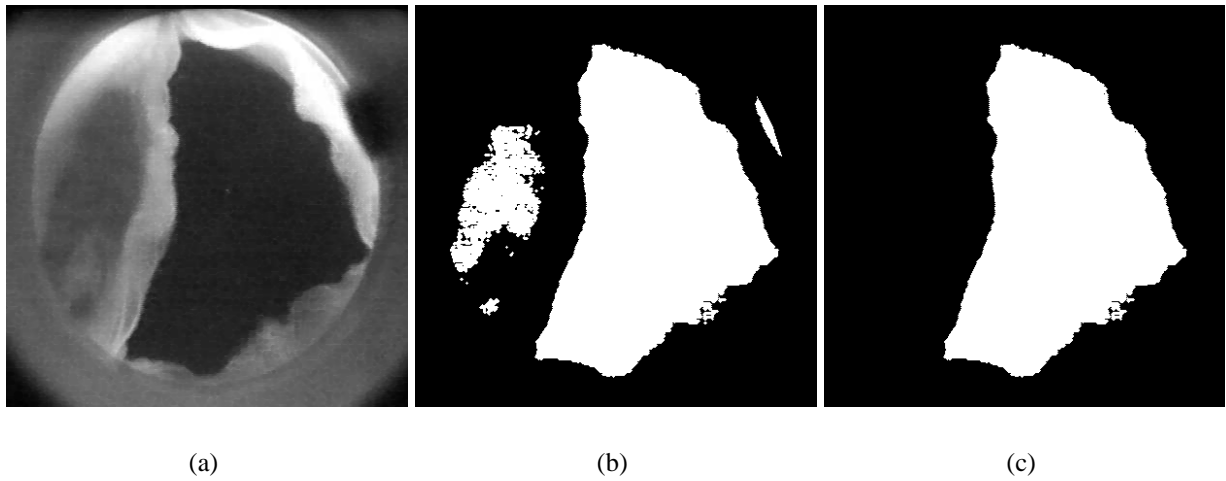
Figure 5.7: Valve image (a), thresholding result (b), and final result after center-based selection (c).

sequences the results achieved by a global threshold include relatively often other image structures besides the orifice, e.g., parts of a leaflet, as shown in Figure 5.7 (b). These artefacts are usually unconnected to the orifice and can be eliminated from the segmentation in a *postprocessing* step.

Manual threshold.   If the threshold is selected by the user, then this is done by trial and error. He may first observe the global histogram and guess a certain value for the threshold, then inspect the results and adapt his decision until satisfied with the results obtained. Clearly, in this case the procedure is tedious and the results not reproducible.

Automatic threshold.   An automatic procedure can solve such drawback. Assuming that only two pixel classes are observable in the analyzed sequences, one can set a global threshold automatically by the Otsu-method [155] (see also Appendix A.6.2).

Postprocessing.   We know that the orifice is a connected structure. As the valve is imaged in a centered position and thus the image center will be practically always contained within the orifice, we use the center-pixels – i.e. the pixels situated in a circle of radius five pixels centered at the image center – to select from the initial raw segmentation only those connected structures containing at least one of them. We use such a circle to account for the variability in the positioning of the valve. This plays a major role particularly in images where the orifice is very small, i.e., at the beginning and at the end of a sequence. The rationale behind this approach are contained within the hysteresis paradigm. This is shown in Figure 5.7. The initial thresholding result contains also other structures besides the orifice (Figure 5.7 (b)). After center-based selection, only the orifice remains (Figure 5.7 (c)).

Thresholding with a multi-class model

Clearly such an approach – i.e. global threshold followed by postprocessing – will fail if the spurious structures are linked to the orifice. Such an example is shown in Figure 5.8 (b). Such

(a)   (b)   (c)
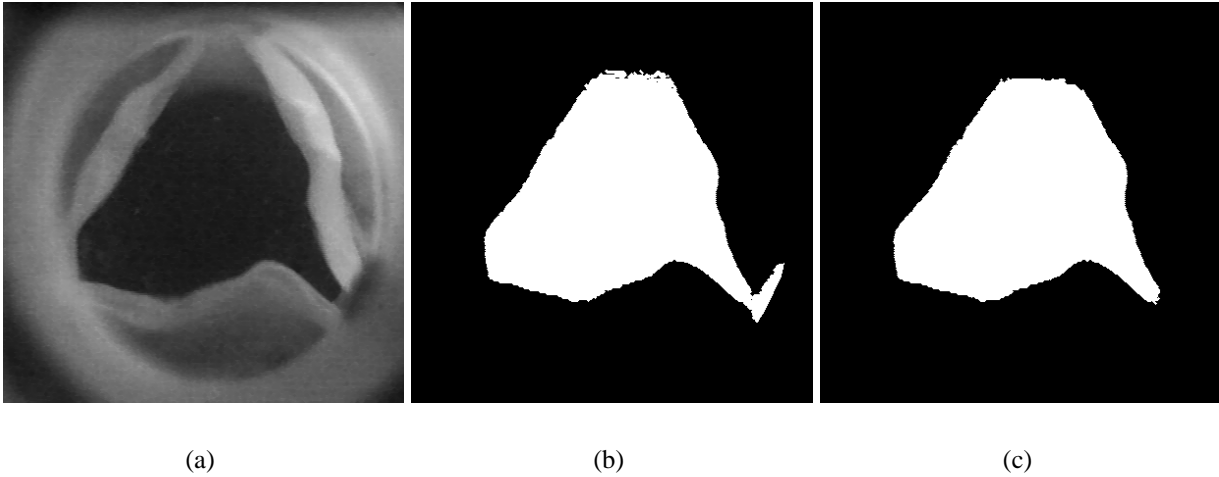
Figure 5.8: Original image (a) segmentation result by the Otsu-method (b) and by the FMM-method (c).

over-segmentations show also that our initial assumption, that there are only two pixel-classes in the analyzed images, is sometimes incorrect and there are actually *several classes present*. However, one can usually safely assume that the darkest pixels in the valve-ROI will always be orifice pixels. Therefore, if we use a *mixture model* of the global-histogram, to *segment the orifice* we have to separate the mode with the darkest mean gray-level from the rest.

Mixture-models.   We would like thus to model a multi-modal histogram and select the darkest mode. We know a priori neither the number of modes nor the type of the mixture. To proceed, one can assume a certain type of mixture (e.g. Gaussian) and determine its parameters by the EM algorithm [65], [20], [21]. One needs to determine also the optimal number of modes k. The maximized likelihood cannot be used directly for this purpose because it is a nondecreasing function of k [111]. However, one can proceed by trying several values and choosing then the best suited one according to coding-theory-based mixture criteria such as: the minimum message length (MML) [152], [193], the minimum description length (MDL) [16], [163], Akaike's information criterion [13] and others [124]. There are also other criteria to find k, but which are less popular (see [79] and the references therein).

Such unsupervised classification is called mixture decomposition [111]. One needs to specify only the type of the mixture components. This separates it from other standard non-syntactical clustering algorithms, which usually need to know a priori the number of clusters. It is also readily applicable to a large class of problems, as it does not need any heuristic to form clusters – as in the case of many syntactical clustering methods.

The key element in mixture decomposition is the estimation of the number of modes k. This is done according to:

$$\hat{k} = arg \min_k \{C(\hat{\theta}(k), k), k = k_{min}, ..., k_{max}\} \tag{5.1}$$

with $\hat{\theta} = \theta_1, \ldots, \theta_k, \alpha_1, \ldots, \alpha_k$ the vector of parameters with α_i the mixing probability and θ_i the set of parameters of the i'th mixture component. The criterion C is a function including two components: (i) the maximized log-likelihood function and (ii) a term to penalize large values of k [111]:

$$C(\hat{\theta}(k), k) = -log(p(\mathcal{Y}|\hat{\theta}(k))) + P(k) \tag{5.2}$$

where $\mathcal{Y} = y_1, \ldots, y_n$ is a set of n independent and identically distributed data samples and $log(p(\mathcal{Y}|\theta(k)))$ is the log-likelihood function.

The rationale behind coding-theory criteria is that an optimal (i.e. short) code for transmitting some data can be achieved only with a good data-generation model [180]. If the data is generated according to $p(\mathcal{Y}|\hat{\theta})$ then the shortest code-length is proportional to the entropy of the source: $-log(p\mathcal{Y}|\theta)$. However, to establish a communication both parties need to know the code. One way to do this is for one party, e.g., the transmitter to start the communication by sending the characteristics of the source (i.e. θ) to the receiver so that it can build the code by itself. Then the length of the first message is:

$$L(\theta, \mathcal{Y}) = L(\theta) + L(\mathcal{Y}|\theta(k)) \tag{5.3}$$

All coding-theory criteria require to find $\hat{\theta}$ such that $L(\hat{\theta}, \mathcal{Y})$ is minimal. So, with respect to classification theory, one requires the optimal fit to the training data for a minimum of complexity of the fitting function, or the best classification results on a training set for the simplest classifier. Such principles, among others, lay the foundations for an entire theory on classifier design, culminating with the introduction of the SVM.

Thus, to segment some data, one has to compute first by the EM algorithm the parameters for each mixture-model from a set of candidate mixture-models, which differ from one another by the number of modes and choose the best one according to the model-selection criterion $C(\hat{\theta}(k), k)$. The set of mixtures to be tried should be defined based on some estimation of the minimal and maximal possible number of classes. Clearly, if the maximal number is very large, the procedure will take a very long time. Therefore, an optimal balance has to be found and prior information plays a major role here.

Finite mixture models. With respect to the way the optimal mixture decomposition is found, this approach has the disadvantage that each candidate mixture has first to be estimated by the EM algorithm, which is sensible to the initialization and whose convergence is difficult – in particular if the number of modes is larger than the true one.

Alternatively, it has been proposed to directly implement a mixture-criterion via EM. For this purpose, a modified MML criterion is used. The algorithm thus obtained is called the Finite Mixture Model (FMM) and it was first proposed in [79].

For a one dimensional Gaussian-mixture with μ the set of means, σ^2 the set of variances, α the set of mixing probabilities and $\mathcal{Y}$ the data, the criterion is:

$$L(\mu, \sigma^2, \alpha, \mathcal{Y}) = \frac{N}{2} \sum_{i:\alpha_i > 0} (\log(\frac{n\alpha_i}{12})) + \frac{k_{nz}}{2} \log \frac{n}{12} + \frac{k_{nz}(N+1)}{2} - \log(p(\mathcal{Y}|\mu, \sigma^2)) \tag{5.4}$$

with k_{nz} the number of non-zero (i.e. $\alpha_i > 0$) distributions, n the number of data-points and N the number of parameters specifying each component (i.e. two for 1D Gaussian distributions). Then, starting from m initial Gaussian modes, the algorithm fits the mixtures iteratively to the

data in the global histogram. At each iteration step t, the means and variances are updated for each distributions according to:

$$\mu_i(t+1) = (\sum_{j=1}^{n} w_i^j)^{-1} \sum_{j=1}^{n} y_j w_i^j \tag{5.5}$$

and

$$\sigma_i^2(t+1) = (\sum_{j=1}^{n} w_i^j)^{-1} \sum_{j=1}^{n} (y_j - \mu_i(t+1))^2 w_i^j \tag{5.6}$$

with $w_i^j = \alpha_i p_i(y_j) \cdot (\sum_{m=1}^{k_{max}} \alpha_m p_m(y_j))^{-1}$ and $p_i(y_j) = \frac{1}{\sigma_i \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y_j - \mu_i}{\sigma_i}\right)^2}$. α is upgraded also. The iterations continue until a minimum number of non-zero distributions is reached (usually two). Then the optimal parameter set is chosen as:

$$\hat{\theta} = arg \min_{\theta} L(\mu, \sigma^2, \alpha, \mathcal{Y}) \tag{5.7}$$

This algorithm is less initialization-dependent and shows a better convergence [79]. We justify the modeling by a Gaussian mixture with the prior knowledge that the orifice is characterized by similar gray-levels in each sequence, with some deviations from the expected orifice gray-level generated by varying illumination conditions. The same is valid for all other objects present in the analyzed images.

Segmentation of the orifice area.   After computing the optimal mixture-model, the threshold is found by a likelihood ratio test (see Section 2.4.1) using the two mixture components with the smallest means – thus separating the darkest mode from the rest, as orifice pixels have the smallest gray-levels in the analyzed circular ROI. This is then practically identical to the overall optimum, as the conditional probabilities of other mixture models around the threshold are always close to zero. This is shown for two examples in Figure 5.9.

When using the mixture-decomposition method, the over-segmentation is no longer an issue in the segmentation of the orifice of a heart-valve (see Figure 5.8 (c)). Consequently, one needs no center-based selection anymore. However, in some cases under-segmentation becomes now an issue. Such an example is shown in Figure 5.10.

5.2.2   Snakes for the segmentation of the orifice area

Threshold-based methods may fail because the orifice does not have a homogeneous gray-level representation, as it is not defined over a certain reflectivity, which can be then translated into an eigen gray-level distribution, but it is defined by the borders of the leaflets. The orifice is actually transparent and it will thus take over the gray-level of the structures against which it is imaged. Practice shows that it is rather difficult to achieve a homogeneous background to project the orifice against. There are also extreme cases, when some mobile parts of the imaging setup can be seen through the orifice, as shown in Figure 5.11. However, the imaging-setup can be easily modified such that these parts are no longer observable.

For a better segmentation one should thus start from the leaflets and select the orifice as defined by their borders, i.e., as the area enclosed within these borders. The border-shape curves cannot be described a priori as they have a large variability. Therefore, we use active

(a)   (b)
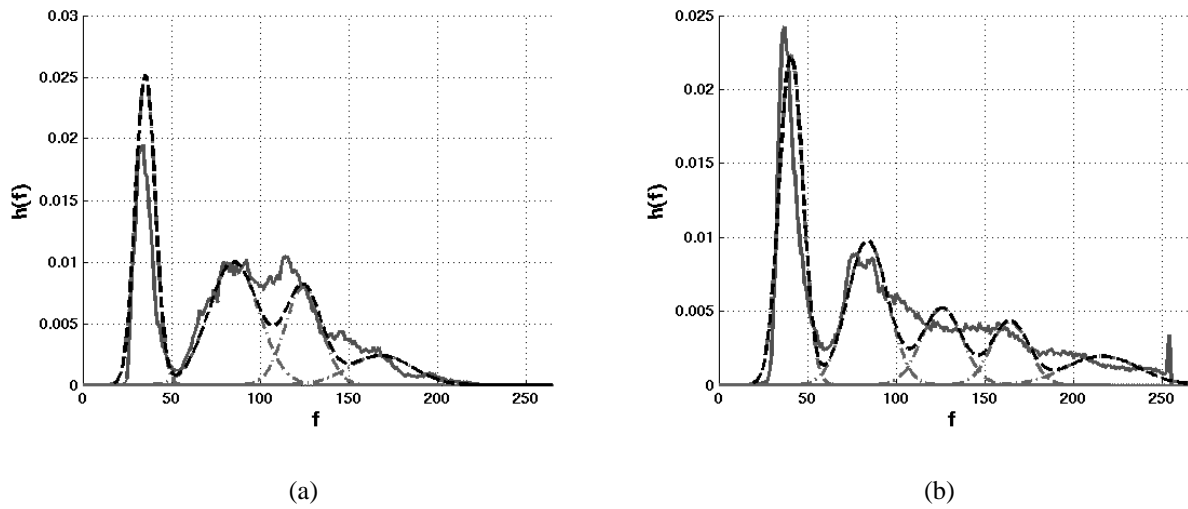
Figure 5.9: Outline of the normalized global-histogram (continuous line), estimated mixture (dashed line) and estimated mixture components (dashed-dotted line) for two sequences from our data-base. The threshold to segment the orifice is set such as to separate the darkest mode from the next darkest.

(a)   (b)   (c)
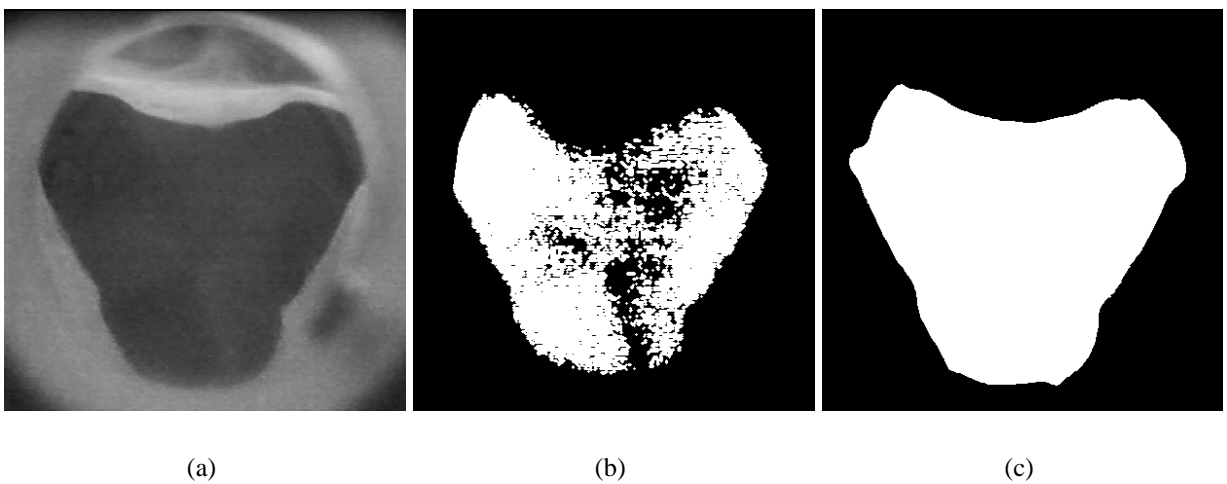
Figure 5.10: Original image (a), FMM-based segmentation (b) and snake-based segmentation (c).

(a)   (b)
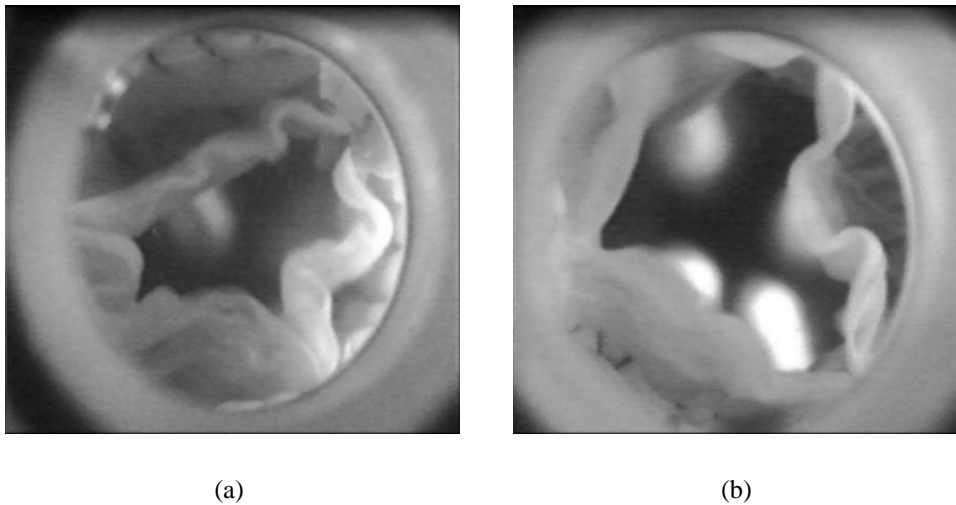
Figure 5.11: Examples of fluttering valves with an inhomogeneous orifice. These types of inhomogeneities can be easily corrected by slightly modifying the test setup.

contours/snakes [119] (see also Appendix A.3), which are well suited to track them over the analyzed sequence [142]. As external force we use a Gradient Vector Flow (GVF) [197], [198] (see also Appendix A.3), which has an improved attraction range in comparison to other external forces [114].

The snake has to be *initialized* automatically and only after the valve has opened enough to see the orifice. We initialize the snake from the leaflet *base-lines*, which are the lines connecting each of the anchor points (see Figure 5.12 (a)). *Anchor points* are the points where two leaflets meet (see Figure 5.16 (f)). After initialization, the snake will *track the outline of the leaflets* (see Figure 5.13 (b)) until the valve closes.

Initialization

The active contours need to be initialized after the moment when the valve starts to open, when enough of the orifice is observable. Initialization can be done automatically, e.g., based on the border line of the orifice, as segmented by thresholding in an first image, or starting from the *base lines* of the leaflets.

The snake can track the outline of the leaflets only in images where this is well observable. Therefore, for a successful initialization one has to wait for the valve to *open enough* and then halt the snake at the end of the sequence, when the opening is *no longer large enough*. To *reduce the time interval* since the moment when the valve opens until the snake is initialized, we use a balloon pressure force [42] (see also Appendix A.3) to push the base line towards the borders of the leaflets.

Initialization from the base line. Clearly, the initialization from the border line of a segmentation result may be erroneous, if the segmentation result is erroneous. The initialization from the base line has the additional advantage of placing snakels already in regions where the borders of the leaflets are not well defined, e.g., they are imaged against parts of the tube

(a)   (b)
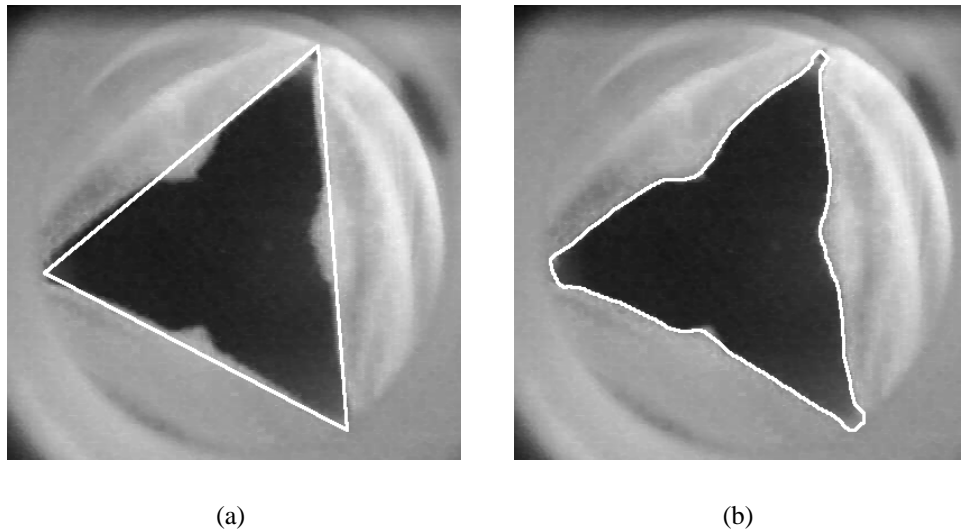
Figure 5.12: Initialization from the base line and the corresponding snake result.

seen through the orifice (see Figure 5.20 (a)). In our experiments we have used the base line to initialize the snake. This is shown in Figure 5.12 (a).

Detection of the frames where to start and stop the tracking.   To find out when to initialize the snake, we follow the mean-value of the gray-levels within a circle, defined such that it is maximally enclosed in the triangle given by the base lines. If this value is beneath a certain threshold, then we conclude that the valve has opened and initialize the snake. The threshold is chosen as the median of the center mean-values recorded over the first 25 frames of a sequence. At the end of the sequence, if the mean value is above the threshold, the snake-based tracking is halted.

Timely initialization of the snake.   The initialization from the base line can be done only relatively late after the valve has open, such that the base lines fall in the attraction range of the leaflets. This can be improved by using a balloon snake pointing towards the center of the image. However, such a balloon has to be carefully parameterized to avoid a collapse of the snake. After the base line initialization, the position of the snake in the previous image is used as initialization for the current image.

Snakes-based tracking of the leaflet border

After the snake has converged in the initialization frame, its position in the past image is used as initialization for the next one and so on. To avoid a collapse of the snake, it tracks the leaflets until shortly before the valve closes. The GVF is used as external force field, conferring robustness to the segmentation through its improved attraction range and concavity behavior [198].

During tracking the snake cannot follow the valve orifice in regions of small contrast, i.e., where portions of the pipe behind the valve are visible through the opening. This is shown in Figure 5.13 (a). As *balloons* are unsuited to solve such problems in this case, we show how to

(a)   (b)
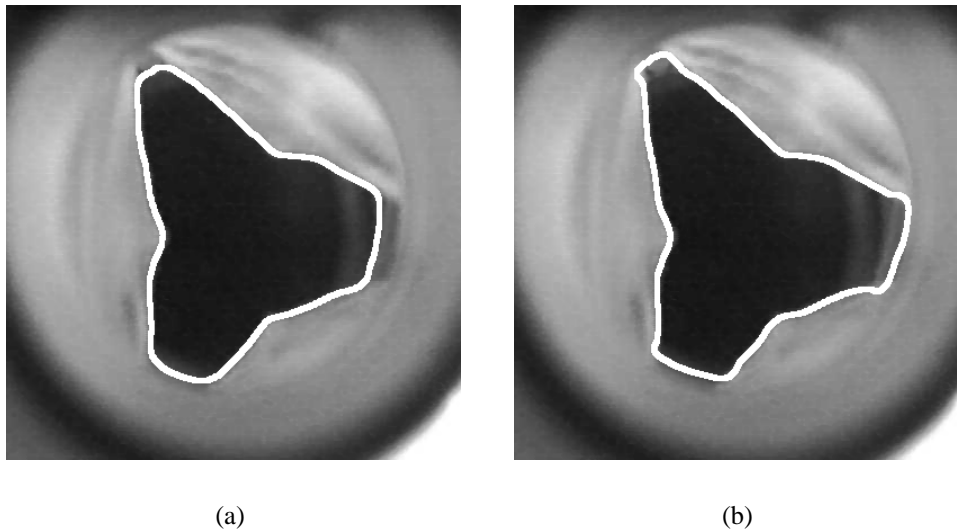
Figure 5.13: Snake segmentation result without (a) and with (b) attractor.

build *attractors* for the snake around the *anchor points*, such that the snake tracks the borders of the leaflets correctly, as in the example shown in Figure 5.13 (b).

Balloon-supported tracking.   The usual solution for solving such convergence problems, is to use a balloon, if the GVF is not enough. This time, the balloon should be pointed outwards from the image center. However, such an approach is problematic when the valve starts closing. It may happen that due to the balloon in some cases the snake loses the orifice. This will lead to a complete failure of the analysis, due to the way the snake is initialized, i.e., from the previous frame. Initializing the snake each time from a small circle centered at the image center has the potential to solve such problems at least partially, as in this case the tracking will fail only for single frames. Generally, balloons need to be used very carefully and usually only when the sought features are strong enough to yield a good attraction basin.

Anchor-points-based attractors for better tracking.   Alternatively, we propose to build artificial attractors around the anchor points. For this purpose we use the circular outline, which was segmented during initialization. The artificial attractors are actually edges, constructed from parts of the circular outline in the vicinity of anchor points. These vicinities are circular in shape and have a radius of 125 pixels.

To begin with, the circular outline is dilated slightly, starting from its initial one pixel width, to increase its attraction range. Then, parts of the outline around each anchor point are selected. The selection includes those parts found in a circle of radius 125 pixels around each anchor point. Finally, they are subtracted from the original image such that their mean gray-level is equal to the mean gray-level in the dark parts of the orifice. An example is shown in Figure 5.14.

We do not use the entire circular outline as one attractor, because such a strategy may lead to failure, e.g., if one of the leaflets can no longer be observed as it travels behind the circular-outline, then the snake will reach and possibly remain there even if the valve starts closing. Such an example is shown in Figure 5.15.

(a)   (b)
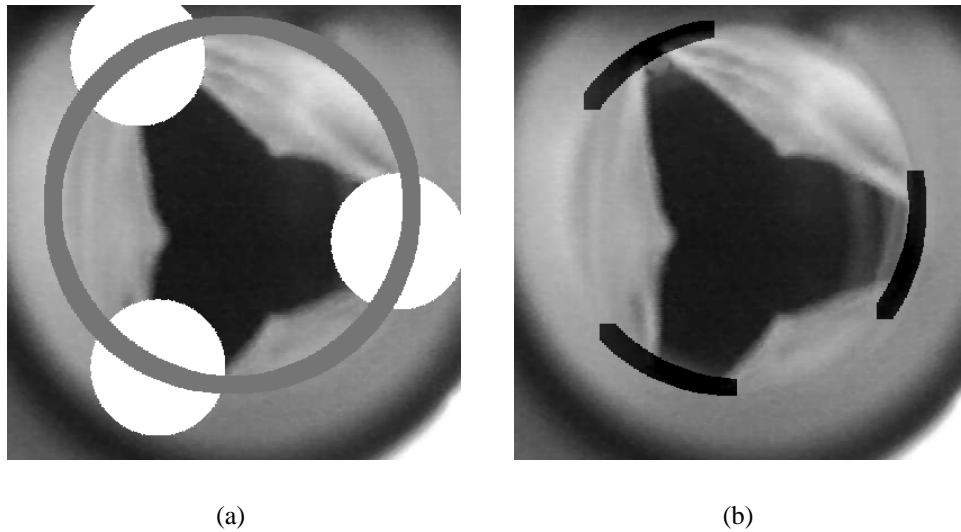
Figure 5.14: Construction of attractors (a) and final result (b).

Automatic determination of the anchor points

The anchor points are needed during initialization to define the base lines, but also during track-
ing to build the attraction basins which constrain it. The anchor points are established auto-
matically. For this purpose, we use the first 11 frames after the valve has opened, i.e., before
the valve is completely opened. For these images we can safely assume that the orifice has a
homogeneous and unique gray-level representation and segment it by thresholding. One such
example is shown in Figure 5.16 (b).

We start by segmenting the circular outline. We then find the center-lines of the orifice by
dilating and then thinning the segmentation result (see Figure 5.16 (c)). The dilation is necessary
to ensure that only the centerline remain after thinning. Then, we use the Hough transform for
lines to detect the three center-line rays and compute the points where these rays intersect the
circular outline (see Figure 5.16 (d)). Then, for each ray and for every point where it intersects
the circular outline, the distances from the intersection point to all center-line points along the
ray are computed. For each intersection point only the minimum is considered. Anchor point is
then the intersection point with the minimum value (see Figure 5.16 (e)). The final positions of
the anchor points are determined by taking the mean position over 11 frames for each anchor
point. This is shown in Figure 5.16 (f).

5.2.3   The orifice curve

The orifice curve is the curve of orifice area over frame index (see Figure 5.21). During one
heart-beat, a valve goes through a complete cycle: from opening to closing. As the analyzed
sequences show the evolution of the valve during one simulated heart beat, the orifice curve
reflects the dynamic evolution of the area of the orifice in one *valve cycle*.

There are two quality criteria which can be measured on such curves. First is the *maximum-
value* and then the *dynamic evolution* of the orifice. For a healthy valve, the maximum value
has to be within a certain predefined interval, such that a proper blood flow is obtained without

(a)   (b)

(c)   (d)
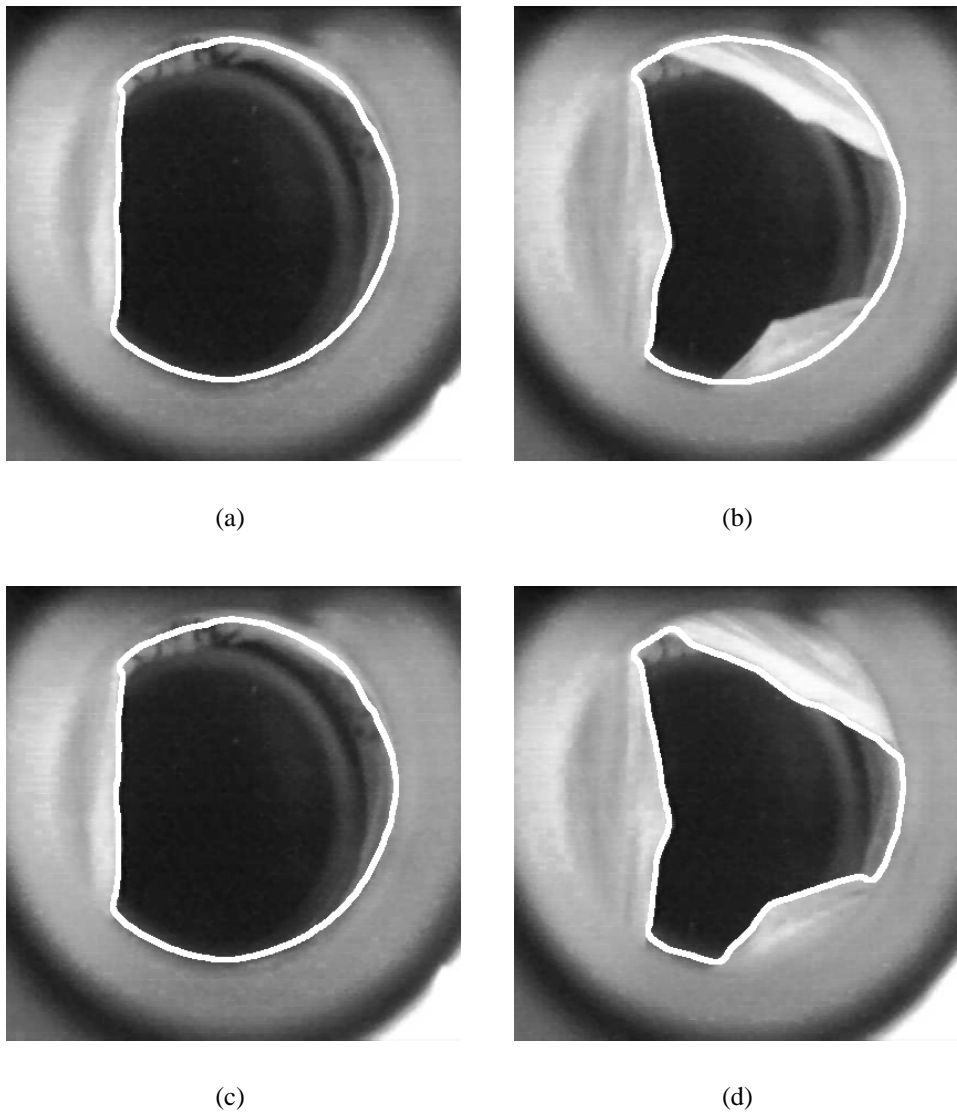
Figure 5.15: Snake result for two frames using the entire circular outline as attractor (a) and (b) and the same images using only the portions of the circular outline around the anchor points as attractor (c) and (d).

(a) Original image.

(b) Segmented orifice.

(c) Result of morphological processing.

(d) Intersection of center-line rays with the outline of the valve-ROI.

(e) Minimum-distance points.

(f) Final result.
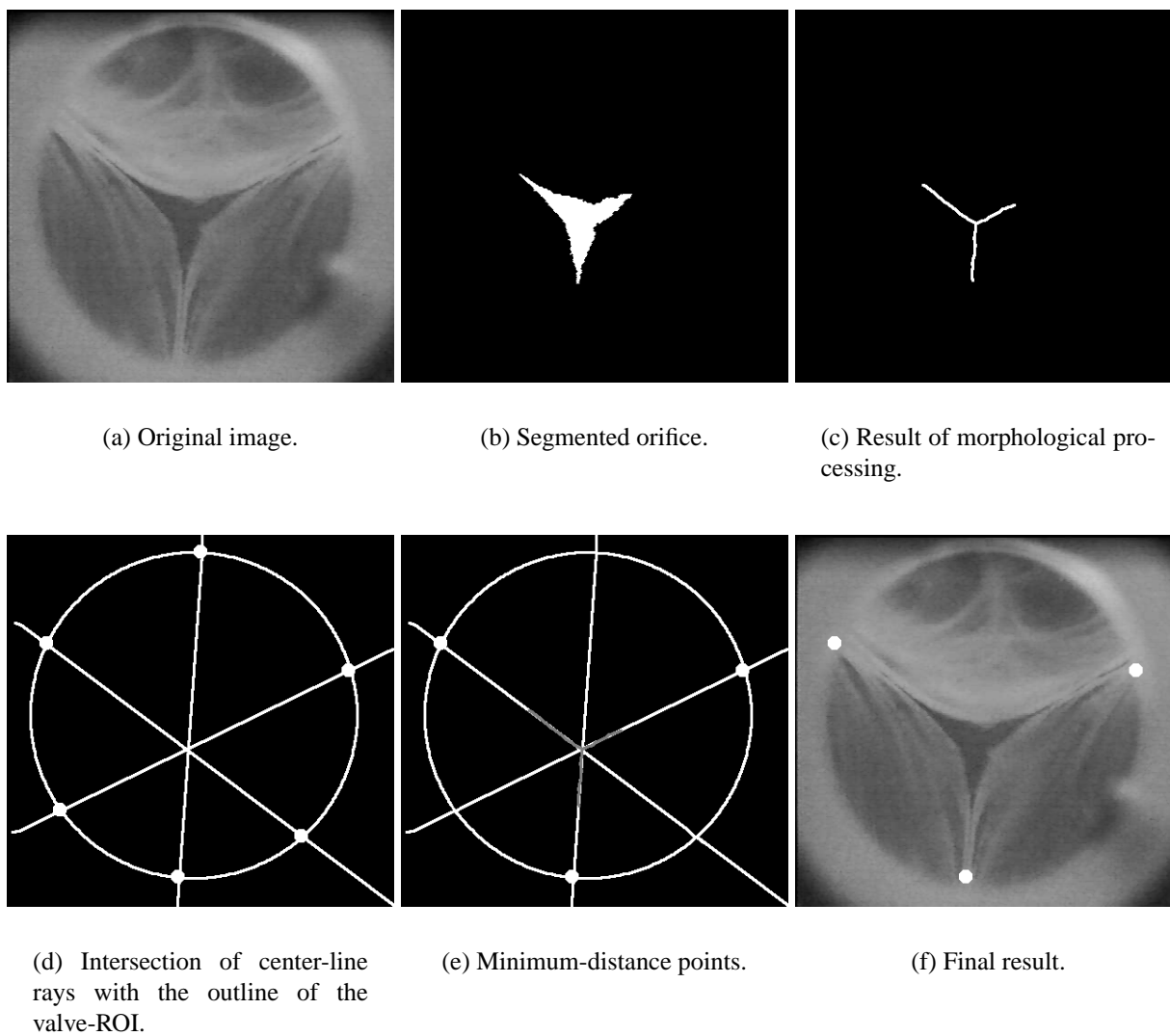
Figure 5.16: Automatic anchor points detection.

additional stress for the heart. With respect to the dynamic evolution, a valve cycle has to show *three phases* and the orifice curve of a healthy valve should accordingly show three regions. The first phase is the rapid increase phase, when the orifice increases rapidly to its maximum value as the valve is opening, during the second phase, the area of the orifice should decrease slowly, the valve being now opened and finally in the third phase, the area decreases sharply and the valve closes completely. Such a typical behavior is shown Figure 5.21 (a). Also, during one valve cycle, the values of the orifice area in each phase have to fall in certain predefined intervals.

The dynamic behavior of the valve during the three phases can be measured by the derivative of the curve in each region. The three orifice curve regions corresponding to the three phases of the valve cycle can be segmented as with the methods used to select the complete state in Section 3.2.1.

5.3   Analysis of the leaflet fluttering

The effects of the fluttering can be observed both in time and space. Leaflets of a fluttering valve will exhibit a jerky motion somewhat similar to the way a flag moves in the wind. At the same time, at each time-instance while they move during the second phase of a valve cycle, their borders will appear with an irregular form, as they bend and wriggle in a plane perpendicular to that of the image (see Section 5.3.1). Implants which exhibit such leaflet deformation are shown, e.g., in Figure 5.7 (a), in Figure 5.8 (a) and in Figure 5.4 (b).

The analysis of fluttering includes detection and quantification. We describe how to detect and quantify the fluttering in the time domain on the orifice curves (see Section 5.3.2) and in the space domain in each image (see Section 5.3.3). We also describe composed methods, taking into consideration information from both time and space (see Section 5.3.4).

5.3.1   Fluttering characteristics

To measure the fluttering successfully, one has to understand how does fluttering appear and also *how* and *where* can it be *seen in the test-sequences*. We seek to quantify only jerky fluttering, which is from now-on called simply fluttering and we seek to measure fluttering for each leaflet. We measure fluttering only in the second phase of a valve cycle, because otherwise, i.e., during opening and closing of the valve, the leaflets are tensed and do not flutter.

The motion of the leaflets and its effect on the leaflet projection in the image plane.   For most of the time during a valve cycle, the borders of the leaflets delineate the orifice and thus their fluttering may be practically observed on the borders of the segmented orifice. However, in some cases, the leaflets bend over their vertical position and parts of their border can no longer be seen, as their contrast almost vanishes (see the left leaflet in Figure 5.15 (a)). There are also cases when the leaflets disappear behind the tube through which the camera observes the valve (see Figure 5.20 (a) and the discussion in Section 5.2.1). Nevertheless, the fluttering for an entire sequence can be estimated, as usually the well-conditioned images represent a large majority.

Segmentation of the second phase of a valve cycle.   While the heart-valve is opening the leaflets are usually tensed and thus will not flutter. Consequently, fluttering should be analyzed only when the valve is fully open, i.e., in the phase two of a valve cycle. This phase can be segmented on the orifice curves using methods similar to those developed to segment the filled state in a sequence of coronary angiograms (see Section 3.2.1).

5.3.2   Fluttering analysis in time. Orifice curve-based fluttering measures

A time-based analysis of fluttering seeks to quantify the "jerkiness" in the motion of the leaflets. The motion of the leaflets directly influences the area of the orifice and therefore *fluttering can be also measured on the orifice curves*.

Fluttering conform orifice curves.   To conduct an analysis of the fluttering based on the orifice-area curves, one needs to ensure that changes of the area are generated alone by the opening and closing of the valve. In comparison to threshold-based segmentation, the snake-based approach is more likely to – and practically does – fulfill this condition.

Threshold-based segmentation is sensible to illumination changes. This is shown in particular in Figure 5.21 (a). The see-saw appearance of the orifice curve computed by thresholding – irrespective which method – is due to the varying illumination and have therefore nothing to do with variation of the orifice area. Also, the spike, at the end of the threshold-based orifice curve shown in Figure 5.21 (d), is due to a bad segmentation result and not to change in the area of the orifice. Conversely, snake-based orifice curves reflect the true evolution of the orifice.

Detection of fluttering

The influence of fluttering on the orifice curve.   Fluttering causes the area of the orifice to change abruptly and often over time, as the leaflets move back and forth. Thus, it can be detected on the orifice curves as a high-frequency signal, modulating the portion of the curve corresponding to the second phase of a valve cycle (see the snake-based results in Figure 5.21).

Quantification of fluttering.

After segmenting the slow-decrease phase on the orifice curves, the energy in the resulting 1D signal can be used as a measure of fluttering. For this purpose, the DC component and the low-frequency components introduced by the slowly decreasing orifice area, need to be first eliminated by high-pass filtration, such that the result is related solely to the fluttering. Also, frequency components related to low-frequency eigen-fluttering have to be eliminated. Empirically, we have set the pass-frequency at 6 Hz. We then measure fluttering by the mean energy in the filtered signal. Practice shows that the more fluttering, the larger the measure.

5.3.3   Fluttering analysis in space. Measuring fluttering by the deformation of leaflets

The fluttering is a time-property of the valves. Therefore, theoretically, a valve whose leaflets have a very irregular form, but which is constant in time, is not considered to flutter. However,

practically, *leaflets that have an irregular form do flutter*. A *space-based analysis* of the fluttering seeks to *measure how irregular* is the shape of the border of the leaflets. The border of the leaflets is obtained directly from the result of the orifice-area segmentation.

Detection of fluttering

The connection between fluttering and the irregular pattern of a leaflet border.   Leaflets of tricuspid heart-valves have a natural form whose outline is low-frequent. Also, the flow of liquid through a valve is turbulent. Thus, as liquid goes through and if the valve has poor elastic properties, its natural form will appear modulated by different flow-related higher-frequency components – i.e. the border of the leaflet has an irregular pattern – and it will also change in time. Therefore, the irregularity of a leaflet borders is related to flutter.

Theoretically, one can imagine a valve of irregular border but with a pattern which remains constant in time. Such a valve does not flutter, but nevertheless a space-based analysis as the one proposed here will still detect a fluttering valve. To differentiate in such a case, one has to observe also how does the irregular pattern of the border change in time. Practically, we never encountered such a case.

Quantification of fluttering

An irregular border-curve will contain relatively large amounts of energy in the high-pass part of its spectrum and it will also have a high curvature. We introduce therefore *two measures* to quantify how irregular is the border of the orifice: one is related to the *energy* in the high-pass part of the spectrum and the other one to the *curvature*.

The border of the orifice of a tricuspid heart-valve usually exhibits three regions of relatively high-curvature (one at each anchor point), which are not related to the fluttering. These clearly decrease the sensibility of our fluttering-analysis. Therefore, we propose to analyze each leaflet alone. An additional reason for analyzing a leaflet alone is, that we would like to be able to characterize each leaflet independently. Therefore, we measure the fluttering for each leaflet and in each frame. For this purpose, we need to *separate the border of each leaflet* from the border of the entire orifice.

We can measure the fluttering of an entire valve in an image by computing the mean of a single-leaflet fluttering-measure, over all leaflets. We can observe thus also the dynamics of fluttering. The mean over an entire sequence gives then the fluttering for the analyzed valve. Practice shows that the more fluttering, the larger the measure.

Finding the border of each leaflet.   The valve orifice is defined by the borders of the leaflets, which can be segmented on the orifice border. For this purpose we use the anchor points, which have been already computed. As the fluttering is analyzed only over the second phase of a valve cycle, these will be situated always on the border of the orifice and will define thus the borders of each leaflet. For a robust segmentation of the leaflet borders, we use the anchor-points together with the center of the valve to divide the image in three regions. These regions give then an implicit partition of the orifice border into curves corresponding to each leaflet. To make sure that the analysis focuses only on leaflet borders and not on other artefacts, like e.g., the circular outline, after segmenting the leaflet border, we keep only the two thirds around the center and ignore the rest (see Figure 5.17 (a) and (b)).
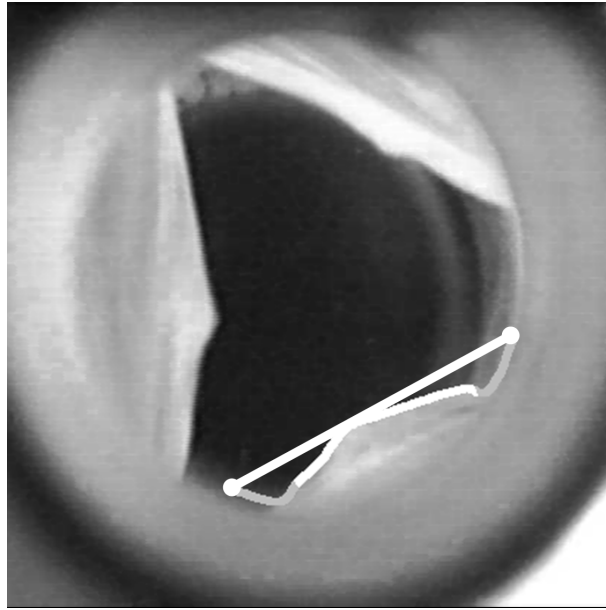
(a)
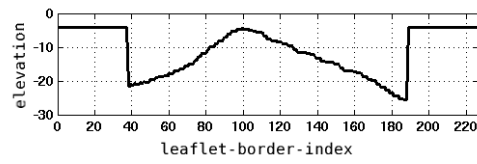
(b)
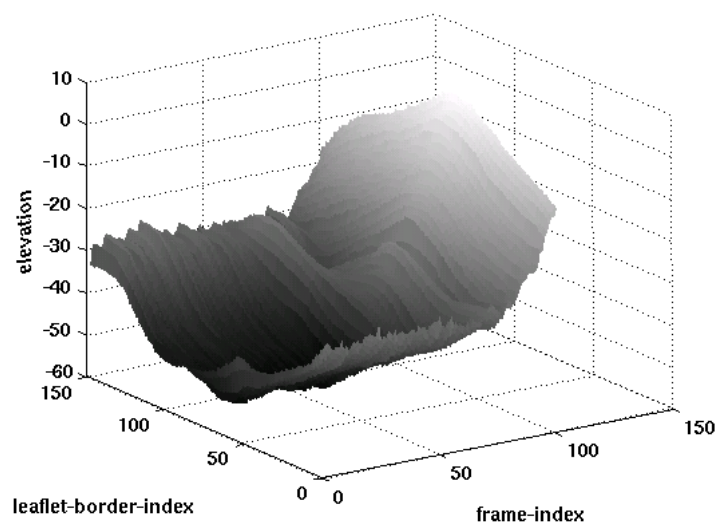
(c)

Figure 5.17: Leaflet selection (a) the corresponding 1D signal (b) and the time-space image (c).

Energy-based measure.   We propose to measure the fluttering by the high-frequency energy-content of the 1D signal given by the border of the valve orifice. To analyze the fluttering alone, we first eliminate the DC component, as well as the low-part of the spectrum of a leaflet-border signal, which contains frequencies linked to the natural shape of the leaflet and other artefacts. The energy in the remaining components is the fluttering measure for the analyzed leaflet.

The high-pass filtration takes place in the Fourier domain by setting to zero all spectral components up to the pass-frequency. In our experiments we choose a pass-frequency of one tenth of the sampling rate. We justify the filtration in the Fourier domain by observing that we are primarily interested in the energy in the high-pass band of the signal and not in reconstructing the signal after filtration. Therefore, we can use an optimal transfer characteristic.

Curvature-based measure.   Theoretically, an irregular valve-border should exhibit large curvature values and thus we propose to use the curvature of the valve-border as an alternative fluttering measure. We approximate the curvature by the second term of the internal energy of a snake (see Appendix A.3).

5.3.4   Quantification of fluttering by spatio-temporal processing

The fluttering is a time-property of leaflets, which however, is also related to the appearance of the leaflets borders in each image, i.e., to space. We propose therefore, to *measure fluttering* by *combining information from both domains*.

Detection of fluttering

Leaflet evolution in time-space images.   The border of a leaflet is obtained by sampling over the base line and we keep only the two thirds around the center, ignoring the rest (see Figure 5.17 (a)). These are then used to compute a time-space image. For this purpose each leaflet-border-curve is first sampled over the base line linking the two corresponding anchor points, obtaining thus for each image a 1D signal of standard length(see Figure 5.17 (b)). We then gather all these signals as columns of a time-space image. An example is shown in Figure 5.17 (c). Clearly, fluttering is related to the energy content in the higher-frequencies of such images.

Such images capture both the time evolution of the leaflets and the form of the leaflet border at each time-instance. Therefore, a fluttering measure using these images has the potential to return correct results even for the theoretical case we have constructed previously, i.e., about a leaflet with irregular border, but with a pattern of irregularity which is constant in time.

Quantification of fluttering.

We propose to use as measure of fluttering for each leaflet, the mean high-pass energy in such time-space images. The mean over leaflets gives then a fluttering-measure for an entire sequence. By high-pass filtration we eliminate the influences given by the natural tendency of the valve to close slowly during the second phase of a valve cycle, as well as the low-frequent eigen-vibration, which we do not want to measure. This is shown in Figure 5.18. Practice shows that the more fluttering, the larger the measure.

(a)

(b)

(c)

(d)
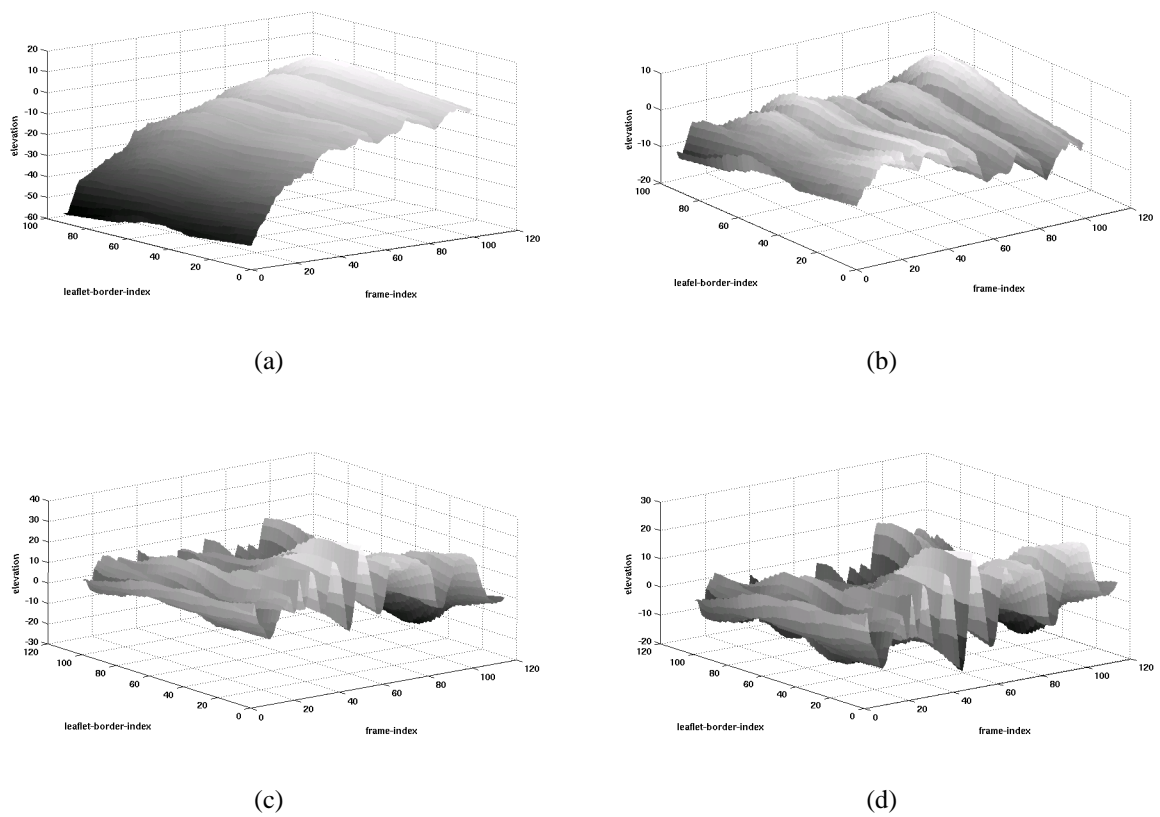
Figure 5.18: Time-space images before (left column: (a) and (c)) and after (right column: (b) and (d)) elimination of the influences given by the natural tendency of the valve to close for a non-fluttering (first line: (a) and (b)) and a fluttering (second line: (c) and (d)) valve.

Sequence	Fluttering	Observations
A	no	eigen-vibration of low amplitude and high frequency
B	strong	no eigen-fluttering observable
C	low	eigen-fluttering of high frequency; strong fluttering for one leaflet
D	strong	no eigen-fluttering observable
E	low	eigen-fluttering of high amplitude and low frequency

Table 5.1: Sequences in the data base.

5.4 Experiments and discussion

We have tested the algorithms proposed here on five valve sequences chosen such that they cover – to the best of our knowledge to date – the entire spectrum of problems encountered while analyzing valves. All sequences show valves with three leaflets. They are described in Table 5.1.

Segmentation of the orifice area

To determine the orifice area of the valve and to find its edges, whose shapes are needed for the fluttering analysis, we have followed two approaches. The first approach is threshold-based and the other is snake-based.

Within the first approach, to find the orifice area we compare each image to an adaptively determined threshold. We have investigated two automatic thresholds: the first one assumes a bi-class model of the global histogram and sets the threshold by the Otsu-method and the second, assumes a multi-class model and finds the threshold by likelihood ratio test between the two darkest components of a mixture whose parameters are found using the FMM. (see Figure 5.19). The edges are obtained from the segmentation result by morphological processing.

The second approach is based on an appropriately designed snake. The orifice area is obtained as the area enclosed by the snake and its edge is directly obtained from the snake. We have improved the external energy term such that the snake follows the leaflet everywhere.

In Figure 5.21 are shown orifice-area curves plotted over time-index for each of the sequences in the data-base and by each method. The results for the Otsu-threshold have been achieved with center-selection, while the results for the mixture decomposition without.

The Otsu-threshold. The Otsu-threshold is usually larger than the mixture-threshold. This has the disadvantage of returning over-segmentations (see the spikes in Figure 5.21 (d)), but at the same time it makes it less sensible to illumination variations (see the ripples in figure 5.21 (a)). In many cases the oversegmentation problem of the Otsu-threshold can be solved by center-based selection and it usually affects only some images in a sequence.

The FMM-based threshold. The FMM-threshold returns *better results* from the point of view of a classification algorithm, since it separates better the darkest cluster on the global-histogram for each sequence. This is shown in Figure 5.19.
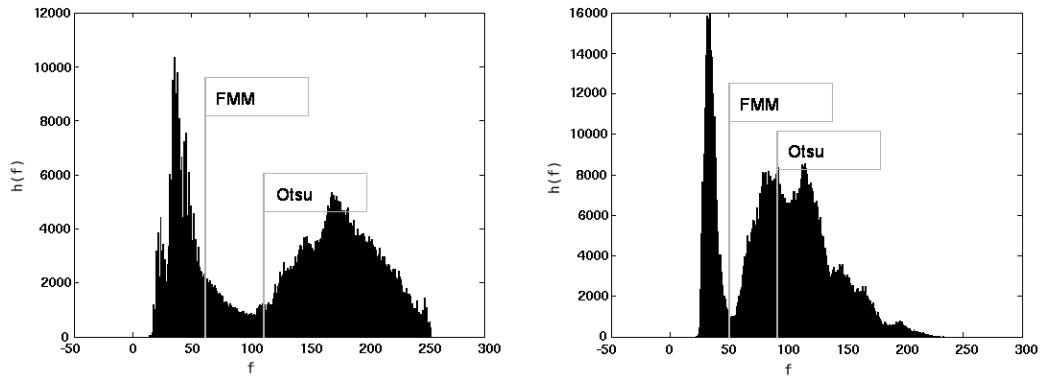
(a) Sequence A

(b) Sequence B
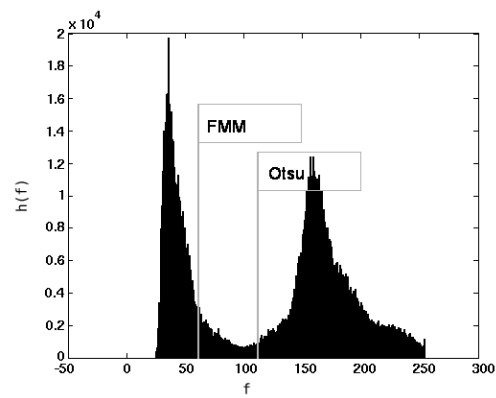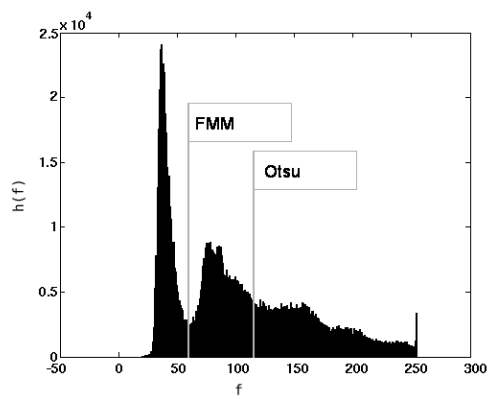
(c) Sequence C
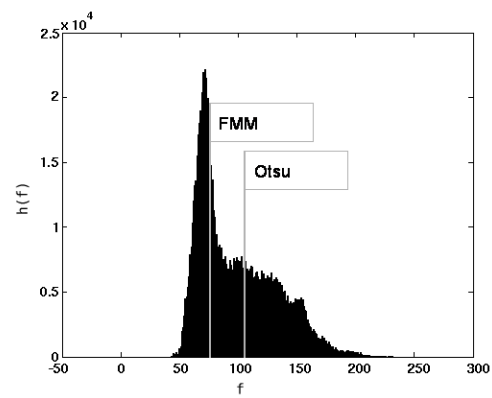
(d) Sequence D

(e) Sequence E

Figure 5.19: Global histogram and thresholds for five different sequences.
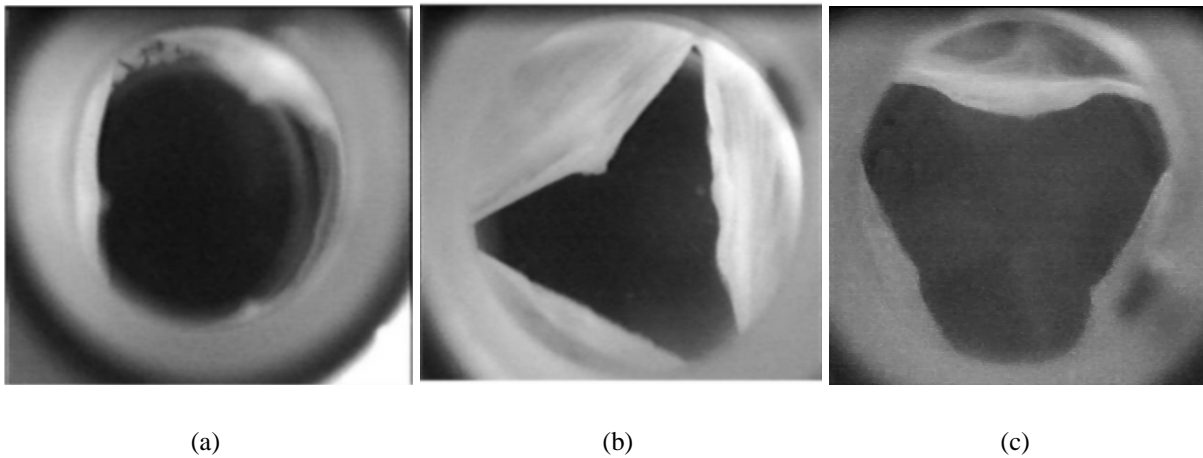
Figure 5.20: Examples from sequences where the valves show an inhomogeneous orifice.

It appears however, that the mixture-threshold is actually too conservative for some sequences. The reason for this is that the orifice gray-levels do not cluster well, or in other words the orifice surface does not contain a single type of dark pixels. This is due, e.g., to the light coming sideways to the valve and making thus some orifice parts lighter than others (see Figure 5.20 (b) and (c)), but also to the fact that reflections appear or even small parts from the test-setup-tube downwards from the position where the frame with the valve is placed. This happens particularly when the valve is fully-opened (see Figure 5.20 (a)). If the FMM-method fails, then it will be likely to fail for an entire sequence (see Figure 5.21 (e)).

Snake-based tracking of the orifice border.   The assumption about an orifice darker than its background holds practically always. It did so in all analyzed images. However, the orifice and background are in many cases not perfectly homogeneous (see Figure 5.20). Also, all sequences used here to test the proposed algorithms and which have been acquired with the currently available test setup are affected by illumination changes.

Snakes are robust against illumination changes and offer good orifice segmentation results, as long as the edges they follow are still distinguishable. They are also able to follow the leaflet border even in regions where the contrast between the leaflet and the background is weak and are thus superior to threshold-based approaches, as shown also in Figures 5.10 (c) and 5.21. The orifice curves computed by snakes usually show neither ripples nor spikes and return results in good agreement with the expectations of a human expert.

The result of the segmentation depends on the attraction range of the external energy and also on the way the energy is defined around the anchor points. A GVF offers a large attraction range and good convergence in regions of high border-curvature [198]. The attractors make sure that the external-energy field converges properly around anchor points.

Because the images are acquired at high frame-rate, the distance traveled by the leaflets between two consecutive images is small enough such that the converged snake from the previous image is practically always in the attraction sink of the leaflet borders in the current image.

Comparison between methods.   Comparing the two automatic thresholds with respect to how good they segment the orifice, it seems that the Otsu-threshold functions better for some se-
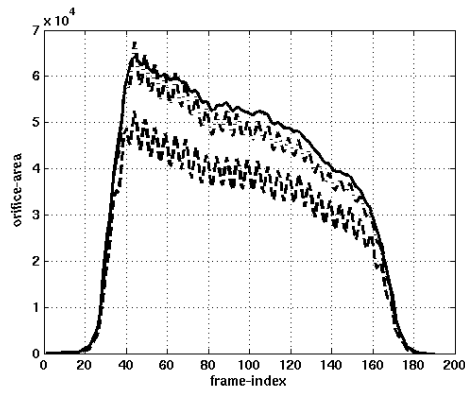
(a) Sequence A
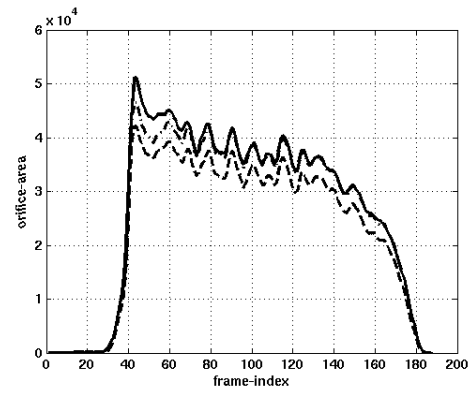
(b) Sequence B
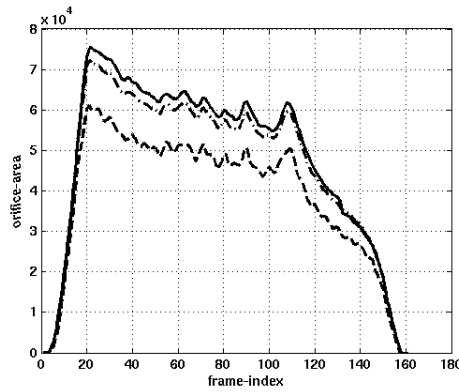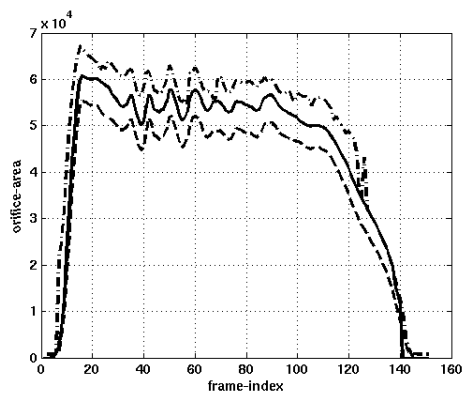
(c) Sequence C
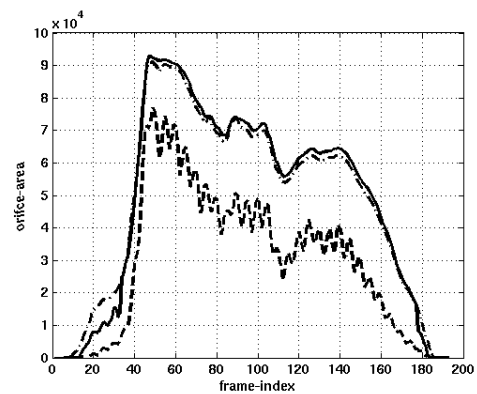
(d) Sequence D

(e) sequence E

Figure 5.21: Orifice curves computed by different methods: using snakes (continuous line), the Otsu-threshold (dashed-dotted line) and the FMM-threshold (dashed line).

quences (e.g. the one shown in Figure 5.21 (e)) and the FMM-threshold for others (e.g. the one shown in Figure 5.21 (d)).

The FMM-threshold returns better results from the point of view of a classification algorithm. It separates better the darkest dome of the global-histogram, which is usually associated with the orifice (see Figure 5.19). This however is not always enough in practice, due to the inhomogeneity of the gray-level representation of the orifice. The FMM-threshold is usually too conservative.

Apart from the fact that they have to be initialized when the valve has already opened, active contours with GVF for the external-energy term are well suited for the segmentation of the valve orifice. The extended energy-term, with attractors near the anchor points, ensures a precise segmentation of the valve border over its entire length. They give better, more robust results than thresholds. This is particularly true for the sequences from our data-base, which have been recorded using a flickering illumination source. As shown, e.g., in Figure 5.21 (a) and Figure 5.21 (d) the threshold-based orifice curves capture also the flickering, which in turn can be misinterpreted for fluttering of the leaflets. The snake-based orifice curves do not exhibit flickering artefacts.

Fluttering detection and quantification

We have described three approaches to the detection and quantification of the fluttering: the time-based, the space-based and the time-space image-based approach respectively. The results obtained by the fluttering-quantification methods proposed here are shown in Table 5.2.

Sequence	Time-based	Space-based		Time and space-based
	energy	energy	curvature	energy $\cdot 10^3$
A	153.03	191.11	27.79	60.93
B	271.43	1002.61	26.50	444.99
C	220.21	280.36	31.20	118.34
D	266.70	684.25	29.02	396.28
E	223.01	257.13	33.37	215.31

Table 5.2: Fluttering measured in the sequences from our data base by different methods. The orifice area was segmented using the snake-based approach.

Time-based measures. The simplest and fastest way to measure fluttering for an entire sequence is by a frequency analysis of the orifice curves. As shown in Table 5.2, the results achieved this way correlate well with the ground-truth. However, there is no possibility to measure the fluttering of each leaflet of the valve.

Space-based measures. The curvature is only of limited use in fluttering-analysis for valves with three leaflets, i.e., the aortic, pulmonary and tricuspid valves. The reason for this is the relatively sharp bend which the border of a leaflet makes in the middle. This bend is usually visible in many of the images acquired during the second phase of a valve cycle and it keeps its form for all but the worst valves. As a consequence, the curvature is equally large for both

(a) F = 68.76 $\cdot 10^3$ (b) F = 213.10 $\cdot 10^3$ (c) F = 73.16 $\cdot 10^3$
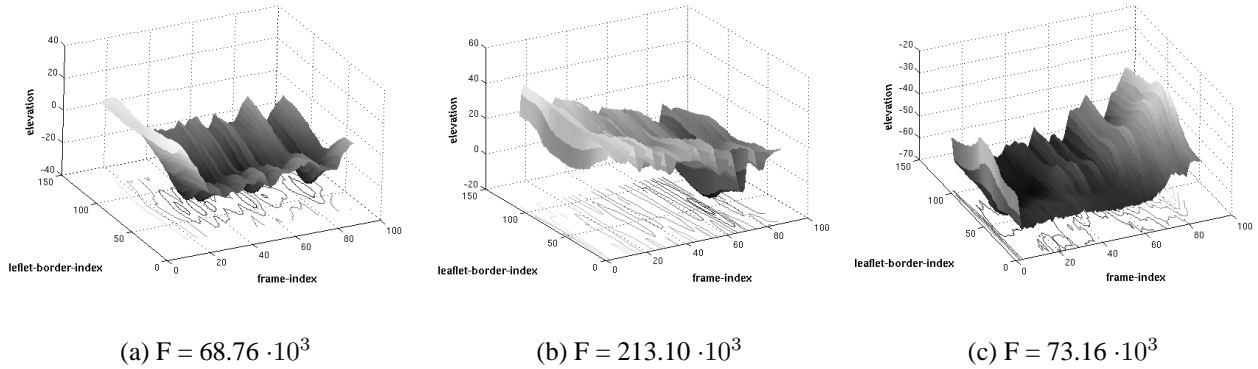
Figure 5.22: Time-space images for the three leaflets of sequence C with the corresponding fluttering measure (F).

fluttering and non-fluttering valves and it becomes smaller for valves showing a very strong flutter. The curvature-based measure shows, to a certain extent, a negative correlation to the ground truth given in Table 5.1, i.e., the more fluttering the smaller the curvature, against our initial expectations.

The energy-based measure returns results which do correlate with the ground truth. This may be seen also as a confirmation of the fact that in the sequences from our data base, an irregular pattern of the border of a leaflet implies that the leaflet flutters.

Space-based measures return results for each leaflet. However, one cannot differentiate between valves which keep the pattern of their borders and those which do not.

Time-space images. We propose to analyze each leaflet independently by means of a high-pass frequency analysis of time-space images. Such images can be obtained by introducing the outline of each border for each analyzed image in a matrix. They can then be used to characterize both the time and space evolution of a leaflet and therefore can additionally differentiate between valves which keep the pattern of their borders and those which do not. An example is shown in Figure 5.22. These results do correlate very well with the available ground truth.

Comparison between fluttering measures. All methods proposed here, with the exception of the space-based curvature measure, return results which are in agreement with our ground truth. Time-based measures are fast to compute, but they cannot return information for each leaflet alone. Space-based measures on the other hand measure only a consequence of the fluttering, i.e., the irregular shape of the border of the leaflets.

A time-space image offers the most information on all aspects of fluttering and can be also computed for each leaflet alone. Therefore, we believe that measuring the fluttering by means of time-space images is the best way.

5.5 Conclusions and outlook

We have described methods for the automatic computation of quality measure for heart valves. We expect the automatic measuring of the orifice area to improve on the current methods mainly

in two ways:

1 **More precise**, because usually humans are prone to errors during such painstaking jobs – in this respect similar to the manual segmentation of vessels. Also, during manual analysis only a few images from the available sequence are actually analyzed and the sought measures are then computed from data interpolated from these measurements. Automatic measurements use all images in the sequence.

2 **Faster**, such that more implants can be inspected and the chances of finding a better one rise.

Segmenting the orifice in each image of the analyzed sequences leads to the computation of an orifice curve and of a fluttering measure.

The orifice curve describes the evolution of the area of the orifice during a valve cycle. This should exhibit a specific form – showing all typical phases: rapid-increase, slow-decrease, fast-decrease – which permits then the calculation of features linked to the dynamic properties of the heart valve as, e.g., opening and closing speed. Also, the values of the area of the orifice during each phase and in particular the maximal value, has to fall in a certain interval.

By fluttering we mean the dynamic deformations of the leaflets during the valve cycle. We can detect both jerky fluttering and eigen-fluttering. We have shown how to measure the jerky fluttering, while ignoring the eigen-fluttering, which is considered uncritical. Measuring the fluttering is a novel approach, which has not yet been proposed. Currently, the fluttering is assessed only visually – in all but a few special cases related to medical research. By the additional analysis of fluttering we provide new ways to better assess the heart valve quality.

We have applied these methods to the quantification of the quality of heart valves with three leaflets. This category includes the aortic, the pulmonary and the tricuspid valves. However, we believe that such methods can be also successfully used for bicuspid valves, thus constituting a tool for heart valves analysis in general.

We consider that a test image-sequence is suited for analysis only if it shows one entire valve cycle, from the moment the valve opens until it closes.

5.5.1 Automatic computation of orifice curves

Orifice curves are obtained from the evolution of the segmented area of the orifice over the frame index. We consider that the snake-based computation of the orifice area is the most appropriate method.

Segmentation by thresholding

As the lighter valves are imaged against a darker background, we first tried to segment the orifice by a gray-level threshold. We do not expect to see the orifice in each analyzed image. Therefore, we use a single threshold for the entire sequence. This threshold is set on the global histogram of the test sequence.

The sequences from our data base have been acquired with a neon-lamp and thus illumination changes appear between different images, which influence the results computed by thresholding the global histogram. To attenuate this flickering, we compute a mean gray-level over the entire sequence in a region of the image with constant gray-levels and modify then each image, such that its mean in the same region is the same as the sequence mean.

Because there are more than two pixel-gray-level classes in the analyzed images, a simple threshold usually returns many false positives. However, the segmentation can be improved using the additional prior knowledge that the valves are imaged in a centered-position and thus the image center will be contained in the orifice. Then, we select from the initial threshold-results only those pixels connected to a center pixel. A center pixel is defined as any pixel situated in a small disk around the image center. We consider a small disk rather than the image center alone to be able to proceed even in the cases when the valve is not perfectly centered.

If falsely segmented background items are linked to the orifice, then the segmentation by such a bi-class procedure will fail. To reduce the rate of false positives in the initial segmentation, we improved the thresholding procedure by accounting for more than two pixel-classes in our model of the global sequence-histogram. For this purpose, we have used a mixture-decomposition algorithm, which can compute the number of mixture modes automatically. Then, we have selected the orifice as the darkest mode. Although from a classification point of view the segmentation improves – as the darker dome of the global histogram usually associated with the orifice is better segmented – the results achieved are sometimes even worse than before. The reason for this is that the orifice, although dark, does not have a homogeneous gray-level representation as expected and at least in some cases it cannot be modeled as a single mixture-mode. Also, sometimes parts of the orifice, in the vicinity of those points where the leaflets meet, are not imaged against a dark background, but against a lighter one and thus will certainly not be segmented.

Segmentation by active contours

To improve the segmentation results we propose to use active contours to track the leaflet border over the entire sequence. They are insensible to illumination variation and to the inhomogeneities of the orifice gray-levels. Also, we showed how to make the snake follow the leaflets even in regions of weak contrast – as long as some leaflet edge is still visible – guided by the external energy field with additional conditioning around anchor points. Snakes are therefore best suited to compute the orifice-curves. The disadvantages of the snakes are that practically they need to be initialized after the valve has opened and can not follow the orifice until the valve is completley closed.

Practical computation of orifice curves. To analyze an entire sequence, we propose to use a global-threshold to segment the orifice in those images in the beginning and the end of a sequence, when the snake cannot be properly defined. One can be reasonably certain that a gray-level threshold segments the orifice properly in the initial and end phases of the valve cycle, because then it can be safely assumed that the orifice, due to its small size, is imaged against an uniform-background . Even if in some rare cases the threshold can fail, most of the sequence will be correctly segmented nevertheless by the snake, as this is initialized from the leaflet baseline and not based on the result achieved by the threshold. Using this combination between gray-level threshold and snake, one can successfully segment the orifice over an entire valve cycle. We use the multi-class-model threshold.

5.5.2   Fluttering analysis.

Fluttering appears as a consequence of the poor elastic properties of valves and is best observable during the second phase of a valve cycle, i.e., when the valves are fully open. This phase can be segmented on the orifice curves as described in Section 3.2.1.

The results show that the fluttering of a valve can be correctly measured by all methods proposed with the exception of the curvature-based measure. However, we consider the time-space images-based fluttering quantification to be the most appropriate method, as it offers the largest amount of information and enables thus an improved analysis in comparison to the other methods.

Background of fluttering.   The natural form of a leaflet acts as a regularizer on its behavior in the flow and it is maintained during the entire second phase of a valve cycle only if the leaflet has good elastic properties. If their elastic properties are poor the leaflets cannot hold their natural form. Consequently, their motion will be jerky and their appearance irregular. In the analyzed image sequences this can be observed both in time and space on the border of the valve orifice, i.e., on the borders of the leaflets. In time the leaflets of a fluttering valve have an abrupt and jerky motion and in space the outline of the leaflets has a highly irregular appearance. These two conditions are linked to one another. The valve can be seen as a system which is excited by a broad-band signal given by the flow of liquid – the flow of blood through large vessels is turbulent and not laminar. This sets the leaflets in motion. If their elastic properties are good – i.e. the valve material is elastic and it is properly spanned – the leaflets will keep their natural form while open and act thus mostly as a band-pass filter, resonating only with certain components of the input signal – the resonance is manifested as a rhythmic motion, which we call eigen-vibration. In the ideal case, they will act as a stop-filter and show no resonance.

Measuring the fluttering.   We have described both time-based and space-based methods to measure the fluttering, as well as methods in both time and space. Time-based methods are related to the evolution of the leaflets during the sequence and space-based methods analyze the appearance of the leaflet border. Time-based methods measure fluttering correctly, but they cannot be used to measure each leaflet individually. From the space-based methods, only those which measure the fluttering by the energy in the upper part of the spectrum of a the leaflet border return good results.

Time-space images in the measuring of fluttering.   Good results have been achieved by the time-space images-based analysis, which considers both time- and space-related aspects. By segmenting the border of each leaflet and in each analyzed image, we compute time-space images of leaflet borders. Time-space images of fluttering-valves will have high amounts of energy in the high-pass part of their spectrum. This way one can measure each leaflet and can also differentiate between different types of fluttering, providing thus the basis for an improved analysis. We consider this to be the best fluttering-quantification method from those proposed here.

5.5.3 Outlook

We have described automatic methods to measure different quality criteria of heart valves. These include measurements of the maximal orifice area and of the evolution of the orifice area as well as quantification of the fluttering of the valve leaflets.

Currently, such measures are used by a human operator to decide on the quality of the heart valve. Assuming a sufficient number of valves is inspected one can use this data to design classification algorithm to automatically decide if a heart valve is good or not. For this purpose we propose to build a feature-vector, containing all the needed information with respect to the quality of a valve. This valve-quality vector would have the following components: the maximum orifice area, the slopes of the three regions of the orifice-curve and the value of the fluttering measure. For maximal safety, the implants considered suitable by the system, together with their corresponding measures, would then be presented to an expert for the final decision about their employment.

The fluttering is computed from a Fourier analysis of the time-space images, which is sufficient to differentiate between fluttering and non-fluttering valves. However, the fluttering analysis could be enhanced if one knows not only what frequency components are present, but also when and where do they appear. Such informations can be obtained, e.g., by a wavelet analysis [137], [138], [63] of these images.

We have described how to evaluate the quality of heart valves based only on 2D-projection images. Clearly, in this case information about the leaflets is lost, as most of the time they are practically perpendicular to the image plane and therefore they are projected there on a curve. This forces us to make inferences about their properties based only on their borders. More informations can be collected and an enhanced analysis made if instead of a 2D-camera one uses a 3D vision-system [81] to record the valve during a cycle.

Chapter 6

Summary

"Computers are incredibly fast, accurate and stupid.
Human beings are incredibly slow, inaccurate and brilliant.
Together they are powerful beyond imagination."
– Albert Einstein

The information technology has already found its way into the hospital where it offers complete solutions for the storage, retrieval and transmission of data. For example, electronic patient-files are instantly available on the laptops of physicians in a hospital and he can then confer over an transmission channel with a colleague on the other side of the Atlantic about the best treatment. However, the major breakthrough takes place currently in the field of processing the information, i.e., the introduction of tools to help analyze data and enhance the decision-taking process.

The field of imaging and machine-vision support has a large contribution to this breakthrough as in many diagnostic procedures and indeed all interventions the physician has to visually-observe the ill tissue, analyze the data and take a decision. This dynamic is reflected also in the continuous growth that the branch of medical image technology enjoyed in the last years. This growth is directly related to an enlarging market reflecting in turn the increased interest of the society as a whole in improved healthcare and its confidence that such technology is the right answer in this case.

Mainstream machine vision has arguably as purpose automatization, i.e., the replacement of humans by machines, like e.g., for quality inspection or for controlling a robotic arm which is supposed to solder a circuit board in the place of a human. Medical machine vision however follows a more synergetic way as it supports the medical staff such as to increase their efficiency and accuracy and it does not aim at replacing them.

Such supportive technologies are possible only if a bridge is build between the medical and the computer vision world. We consider this contribution to be a part of this bridge and at the same time a plea for a senseful division of work between machines and humans such that both do what they can best do: the former should be used at painstaking repetitive and analytical work and the latter for creative and decisional work.

This collaboration requires from the part of the computer vision specialist an engineering approach to solve the practical problems presented by his medical counterpart. From our current experience, an optimal solution can be achieved only with the consideration of all prior information available, which implies again a close collaboration. The accent lies in this case on the solution of the problem – and each particular problem has its own particular solution – this
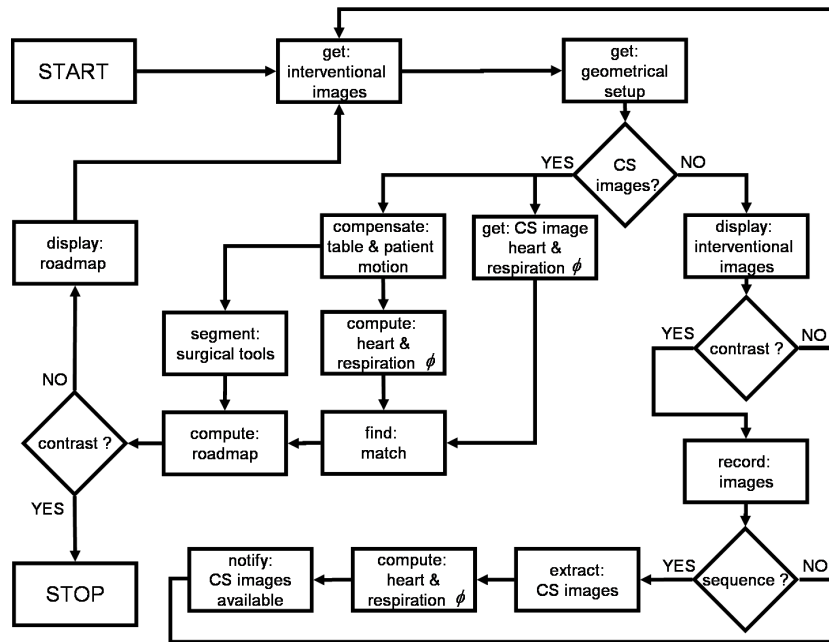
Figure 6.1: Dynamic-roadmap flow-chart. CS stands for complete state.

represents the novelty element, if the solution includes novel approaches or a combination of old approaches specifically adapted is of secondary importance – although a novel well justified approach clearly increases the value of the solution.

This contribution is related to cardiovascular medicine and to imaging techniques which return 2D-projection images. We have described here machine vision methods to enhance the catheter intervention (Section 6.1), for vessel segmentation (Section 6.2) and for the quality control of xenograft heart-valve implants (Section 6.3). Although, in some sense the methodology used in each case exhibits commonalities given mainly by the imaging and physiological background, we believe this work to be another proof of the principle that particular problems need particular solutions. Such particular solutions can be achieved both by new algorithms and by new parametrizations and/or developments of dedicated algorithms.

6.1   Machine vision methods for the improvement of the catheter intervention

We have described computer-based supportive technologies for the interventional diagnostic, PTCA-based-treatment and post-interventional evaluation of the coronary artery disease and of the myocardial infarction [46].

For PTCA we have described methods to support the computation and synchronous display of a dynamic-roadmap with the purpose to facilitate navigation during the intervention. These methods are fully automatic and therefore need no additional knowledge from the medical stuff. They do not affect the interventional routine and there is no need for additional hardware. Also, the physician remains in full controll during the entire intervention and can start and stop the display. A chart-flow of the dynamic roadmap as described here is shown in Figure 6.1.

Some of the algorithms developed here – i.e. the complete-state segmentation and the de-

tection of the contrast-agent in live-images – are about to find their way into the clinical-routine [47], [49]. They are already part of a clinical demonstrator being currently tested under typical conditions in a hospital.

Selection of complete-state images. For the roadmap one needs a set of coronary angiograms showing the entire vessel-tree – i.e. complete coronary-angiograms. We have described methods to segment the complete-state of a sequence of coronary angiograms (also in [6]). For this purpose, a feature related to the area covered by vessels is computed in each angiogram. The complete-state is segmented on the curve of feature value over frame-index – i.e. the feature-curve – by a MAP-approach with an ECG-based coherency constraint.

Signaling of arbitrary contrast-agent injections. If no complete coronary angiograms are available they need to be recorded on the spot. The physician needs then to inject a contrast agent to obtain the angiograms. The moment when the contrast agent appears in the live-images needs to be detected and starting then a sequence of images are recorded from which the needed complete coronary-angiograms are selected.

When the dynamic roadmap is active, if the physician would like to see live contrasted images and also halt the roadmap he just needs to introduce a small burst of contrast agent. Again as soon as contrast agent appears in the live-images the computation of the dynamic-roadmap is halted.

We have described also methods to detect the first live contrast image after a contrast agent injection (also in [48]). For this purpose we extract another feature related to the area covered by vessels. The sought initial contrast image is then found by a significance test, which uses for training the features computed from the images recorded in the first few seconds of the current viewing session.

Sequence matching and registration of surgical tools. To compute the roadmap, one needs also to match the complete coronary-angiograms with the interventional images and to super-impose the interventional tools. Matching is done using a heart-position vector including in-formation about the heart and respiration phases. We have shown how to build this vector by using the diaphragm to compute the respiration phase (also in [50]). If the diaphragm is not observable, image-based similarity measures should be used for this purpose. The diaphragm is detected and tracked by the Hough transform. If the diaphragm is not correctly found, this situation is detected by a confidence measure and a snake is used to refine the fit.

We have described also methods to segment the interventional tools, like e.g., guidewire tip and balloon markers as well as how to detect potential sewing wires present in the analyzed-images. The interventional tools are segmented by a percentile threshold on an enhancement-result. We use the compactness to separate balloon markers from the guidewire tip. To find the sewing wires we use the fact that they remain static in comparison to other interesting items in the analyzed images.

Automatic estimation of the MBG. The MBG is used for the diagnostic of the myocardial infarction, but also to evaluate the success of a PTCA-based treatment. We describe methods for the "quantitative" assessment of the MBG from sequences of X-ray images of the blush-region.

As the physician needs to define per hand in an initialization step the target blush-region, the assessment is semiautomatic. The target region is then tracked automatically during the rest of the analyzed sequence (also in [54]). Both human and machine then do what they can do best. The physician uses his vast domain-knowledge and defines a region of interest and then the painstaking job of following and analyzing it in each image of the investigated sequence is taken over by the machine.

The purpose of tracking is to compensate the variation in position of the target blush-region. The variation appears as a consequence of the heart motion. The motion of the heart is estimated from complete-state images by tracking vessel junctions between consecutive images. The junctions are then used to fit a physically-motivated ROI model, which evolves under the influence of internal and external constraints to an optimal position. The constraints are expressed as energy-terms and the optimal position has minimal energy.

From the gray-levels of the pixels in the tracked ROI we compute a blush-feature as a the 25-percentile of their histogram. We then use this feature to obtain a blush curve for each analyzed sequence. The MBG is computed from two blush curves – one from the ill artery and one from a reference healthy-artery – as the difference between the areas under the curves.

6.2 Vessel segmentation in 2D-projection images: methods and applications

How should vessels be segmented in 2D-projection images such that one can quantify their, e.g., area and length? We believe that there is no single algorithmic answer to this question. However, there exist a set of principles which can be followed to devise the best suited solution for each particular segmentation problem.

The key principle is that one should use any available prior knowledge to the largest possible extent to find particular solutions to particular problems. Yet another principle is that one should achieve the segmentation by deciding for each pixel individually, as only then acceptable results can be achieved, particularly when segmenting vessels of a very weak contrast. Then, a solution should be achieved in two steps: (i) enhancement and (ii) segmentation.

Vessel enhancement

The enhancement step has the purpose to increase the contrast and the homogeneity of the gray-level representation of vessels and background. After enhancement vessels appear dark on a light background and in higher contrast. Enhancement decouples the imaging part from the segmentation. Therefore, similar segmentation algorithms can be developed for the vasculature of different organs and under different imaging modalities.

Enhancement is also done in two steps: (i) background suppression and (ii) vessel augmentation. Each time with the purpose of increasing both the homogeneity of the gray-level representation of vessels and background (decreasing their variance) and the contrast separating them. Usually, prior information about the size of the vessels is used during background suppression and prior information about the elongated tubular structures of the vessels is used during vessel augmentation. We showed that the optimal background suppression method is given by the Bothat-transform. Because the vessel contrast is dependent on the vessel-size, the best vessel augmentation methods use a multiscale approach. We have experimented with an

adapted form of the Laplacian pyramid for vessel augmentation and introduced a novel Hessian-based multiscale method (also in [53], [43], [52]).

Each of the vessel enhancement methods has a set of parameters. These parameters are usually imposed by the user based again on some prior knowledge and they can be eventually adapted in a trial and error setup by visualizing some results and modifying the parameters until they correspond to the expectation of a human operator. However, if a set of labeled vessel-images is available, they can be used to compute the parameters in an optimal way during a calibration step.

Vessel segmentation

Depending on the application the segmentation of vessels is done either in a automatic or in a semiautomatic manner. Semi-automatic segmentation methods are only feasible for applications where few images had to be segmented, but the results have to be of a higher quality and there is a human operator which is ready to invest sufficient time for this. If the segmentation algorithm needs to work automatically – e.g. as part of a medical-image-analysis tool – and/or if the number of images that had to be segmented is very high, then automatic methods need to be used. The automatic methods can be either supervised or unsupervised. Unsupervised segmentation methods have currently a larger applicability than supervised ones. The latter can be successfully used to analyze many images, which are acquired under very similar conditions, like e.g., during a screening campaign of a larger population sample.

Semi-automatic segmentation. During semi-automatic segmentation a human operator can influence and/or support the procedure to achieve better results than that obtained by an automatic algorithm. In this case, one can observe an optimal synergy between human and machine, each doing the task for which they are best suited. The computer does the painstaking job of deciding on a label for each pixel and the human can then use such information to extract only true vessel-structures.

The semi-automatic vessel segmentation algorithms described in Section 4.4 constitute the basis for a software package currently being used in medical research (see [132]). The Graphical User Interface (GUI) of this software package is shown in Figure 6.2. The user has the possibility to choose between fuzzy-clustering (also in [52], [51]) and hysteresis-based (also in [43]) vessel segmentation. He can then edit the results to achieve an optimal segmentation for the investigated vessel structures. As final results he can obtain the area and length of the vessels in an image, as well as the vessel density over a certain area – which he selects – in percents of the defined surface covered by vessels.

Automatic segmentation. Prior information can again be used to achieve a high-quality automatic segmentation. In this case, we made use of the vessel connectivity and their compact shape as well as their curly appearance and the fact that usually vessels of different sizes are connected to one another, as large vessels flow into smaller ones up to the capillaries [53].

Hysteresis classification. Such observations justify the application of the hysteresis classification paradigm to vessel segmentation (also in [53], [44]). Hysteresis thresholding as a bi-threshold procedure for edge detection is widely known since the seminal paper by Canny [34].
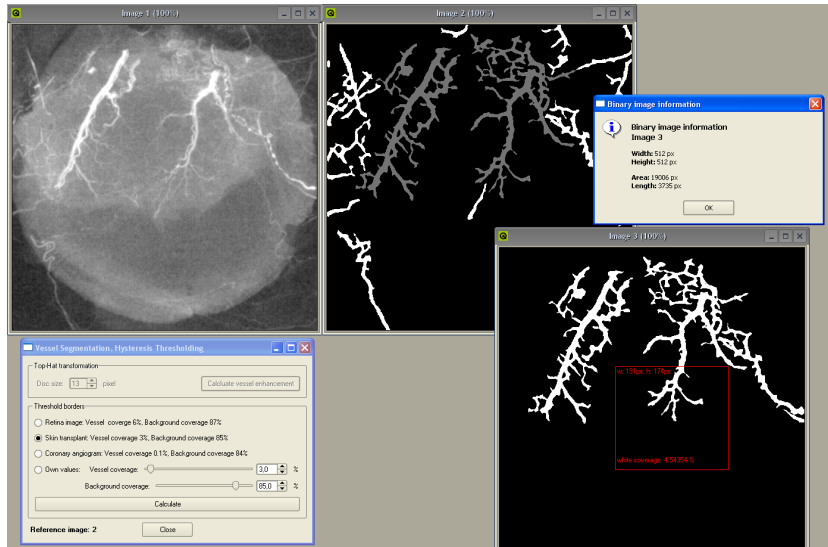
Figure 6.2: GUI of the semiautomatic vessel analysis tool (VesSeg).

In Section 4.3.2 we extend the initial hysteresis threshold to a hysteresis-classification methodology, including supervised and unsupervised approaches both for uni- and multi-dimensional data. We exemplify it on vessel segmentation but theoretically it can be used for many other similar problems as well [45].

We then introduced hysteresis classification as a bi-classifier procedure: one classifier, the "pessimist", selects only vessel points which are then used to choose true vessels from the segmentation result of the other classifier, the "optimist", which contains all vessel points. This implicitly assumes that vessels and background-structures are well separated from one another. If this is not the case, then background structures will also be selected in the segmentation. For example, in the case of coronary angiograms, for certain projections some background artefacts, like e.g., ribs, are well visible and as their appearance is similar to that of the vessels, they are also enhanced together with the vessels and if the vessels are projected against then they appear connected also to the vessels and will be included in the segmentation.

A hysteresis classification procedure is always used to separate two classes and it includes two base-classifiers. From this point of view the MRF-based adaptive threshold together with the junction-based selection of vessels represents also a hysteresis classification procedure. The MRF-threshold is the "optimist", while the method to segment the junction-points yields the "pessimist". We have also introduced a hysteresis classifier where both the "optimist" and the "pessimist" work on the same input and shown how to train it using the ROC curve. We have also used the ROC for classifier-specific feature selection as well as parameter estimation for various vessel maps. We believe that in the case of vessel segmentation and indeed for classification in many medical applications as well, the main purpose should be the optimization of the ROC rather than the optimization of the absolute error rate as in such cases the fp rate and the dynamic of the cc rate against the fp rate is also of major concern.

Vessel segmentation is characterized by major class-skew, i.e., the number of vessel pixels is far lower than that of background pixels. Such a skew is difficult to handle by many standard classification algorithms as they are designed to optimize accuracy [69]. Techniques based on an ROC-analysis as the supervised hysteresis classification procedures introduced here are better suited in this case [182], [80].
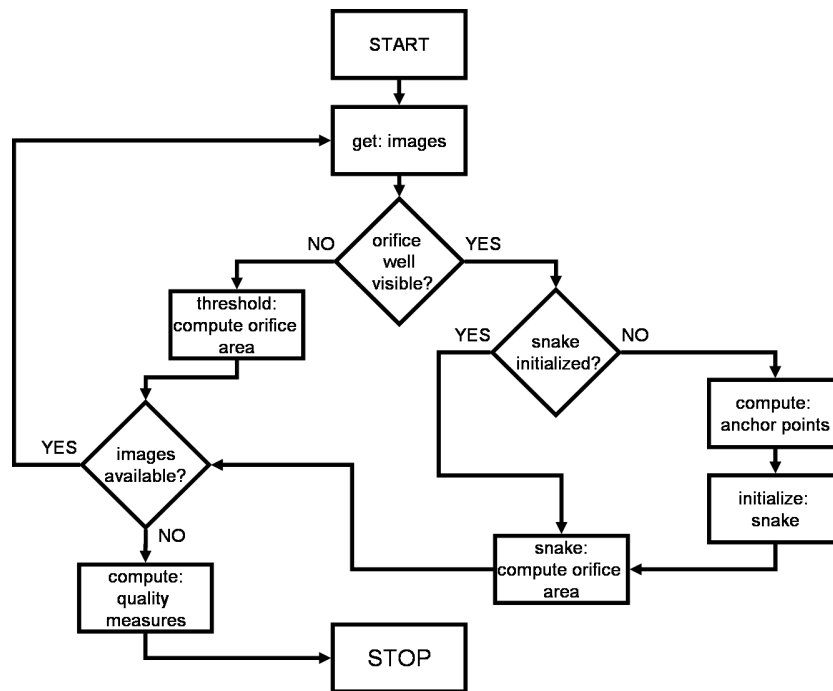
Figure 6.3: Flow-chart of the algorithm for the automatic computation of quality measures for heart valves.

6.3 Machine vision-enhanced quality inspection of xenograft heart valve implants

If a heart valve is ill, then in certain cases it is replaced by an implant during open-heart surgery. These implants can be either mechanical or biological. Some biological implants are valves collected from animals – i.e. xenograft valves – and placed on a supporting frame. The assessment of the quality of such implants is a major issue in heart-related health-care. A quality-conform valve is needed for a good postoperative evolution of the patient implying smaller treatment costs and longer periods of time to an eventual consecutive intervention. We have described automatic methods for the measurement of quality criteria for xenograft-heart-valve-implants (also in [56]).

For quality controll, an implant is placed in an artificial circulatory system and a valve cycle is initiated. The valve is observed by a high-speed camera whose image-plane is parallel to the surface of the closed valve. Image sequences acquired in such a test bed are used to evaluate the quality of the implant. For this purpose the evolution of the orifice-area over the valve-cycle and the fluttering of the valve-leaflets need to be measured. We have described here methods for the automatic measuring of such quality-parameters of heart-valve-implants, using only 2D-projection images acquired in the test-bed. Such measures provide the basis for a quantitative analysis of the quality of heart-valve implants with less inter and intra-observer variability. Although the experiments have been conducted only on tricuspid valves, we believe that these methods are applicable to bicuspid valves also. A flow-chart of the automatic measuring of quality parameters for heart valves, as it was described herein is shown in Figure 6.3.

Automatic measurement of the valve orifice. To segment the orifice in a sequence of valve-images, we have proposed to combine the advantages of two different methods: (i) for a few images, before the valve is fully open and also before it is fully closed, the orifice is segmented by a gray-level threshold, then as soon as the orifice is large enough, (ii) a snake is initialized starting from the base-lines of leaflets, which is then used to find the border of the orifice in the current image. Then, the converged snake is used as initialization for the next image.

To ensure that the snake follows the leaflets, which delineate the orifice, attraction basins are defined around the anchor points of a heart-valve, i.e., the points where two leaflets meet. This attraction basins follow the circular outline of the valve. The outline is detected by a Hough transform. The anchor points are computed staring from the centerlines of the orifice detected in some frames at the beginning of the sequence, when the valve opens.

Automatic measurement of the fluttering of leaflets. Besides the computation of the orifice curve, we have described also methods for an automatic quantification of the fluttering of the leaflets of the inspected valve providing thus the means for an enhanced analysis.

The fluttering can be observed under optimal conditions during the time when the valve is fully open, i.e., during the second phase of a valve-cycle. This phase is segmented on the orifice-curve. To measure the fluttering, we use the outline of each leaflet. The outline of a leaflet is segmented from the outline of the orifice using the anchor points. Then, for each leaflet, time-space images of its border are computed by gathering in an image the outline of the leaflet in each frame of the analyzed sequence. Such time-space images contain information both over the time and space behavior of the leaflet. The fluttering is measured as the energy-content in the high-pass part of the spectrum of such images.

Appendix A

Theoretical background

A.1 Morphological image processing

The purpose of this section is to review some well-known notions of morphological image processing which are used in this book. Details can be found in [70], [67], [68], [90], [179], [144], [89].

A.1.1 Binary morphology

Morphological image processing is a powerful tool in the analysis of segmentation results, i.e., binary images and is thus useful in many practical applications.

Basic definitions

Let $A = \{a_1, a_2, ..., a_{N^a}\}$ and $B = \{b_1, b_2, ..., b_{N^b}\}$ be two subsets of Z^2 and $\emptyset$ the empty set. Usually B is called the structuring element. The translation of B by $x \in Z^2$ is:

$$B_x = \{p | p = b + x, b \in B\} \tag{A.1}$$

The reflection of B is:

$$\hat{B} = \{x | x = -b, b \in B\} \tag{A.2}$$

The complement of A is:

$$\overline{A} = \{x | x \notin A\} \tag{A.3}$$

The difference of A and B is:

$$A - B = \{x | x \in A, x \notin B\} = A \cap \overline{B} \tag{A.4}$$

Dilation

The dilation of A by B is:

$$A \oplus B = \{x | \hat{B}_x \cap A \neq \emptyset\} \tag{A.5}$$

Erosion

The erosion of A by B is:

$$A \ominus B = \{x | B_x \subseteq A\} \tag{A.6}$$

Dilation and erosion are duals, which means that:

$$A \ominus B = \overline{\overline{A} \oplus \hat{B}} \tag{A.7}$$

Opening

The opening of a A by B is:

$$A \circ B = (A \ominus B) \oplus B \tag{A.8}$$

The opening can be used, e.g., to eliminate from A shapes smaller than B.

Closing

The closing of A by B is:

$$A \bullet B = (A \oplus B) \ominus B \tag{A.9}$$

The closing can be used to cover – close – holes smaller than B in larger objects from A.

The Hit-or-Miss transform

The Hit-or-Miss is used to detect objects with specific shapes in a segmentation result. The parameters of the transform are the sought shape X, its local background G, which is a small window that encloses X and the set of object-pixels in the segmentation result A. Considering the set $S = \{B_1, B_2\}$, with $B_1 = X$ and $B_2 = (G - X)$ the transform is computed as:

$$A \odot S = (A \ominus X) \cap [\overline{A} \ominus (G - X)] = (A \ominus B_1) \cap (\overline{A} \ominus B_2) \tag{A.10}$$

Thus, the result contains all points where simultaneously X found a match in A and $(G - X)$ found a match in $\overline{A}$.

Thinning

Thinning is one of the basic morphological algorithms. It is an iterative procedure and one iteration step can be defined in terms of the Hit-or-Miss transform as:

$$A \otimes S = A - (A \odot S) \tag{A.11}$$

The thinning of a set A using a set of shapes $\tilde{S} = \{S_1, S_2, ..., S_{N^s}\}$ is then:

$$A \otimes \tilde{S} = (...((A \otimes S_1) \otimes S_2)... \otimes S_{N^s}) \tag{A.12}$$

The thinning can be used to compute the centerline of segmented objects.

Selection of connected components

Assume that one knows a priori that a point p belongs to a connected object Y. At the same time, in the segmentation result several items have been selected, including Y. Then, using p and the connectivity of Y, one can select Y in the following way:

$$X_k = (X_{k-1} \oplus B) \cap A \tag{A.13}$$

where $X_0 = p$, $X_k, k = \{1, 2, ..., N\}$ are subsets of Y such that $X_N = Y$ and B is a suitable structuring element.

A.1.2 Gray-level morphology

Initial research in the field of morphological image processing was done exclusively on binary images [141]. However, later research recognized the possibilities which lie in the extension of such methods to gray-level images and reported on modalities suited for such a task [96].

Basic definitions

Let $f(x, y) : Z^2 \rightarrow R$ and $b(k, l) : Z^2 \rightarrow R$ be two image functions with b being usually called the structuring element.

Dilation

The dilation of f by b is defined as:

$$f \oplus b = max\{f(x, y) + b(x - k, y - l) | (x, y) \in D_f, (k, l) \in D_b\} \tag{A.14}$$

with D_f the domain of f and D_b the domain of b. This is equivalent to taking the maximum of $f + b$ in a sliding window whose shape is defined by D_b. Consequently, by dilation dark details smaller than D_b are attenuated.

Erosion

The erosion of f by b is defined as:

$$f \ominus b = min\{f(x, y) - b(x - k, y - l) | (x, y) \in D_f, (k, l) \in D_b\} \tag{A.15}$$

This is equivalent to taking the minimum of $f - b$ in a sliding window whose shape is defined by D_b. Consequently, by erosion bright details smaller than D_b are attenuated. Erosion and dilation are dual operations.

Opening

The opening of f by b is defined as:

$$f \circ b = (f \ominus b) \oplus b \tag{A.16}$$

Closing

The closing of f by b is defined as:

$$f \bullet b = (f \oplus b) \ominus b \tag{A.17}$$

Opening and closing are dual operations. They are also idempotent. Thus applying a second time the opening with the same structuring element upon the result of a first opening returns the same result.

The Tophat transform and the Bothat transform

The Tophat transform is defined as:

$$Th = f - (f \circ b) \tag{A.18}$$

and it can be used to attenuate a dark background and select bright objects comparable in shape and size with D_b.

The Bothat transform is defined as:

$$Bh = (f \bullet b) - f \tag{A.19}$$

and it can be used to attenuate a bright background and to select dark objects comparable in shape and size with D_b. The Tophat and the Bothat are duals.

A.2 Linear shift-invariant systems

The purpose of this section is to review some aspects related to the design and particularities of filters and of FIR filters in particular. Further details can be found in [154], [90], [110].

The following discussion is for linear time-invariant systems, i.e., for 1D signals. The extension to 2D-signals and thus to linear shift-invariant systems is straightforward.

A system is called linear if it satisfies the principle of superposition, i.e., is characterized by the properties of additivity:

$$T\{x_1 + x_2\} = T\{x_1\} + T\{x_2\} \tag{A.20}$$

and scaling:

$$T\{ax\} = aT\{x\} \tag{A.21}$$

with a an arbitrary constant and the output y computed by modifying the input x according to T.

A system is called time-invariant if T remains constant in time. Thus, if the same input is applied at the time instance τ_1 and then again later at τ_2, the same output is recorded both at τ_1 and τ_2.

The most important parameter of a filter is its frequency behavior. The filter's impulse response can then be directly computed using the Fourier transform. Ideally the filter should have infinitely steep transitions between its pass and its cut bands which would require an infinite impulse response in form of a sinc-function. As such a solution is impracticable different

approximations to of the impulse response can be used. There are two main classes of such approximations which result in two types of filters: approximating functions of infinite support and recursive implementation result in Infinite Impulse Response (IIR) filters and approximating functions of finite support result in Finite Impulse Response (FIR) filters.

IIR filters have in some applications problems with the stability of their answer and with the linearity of their phase. Such problems are usually a consequence of the recursive implementation of such filters. However, they are fast to compute and can give very good approximations of an ideal filter. Examples of IIR filters include the Chebyshev filter and the Butterworth filter.

FIR filters are always stable and have a linear phase, but require more resources to obtain an approximation of the ideal filter comparable with the IIR filters. They are also easy to implement by convolution. Practically they are more often used than IIR filters. FIR filters can be designed by different methods, like e.g., the window method. Knowing the limits of the filter-pass-band, it is desired to find the optimal form of the filtration window such that there are as good as no oscillations in the pass and cut bands and the transition between them is as steep as possible. To avoid disadvantageous phase-shifts – i.e. phase-shifts that can not be compensated by shifting the signal over an integer number of samples – the number of samples in the time-domain filtration window – i.e. the size of the window – should be odd.

Based on their frequency-behavior, the filters can be divided in three large classes:

- *Low-Pass.* Such filters attenuate high-frequency components, letting the low frequent components pass. They will blur an image. The sum of their coefficients in the spatial domain is one.

- *Band-Pass.* Such filters attenuate both low and high-frequency components between certain limits. They can select, e.g., ridges.

- *High-Pass.* Such filters are the counterpart of the Low-Pass filters, suppressing the low-frequency components. They can be used to select, e.g., edges. The sum of their coefficients in the spatial domain in zero.

Due to the linearity property, e.g., the result of the high-pass filtration can be computed also as the difference between the original image and a low-pass filtration result with a properly chosen cut-off frequency.

As practical filters are only approximations to the ideal case, sometimes artefacts from frequency components, which have not been properly attenuated, can be seen.

A.3   Active Contours

There are two large classes of snakes: parametric [119] and geometric [171], [136]. The purpose of this Section is to review some theoretical aspects related to parametric closed snakes, open snakes can be modeled in a similar fashion – further details can be found in the bibliographic references mentioned herein.

Active-contours have been brought to the attention of the computer vision community by the seminal contribution of Kass, Witkin and Terzopoulos [119]. The following definition of snakes can be found in there, "a snake is an energy-minimizing curve guided by external constraint forces and influenced by image forces that pull it towards features such as lines and edges."

Since then they have been extended [22], [188], [150], [59], [58], [42], [196], [198], [171] and applied in a variety of domains including medical machine vision [142], [75], [55].

A snake is a curve $\mathbf{v}(s) = [x(s), y(s)]$, $s \in [0, 1]$ evolving over the spatial domain of an image under the influence of forces derived from internal and external energy terms to a position of minimum energy.

The internal energy term governs the behavior of the snake during its evolution and it is defined to penalize stretching and bending, encouraging thus smooth solutions:

$$E_{int} = \int_0^1 \frac{1}{2}(\alpha(s)|\frac{\partial \mathbf{v}(s)}{\partial s}|^2 + \beta(s)|\frac{\partial^2 \mathbf{v}(s)}{\partial s^2}|^2)ds \tag{A.22}$$

The first-order term controls stretching and the second order bending. The weights $\alpha(s)$ and $\beta(s)$ adjust the importance of the two terms and favor thus a specific snake behavior.

The external energy term is computed from the image such that it reaches a minimum over the sought image features:

$$E_{ext} = -\int_0^1 \mathbf{P}(\mathbf{v}(s))ds \tag{A.23}$$

with e.g.:

$$\mathbf{P}(x, y) = -|\nabla(G_\sigma(x, y) * I(x, y)))|^2 \tag{A.24}$$

to attract the snake towards edges. In this case, $G_\sigma(x, y)$ is a 2D Gaussian kernel used to extend the attraction range of the sought features. As usual, the detection-localization tradeoff is controlled by the standard deviation σ. Depending on the application other external energy terms can be also defined to push the snake to other image features such as lines or terminations [119].

The snake energy-functional can then be written as $E = E_{int} + E_{ext}$. The snake will evolve such that E is minimized. One actually searches for a minimum of E_{ext}, but this is usually a non-convex problem with a lot of local-minima, then the convex energy-term E_{int} can formally be regarded as regularizing the problem such that a solution is found at a reasonable local minima. This minimization problem is solved by variational calculus, i.e., by solving the corresponding Euler-Lagrange equation:

$$\frac{dE}{ds} = 0 \implies \alpha(s)\frac{\partial^2 \mathbf{v}(s)}{\partial s^2} - \beta(s)\frac{\partial^4 \mathbf{v}(s)}{\partial s^4} - \nabla P(\mathbf{v}(s)) = 0 \tag{A.25}$$

which can be viewed as a force-balance equation: $F_{int} + F_{ext} = 0$.

Usually $\alpha(s)$ and $\beta(s)$ are constant over the snake and the solution is sought in an iterative manner, i.e., the snake is made dynamic by treating it as a function of time also: $\mathbf{v}(s, t)$. Then the minimization equation becomes:

$$\alpha\frac{\partial^2 \mathbf{v}(s, t)}{\partial s^2} - \beta\frac{\partial^4 \mathbf{v}(s, t)}{\partial s^4} - \nabla P(\mathbf{v}(s)) = \frac{\partial \mathbf{v}(s, t)}{\partial t} \tag{A.26}$$

and a solution is found when the snake stabilizes and the partial derivative of $\mathbf{v}$ with respect to time vanishes.

Sometimes, additional energy terms are used to further influence the behavior of the snake in the search of better solutions. Such terms include, e.g., a damping/viscosity term to take into account energy dissipation during the displacement of the snake:

$$D = \frac{1}{2}\int_0^1 \gamma(s)\|\frac{\partial \mathbf{v}(s, t)}{\partial t}\|^2 ds \tag{A.27}$$

with $\gamma(s)$ the damping coefficient which again is usually constant over the snake. Practically, this term constrains the snake to evolve in relatively small steps. Yet another largely used additional energy term is the pressure or balloon term [42]:

$$E_b = \int_0^1 \kappa(s)\mathbf{n}(s)ds \qquad (A.28)$$

with $\mathbf{n}(s)$ the normal unitary vector of the curve, gives the direction of the balloon force. The sense of the balloon force is practically given by the sign of the weight κ, which again is usually constant over the curve. This term forces the snake to evolve in a certain direction even if it is placed outside the attraction range of the sought image features. This term can compensate for external energy fields with relatively small attraction basins.

Usual external energies have a rather limited capture range and show poor convergence in regions where the sought feature exhibits a concavity. Balloons on the other hand will simply role over features characterized by a poor SNR. Gradient vector flow snakes [198] offer a solution to such problems. A GVF $\mathbf{g}(x,y) = (u(x,y), v(x,y))$ is a static external force field. If $f(x,y) = -P(x,y)$, then $\mathbf{g}$ is defined as the vector field that minimizes the energy functional:

$$\mathcal{E} = \int\int \mu(\frac{\partial u}{\partial x}^2 + \frac{\partial u}{\partial y}^2 + \frac{\partial v}{\partial x}^2 + \frac{\partial v}{\partial y}^2) + \|\nabla f\|^2\|\mathbf{g} - \nabla f\|^2 dxdy \qquad (A.29)$$

Consequently, $\mathbf{g}$ is equal to the gradient of the external energy, when this is large, and is defined by its own partial derivatives in the rest, thus working against abrupt changes in the external energy and forcing the field to be slow varying in homogeneous regions, thereby greatly extending the attraction range.

The solution to this minimization problem is again found iteratively:

$$\frac{d\mathbf{g}}{dt} = \mu\nabla^2\mathbf{g} - (\mathbf{g} - \nabla f)\|\nabla f\|^2 \qquad (A.30)$$

Such active contours are also called parametric snakes. Clearly it is rather difficult to force them to split or merge during their evolution. There are also geometric snakes [171], [172], [135], [136], [37], [36], [157], which are defined as the level-sets of a multidimensional function evolving under the guidance of a speed term defined in relation to the sought image features.

A.4 The Hough transform

The standard Hough-transform has been first introduced by Hough in [102]. Later Ballard showed how to compute a generalized Hough transform [15]. The purpose of this section is to revisit some theoretical aspects related to the standard Hough transform – further details can be found in the references mentioned herein.

The Hough transform was originally developed to detect straight lines in black and white images [102]. Since then, it has become a general accepted tool for detecting parametric and also arbitrary curves in images [106], [15], [92], [201], [125], [126]. The key idea is to replace the complicated problem of finding the instances of a certain curve among contour points obtained, e.g., at the output of an edge detector, with the more simple problem of detecting a peak in a transformed space. The axis of the transformed space (accumulator) are the free parameters of the curve.

The transform can be used theoretically with all parameterizable curves but without additional restrictions, curves with more than two parameters make the accumulator impractically large, therefore the Hough transform is usually used to find lines and circles.

A.4.1   Hough transform for lines

From line detection to intersection. Any line $y = mx + n$ is identified by a unique parameter pair (m, n) and therefore, all points on the line are mapped to a point in the parameter space. Conversely, for each point on the line one can write: $n = y - mx$ and thus a point is mapped into a line in the parameter space. Then the parameter-space-point of the sought line will be found at the intersection of the parameter-space-lines corresponding to each of the collinear-points.

From intersection to peak detection. The plane which defines the transformed space is divided into a finite grid of cells, with a counter – initialized to zero – attached to each cell. The counter is incremented by a unit each time a parameter-space-line passes trough the cell. Thus, the counter of the cell containing the intersection of several lines will have a higher count. To find a line, one has simply to set a threshold over the counters, i.e., in the parameter space. Clearly, the precision of the transformation depends on the size of the accumulator cells.

Practical computation. Both m and n take values on the real axis, therefore a complete partition of the parameter space is impossible. One can either limit m and n and/or use cells of very large size, thus losing interesting parameter-ranges. Also, the (m, n) parameterization does not capture the bundle $x = k$ with k a constant. To solve at least partially such problems a polar representation of the line is usually used: $d = xcos(\alpha) + y(sin(\alpha))$ with d the distance from the image-origin to the line. In this case α is finite $0 \leq \alpha \leq \pi$ and d facilitates a denser sampling as before: $0 \leq d \leq d_{max}$ With this parameterization, a point is represented by a parameter-sinusoid.

A.4.2   Hough transform for circles

The parametric representation of a circle is: $(x - a)^2 + (y - b)^2 = r^2$ where (a, b) are the center cartesian coordinates and r is the radius.

From circle detection to intersection. The Hough accumulator for detecting circles will be a 3D space whose axes are the two center coordinates and the radius. To each presumptive circle point in the original plane, corresponds a cone in the Hough transformed space. The cone contains the parameters of all possible circles passing through that point.

From intersection to peak detection. Under the simplifying assumptions that the Hough space is discrete and finite, a point on the circle in the original plane casts a vote on each accumulator cell which belongs to its cone. Thus, the accumulator cell corresponding to the parameters of the circle passing through a majority of the investigated points will get the highest number of votes. The desired circle is represented by the maximum in the accumulator.

A.5   Orientation estimation

The purpose of this Section is to review the basic tools for the estimation of a single local orientation in images. Further details can be found in [12], [118], [108], [19], [188].

The 2D differential operator in direction ϕ with respect to the horizontal axis is defined as:

$$\alpha(\phi) = cos(\phi)\frac{\partial}{\partial x} + sin(\phi)\frac{\partial}{\partial x} \tag{A.31}$$

If the image $f(x, y)$ is ideally oriented at (x, y) under an angle θ, then its derivative in along θ should be zero:

$$\alpha(\theta)f(x, y) = \mathbf{v}\nabla f = 0 \tag{A.32}$$

with $\mathbf{v}^T = (cos(\theta), sin(\theta))$. When this condition is met for θ is met also for $\theta + \pi$ and for $\theta - \pi$, thus by convention $\theta \in (-\frac{\pi}{2}\frac{\pi}{2})$.

Local orientation is evaluated over a neighborhood Ω within which it is assumed to be constant. Minimizing the square error in the case of Equation(A.32) with respect to the angle yields:

$$\theta = \min_{\phi} Q(\phi) = \min_{\phi}\left[\int_{\Omega}(\mathbf{v}^T\nabla f)^2)d\Omega\right] \tag{A.33}$$

with $\mathbf{v}^T\mathbf{v} = 1$. This can be rewritten as:

$$Q(\phi) = \mathbf{v}^T\mathbf{T}\mathbf{v} \tag{A.34}$$

with $\mathbf{T}$ being a 2×2 tensor computed as:

$$T = \int_{\Omega}\nabla f(\nabla f)^T d\Omega = \int_{\Omega}\begin{bmatrix} f_x^2 & f_x f_y \\ f_y f_x & f_y^2 \end{bmatrix} d\Omega \tag{A.35}$$

Minimizing the composite criterion:

$$L(\mathbf{v}) = \mathbf{v}^T\mathbf{T}\mathbf{v} + \lambda(\mathbf{v}^T\mathbf{v} - 1) \tag{A.36}$$

including the condition $\mathbf{v}^T\mathbf{v} = 1$ over the Lagrange multiplier λ, is equivalent to finding $\mathbf{v}$ such that:

$$\mathbf{T}\mathbf{v} = \lambda\mathbf{v} \tag{A.37}$$

i.e. finding the normalized eigenvector of $\mathbf{T}$ corresponding to the lower eigenvalue λ. As $Q(\phi)$ is a measure of variation of $f(x, y)$ in the direction ϕ, then the minimum quantity of variation in Ω is given by λ:

$$Q(\phi) = \mathbf{v}^T\mathbf{T}\mathbf{v} = \mathbf{v}^T\lambda\mathbf{v} = \lambda \tag{A.38}$$

$\mathbf{T}$ is also known as the structure tensor and its eigenvalues can be used to analyze the local orientation [108]:

Practically, $\mathbf{T}$ is computed as:

$$\begin{bmatrix} S(D_x \cdot D_x) & S(D_x \cdot D_y) \\ S(D_y \cdot D_x) & S(D_y \cdot D_y) \end{bmatrix} \tag{A.39}$$

with S a smoothing operator which defines the size of the neighborhood and $D_{x,y}$ the derivative operator on direction x respectively y. The results obtained by derivation are multiplied pixelwise.

$\lambda_1 = \lambda_2 = 0$	The neighborhood has constant values.
$\lambda_1 > 0, \lambda_2 = 0$	No variation in the neighborhood in the direction of the eigenvector corresponding to λ_2. The neighborhood is ideally oriented.
$\lambda_1 > 0, \lambda_2 > 0$	Variations in both directions. If $\lambda_1 > \lambda_2$, there is one dominant orientation. A large λ_2 is an indicator for a possible corner in the neighborhood.

Table A.1: Structure-tensor-eigenvalues-based orientation analysis in images.

A.6   Discriminant analysis

The quality of a feature space is defined by its ability to build classes which can then be correctly separated by a classification algorithm. The simpler the classification algorithm needed for this task, the better the feature space. Statistical measures of separability in the sense of [87], [184], [149] use the *within-class* and the *between-class* scatter matrices. Such measures can be used both for feature selection and for unsupervised segmentation.

A.6.1   Statistical measures of separability between classes

The *within-class* scatter matrix is defined as:

$$S_w = \sum_{i=1}^{N} P_i \Sigma_i \tag{A.40}$$

with N the number of classes, P_i the a priori class-probability and Σ_i the covariance matrix for class i.

The *between-class* scatter matrix is defined as:

$$S_b = \sum_{i=1}^{N} P_i (\mu_i - \mu_0)(\mu_i - \mu_0)^T \tag{A.41}$$

where μ_i is the mean of class i and μ_0 is the mean of all components of the feature space.

A *mixture scatter matrix* can as well be defined as:

$$S_m = S_w + S_b \tag{A.42}$$

being actually the covariance of the entire feature space regardless of class-assignments.

Using these matrices – which are invariant under coordinate shifts – several separability measures can be defined. These measures yield extreme values for feature spaces where the *between-class* scatter matrix is large – i.e. well-separated means of the class-conditional distributions – and/or the *within-class* scatter matrix is small – i.e. small variances around the means of each class. One of the most often used criteria is:

$$J_1^a = tr(S_w^{-1} S_b); \tag{A.43}$$

which is also invariant under non-singular linear transformations. In this case J_1^a is an increasing function of the separability of the feature space thus reaching a maximum for the most separable feature space.

In some applications it is interesting to assess the two components of separability – i.e. the within and between-class scatter matrices – independently. For this purpose, J_1 should be defined as:

$$J_1^b = tr(S_m^{-1} S_b);$$
(A.44)

to investigate only how well-separated are the classes by their means. Again J_1^b is a an increasing function of the separability. To investigate how well do the feature space components cluster around their means the J_1 should be defined as:

$$J_1^c = tr(S_m^{-1} S_w);$$
(A.45)

in which case J_1^c is a decreasing function of the separability, thus reaching a minimum for the most separable feature space.

A.6.2 Automatic threshold selection for two-class segmentation

The purpose of this section is to revisit a method for automatic threshold selection (two-class unsupervised segmentation) in images first proposed by Otsu [155]. It uses discriminant analysis to evaluate several candidate-thresholds and chooses then the optimal one.

For an image with N pixels and L gray-levels, the normalized histogram can be regarded as an approximation of the gray-level pdf. For each gray-level i the corresponding pdf value is:

$$p_i = \frac{n_i}{N}$$
(A.46)

with $p_i \geq 0$ and $\sum_{i=1}^{L} p_i = 1$.

For a certain threshold $T = k$, the gray-levels $[1, k]$ belong to class C_0 and $[k+1, L]$ to class C_1. Then, the prior probabilities for each class can be computed as:

$$\omega_0 = \sum_{i=1}^{k} p_i = \omega(k) \qquad \omega_1 = \sum_{i=k+1}^{L} p_i = 1 - \omega(k)$$
(A.47)

the means as:

$$\mu_0 = \sum_{i=1}^{k} i P(i|C_0) = \frac{\mu(k)}{\omega(k)} \qquad \mu_1 = \sum_{i=k+1}^{L} i P(i|C_1) = \frac{\mu_T - \mu(k)}{1 - \omega(k)}$$
(A.48)

and the class-variances as:

$$\sigma_0^2 = \sum_{i=1}^{k} (i - \mu_0)^2 P(i|C_0) \qquad \sigma_1^2 = \sum_{i=k+1}^{L} (i - \mu_1)^2 P(i|C_1)$$
(A.49)

with $\mu(k) = \sum_{i=1}^{k} i p_i$ the first-order moment up to histogram-level k and $\mu_T = \sum_{i=1}^{L} i p_i$ the mean of the image.

The quality of the threshold T can be established by evaluating how separable are the two classes which it yields, using the *within-class* σ_w^2 variance and the *between-class* variance σ_b^2:

$$\sigma_w^2 = \omega_0 \sigma_0^2 + \omega_1 \sigma_1^2 \qquad \sigma_b^2 = \omega_0 (\mu_0 - \mu_T)^2 + \omega_1 (\mu_1 - \mu_T)^2$$
(A.50)

Suitable goodness measures are (see also Appendix):

$$\lambda = \frac{\sigma_b^2}{\sigma_w^2} \kappa = \frac{\sigma_T^2}{\sigma_w^2} \eta = \frac{\sigma_b^2}{\sigma_T^2} \tag{A.51}$$

with

$$\sigma_T^2 = \sum_{i=1}^{L} (i - \mu_T)^2 p_i \tag{A.52}$$

being the total-variance. They are equivalent, but η is the simplest with respect to k, as it contains only first-order statistics and σ_T^2 is independent of k.

The optimal threshold is selected by sequential search to maximize the chosen goodness measure:

$$T = \arg \max_{1 \leq k \leq L} (\eta(k)) \tag{A.53}$$

Appendix B

Glossary of medical terms

A

Angina pectoris	-	chest-pain
Angiogenesis	-	the formation of new blood vessels
Angiography	-	examination of the blood vessels using X-rays following the injection of a radiopaque substance
Angioplasty	-	the surgical repair of a blood vessel – plastic deformation of a blood vessel
Aorta	-	the main trunk of the systemic arteries, carrying blood from the left side of the heart to the arteries of all limbs and organs except the lungs
Aortic valve	-	heart valve situated between the left ventricle and the aorta. It has three leaflets
Arrhythmia	-	changes in the normal rhythm of the heartbeats
Arteries	-	any of the muscular elastic tubes that form a branching system and that carry blood away from the heart to the cells, tissues, and organs of the body
Atherosclerosis	-	the deposition of atheromatous plaques containing cholesterol and lipids on the innermost layer of the walls of large and medium-sized arteries
Atria	-	the chambers of the heart that collect blood and allow it to return to the heart.

B

Balloon - PTCA	-	a small balloon, which is advanced on the guidewire to the site of the blockage and inflated there to reopen the artery
Balloon-markers	-	small blobs of an X-ray absorbing material placed at the two ends of a PTCA-balloon
Blood clot	-	a clump of blood that forms in or around a vessel as a result of coagulation

C

Capillaries	-	the tiny blood vessels throughout the body that connect arteries and veins. Capillaries form an intricate network around body tissues in order to distribute oxygen and nutrients to the cells and remove waste substances
Chordae tendinae	-	cord-like tendons that connect the papillary muscles to the tricuspid valve and the mitral valve in the heart
Coronary arteries	-	the arteries supplying blood to the heart

D

Diaphragm	-	a muscular membranous partition separating the abdominal and thoracic cavities
Diabetic retinopathy	-	damage to the retina caused by complications of diabetes, which could eventually lead to blindness
Diabetes	-	a chronic disease characterized by the inability of the body to produce or respond properly to insulin, a hormone required by the body to convert glucose to energy

E

Endocarditis	-	an inflammation of the endocardium. The most common structures involved are the heart valves
Endocardium	-	the inner layer of the heart

F

Fascia	-	specialized connective tissue-layer, which surrounds muscles, bones, and joints, providing support and protection and giving structure to the body
Fasciocutaneous flap	-	flap comprising skin, subcutaneous fat and fascia but no muscles

G

Guidewire	-	a thin wire threaded into the ill coronary artery and used to guide the PTCA-balloon and the stent to the ill site
Guidewire tip	-	the maneuverable tip of the guidewire

H

Heart	-	the chambered muscular organ in vertebrates that pumps blood received from the veins into the arteries, thereby maintaining the flow of blood through the entire circulatory system. The heart has four chambers two atria and two ventricles
Heart failure	-	when the heart is not able to pump blood to the rest of the body effectively
Heart attack	-	sudden interruption or insufficiency of the supply of blood to the heart, typically resulting from occlusion or obstruction of a coronary artery and often characterized by severe angina pectoris
Heart valve	-	valves in the heart that maintain the unidirectional flow of blood by opening and closing depending on the difference in pressure on each side. There are four heart valves: the mitral valve, the tricuspid valve, the aortic valve and the pulmonic valve
Hypertension	-	chronic high blood pressure

I

Ischemia	-	a decrease in the blood supply to a bodily organ, tissue, or part caused by constriction or obstruction of the blood vessels

L

Leaflet	-	a flap, which is a constituting part of a heart-valve. All heart-valves have three leaflets with the exception of the mitral valve which has only two

M

Metabolism	-	the chemical processes occurring within a living cell or organism that are necessary for the maintenance of life
Mitral valve	-	heart valve situated between the left atrium and the left ventricle. It has two leaflets
Myocardium	-	the muscular tissue of the heart
Myocardial infarction	-	see Heart attack

N

Nutrient	-	any element or compound necessary for or contributing to an organism's metabolism, growth, or other functioning

O

Ostium	-	a small opening or orifice, as in a body organ or passage

P

Papillary muscles	-	serve to limit the movements of the mitral and tricuspid valves and prevent them from being inverted
Plaque	-	fatty deposits on the inner wall of a blood vessel
Pathology	-	the scientific study of the nature of disease and its causes, processes, development, and consequences
Percutaneous	-	passed trough the skin
Peripheral Angiography	-	angiography of the arteries supplying blood to the arms, legs stomach, kidneys.
Physiology	-	the functions of living organisms and their parts
Prosthetic	-	serving as or relating to a prosthesis
Pulmonary artery	-	an artery that carries oxygen depleted blood from the right ventricle of the heart to the lungs
Pulmonary veins	-	vein that carries oxygen rich blood from the lungs to the left atrium of the heart
Pulmonic valve	-	heart valve situated between the right ventricle and the pulmonary artery. It has three leaflets

S

Septum	-	a partition separating two cavities or two spaces containing a less dense material
Stenosis	-	a constriction or narrowing of a vessel
Stent	-	an expandable wire form or perforated tube that is inserted into a vessel to prevent or counteract a stenosis

T

Tendon	-	a tough band of fibrous connective tissue that connects muscle to bone
Thrombolytic	-	dissolution or destruction of a thrombus
Thrombus	-	a blood clot formed in a blood vessel and remaining attached to its place of origin
Transluminal	-	passing across a lumen, as of a blood vessel
Tricuspid valve	-	heart valve situated between the right atrium and the right ventricle. It usually has three leaflets, but it may occur that it has two or four

V

Vascular disease	-	the disease which is mainly caused by atherosclerosis
Veins	-	any of the membranous tubes that form a branching system and carry blood to the heart
Vena cava superior	-	a large but short vein, that carries oxygen depleted blood from the upper half of the body to the right atrium
Vena cava inferior	-	a large vein, that carries oxygen depleted blood from the lower half of the body into the heart
Ventricles	-	the chambers of the heart, which collect blood from the atria and pump it out of the heart

List of Figures

List of Tables

Bibliography

[1] T. Aach. *Bayes-Methoden zur Bildsegmentierung, Änderungsdetektion und Verschiebungsvektorschätzung.* VDI Verlag, 1993.

[2] T. Aach. Lecture on computer vision. University of Lübeck, 2002-2004.

[3] T. Aach. Quantitative comparison of shift variance and cyclostationarity in multirate filter banks. In *Proceedings of the international workshop SMMSP-2004*, pages 7–14, Vienna, Austria, September 11–12 2004. TICSP Series.

[4] T. Aach. New criteria for shift variance and wide-sense cyclostationarity in multirate filter banks. In *Proceedings of the international Workshop SMMSP-2006*, pages 7–13, Florence, September 2–3 2006. TICSP Series.

[5] T. Aach and A. P. Condurache. Transformation of adaptive thresholds by significance invariance for change detection. In *Proceedings of the 13'th workshop SSP-2005*, Bordeaux, France, July 17–20 2005. IEEE.

[6] T. Aach, A. P. Condurache, K. Eck, and J. Bredno. Statistical-model based identification of complete vessel-tree frames in coronary angiograms. In *Electronic Imaging 2004: Computational Imaging II*, volume 5299, pages 283–294, San Jose, CA, USA, January 18–22 2004. SPIE.

[7] T. Aach, U. Franke, and R. Mester. Top-down image segmentation using object detection and contour relaxation. In *Proceedings of ICASSP-1989*, pages 1703–1706, Glasgow, UK, May 23–26 1989. IEEE.

[8] T. Aach and A. Kaup. Bayesian algorithms for adaptive change detection in image sequences using Markov random fields. *Elsevier Signal Processing: Image Communication*, 7:147–160, 1995.

[9] T. Aach, C. Mayntz, P. Rongen, G. Schmitz, and H. Stegehuis. Spatiotemporal multiscale vessel enhancement for coronary angiograms. In *Medical Imaging: Image Processing*, volume 4684, pages 1010–1021, San Diego, CA, USA, February, 23–28 2002. SPIE.

[10] T. Aach, C. Mota, I. Stuke, M. Mühlich, and E. Barth. Analysis of superimposed oriented patterns. *IEEE Transactions on Image Processing*, 25(11):2–16, 2006.

[11] T. Aach, U. Schiebel, and G. Spekowius. Digital image acquisition and processing in medical X-ray imaging. *Electronic Imaging, Special Section on Biomedical Image Representation*, 8(1):7–22, 1999.

[12] T. Aach, I. Stuke, C. Mota, and E. Barth. Estimation of multiple local orientations in image signals. In *Proceedings of ICASSP-2004*, pages III 553–556, Montreal, Canada, May 17–21 2004. IEEE.

[13] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19:716–723, 1974.

[14] E. Backer and A. K. Jain. A clustering performance measure based on fuzzy set decomposition. *IEEE Transactions on Pattern Anaysis and Machine Intelligence*, 3(1):66–75, 1981.

[15] D. H. Ballard. Generalizing the Hough transform to detect arbitrary edges. *Pattern Recognition*, 13(2):111–122, 1981.

[16] A. R. Barron, J. Rissanen, and B. Yu. The minimum description length principle in coding and modeling. *IEEE Transactions on Information Theory*, 44(6):2743–2760, 1998.

[17] E. Barth, T. Caelli, and C. Zetzsche. Image encoding, labeling, and reconstruction from differential geometry. *Graphical Model and Image Processing*, 55(6):428–446, 1993.

[18] B. Belaroussi, H. Benoit-Cattin, and C. Odet. Scalable discrepancy measures for segmentation evaluation. In *Proceedings of ICIP-2002*, pages 785–788, Rochester, NY, USA, September 22–25 2002. IEEE.

[19] J. Bigün and G. H. Granlund. Optimal orientation detection of linear symmetry. In *Proceedings of ICCV-1987*, pages 433–438, London, UK, June 8–11 1987. IEEE.

[20] J. Bilmes. *A gentle tutorial on the EM algorithm and its application to parameter estimation for Gaussian mixture and hidden Markov models.* Technical Report, University of Berkeley, ICSI-TR-97-021, 1997., 1997.

[21] C. M. Bishop. *Neural networks for pattern recognition.* Oxford University Press, 1995.

[22] A. Blake and M. Isard. *Active Contours.* Springer, 1998.

[23] C. Blondel, G. Malandain, R. Vaillant, and N. Ayache. 4D deformation field of coronary arteries from monoplane rotational X-ray angiography. In *Proceedings of CARS-2003*, pages 1073–1078, London, UK, June 25–28 2003.

[24] F. L. Bookstein. Principle warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989.

[25] A. P. Bradley. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7):1145–1159, 1997.

[26] J. Bredno, B. Martin-Leung, and K. Eck. Algorithmic solutions for live device-to-vessel match. In *Medical Imaging: Image Processing*, volume 5370, pages 1486–1497, San Diego, CA, USA, February 14–19 2004. SPIE.

[27] J. D. Bronzino. *The biomedical engineering handbook.* CRC-Press, 2000.

[28] H. Bunke and A. Sanfeliu. *Syntactic and Structural Pattern Recognition.* World Scientific, 1990.

[29] C. J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2):121–167, 1998.

[30] P. J. Burt and E. H. Adelson. The Laplacian pyrmaid as compact image code. *IEEE Transactions on Communications*, 31(4):532–540, 1983.

[31] T. M. Buzug, J. Weese, C. Lorenz, and W. Beil. Histogram-based image registration for digital subtraction angiography. In *Proceedins of ICIAP-1997*, volume 2 of *Lecture Notes in Computer Science*, pages 380–387, Florence, Italy, September 17–19 1997. Springer.

[32] T. M. Buzug, J. Weese, and K. C. Strasters. Motion detection and motion compensation for digital subtraction angiography image enhancement. *Philips Journal of Research*, 51(2):203–229, 1998.

[33] A. Can, H. Shen, J. N. Turner, H. L. Tanenbaum, and B. Roysam. Rapid automated tracing and feature extraction from retinal fundus images using direct exploratory algorithms. *IEEE Transactions on Information Technology in Biomedicine*, 3(2):125–138, 1999.

[34] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.

[35] J. S. Cardoso and L. Corte-Real. Toward a generic evaluation of image segmentation. *IEEE Transactions on Image Processing*, 14(11):1773–17872, 2005.

[36] V. Caselles, F. Catte, T. Coll, and F. Dibos. A geometric model for active contours. *Numerische Mathematik*, 66:1–31, 1993.

[37] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision.*, 22(1):61–79, 1997.

[38] P.-H. Chen, C.-J. Lin, and B. Schölkopf. A tutorial on ν-support vector machines: Research articles. *Applications of Stochastical Models in Business and Industry*, 21(2):111–136, 2005.

[39] Z. Chen and S. Molloi. Multiresolution vessel tracking in angiographic images using valley courses. *Optical Engineering*, 42:1673–1682, 2003.

[40] Z. H. Cho, J. P. Jones, and M. Singh. *Foundations of medical imaging.* J. Willey & Sons Inc., New York, 1993.

[41] E. Clardige, D. Hidovic-Rowe, F. O. Espina, and I. Styles. Quantifying composition of human tissue from multispectral images using a model of image formation. Technical report, October 6 2006.

[42] L. D. Cohen. On active contour models and balloons. *Computer Vision, Graphics and Image Processing: Image understanding*, 53(2):211–218, 1991.

[43] A. P. Condurache and T. Aach. Vessel segmentation in angiograms using hysteresis thresholding. In *Proceedings of MVA-2005*, pages 269–272, Tsukuba, Japan, May 16–18 2005.

[44] A. P. Condurache and T. Aach. Vessel segmentation in 2D-projection images using a supervised linear hysteresis classifier. In *Proceedings of ICPR-2006*, volume 1, pages 239–243, Hong Kong, China, August 20–24 2006. IEEE.

[45] A. P. Condurache, T. Aach, and E. Barth. Hysteresis classification with application to vessel segmentation in retina images. *submitted to the Pattern Recognition Letters*.

[46] A. P. Condurache, T. Aach, and K. Eck. Signaling of contrast agent injection and detection of complete vessel-tree frames in cardio-vascular X-ray imaging. *submitted to the IEEE Transactions on Medical Imaging*.

[47] A. P. Condurache, T. Aach, K. Eck, and J. Bredno. Automatic contrast-medium control in images. Technical report, European patent - App. nr. EP04100119.9, 2003.

[48] A. P. Condurache, T. Aach, K. Eck, and J. Bredno. Fast detection and processing of arbitrary contrast agent injections in coronary angiography and fluoroscopy. In *Proceedings of BVM-2004*, pages 5–9, Berlin, Germany, March 29–30 2004. Springer.

[49] A. P. Condurache, T. Aach, K. Eck, and J. Bredno. Automatic contrast-medium control in images. Technical report, US patent - App. nr. PCT/IB2005/050077, 2005.

[50] A. P. Condurache, T. Aach, K. Eck, J. Bredno, and T. Stehle. Fast and robust diaphragm detection and tracking in cardiac X-ray projection images. In *Medical Imaging 2005: Image Processing*, volume 5747, pages 1766–1775, San Diego, CA, February 12–17 2005. SPIE.

[51] A. P. Condurache, T. Aach, S. Grzybowski, and H.-G. Machens. Imaging and analysis of angiogenesis for skin transplantation by microangiography. In *Proceedings of ICIP-2005*, pages II/1250 – II/1253 (also on CD–ROM: ISBN 07803 9135–7), Genoa, Italy, September 11–14 2005. IEEE.

[52] A. P. Condurache, T. Aach, S. Grzybowski, and H.-G. Machens. Vessel segmentation and analysis in laboratory skin transplant micro-angiogram. In *Proceedings of the 18'th IEEE Symposium on CBMS-2005*, pages 21–26, Dublin, Ireland, June 23–24 2005. IEEE.

[53] A. P. Condurache, T. Aach, S. Grzybowski, and H.-G. Machens. Vessel segmentation for angiographic enhancement and analysis. In *Proceedings of BVM-2005*, pages 173–177, Heidelberg, Germany, March 13–15 2005. Springer.

[54] A. P. Condurache, T. Aach, A. Kaiser, and P. Radke. User-defined roi tracking for estimation of the myocardial blush grade. In *Proceedings of SSIAI-2006*, pages 66–70, Denver, USA, March 26-28 2006. IEEE.

[55] A. P. Condurache, E.A. Essah, A. Reske, M. Seiwerts, H. Busse, T. Aach, and U.G. Hofmann. Robust rib cage segmentation in CT image series using active contour models. In *Proceedings of BMT-2005*, pages 421–422, Nürenberg, Germany, September 14–16 2005. VDE.

[56] A. P. Condurache, T. Hahn, M. Scharfschwerdt, and T. Aach. Automatic measurements of quality parameters for tricuspid xenograft heart-valve implants. *submitted to the IEEE Transactions on Biomedical Engineering.*

[57] A. P. Condurache, T. Hahn, M. Scharfschwerdt, M. Misfeld, U. G. Hofmann, and T. Aach. Automatic measuring of quality criteria for heart valves. In *Medical Imaging 2007: Image Processing*, San Diego, CA, February 17–22 2007. SPIE.

[58] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. In *Proceedings of the 5'th ECCV-1998*, volume 2, pages 484–498, Freiburg, Germany, June 2–6 1998. Springer.

[59] T. F. Cootes and C. J. Taylor. Active shape models – smart snakes. In *Proceedings of BMVC-1992*, pages 266–270, Leeds, UK, September 21–24 1992. Springer.

[60] N. Cristianini and J. Shawe-Taylor. *An introduction to support vector machines.* Cambridge University Press, 2000.

[61] V. Croitoru, M. Kizik, and S. Stoica. *Comunicatii digitale.* Editura Presa Nationala, Bucuresti, 1997.

[62] D. Dascalu, A. Rusu, M. Profirescu, and I. Costea. *Dispozitive si circuite electronice.* Editura Didactica si Pedagogica, Bucuresti, 1987.

[63] I. Daubechies. *Ten lectures on wavelets.* SIAM, Philadelphia, 1992.

[64] T. E. David. Aortic valve replacement with stentless porcine bioprostheses.. *Journal of Cardiac Surgery*, 13(5):344–351, 1998.

[65] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society*, B(39):1–38, 1977.

[66] S. Dippel, M. Stahl, R. Wiemker, and T. Blaffert. Multiscale contrast enhancement for radiographies: Laplacian pyramid vs. fast wavelet transform. *IEEE Transations Medical Imaging*, 21(4):343–353, 2002.

[67] E. R. Dougherty. *An introduction to morphological image processing.* SPIE Optical Engineering Press, 1992.

[68] E. R. Dougherty. *Mathematical morphology in image processing.* Marcel Dekker, 1993.

[69] E. R. Dougherty. *Pattern classification.* J. Willey & Sons Inc., New York, 2000.

[70] E. R. Dougherty and C. R. Giardina. *Mathematical methods for artificial intelligence and autonomous systems.* Prentice-Hall, Inc., 1988.

[71] R. P. W. Duin. The combining classifier: to train or not to train? In *Proceedings of the ICPR-2002*, pages 765–770, Quebec, Canada, August 11–15 2002. IEEE.

[72] R. P. W. Duin, F. Roli, and D. de Ridder. A note on core research issues for statistical pattern recognition. *Pattern Recognition Letters*, 23(4):493–499, 2002.

[73] J. S. Duncan and F. Ayache. Medical image analysis: progress over two decades and the challanges ahead. *IEEE Transactions on Pattern analysis and Machine Intelligence*, 22(1):85–106, 2000.

[74] K. Eck, I. Wächter, and J. Bredno. Synthesis of angiographic images using iterative approximation. In *Medical Imaging 2004: Image Processing*, volume 5370, pages 163–171, San Diego, CA,USA, February 14–19 2004. SPIE.

[75] M. Sonka J. M. Fitzpatrick eds. *Handbook of Medical Imaging.* SPIE Press, 2000.

[76] J. Egan. *Signal detection theory and ROC analysis.* Academic Press, New York, 1975.

[77] K. H. Englmeier, S. Bichler, K. Schmid, M. Maurino, M. Porta, T. Bek, , B. Ege, O. V. Larsen, and O. K. Hejlesen. Multi-resolution retinal vessel tracker based on directional smoothing. In *Medical Imaging 2002: Physiolgy and Function from Multidimensional Images*, volume 4683, pages 230–237, San Diego, USA, February 23–28 2002. SPIE.

[78] A. Erasmi, H.-H. Sievers, M. Scharfschwerdt, T. Eckel, and M. Misfeld. In vitro hydro-dynamics, cusp-bending deformation, and root distensibility for different types of aortic valve-sparing operations: remodeling, sinus prosthesis, and reimplantation. *Thoracic and Cardiovascular Surgery*, 130(4):1044–1049, 2005.

[79] M. A. T. Figueiredo and A. K. Jain. Unsupervised learning of finite mixture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(3):381–396, 2002.

[80] G. Forman. An extensive empirical study of feature selection metrics for text classification. *Journal of Machine Learning Research*, 3:1289–1305, 2003.

[81] D. A. Forsyth and J. Ponce. *Computer vision: a modern approach.* Prentice Hall Professional Technical Reference, 2002.

[82] A. F. Frangi, A. A. Amini, and E. Bullitt. Vascular imaging - editorial. *IEEE Transcations on Medical Imaging - Special Issue on Vascular Imaging*, 24(4):433–434, 2005.

[83] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever. Multiscale vessel enhancement filtering. *Lecture Notes in Computer Science*, 1496:130–137, 1998.

[84] M. Franz and R. Schüffny. Segmentation of blood vessels in subtraction angiographic images. In *Proceedings of DICTA-2003*, pages 215–224, Sydney, Australia, December 10–12 2003. CSIRO.

[85] W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(9):891–906, 1991.

[86] Y. Freund and R. Schapire. A short introduction to boosting. *Journal of Japanese Society for Artificial Intelligence*, 14(5):771–780, 1999.

[87] K. Fukunaga. *Introduction to statistical pattern recognition.* Academic Press, 1990.

[88] S. Geman and D. Geman. Stochastic relaxation, Gibbs distribtions, and the Bayesian restauration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721–741, 1984.

[89] J. Gil and R. Kimmel. Efficient dilation, erosion, opening, and closing algorithms. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 24(12):1606–1617, 2002.

[90] R. C. Gonzales and R. E. Woods. *Digital image processing.* Addison-Wesley, 1993.

[91] H. Gray. *Gray's anatomy.* Running Press, 1901.

[92] W. E. L. Grimson and D. P. Huttenlocher. On the sensitivity of the Hough transform for object recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(3):255–274, 1990.

[93] S. Grzybowski, B. Bucsky, B. Stöckelhuber, T. Aach, A. P. Condurache, W. Lindenmaier, P. Mailänder, and H.-G. Machens. A microangiography technique to quantify fasciocutaneous blood vessels in small laboratory animals. *Langenbeck's Archive of Surgery*, 389(10), 2004.

[94] S. Grzybowski, B. Bucsky, B. Stöckelhuber, T. Aach, A. P. Condurache, W. Lindenmaier, P. Mailänder, and H.-G. Machens. Model for induction of angiogenesis by synergetic effects of BFGF and VEGF[165] by bioactive dermal matrices. *Langenbeck's Archive of Surgery*, 389(10), 2004.

[95] T. Hahn, A. P. Condurache, T. Aach, M. Scharfschwerdt, and M. Misfeld. Automatic in-vitro orifice area determination and fluttering analysis for tricuspid heart-valves. In *Proceedings of BVM-2006*, pages 21–25, Hamburg, Germany, May 19–21 2006. Springer.

[96] R. M. Haralick, S. R. Sternberg, and X. Zhuang. Image analysis using mathematical morphology. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(4):532–550, 1987.

[97] S. Haykin. *Neural networks a comprehensive fundation.* Prentice Hall, 1999.

[98] J. P. S. Henriques, F. Zijlstra, A. W. J. van't Hof, M.-J. de Boer, J.-H. E. Dambrink, M. Gosselink, J. C. A. Hoorntje, and H. Suryapranata. Angiographic assessment of reperfusion in acute myocardial infarction by myocardial blush grade. *Circulation*, 107(16):2115–2119, 2003.

[99] R. B. Hill. Apparatus and method for identifying individuals trough their retinal vascular patterns. Technical report, US patent - nr. 4109237, 1978.

[100] A. Hoover, V. Kouznetsova, and M. H. Goldbaum. Locating blood vessels in retinal images by piece-wise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, 19(3):203–210, 2000.

[101] A. Hoover, V. Kouznetzova, and et al. Locating blood vessels in retinal images by pice-wise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, 19(3):203–210, 2000.

[102] P. V. C. Hough. Methods and means for recognizing complex patterns. In *US patent nr. 3069654*, 1962.

[103] J. Hsieh. *Computed tomography.* SPIE-The International Society for Optical Engineering, 2003.

[104] Y. Z. Hsu, H.-H. Nagel, and G. Rekers. New likelihood test methods for change detection in image sequences. *Computer Vision, Graphics, and Image Processing*, 26(1):73–106, 1984.

[105] S. V. Huffel and P. Lemmerling. *Total least squares and errors-in-variables modeling: analysis, algorithms and applications.* Dordrecht, The Netherlands: Kluwer Academic Publishers, 2002.

[106] J. Illingworth and J. Kittler. A survey of the Hough transform. *Computer Vision Graphics and Image Processing*, 44(1):87–116, 1988.

[107] M. Jacob and M. Unser. Design of steerable filters for feature detection using Canny-like criteria. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(8):1007–1019, 2004.

[108] B. Jähne. *Digital image processing.* Springer, 1997.

[109] B. Jähne. *Image processing for scientific applications.* CRC Press RLC, 1997.

[110] A. K. Jain. *Fundamentals of digital image processing.* Prentice-Hall Inc., 1989.

[111] A. K. Jain and P. W. Duin. Statistical pattern recognition: a review. *IEEE Transactions on Pattern Anaysis and Machine Intelligence*, 22(1):4–37, 2000.

[112] A. K. Jain, M. N. Murty, and P. J. Flynn. Data clustering: a review. *ACM Computing Surveys*, 31(3):264–323, 1999.

[113] A. K. Jain, A. Ross, and S. Prabhakar. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):4–20, 2004.

[114] A. K. Jain, Y. Zhong, and M.-P. Dubuisson-Jolly. Deformable template models: a review. *Signal Processing*, 71(2):109–129, 1998.

[115] R. Jain, R. Kasturi, and B. G. Schunk. *Machine vision.* McGraw-Hill, Inc., 1995.

[116] P. Jannin, J.M. Fitzpatrick, D.J. Hawkes, X. Pennec, R. Shahidi, and M.W. Vannier. Validation of medical image processing in image-guided therapy. *IEEE Transactions on Medical Imaging*, 21(12):1445–1450, 2002.

[117] X. Jiang and D. Mojon. Adaptive local thresholding by verification based multithreshold probing with application to vessel detection in retinal images. *IEEE Transactions on Pattern Anaysis and Machine Intelligence*, 25(1):131–137, 2003.

[118] M. Kass and A. Witkin. Analyzing oriented patterns. *Computer Vision, Graphics and Image Processing*, 37(3):362–385, 1987.

[119] M. Kass, A. Witkin, and D. Trezopoulos. Snakes: active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988.

[120] S. B. King. Angioplasty from bench to bedside to bench. *Circulation*, 93(9):1621–1629, 1996.

[121] C. Kirbas and F. K. H. Quek. A review of vessel extraction techniques and algorithms. *ACM Computing Surveys*, 36(2):81–121, 2004.

[122] A. K. Klein, F. Lee, and A. A. Amini. Quantitative coronary angiography with deformable spline models. *IEEE Transactions on Medical Imaging*, 16(5):468–482, 1997.

[123] T. Kohonen. The self-organizing map. *Proceedings of the IEEE*, 78(9):1464–1480, 1990.

[124] A. Lanterman. Schwarz, Wallace, and Rissanen: Intertwining themes in theories of model order estimation. *International Statistical Review*, 69:185–212, 2001.

[125] V. F. Leavers. The dynamic generalized Hough transform. In *Proceedings of ECCV-1990*, pages 592–594, Antibes, France, April 23–27 1990. Springer.

[126] V. F. Leavers. The dynamic generalized Hough transform: its relationship to the probabilistic Hough transforms and an application to the concurrent detection of circles and ellipses. *CVGIP: Image Understanding*, 56(3):381–398, 1992.

[127] M. D. Levine and A. M. Nazif. An experimental rule based system for testing low level segmentation strategies. *Multicomputers and Image Processing: Algorithms and Programs*, pages 149–160, 1982.

[128] M. D. Levine and A. M. Nazif. Dynamic measurement of computer-generated image segmentations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 7(2):155–164, 1985.

[129] Q. Li, J. You, L. Zhang, D. Zhang, and P. Bhattacharya. A new approach to automated retinal vessel segmentation using multiscale analysis. In *Proceedings of ICPR-2006*, volume 4, pages 884–888, Hong Kong, China, August 20–24 2006. IEEE.

[130] D. Litwiller. CCD vs. CMOS: facts and fiction. *Photonics Spectra*, 35(1):154–158, 2001.

[131] G. Liu and R.M. Haralick. Two practical issues in Canny's edge detector implementation. In *Proceedings of ICPR-2000*, pages 680–682, Barcelona, Spain, September 3–8 2000. IEEE.

[132] H.-G. Machens, S. Grzybowski, B. Bucsky, T. Spanholtz, C Niedworok, A. Maichle, B. Stockelhuber, A. P. Condurache, F. Liu, J. T. Egana, M. Kaun, P. Mailander, and T. Aach. A technique to detect and to quantify fasciocutaneous blood vessels in small laboratory animals ex vivo. *Journal of Surgical Research*, 131(1):91–96, 2006.

[133] R. Maclin and D. Opitz. An empirical evaluation of bagging and boosting. In *Proceedings of NCAI-1997*, pages 546–551, Menlo Park, CA, USA, July 27–31 1997. AAAI Press.

[134] W. Maier, A. Zeiher, M. Rotter, M. Togni, S. Windecker, and B. Meier. The european registry of catheter interventions 1998. *Journal für Kardiologie*, 11(4):145–148, 2004.

[135] R. Malladi, J. A. Sethian, and B. C. Vemuri. Evolutionary fronts for topology-independent shape modeling and recovery. In *Proceedings of ECCV-1994*, volume 1, pages 3–13, Stockholm, Sweden, 1994.

[136] R. Malladi, J. A. Sethian, and B. C. Vemuri. Shape modelling with front propagation: a level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2):158–175, 1995.

[137] S. G. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674–693, 1989.

[138] S. G. Mallat. *A wavelet tour of signal processing.* Academic Press, San Diego, 1998.

[139] D. Marr and E. Hildreth. Theory of edge detection. In *Proceedings of the Royal Society London B*, volume 207, pages 187–217, 1980.

[140] B. Martin-Leung, K. Eck, I. Stuke, J. Bredno, and T. Aach. Mutual information based respiration detection. In *Proceedings of CARS-2003*, pages 1085–1092, London, UK, June 25–28 2003.

[141] G. Matheron. *Random sets and integral geometry.* J.Wiley & Sons Inc., New York, 1975.

[142] T. McInerney and D. Terzopoulos. Deformable models in medical images analysis: a survey. *Medical Image Analysis*, 1(2):91–108, 1996.

[143] B. Meffert and O. Hochmut. *Werkzeuge der Signaverarbeitung.* Pearson Studium, 2004.

[144] V. Metzler. *Morphologishe Dekomposition als universelle Methodik der medizinischen Bildverarbeitung: Grundalgen und praktische Umsetzung.* VDI Verlag - Düsseldorf, 2002.

[145] M. Misfeld. Die Aortenklappenchirurgie – gestern, heute, morgen. *Focus MUL*, 23(2):82–89, 2006.

[146] R. Mueller and T. Sanborn. The history of interventional cardiology: Cardiac catheterization, angioplasty, and related interventions. *American Heart Journal*, 129(1):146–172, 1995.

[147] Z. Nagy, J. Fisher, P. G. Walker, and K. G. Watterson. The effect of sizing on the in vitro hydrodynamic characteristics and leaflet motion on the Toronto SPV stentelss valve. *Thoracic and Cardiovascular Surgery*, 117(1):92–98, 1999.

[148] Z. Nagy, J. Fisher, P. G. Walker, and K. G. Watterson. The effect of sizing on the hydrodynamic parameters of the medtronic freestyle valve in vitro. *Annals of Thoracic Surgery*, 69(1):1408–1413, 2000.

[149] V. Neagoe and O. Stanasila. *Recunoasterea formelor si retele neurale.* Matrixrom, 1999.

[150] W. Neuenschwander, P. Fua, G. Szekely, and O. Kuebler. Initializing snakes. In *Proceedings of CVPR-1994*, pages 658–663, Seattle, USA, June 21–23 1994. IEEE.

[151] A. Niemistö, V. Dunmire, O Yli-Harja, W. Zahng, and I. Shmulevich. Robust quantification of in vitro angiogenesis through image analysis. *IEEE Transactions on Medical Imaging*, 24(4):549–553, 2005.

[152] J. J. Oliver, Rohan A. Baxter, and C. S. Wallace. Unsupervised learning using MML. In *Proceedings of ICML-1996*, pages 364–372, Bari, Italy, June 3–6 1996. Morgan Kaufmann Publishers.

[153] A. V. Oppenheim, R. W. Schafer, and T. G. Stockham. Nonlinear filtering of multiplied and convolved signals. *Proceedings of the IEEE*, 56(8):1264–1291, 1968.

[154] A. W. Oppenheim and R. W. Schafer. *Discrete-time signal processing.* Prentice-Hall International, Englewood Cliffs, 1989.

[155] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-9(1):62–66, 1979.

[156] A. Papoulis. *Probability & statistics.* Prentice-Hall International, Englewood Cliffs, 1990.

[157] N. Paragios, O. Mellina-Gottardo, and V. Ramesh. Gradient vector flow fast geometric active contours. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 26(3):402–407, 2004.

[158] A. Pentland and S. Sclaroff. Closed-form solutions for physically base shape modelling and recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):715–729, 1991.

[159] R. Poli and G. Valli. An algorithm for real time vessel enhancement and detection. *Computer Methods and Programs in Biomedicine*, 52(1):1–22, 1997.

[160] P. Pudil, J. Novovicova, and J. Killtler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15(11):1119–1125, 1994.

[161] M. Rabbani, R. Shaw, and R. Van Metter. Detective quantum efficiency of imaging systems with amplifying and scattering mechanisms. *Journal of the Optical Society of America A*, 4(5):895–902, 1987.

[162] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. *IEEE Proceedings*, 77(2):257–286, 1989.

[163] J. Rissanen. *Stochastic complexity in statistical inquiry theory.* World Scientific Publishing Co., Inc., River Edge, NJ, USA, 1989.

[164] E. M. F. Riverón, I. Castellanos-Bisset, and L. Sánchez-Cuello. Segmentation of tiny objects in very poor-quality angiogenesis images. *Pattern Recognition Letters*, 26(16):2579–2587, 2005.

[165] K. Rohr, M. Fornefett, and H. S. Stiehl. Approximating thin-plate splines for elastic registration: integration of lensmarks errors and orientation attributes. *Lecture Notes in Computer Science*, 1613:252–265, 1999.

[166] P. K. Sahoo, S. Soltani, A. K. C. Wong, and Y. C. Chen. A survey of thresholding techniques. *Computer Vision, Graphics, and Image Processing*, 41(2):233–260, 1988.

[167] M. Scharfschwerdt, M. Misfeld, and H.-H. Sievers. The influence of a nonlinear resistance element upon in-vitro aortic pressure tracings and aortic valve motions. *American Society for Artificial Internal Organs*, 50(5):498–502, 2004.

[168] M. Schrijver. *Angiographic image analysis to assess the severity of coronary stenoses,*. Twente University Press, 2002.

[169] F. W. Sears, M. W. Zemansky, and H. D. Young. *College Physics*. Person Higher Education, 1990.

[170] D. Selle, B. Preim, A. Schenk, and H.-O Peitgen. Analysis of vasculature for liver surgical planning. *IEEE Transactions on Medical Imaging*, 21(11):1344–1357, 2002.

[171] J. A. Sethian. *Level set methods and fast marching methods*. Cambridge University Press, 1999.

[172] J. A. Sethian and J. Strain. Crystal growth and dendritic solidification. *Journal of Computational Physics*, 98(2):231–253, 1992.

[173] M. Sezgin and B. Sankur. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1):146–165, 2004.

[174] G. Shechter, F. Devernay, E. Coste-Maniere, A. Quyyumi, and E. McVeigh. Three-dimensional motion tracking of coronary arteries in biplane cineangiograms. *IEEE Transactions on Medical Imaging*, 22(4):493–503, 2003.

[175] G. Shechter, C. Ozturk, J. Resar, and E. McVeigh. Respiratory motion of the heart from free breathing coronary angiograms. *IEEE Transactions on Medical Imaging*, 23(8):1046–1056, 2004.

[176] G. Shechter, B. Shechter, J. R. Resar, and R. Beyar. Prospective motion correction of X-ray images for coronary interventions. *IEEE Transactions on Medical Imaging*, 24(4):441–450, 2005.

[177] P. M. Shikhaliev, T. Xu, and S. Molloi. Evaluation of a photon-counting X-ray imaging detector based on microchannel plates for mammography applications. In *Medical Imaging: Physics of Medical Imaging*, volume 5368, pages 726–733, San Diego, CA, USA, February 14–19 2004. SPIE.

[178] A. J. Smola and B. Schölkopf. A tutorial on support vector regression. *Statistics and Computing*, 14(3):199–222, 2004.

[179] P. Soille. *Morphologische Bildverarbeitung*. Springer Verlag, 1998.

[180] A. Spataru. *Teoria transmisiunii informatiei*. Editura didactica si pedagogica, Bucuresti, 1983.

[181] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken. Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509, 2004.

[182] M. Stäger, P. Lukowicz, and G. Tröster. Dealing with class skew in context recognition. In *Proceedings of IWSAWC-2006*, Lisabon, Portugal, July 4–5 2006.

[183] Y. Sun, R. J. Lucariello, and S. A. Chiaramida. Directional low-pass filtering for improved accuracy and reproducibility of stenosis quantification in coronary angiograms. *IEEE Transactions on Medical Imaging*, 14(1):242–248, 1995.

[184] C. W. Therrien. *Decision, estimation and classification.* J.Wiley & Sons Inc., New York, 1989.

[185] K. D. Tönnies. *Grundlagen der Bildverarbeitung*. Pearson Studium, 2005.

[186] V. Torre and T. Poggio. On edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(2):147–163, 1986.

[187] D. Toth, A. P. Condurache, and T. Aach. A two-stage-classifier for defect classification in optical media inspection. In *Proceedings ICPR-2002*, volume 4, pages 373–376, Quebec, Canada, August 11–15 2002. IEEE.

[188] E. Trucco and A. Verri. *Introductory techniques for 3D computer vision.* Prentice Hall, 1998.

[189] P. A. van den Elsen, J. B. A. Maintz, E. J. D. Pol, and M. A. Viergever. Medical image matching – a review with classification. *IEEE Engineering in Medicine and Biology*, 12(1):26–39, 1993.

[190] A. W. J. van't Hof, A. Liem, H. Suryapranata, J. C. A. Hoorntjem, M.-J. de Boer, and F. Zijlstra. Angiographic assessment of myocardial reperfusion in patients treated with primary angioplasty for acute myocardial infarction: myocardial blush grade. *Circulation*, 97(23):2302–2306, 1998.

[191] T. Vesely, D. Boughner, and T. Song. Tissue buckling as a mechanism of bioprosthetic valve failure. *Annals of Thoracic Surgery*, 46(3):302–308, 1988.

[192] W. Vongpatanasin, L. D. Hillis, and R. A. Lange. Prosthetic heart valves. *Medical Progress*, 355(6):407–416, 1996.

[193] C. S. Wallace and D. L. Dowe. Minimum message length and Kolmogorov complexity. *The Computer Journal*, 42(4):270–283, 1999.

[194] B. A. Wandell. *Foundations of vision.* Sinauer Associates Inc., 1995.

[195] A. Weissler, W. Harris, and C. Schönfeld. Systolic time intervals in heart failure in man. *Circulation*, 37:149–159, 1968.

[196] C. Xu and J. L. Prince. Gradient vector flow: a new external force for snakes. In *Proceedings of CVPR-1997*, pages 66–71, San Juan, June 17–19 1997. IEEE.

[197] C. Xu and J. L. Prince. Generalized gradient vector flow external forces for active contours. *Signal Processing*, 71(2):131–139, 1998.

[198] C. Xu and J. L. Prince. Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, 7(3):359–369, 1998.

[199] Y. Yakimovski. Boundary and object detection in real world images. *Journal of the Association for Computing Machinery*, 23(4):599–618, 1976.

[200] W. A. Yasnoff, J. K. Mui, and J. W. Bacus. Error measures for scene segmentation. *Pattern Recognition*, 9(4):217–231, 1977.

[201] H. K. Yuen, J. Illingworth, and J. Kittler. Detecting partially occluded ellipses using the Hough transform. *Image and Vision Computing*, 7(1):31–37.

[202] Y. J. Zhang. A survey on evaluation methods for image segmentation. *Pattern Recognition*, 29(8):1335–1346, 1995.

[203] Y. J. Zhang and J. J. Gerbrands. Objective and quantitative segmentation evaluation and comparison. *Signal Processing*, 39(1-2):43–54, 1994.

Abbreviations and acronyms

AGC	-	Automatic Gain Control
APS	-	Active Pixel Sensor
AROC	-	Area under the Receiver Operating Characteristic
BDI	-	Bending Deformation Index
CAD	-	Coronary Artery Disease
CC	-	Correct Classifications
CV	-	Coefficient of Variation
CCD	-	Charge Coupled Device
CMOS	-	Complementary Metal Oxide Semiconductor
DoG	-	Difference of Gaussians
DSA	-	Digital Subtraction Angiography
ECG	-	ElectroCardioGram
EM	-	Expectation Maximization
FEM	-	Finite Element Method
FMM	-	Finite Mixture Models
FN	-	False Negatives
FP	-	False Positives
fps	-	Frames Per Second
FPD	-	Flat Panel Detector
GUI	-	Graphical User Interface
GVF	-	Gradient Vector Flow
LoG	-	Laplacian of Gaussian
MAP	-	Maximum A Posteriori
MB	-	Myocardial Blush
MBG	-	Myocardial Blush Grade
MI	-	Myocardial Infarction
ML	-	Maximum Likelihood
MRE	-	Mean Relative Error
MRF	-	Markov Random Field
MSE	-	Mean Square Error
OTF	-	Optical Transfer Function
pAROC	-	partial Area under the Receiver Operating Characteristic
PCTA	-	Percutaneous Transluminal Coronary Angioplasty
pdf	-	Probability Density Function
PSF	-	Point Spread Function
ROC	-	Receiver Operating Characteristic
SVM	-	Support Vector Machine
TFT	-	Thin Film Transistor array
TN	-	True Negatives
TP	-	True Positives

CURRICULUM VITAE

PERSONAL DETAILS

Name	Alexandru Paul Condurache
Date of birth	4. September 1976
Born in	Bucharest, Romania

EXPERIENCE

10/2001 – today	University of Luebeck, Institute for Signal Processing

EDUCATION

09/2000 – 09/2001	Postgraduate study in Biomedical Engineering (equivalent of a french "Diplome d'Etude Approfondie" (DEA)) at the "Politehnica" University in Bucharest
09/1995 – 09/2000	Study of Electronics at the "Politehnica" University in Bucharest
09/2000	Engineer degree in electronics

SECONDARY SCHOOL

09/1991 – 07/1995	Lyceum "Alexandru Ioan Cuza", Bucharest