TECHNISCHE UNIVERSITÄT MÜNCHEN

Fakultät für Informatik
Computer Aided Medical Procedures & Augmented Reality / I16

Statistical Image Processing of Medical Ultrasound Radio Frequency Data

Tassilo Johannes Klein

Vollständiger Abdruck der von der Fakultät für Informatik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. M. Bader
Prüfer der Dissertation:
1. Univ.-Prof. Dr. N. Navab
2. Prof. A. Noble, DPhil, University of Oxford, UK

Die Dissertation wurde am 29.05.2012 bei der Technischen Universität München eingereicht und durch die Fakultät für Informatik am 21.11.2012 angenommen.

Abstract

Ultrasonography is often the modality of choice in terms of e.g. diagnostic and intra-operative applications due to reasons such as safety, mobility and inexpensiveness compared to other imaging techniques. This together with the steady increase in image quality has made it a widely used diagnostic method in various medical disciplines in recent years. However, the interpretation of ultrasound imagery is typically of quite subjective nature and further data processing not straightforward. This largely stems from the physical mechanisms at the heart of ultrasound imaging. The underlying process yields imagery that is above all view-dependent and heavily subject to noise as well as prone to containing various types of artifacts. Conventional ultrasound images, commonly referred to as B-Mode, are the result of many processing steps optimizing data for visual assessment by physicians. However, at the core of ultrasound imaging pipeline lies the radio frequency (RF) data. Just lately, RF data has become more readily available to the research community such that its potential has not fully unveiled yet. From a data processing standpoint using RF data over B-Mode suggests many advantages. First of all, it is generally much richer in information due to the comparably higher resolution. Furthermore, it is not affected by non-linear post-processing steps such as log-compression and proprietary filter algorithms that change the noise statistics for reasons of improved visual appeal. In addition, it has nice probabilistic properties facilitating various ways of distributional modeling of ultrasound specific texture patterns, referred to as speckle noise.
In this thesis various methods are investigated of how to take advantage of this rich pool of information. Thereby special focus is devoted to the applications of correct spatial alignment of ultrasound data, referred to as registration. At first methods are presented that reside in 2D domain addressing issues of textural and probabilistic description with possible applications in tissue classification, elastography, speckle tracking and motion recovery. Later this is expanded to 3D ultrasound freehand data and its reconstruction with a particular view on transcranial scans that are recorded in order to perform an early diagnosis of Parkinson's disease.

Zusammenfassung

Ultraschall ist häufig die Modalität der Wahl für diagnostische und intraoperative Anwendungen. Dies kann zurückgeführt werden auf seine vielen positiven Eigenschaften wozu Faktoren zählen wie Sicherheit, Mobilität und niedriger Anschaffungs- und Unterhaltspreis. Dies hat in Verbindung mit der stetig verbesserten Bildqualität dazu geführt, dass Ultraschall mittlerweile ein breites Anwendungsgebiet aufweist. Allerdings ist die Ultraschallbildgebung mit dem großen Nachteil verbunden, dass die Interpretation der erzeugten Bilddaten sehr subjektiv und die Weiterverarbeitung der Daten im Allgemeinen nicht einfach ist. Dies ist weitestgehend auf die physikalischen Mechanismen zurückzuführen, auf die der medizinische Ultraschall basiert. So zeichnen sich Ultraschallbilder vor allem durch die Eigenschaft der Blickwinkelabhängigkeit sowie der starken Rauschanfälligkeit aus. Darüber hinaus beinhalten die Bilder häufig eine Vielzahl an Artefakten.
Gewöhnliche Ultraschallbilder, auch als B-Mode Bilder bekannt, sind dabei das Produkt einer Vielzahl von Verarbeitungsschritten in der Ultraschallmaschine, die darauf abzielen die Signalmessdaten für das menschliche Auge ansprechend und anschaulich zu machen. Diesbezüglich stellen die sog. Radiofrequenz (RF) Daten die weitestgehend unbearbeitete Signalgrundlage in der Processing-Pipeline von Ultraschallmaschinen dar. Erst in jüngster Zeit sind mehr und mehr Hersteller von Ultraschallmaschinen bereit der Wissenschaft Zugang zu diesen Daten zu gewähren, so dass das ganze Potential der Daten noch alles andere als ausgeschöpft und erforscht ist. Vom Datenverarbeitungsaspekt aus bieten RF Daten gegenüber den gewöhnlichen B-mode Daten eine große Reihe von Vorteilen. So ist primär der Informationsgehalt in RF Daten viel größer als bei B-mode Bildern, was auf die höhere Auflösung zurückgeführt werden kann. Zum anderen sind die Daten noch keinen stärkeren Verarbeitungenschritten unterworfen, wie z.B. nicht lineare Filteroperationen, log-Kompression sowie ähnliche i.d.R. proprietäre Datenmodifizierungsschritte, die das Bild für den Betrachter anschaulich machen. Diese Tatsache ist einer der entscheidenden Vorteile von RF Daten gegenüber B-Mode, bei deren Erzeugung oben genannte Operationen zum Einsatz kommen und so die statistischen Eigenschaften der Daten massiv beeinflusst werden.
Im Rahmen dieser Dissertation wurden mehrere Mittel und Möglichkeiten eruiert, wie RF Daten genutzt werden können. Fokus dieser Arbeit liegt dabei in der Entwicklung von Verfahren zur verbesserten räumlich korrekten Überlagerung von Bilddaten, auch als Registrierung bekannt. Zu Beginn werden Methoden zur statistischen Analyse und Beschreibungen von Texturen im zwei dimensionalen Raum beschrieben mit potentiellen Anwendungsgebieten in Gewebeklassifizierung, Elastographie, Speckle-Tracking und Motion-Recovery. Im Anschluss daran wird sich konkreten Problemstellungen von 3D Freehand Ultraschalldaten gewidmet, wie Rekonstruktion und Registrierung, die vorwiegend im Rahmen zur Frühdiagnose von Parkinson erstellt werden.

Acknowledgments

First of all I would like to thank Prof. Nassir Navab for providing me the opportunity to work at his chair for computer-aided medical procedures - خیلی متشکرم. It was an exciting and extremely enriching experience to me. In particular, I really enjoyed the comfortable work and research atmosphere with the special international flavor. Moreover, I am very grateful that I had right away from the beginning at the chair, many opportunities to participate in international conferences, which was encouraged by Prof. Navab. In this respect, I would also like to thank all my friends and colleagues at CAMP chair. Christian Wachinger, the founding father of the 'tea-group',who was always helpful and inspiring. Ahmad Ahmadi, the research dadash, with whom I had the pleasure to go through thick and thin with in the context of brave EU projects, spending a myriad nights in miscellaneous labs. Loren Schwarz, who as a member of the 'tea-group' always contributed to great discussions. Athanasios Karamalsis (BigK), with whom I had the pleasure to share office for many years as well as have a glass of pristine 'Ouzo 12'. A particular great thanks goes an international chair's visitor, Mattias Hansson, with whom I had uncountably many fruitful discussions and who often provided mathematical assistance. Furthermore, a very big thank you goes to our medical collaborators at the neurology department of Klinikum Grosshadern, namely Prof. Dr. Kai Bötzel and Dr. Annika Plate. They were always extremely supportive and committed to our joint research. During our work they helped tremendously and with great patience while providing an insight into their very interesting application domain. Moreover, without them data acquisition for our patient studies would not have been possible. I am also very thankful for the delightful discussions and conversations with Patrick, Philipp, Diana, Tobias, Pascal, Oli, Olivier, Debdoot, Loic, Mehmet, Lejing and Jose.
Finally, I would like to thank for the great support of my beloved parents, family and friends, as all this work would not have been possible without them.

Contents

Contents **ix**

'Certain books seem to be written, not that we might learn from them, but in order that we might see how much the author knows.'

[Johann Wolfgang von Goethe, 1749–1832]

Chapter 1

Introduction

Until the beginning of the twentieth century the ability to investigate structure and functions within an intact human body were largely far beyond imagination. Back then the discovery of the X-ray by the German physicist Wilhelm Röntgen was the first great milestone that made the unimaginable possible and simultaneously was the igniting spark for a new research domain to develop - medical imaging. Now in the twenty-first century, technological inventions and developments have created many new possibilities and similar breakthroughs in the medical domain with an ever increasing number of modalities available. As a result, medical diagnosis and treatment is nowadays unthinkable without the support of medical imagery. In this respect, radiology has long been dominated by the two established imaging technologies, namely magnetic resonance imaging (MRI) and Computed tomography (CT). Simultaneously, ultrasound has for a very long time been the red-haired stepchild, with the result that for years it had established itself probably as the most poorly used imaging technology, in spite of its wide potential applicability.
However, there has been a process of change noticeable in medical treatment due to factors such as the ever increasing greater importance that is attached to cost effectiveness and the strong trend towards minimally invasive surgery, where ultrasound technology has the potential to come into play and promises extensive applicability. This has led to developments such that these days ultrasonography is often the modality of choice in terms of e.g. intra-operative and screening applications, which can be further attributed to its safety, mobility and inexpensiveness compared to other imaging techniques. This together with the steady increase in image quality promises ultrasound in future to become an even more widely used diagnostic method in various medical disciplines.
Lately, there has been drive to make ultrasound equipment more and more compact. This has led to down shrinking of bulky-cart based systems to the size of portable and hand-held devices, such that they are easily accessible, which is of prime importance e.g. in emergencies when seconds count. Given that those devices are still far from widespread use, they are nowadays often referred to as the stethoscopes of the future. However, in contrast to their historical counterpart they do not only allow to listen, but rather to see and thereby expand the spectrum of medical care by a new dimension. This devel-

opment might actually conjure up memories of the famous television series Star Trek, in which the first officer Mr. Spock and the chief medical officer McCoy used a hand-held device called a 'tricorder' to scan people and amongst other things assess their health status - a technology which was back then far beyond imagination.
Further current development trends in the field of ultrasound are the 2D arrays, which allow for the real-time acquisition of 3D images. In such 2D arrays, several thousand piezo elements are typically arranged in a grid. Controlled triggering of the elements then allows focusing and steering of ultrasound beams in both elevation and azimuth, acquiring pyramid-shaped volumes. Although such 2D arrays still do not provide the same image resolution as conventional ultrasound system and are still extremely expensive, this clearly indicates the future path with exciting applications. The bottom line is that ultrasound imaging is a rapidly changing field, with ongoing research of new technologies in various directions. This development will most likely lead to a multitude of new applications for diagnostic purposes in the years to come.
Because of the complex physical processes underlying ultrasound imaging, it is necessary to create methods that take all characteristics of the modality into account. Consequently, this requires a precise understanding of the data acquisition process, ranging from the physical signal properties to the data processing pipeline that generates the output data. As a result, algorithms can be devised that take into account the specifics of the system. Therefore in this dissertation, we consider ultrasound imaging in more detail - especially the use of the raw echo signal data, referred to as radio frequency (RF) data. This is of particular interest, as just lately RF data has become more readily available by ultrasound machine vendors. RF data as opposed to traditional ultrasound images, known as B-mode, comes with the benefit of being unaffected by non-linear post-processing steps. At first glance this might seem not very spectacular. However, the data processing pipeline in conventional machines is made up of a series of steps such as log-compression and application-dependent proprietary filter as well as algorithms that change speckle statistics for reasons of improved visual appeal. But there is a lot of information content relevant for diagnostics that is beyond the easily visualizable content that make up current 2D or 3D images. Hence, employing RF data allows for better and more predictable statistical data analysis and processing, making full use of the available information content in the signal. In spite of its clear advantages, RF data processing has not been paid much attention for a long time. Beside the limited availability, this can also be attributed to its large volume, which is a result of the comparably much higher resolution as B-mode that made the required massive data processing seem intractable for a long time. Given the rise in computing power in desktop machines and the developments in the domain of graphic processing units (GPU) that allow for massive parallelization, there is no longer the bottleneck of data processing. Thus there exists no longer any obvious reason not to exploit the extra information in RF data - in the otherwise information-poor ultrasound images.
In this respect, focus of this thesis is particular on statistical methods of ultrasound data processing of RF data. These methods are able to facilitate the

description of both the image and disturbances underlying in the recording process. This is in particular important for ultrasound imagery with its noise-like image acquisition artifacts. Special significance is thereby attributed to multi-view transcranial ultrasound imaging. That is ultrasound data acquisition using a low-frequency transducer, where images are obtained through the temporal lobe of the skull. These relatively new type of transcranial examinations are performed for amongst others, the early diagnosis of Parkinson's as well as other severe and progressive neurodegenerative diseases. This is of prime interest in medical research against the background that neurodegenerative diseases become more and more prominent within our aging societies and successful treatment depends on early detection.

1.1 Main Topics

The main topics of this thesis are registration and 3D reconstruction of ultrasound imagery as will be briefly outlined in the following. Thereby we will explicitly elaborate on the associated challenges.

Registration

The process of obtaining the correct spatial alignment between two or more datasets is referred to as registration. There are lots of applications that require the registration of ultrasound images. First of all, the limited view of ultrasound images can be resolved by combining multiple images with overlapping image content. Furthermore, there is application for regular staging of structures in tissue, whereby it is necessary to have perfectly registered images to track changes. Similarly, in the domain of elastography and motion recovery the identification of corresponding structures is required. However, in ultrasound, there are several challenges that complicate image registration. On the one hand, the imaging process is view dependent and subject to non-linear processes. On the other hand, ultrasound images are prone to contain a wide range of artifacts, such that imaging of the same structure from differing views can lead to totally different image content. In particular, strong ultrasound reflectors such as bone, ligaments or muscular tissue result in shadow artifacts that reinforce such effects. Moreover, if insonofied orthogonally, those areas are then further prone to contain mirroring and reverberation artifacts.

3D Reconstruction

Given that conventional ultrasound machines are restricted to produce 2D images, there is a desire to add a further dimension by acquiring a series of images with spatial information. This data pool then permits the generation of volumes. However, similar as in the registration case, issues like ultrasound view dependency complicate this process. If an object has a different appearance depending on the point of view, the question is what representation should be taken. Within the context of this thesis, a particular focus is devoted to bilateral transcranial ultrasound reconstructions, which are performed to

facilitate a more objective early diagnosis of Parkinson's disease. This type of imagery is in particular challenging due to the low signal-to-noise ratio, which is a result of the ultrasound transmission through the skull bone at low frequency in order to avoid total signal reflection due to high acoustic impedance difference.

1.2 Contributions

The research activity that was conducted within the scope of this dissertation has led to several contributions to the field of medical image analysis. In the following we will briefly summarize them and indicate those contribution that already have been presented to the scientific community.

- **Advanced 3D Ultrasound Reconstruction:** We present several techniques for improved 3D ultrasound reconstruction. This ranges from multi-view reconstruction for neurological applications to the incorporation of confidence information and the multi-view freehand modeling employing a mixture model, leading to superior image quality and more accurate volume registration. [Plate et al., 2012, Plate et al., 2010, Ahmadi et al., 2011, Klein et al., 2012b, Karamalis et al., 2012b]

- **Ultrasound Specific Feature Descriptor:** We propose a novel type of feature descriptor for radio frequency ultrasound data, with potential applications in segmentation and registration. Use is showcased in the application of shadow detection in presence of mirroring/reverberation artifacts as well as texture-based distance metric learning. [Klein et al., 2011, Klein et al.,]

- **Ultrasound Specific Similarity Measures:** We present similarity measures specifically designed for the alignment of ultrasound images. In this respect, the specific noise distributions of ultrasound radio frequency (RF) images as well as the differentiation between local and global characteristics is taken into account, which leads to superior registration results as compared to conventional methods. [Wachinger et al., 2012b, Klein et al., 2012a]

- **2D Envelope Detection in Ultrasound:** The 2D analytic signal is an extension to commonly applied 1D analytic signal for the step of envelope detection of ultrasound radio frequency data in the processing pipeline. In contrast to the conventional approach, this method considers information from both across scanlines in lateral and axial direction. This allows for a more robust estimation of the local amplitude. We present the 2D analytic signal, where a particular focus is devoted to the statistical analysis, which suggests comparably better modeling properties. [Wachinger et al., 2011, Wachinger et al., 2012a]

1.3 Thesis Outline

Chapter 2: An overview is given about ultrasound physics. This is followed by a detailed explanation of the image formation processes in conventional machines. Furthermore, fundamental ultrasound speckle scenarios and associated statistical models are introduced.

Chapter 3: This chapter deals with the 2D analytic signal, which suggests improved envelope detection. This is due to the fact that the proposed method considers information from across scanlines in both lateral as well as in axial direction, and therefore produces spatially more consistent output. The resulting images further suggest better statistical modeling properties as is confirmed in goodness-of-fit tests. The chapter is based on the papers *The 2D Analytic Signal for Envelope Detection and Feature Extraction on ultrasound images* [Wachinger et al., 2011] and *The 2D Analytic Signal on RF and B-mode Ultrasound Images* [Wachinger et al., 2012a].

Chapter 4: We present an overview about texture modeling in ultrasound. At first we provide an overview of common techniques applied. Then the Markov Random Field texture models are introduced. Several issues are discussed such as texture generation as well as parameter estimation for specific texture models. This is followed by an introduction of Local Binary Patterns. Next, an example application using the Markov Random Field Nakagami Auto-Model is presented in combination with a learning based distance metric to discriminate between different tissues. Finally, an approach of shadow detection containing mirroring and reverberation artifacts is presented. Within this context we also present ultrasound confidence maps and give a detailed explanation of Expectation-Maximization (EM) algorithm. The chapter is based on the papers *Spatial Statistics Based Feature Descriptor For RF Ultrasound Data* [Klein et al., 2011] and *Shadow Detection in Ultrasound RF Data.* [Klein et al.,].

Chapter 5: This chapter deals with similarity metrics. In this respect, two different approaches are presented. The first approach presented is a combination of global and local similarity measures. On the global scale, distribution matching is performed, which is robust towards intensity changes due to noise. On the local scale, a variant of Local Binary Patterns is employed to increase the accuracy by matching small scale textural details. The second method is an adaptive bivariate Nakagami distribution based measure. By local adaptation of the parameters, this approach yields superior registration performance. The chapter is based on the papers *Locally Adaptive Nakagami-based Ultrasound Similarity Measures* [Wachinger et al., 2012b] and *Registration of RF Ultrasound Data Using Hybrid Local Binary Patterns* [Klein et al., 2012a].

Chapter 6: In this chapter we give an overview of 3D freehand ultrasound. At first different reconstruction methods are outlined. Next we present the results of a first patient study for the early detection of Parkinson's

disease by making use of novel type of bilateral 3D reconstructions. This is followed by advanced reconstruction methods. The first is based on confidence maps as a means of weighting intensity information. The second approach presented is a novel type of multi-view reconstruction method based on finite mixture modeling. Mixture modeling facilitates improved reconstruction in terms of more coherent images as well as comparably better registration performance. In this context, an in-depth presentation of Nakagami Expectation-Maximization algorithm is provided. The chapter is based on the papers *Towards a More Objective Visualization of the Midbrain and its Surroundings Using 3D Transcranial Ultrasound* [Plate et al., 2010], *Modeling of Multi-View 3D Freehand Radio Frequency Ultrasound* [Klein et al., 2012b], *Ultrasound Confidence Maps using Random Walks* [Karamalis et al., 2012b] and *3D Sonographic Examination of the Midbrain for Computer-Aided Diagnosis of Movement Disorders* [Plate et al., 2012].

Chapter 2

Fundamentals of Ultrasound

This chapter gives an overview about ultrasound. Starting from the beginning of ultrasound research to an explanation of how conventional US systems work, including the underlying physical processes. In addition, statistical models describing the backscatter signal, particular in the radio frequency domain, are presented.

2.1 Brief History of Ultrasound

The origins of sonography can be traced back in history to the times of the ancient Greek philosophers and scientists when the fundamental principles were developed. It was Pythagoras (c.570–c.495 BC) who conducted experiments on the sounds produced by vibrating strings and invented the so called Sonometer to study sounds [Brooks and Chan, 2002]. Eventually he established a relationship between pitch and frequency. Aristotle (384–322 BC) discovered that the propagation of sound requires a medium and how its properties affect the transmission [Long et al., 2005]. Later it was the Roman Philosopher Boethius (c.480–c.525 AD) who established that sound travels as a wave, comparing sound waves to waves, produced by dropping a pebble into water [Long et al., 2005]. A scenario where it is observable that waves grew fainter as they move away from their origin.
However, the first milestone towards modern ultrasonography was achieved by Pierre Curie (1859–1906) and Jacques Curie (1856–1941) with the discovery of the piezoelectric effect. It describes the creation of electric potential by compression of crystals and the reverse process, crystals deforming upon applying electrical charge. From the discovery of the piezoelectric effect it was only one small further step to the actual development of ultrasonography. However, it was not until two major striking events until the development, in particular of underwater sonography, gained major momentum: the sinking of the Titanic (1912) and World War I (1914).
After all, it was the french physicists Paul Langévin (1872–1946) and the Russian researcher Constantin Chilowsky (1880–1958), who developed a high frequency ultrasonic echo-sounding device called 'hydrophone' that allowed to detect underwater objects and submarines [Raichel, 2006]. Outbreak of World

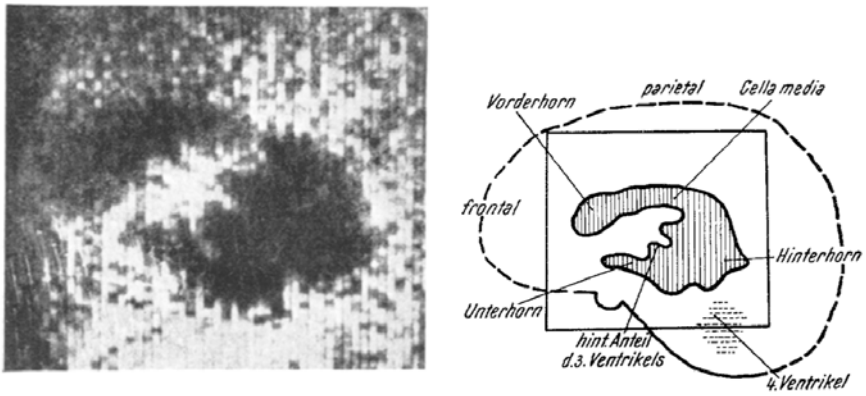


Figure 2.1: Dussik's transmission 'Ventriculograms'. The ultrasound image on the left was initially thought to correspond the lateral ventricles as depicted on the right. Reprinted from [Dussik, 1952], with kind permission from Springer Science and Business Media.

War II further stimulated rapid research activities resulting in great advances in sophisticated naval and military radar sonar as well as radar techniques. Finally, the invention of transistors in 1947 allowed for the miniaturization of electronic devices, which eventually paved the way to applications of ultrasound beyond the military scope.
In spite of that, the first medical applications of ultrasound already took place in the 1940s. The American neurosurgeons William Fry and Russell Meyers performed interventions in which they tried to destroy the basal ganglia of patients suffering from Parkinson's disease for symptom relief. About the same time, the Austrian neurologist Karl Theo Dussik and his brother Friederich Dussik, a physicist began first experiments in the domain of ultrasound diagnosis. They tried to locate brain tumors and the cerebral ventricles as well as visualize intracranial structures [Dussik, 1952, Thomas et al., 2005]. Eventually they developed a through-transmission technique with two transducers placed on either side of the head, yielding so called 'ventriculograms'. This required the immersion of the transducers and part of the patient's head in a water bath. However, the images they produced were later found not to really show cerebral ventricles but merely having been the result of reflection of the skull and attenuation patterns [Thomas et al., 2005] - see Fig. 2.1 for Dussik's ventriculograms. This misinterpretation was exposed by a team of researchers that performed similar acquisitions on an empty skull, yielding identically looking imagery. Although the results obtained by Dussik were of no real diagnostic value, they were the first to obtain rudimentary 2D images. Basically this unveiled the potential of ultrasound for medical diagnostics. However, the ineffectiveness of Dussik's technique, led to the abandoning of the transmission-based ultrasound approach. From now on, research focus was techniques based on the properties of ultrasound reflection [Thomas et al., 2005]. Essentially, this technology is the basis of nowadays modern medical ultrasound systems

found in almost every hospital. In this respect, further developments continued such that within a short time frame technology had advanced to the stage that machines could be built capable of acquiring 2D images. However, this acquisition still took approximately a minute before an image was generated. It was not until the mid 1960s when the first real breakthrough occurred, which heralded a new era in modern ultrasound diagnostics. At that time, largely due to the research work by the German engineer Richard Soldner, the first *real-time ultrasound* imaging device came to the market - the Siemens Vidoson. It was capable in producing images at a respectable frame rate of 15 Hz. By comparison, modern machines of today reach a frequency, depending on the depth, up to approximately 60 Hz. From an application point of view, the Vidoson machine was mainly tailored for ultrasound imaging in the domain of obstetrics and gynecology. However, this soon began to change as this new modality gained acceptance and the underlying technology improved step by step. Basically, the arrival of real-time imaging marked the onset of medical ultrasound as we know it today. Further developments in the years to come then paved the way for the production of ever more powerful and smaller machines. Simultaneously, advances in transducer technology led to better as well as more specific ultrasound probes. These developments taken together then widened the domain of application - a development that has continued until today.

2.2 Ultrasound Physics

Ultrasound refers to sound waves in the range between 20 kHz and 1 GHz, far beyond the capability of the human ear, which is sensitive in the rather narrow interval between 20 Hz and 20 kHz. Conventional medical ultrasound transducers, however, operate in the frequency interval between 1–40 MHz. The propagation of sound can be nicely described given the simple relationship between velocity of sound, frequency and wavelength

$$c = f \cdot \lambda, \tag{2.1}$$

where c denotes the speed of sound, f the frequency and λ the wavelength. Assuming a constant speed of sound $c = 1540 m/s$, the wavelength for medical ultrasound ranges from 1.54 mm to 0.04 mm. As a consequence, ultrasound allows to image objects with sub-millimeter accuracy, which is one of the reasons why it is such a widely accepted clinical imaging modality today. In this respect, important for ultrasound machines is the lateral resolution, which directly depends on the wavelength. The lateral resolution describes the characteristic to distinguish adjacent objects. It is not constant over the image domain and depends on the beam width. The beam width w itself is defined as

$$w = \frac{1.4 \cdot \lambda F c}{2af}, \tag{2.2}$$

where a is the aperture (diameter of the piezo crystal) and F is the focal length (distance from the front face of the transducer to the focal point). In this respect the focal zone is defined as the region where intensity has a value within $3dB$ of the maximum along the transducer axis. Lateral resolution can be improved by focusing the transducer, which is a controlled delay-triggered pulsing and summation of adjacent piezo crystals. Focusing typically leads to a grossly hourglass shaped insonofied beam region, which is most narrow at the focus point. Within the beam, sound propagates through medium as longitudinal pressure wave, inducing local motions. While passing through the propagation medium, ultrasound waves modify the local medium density. This change in density is associated with motion by particle vibration, whereby ultrasound propagates more or less parallel to the direction of vibrating particles. Transversal vibration or shear waves also occur, however, as they do not travel as effectively as longitudinal pressure waves through soft tissue, they are not important for medical ultrasound [Zagzebski, 1996]. Given the basics about wave traveling, next we will derive the basic wave equations. For the following, we typically assume the Euler formalism. Therefore field variables are expressed with respect to a fixed coordinate system in space. As mentioned above, sound propagation induces a change in medium density. The first part in the derivation of the basic wave equation is the motion equation. It relates the force acting on any element of the continuous medium to the pressure gradient on which it depends. In this respect, we employ Newton's second law, which describes the relationship between the forces acting on a body and its motion due to those forces. It can be written as [Hill et al., 2004]

$$-\frac{\partial p}{\partial x} = \rho_0 \frac{u}{Dt}, \tag{2.3}$$

where p denotes the pressure, ρ_0 the ambient density and $u(x,t)$ the velocity in the medium at location x at time t. Beside the motion equation describing the forces, the second relevant principle in the derivation of the wave equation is the conservation of mass. This implies that any change of density within the medium must be associated with an exchange of mass in its surroundings. It can be written as [Hill et al., 2004]

$$-\frac{\partial \rho}{\partial t} = \rho_0 \frac{u}{x}. \tag{2.4}$$

For describing the motion of fluid substances the Navier-Stokes equations are often employed in physics, which basically arise from applying Newton's second law to fluid motion. The Navier-Stokes equation for incompressible viscous flow phenomena in Eulerian form can be written as

$$\rho \left(\frac{\partial u}{\partial t} + u \cdot \nabla u \right) = -\nabla p + \mu \nabla^2 u, \tag{2.5}$$

where μ denotes the viscosity of the medium.

Now combining the Navier-Stokes equation with the equation of motion Eq. 2.3 and the continuity equation Eq. 2.4, one yields the wave propagation equation for inviscid fluids [Hill et al., 2004, Cobbold, 2007].

Furthermore, one is able to obtain a wave equation in p alone defined as

$$\frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} - \nabla^2 p = 0. \tag{2.6}$$

This equation models only the propagation of models ultrasound pulses in a lossless homogeneous medium, particularly neglecting absorption effects. Another restriction is the non-coverage of non-linear wave propagation. That is a phenomena in which the speed depends on the amplitude of the wave. Specifically, while ultrasound waves travel through the body, regions of high pressure move slightly faster than regions of lower pressure and result in a progressive distortion of the transmitted waves. Wave travels faster during the high pressure phase of the oscillation than during the lower pressure phase [Hamilton, 1998]. This is associated with a change of the shape from sinusoidal to non-sinusoidal wave. The non-sinusoidal waves thereby contain additional frequencies, which are multiples of the fundamental originating frequency. In this respect, the even and odd multiples of the fundamental frequency are referred to as harmonic frequencies. This effect is exploited in Tissue Harmonic Imaging (THI), where the fundamental frequency is filtered out and only the second harmonic waves are used for the creation of the image [Boon, 2011]. Among the models incorporating the non-linearity effects are the Khokhlov-Zabolotskaya-Kuznetsov (KZK) [Bakhvalov et al., 1987], Burgers [Burgers, 1948] and notably the Westervelt [Taraldsen, 2001] equation. The former derived from the fluid motion equation by keeping up to quadratic

order terms is defined as [Hamilton, 1998, Taraldsen, 2001]

$$\nabla^2 p - \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2} + \frac{\delta}{c^4}\frac{\partial^3 p}{\partial t^3} = -\frac{\beta}{\rho c^4}\frac{\partial^2 p^2}{\partial t^2}, \tag{2.7}$$

where δ denotes the the sound diffusivity and β is the non-linearity coefficient. Next we will discuss different scenarios of wave interaction. For medical ultrasound there are two important wave interaction phenomena: reflection and transmission. The former being the major interest in diagnostic ultrasound. Assuming that the wavelength is small compared to the object being hit, the wave is redirected upon hitting the object. When the sound beam hits the interface with normal incidence it will be partially reflected towards its origin, whereas the other part continues traveling (transmission). However, when the beam hits the interface with non-normal incidence, the reflection angle equals the angle of incidence and the energy is diffracted away in that direction.
In this regard, Snell's law describes the relationship between the angles of incidence θ_i, transmission θ_t and reflection θ_r, when referring to light or other waves passing through a boundary between two different isotropic media. The law states that the ratio of the sines of the angles of incidence and transmission is equivalent to the ratio of velocities c_1, c_2 in the two media that the waves are passing,

$$\frac{\sin\theta_i}{\sin\theta_t} = \frac{c_1}{c_2}. \tag{2.8}$$

This can be rewritten as

$$\frac{\sin\theta_i}{c_1} = \frac{\sin\theta_r}{c_1} = \frac{\sin\theta_t}{c_2}. \tag{2.9}$$

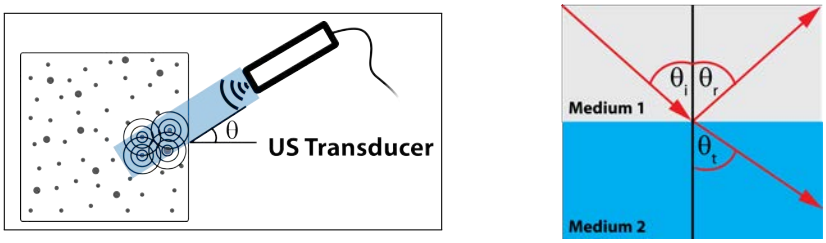
See Fig. 2.2b for an illustration of the reflection scenario with associated angles. A similar relationship can be established for tissue interfaces in medical ultrasound. Thereby different tissues are characterized by their speed of sounds c_i and their respective densities ρ. Their product marks an important property for ultrasound imaging, which is referred to as *acoustic impedance* defined as

$$Z_i = \rho_i \cdot c_i. \tag{2.10}$$

Beside Snell's law, another possibility to describe the relationship between the incident p_i, transmitted p_t and reflected p_r pressure amplitudes is by means of amplitude reflection R_p and transmission T_p coefficients, respectively. They are governed by the two underlying fundamental principles [Postema, 2011]: (1) Continuity of velocity, which states that the normal components of the particle velocity must be equal on either side of the interface. (2) Continuity of pressure, which states that the pressure variation must be equal on either side of the interface. In this respect, the reflection and the transmission coefficients can be written as [Postema, 2011]

$$R_p = \frac{p_r}{p_i} = \frac{Z_2\cos\theta_1 - Z_1\cos\theta_2}{Z_2\cos\theta_1 + Z_1\cos\theta_2} \tag{2.11}$$

$$T_p = \frac{p_t}{p_i} = \frac{2Z_2\cos\theta_1}{Z_2\cos\theta_1 + Z_1\cos\theta_2}, \tag{2.12}$$



(a) (b)

Figure 2.2: Ultrasound wave interaction: (a) Scatterer in resolution cell producing backscatter with wave interference. (b) Ultrasound reflection following Snell's law: θ_i angle of incidence, θ_r angle of reflection, θ_t angle of transmission

where R_p and T_p are the ratios of the amplitudes of the reflected and transmitted waves to the incident wave. Similarly, the intensity reflection R_i and transmission coefficients T_i are defined as [Postema, 2011]

$$R_i = \frac{I_r}{I_i} = R_p^2 = \frac{(Z_2 \cos\theta_1 - Z_1 \cos\theta_2)^2}{(Z_2 \cos\theta_1 + Z_1 \cos\theta_2)} \tag{2.13}$$

$$T_i = \frac{I_t}{I_i} = \frac{4Z_2 Z_1 \cos^2\theta_1}{(Z_2 \cos\theta_1 + Z_1 \cos\theta_2)^2}, \tag{2.14}$$

where I_r, I_t and I_i are the reflected, transmitted and incident intensity, respectively. It should be noted that at normal incidence the expression for R and T simplify significantly, such that we yield the following intensity terms

$$R_i = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1}\right)^2 \quad \text{and} \quad T_i = \frac{4Z_2 Z_1}{(Z_2 + Z_1)^2}. \tag{2.15}$$

The most important scenario occurs, when the interface is larger than the beam width and the beam intersects the interface at normal incidence. Then the interface acts as a specular reflector, reflecting a certain amount of energy directly back to the transducer - depending on the acoustic impedance. It is exactly this phenomena, which is responsible for producing the clear outlines of objects in medical ultrasound. At other angles than normal incidence, the wave will be reflected away from the transducer.

Beside reflection and refraction there are also other important interaction effects that influence US imaging such as attenuation, absorption and scattering.

In this respect, scattering can be considered a specific type of reflection. It occurs at the encounter of ultrasound waves at impedance gradients of objects the size of the wavelength or much smaller [Zagzebski, 1996]. Scattered waves have the tendency to spread in all directions as each scatterer acts as a separate sound source, which leads to a large weakening of the signal as compared to normal reflection. The omni directionality of sound redirection contrasts with the reflection of the echo, which is view dependent. Speckle is responsible for generating textural patterns that are characteristic for ultrasound providing the internal texture of organs. It is noteworthy, that shadow regions result when

scattered waves and the incident wave interfere destructively. See Fig. 2.2a for an illustration of an insonofied resolution cell with scatterers producing backscatter.

While passing through media, the wave loses energy due absorption. More specifically, during the process of wave travel, particles are excited in cycles of medium compression-dilation. This particle vibration and wave motion is linked with a loss of acoustic energy, which is converted into heat or chemical energy. The energy loss in the pressure amplitude is related exponentially to the distance from the source. Generally, all the effects that lead to the gradual loss of wave amplitude such as absorption and scattering are summarized as attenuation. Put more formally, attenuation can be formulated as [Cobbold, 2007, Hedrick et al., 2004]

$$p(x) = p_{max} \cdot \exp(-\alpha x), \tag{2.16}$$

where x denotes the distance travelled by the wave and p_{max} the initial peak of amplitude pressure and α is the absorption coefficient, which depends on frequency, temperature and pressure. Consequently, attenuation necessitates compensation by the ultrasound machine to avoid related intensity gradients over the image. Therefore, in order to obtain images that have an overall consistent brightness, ultrasound machines apply time gain compensation (TGC) during the signal acquisition process. Thereby the signal is amplified increasingly with the time.

Summing up, visible structures are the result of echo, which is generated by scattering or reflection that lead to redirection of some of acoustic energy back to the transducer. Particular, the characteristic texture patterns due to scattering will be investigated in more detail in the following chapters. A special focus is thereby devoted to the statistical modeling of ultrasound backscatter.

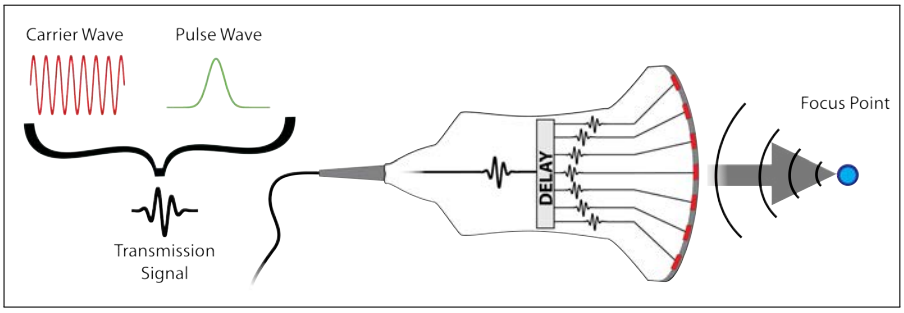


Figure 2.3: Transmission pulse creation from modulated carrier wave and focusing by synchronized delay of pulsation w.r.t. adjacent piezo elements (red).

2.3 Image Formation

Conventional ultrasound transducers are built from an array of up to 250 piezo crystal elements, which altogether yield 2D imagery. Piezoeletric elements can be triggered individually or in groups, which allows focusing and steering of beams by controlled delaying of signal pulsing. Typically, a group of 10-20 elements is trigged synchronized to create a strong and focused beam. Focusing and steering requires that each piezo element fires the transmit signal with an associated time lag with respect to the other elements. The transmitted pulse forming the beam is a signal convolved with a carrier wave created in a modulation process. This in turn requires demodulation upon receiving the echo back at the transducer. Fig. 2.3 shows the process of modulation and sending of a pulse. During modulation, a sinusoidal wave is convolved with a Gaussian modulator.
After crystal excitation the transducers switch mode from sending to receiving. Thereby the transducer stays listening depending on the depth the user set on the machine. More specifically, the receiving period lasts the sufficient amount of time such that echoes from distal areas can be received. Similarly like for the sending process, individual time lags are applied to the crystal elements to compensate for the varying run times of the pulses. The receiving process is also associated with several amplification steps. The first amplifier scales the signal from microvolt/millivolt up to a range of 1-10 Volts [Hedrick et al., 2004]. Afterwards, in order to compensate for the attenuation of the signal a time based correction is applied - referred to time gain compensation (TGC). For more detail on attenuation see Sec. 2.2. Finally, the delayed signals of neighboring elements are summed up to yield a strong echo, producing a beam in the RF image. Optionally, the amplification is followed by a so called frequency-compounding step. Thereby the echo frequency spectrum is subdivided into sub-bands and processed independently. Subsequently, the sub-bands are recombined together, yielding a frequency averaged image of improved quality [Hedrick et al., 2004]. See Fig. 2.5 for an illustration of the receive process in a US transducer.

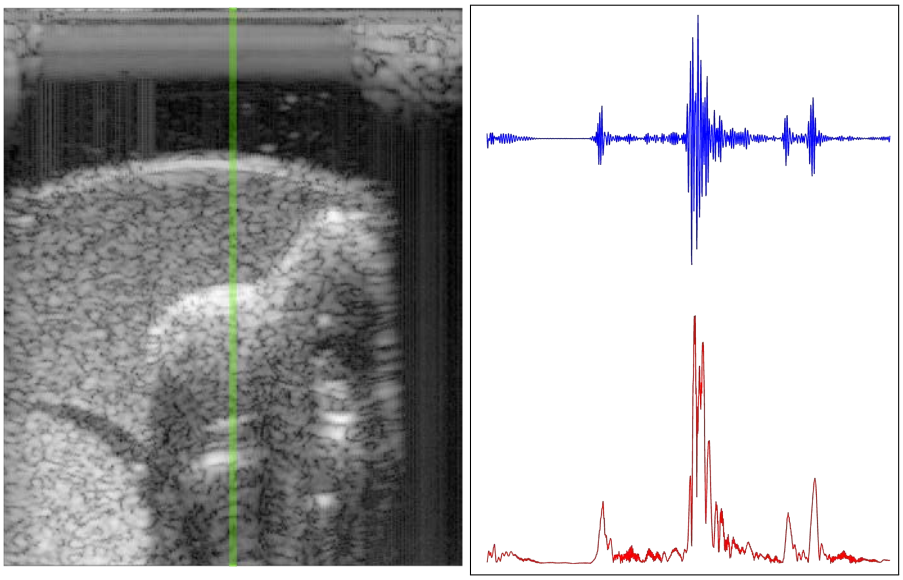


Figure 2.4: Left: RF image build up from 256 A-lines. Right: Corresponding A-line to highlighted zone (green) in the RF image (left). Top-right: RF A-line. Bottom-right: Envelope detected A-line, indicating reduced oscillations as compared to non-enveloped signal.

In order to obtain a 2D image, a series of adjacent beams (A-line) are recorded from which a RF matrix is built. This RF matrix is then further processed in the machine's pipeline, ultimately leading to the B-mode image, which is visualized on the screen. See Fig. 2.4 for an RF matrix and a single beam (A-line) extracted. The real modulated signal as acquired from the US transducer can be written as [Misaridis, 2001]

$$x(t) = A(t) \cdot \cos\left[2\pi f_c t + \phi(t)\right], \tag{2.17}$$

where $A(t)$ corresponds to the amplitude function, f_c to the carrier frequency and $\phi(t)$ to the phase function.

However, often it is advantageous to reformulate Eq. 2.17 to obtain a compact complex representation

$$z(t) = x(t) + jx_H(t), \tag{2.18}$$

which is referred to as the *analytic signal*. Here $x_H(t)$ corresponds to the Hilbert transformed signal of $x(t)$ [Shiavi, 2007, Misaridis, 2001, Oppenheim and Schafer, 2010]. Equation 2.18 can also be reformulated in an expanded manner as

$$z(t) = A(t) \cdot \cos\left[2\pi f_c t + \phi(t)\right] + j \cdot A(t) \sin\left[2\pi f_c t + \phi(t)\right]. \tag{2.19}$$

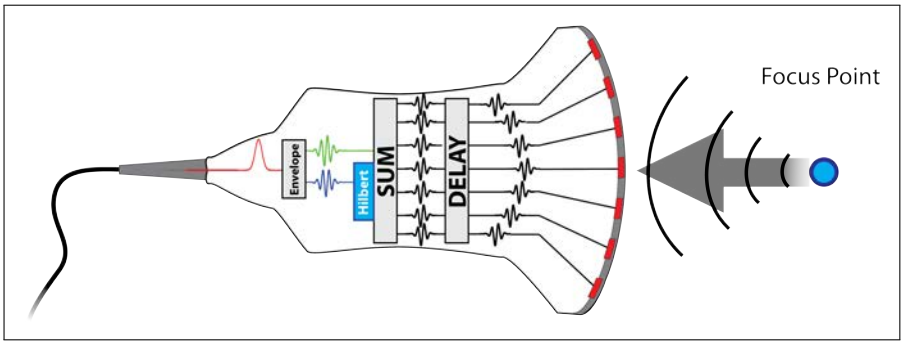


Figure 2.5: Transmission pulse reception. Signal received at adjacent piezo elements (red) are individually delayed then added together. The summed signed is Hilbert transformed and finally envelope detected.

It is noteworthy that the analytic signal has only components in the positive frequency range. All the negative frequency components are deemed superfluous due to symmetry [Shiavi, 2007, Oppenheim and Schafer, 2010].

Equation 2.19 can also be rewritten in a more condensed form yielding the following term

$$z(t) = A(t)e^{[j2\pi(f_c t+\phi(t))]} = c(t)e^{j2\pi f_c}, \tag{2.20}$$

where the complex function $c(t) = A(t)e^{j\phi(t)}$ is referred to as complex envelope, combining the amplitude and phase information of the signal.

Prior to further processing the signal data has to be demodulated, which removes the oscillatory carrier part of the signal. Put in other words, the original information containing signal is separated from the modulated carrier wave that is required to convey the signal pulse. In this respect, ultrasound machines commonly perform this step, referred to as envelope detection, by taking absolute value of the analytic signal given by

$$A_{Envelope} = |z(t)| = \sqrt{x(t)^2 + x_H(t)^2}, \tag{2.21}$$

yielding the envelope detected radio frequency data.

It should be noted that this is performed for each scanline independently. The envelope detected signal is the basis of further processing such as the production of B-mode image as presented to the user on screen.

Enveloped data typically is represented as 16-Bit values on conventional machines. However, computers can usually only display on 8-Bit wide grayscale intensities. Therefore the dynamic range of the data must be reduced to match those of monitors. This is achieved by a non-linear mapping referred to as logarithmic compression. A common way to formalize the log-compression is [Kaplan and Ma;, 1993, Dutt, 1995]

$$f_{compress}(x) = D\ln(x) + G, \tag{2.22}$$

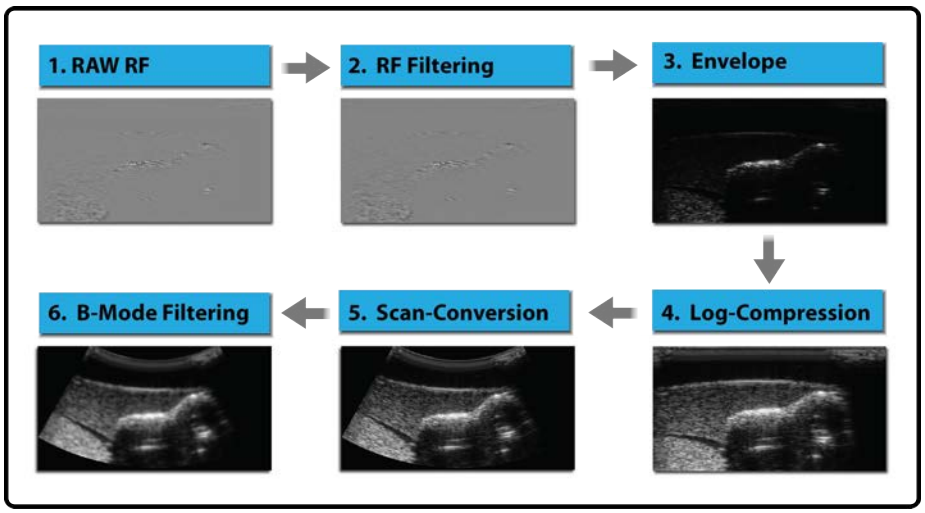


Figure 2.6: Ultrasound data processing chain - from radio frequency (RF) data to B-mode image.

where D is the parameter of the amplifier, representing the input dynamic range to be shown in the output and G is a linear gain parameter of the amplification. Following [Dutt, 1995], given the minimum and maximum input values A_{min} and A_{max} mapped to the corresponding output values X_{min} and X_{max}, respectively, the dynamic range DR_{dB} can be estimated with

$$DR_{dB} = 10 \log \left(\frac{A_{max}^2}{A_{min}^2} \right) = 20 \log \left(\frac{A_{max}}{A_{min}} \right). \tag{2.23}$$

Furthermore, the amplifier parameter D can be computed according to

$$D = \frac{20}{DR_{dB}} \left(X_{max} - X_{min} \right). \tag{2.24}$$

It should be noted that this type of non-linear mapping totally changes the speckle statistics of the envelope detected input signal and in turn makes speckle modeling complicated [Dutt, 1995]. Therefore RF data is quite attractive for directly modeling speckle. Another option to get similarly distributed data, when access to RF data is not possible, is to invert the log-compression of B-mode images in order to obtain an RF data approximation. However, the quality of the approximation is largely dependent on the steps of the processing pipeline.
After the log-compression is applied data must be re-arranged to bring it into geometrical and spatial correct position, which is referred to as scan conversion. That is, data acquired with convex and phased-array transducers must be converted from polar coordinate system to rectangular coordinates as adjacent elements in the RF matrix are not necessarily spatially adjacent. Strictly speaking, spatially adjacency applies only to linear transducers, where beams

are close together and above all parallel. This process entails interpolation steps to fill in the gaps between the beams in order to produce a consistent looking image. At the end of the processing chain typically further proprietary filters are operated on the image data, specific for the application to enhance the visual appeal. Potential filters are speckle reduction and edge enhancement. See Fig. 2.6 for a schematic illustration of the processing chain and results of each step.

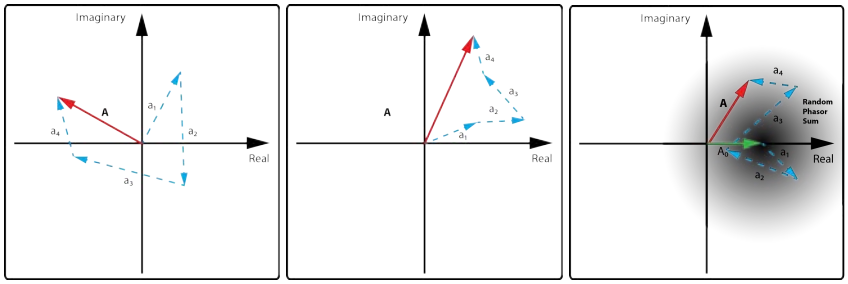


Figure 2.7: Random walk representation for various speckle scenarios. Resultant phasor shown in red; From left to right: Destructive interference, constructive interference and interference with a coherent component/constant phasor (green).

2.4 Ultrasound Image Statistics

Speckle is the result of a random deterministic interference process resulting in characteristic texture patterns [Sanches et al., 2011]. It occurs at structures that are rough in scale as compared to the wavelength - that is inhomogeneities at microscale [Cobbold, 2007]. This noise-like phenomenon affects multiple modalities such as synthetic aperture radar (SAR), optics and ultrasound. For medical ultrasound, the generated speckle patterns are largely characteristic for specific tissue and organ types. Strictly, speaking speckle does not have the characteristics of noise but is rather having noise-like properties [Oliver and Quegan, 2004]. That is, images taken from the same position will look the same, expressing the identical speckle patterns [Soergel, 2010]. This contrasts with the notion of noise, which is associated with an inherent random process.

In the subsequent sections we largely follow the exposition of [Goodman, 2007, Dutt, 1995] in describing the statistics of ultrasound backscatter. Looking at the physical process underlying speckle, those ultrasound characteristic patterns arise in regions where the resolution cell contains a larger number of scatterers. There the ultrasound wave interacts with each scatterer, each in return producing its own backscatter echo; a sinusoidal wave with a phase and amplitude change. The individual amplitudes $\{a_i\}$ and phases $\{\phi_i\}$ of the scatterers are not directly observable, because of their small scale as compared to the resolution cell and wavelength. Therefore the received signal $\boldsymbol{A}$ basically consists of the superposition of all the phases and amplitudes from each scatterer [Goodman, 2007, Dutt, 1995],

$$\boldsymbol{A} = A \cdot e^{j\theta} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \boldsymbol{a_i} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} a_i e^{j\phi_i}, \tag{2.25}$$

where $\boldsymbol{a_i}$ is the i^{th} of N complex phasor components with associated length a_i and phase ϕ_i. In this regard, $\frac{1}{\sqrt{N}}$ serves just a normalization constant needed to preserve second-order moment [Goodman, 2007].

However, since the interference between the scatterers can be either constructive or destructive, the observed signal bears randomness in the phase [Soergel, 2010]. It should be noted that this process of interference itself is purely deterministic, but due the permanent change in interference leads to observable spatial randomness. This in turn is responsible for the noise-like behavior. As the principal source of the noise-like quality of the observed data is the distribution of the phase terms, scatterers at different parts of the resolution cell will contribute very differently, even when having identical backscatter properties [Oliver and Quegan, 2004]. Following this notion, the received backscatter echo can be assumed to behave like a random walk in the complex plane, where each step of length a_i is in a completely random direction ϕ_i. See Fig. 2.7 for a visualization of random walk scatterer interference scenarios. Since the received signal $\boldsymbol{A}$ is a complex random phasor sum, it can be decomposed into a real $\mathcal{R}$ and imaginary $\mathcal{I}$ component, respectively,

$$\mathcal{R} = Re\{\boldsymbol{A}\} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} a_i \cos\phi_i \tag{2.26}$$

$$\mathcal{I} = Im\{\boldsymbol{A}\} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} a_i \sin\phi_i, \tag{2.27}$$

where $Re\{.\}$ and $Im\{.\}$ extract the real and imaginary component.

In practice for describing the speckle by means of random walks it is convenient to adopt certain assumptions about the amplitudes and phases. At first we assume that the phase is uniformly distributed in $[-\pi, \pi]$ and independent of the amplitude, which will be distributed around a specific mean value [Oliver and Quegan, 2004, Soergel, 2010, Goodman, 2007]. The next assumption is that the amplitudes a_i and phases ϕ_i are statistically independent of any a_j and phases ϕ_j with $i \neq j$. Furthermore, beside uniformly distributed phases, we assume independence between the amplitudes and phases. This in turn allows a series of simplifications in modeling. As a direct consequence of the assumptions above, the mean and variance of $\mathcal{R}$ and $\mathcal{I}$ can be written as [Goodman, 2007],

$$E[\mathcal{R}] = E\left[\frac{1}{\sqrt{N}} \sum_{i=1}^{N} a_i \cos\phi_i\right] = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} E[a_i]\, E[\cos\phi_i] = 0 \tag{2.28}$$

$$E[\mathcal{I}] = E\left[\frac{1}{\sqrt{N}} \sum_{i=1}^{N} a_i \sin\phi_i\right] = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} E[a_i]\, E[\sin\phi_i] = 0, \tag{2.29}$$

where $E[.]$ denotes the expectation. This in turns allows to infer the following terms for the variances

$$\sigma_{\mathcal{R}}^2 = \frac{1}{N} \sum_{i=1}^{N} E\left[a_i^2\right] E\left[\cos^2\phi_i\right] = \frac{1}{N} \sum_{i=1}^{N} \frac{E\left[a_i^2\right]}{2} \tag{2.30}$$

$$\sigma_{\mathcal{I}}^2 = \frac{1}{N} \sum_{i=1}^{N} E\left[a_i^2\right] E\left[\sin^2\phi_i\right] = \frac{1}{N} \sum_{i=1}^{N} \frac{E\left[a_i^2\right]}{2}. \tag{2.31}$$

These results will be of use for the following sections.

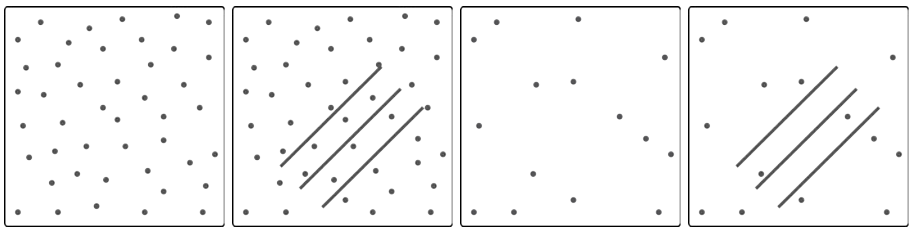


Figure 2.8: Various scattering conditions. From left to right: Large number of scatterers, larger number of scatterers with coherent component, small number of scatterers, small number of scatterers with coherent component.

2.4.1 Distributions for Basic Scattering Scenarios

In the following subsection distributions for various common scattering scenarios will be derived. Derivations are based on the random phasor representation referring to Eq. 2.25, 2.26 and 2.27. See Fig. 2.8 for an illustration of scattering conditions modeled with distributional models.

Large Number of Scatterers

Given the random walk interpretation and assuming the number of scatterers N very large, with scattering cross-sections approximately uniform [Shankar, 1995], Eqns. 2.25 and 2.26 yield very large sums. Applying the Central Limit Theorem with $N \to \infty$ and the independence assumption of each random walk step, the real $\mathcal{R}$ and imaginary parts $\mathcal{I}$ of the summed backscatter phasor $\boldsymbol{A}$ become approximately Gaussian distributed. Therefore the joint distribution for $\mathcal{R}$ and $\mathcal{I}$ can be written as [Goodman, 2007]

$$p_{\mathcal{R},\mathcal{I}}(\mathcal{R},\mathcal{I}) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{\mathcal{R}^2+\mathcal{I}^2}{2\sigma^2}\right) \tag{2.32}$$

with $\sigma^2 = \sigma_{\mathcal{R}}^2 = \sigma_{\mathcal{I}}^2$ due to Eq. 2.30 and 2.31. However, we are rather interested in the distributions for the amplitude A and the phase θ, which have the following relation to the real and imaginary quantities.

$$A = \sqrt{\mathcal{R}^2+\mathcal{I}^2} \tag{2.33}$$

$$\theta = \arctan\left(\frac{\mathcal{I}}{\mathcal{R}}\right) \tag{2.34}$$

and

$$\mathcal{R} = A\cdot\cos\phi \tag{2.35}$$

$$\mathcal{I} = A\cdot\sin\phi. \tag{2.36}$$

The joint distribution for the amplitude A and the phase θ can be derived from the joint distribution $p_{\mathcal{R},\mathcal{I}}$ by transformation of random variables [Flury, 1997],

$$p_{A,\theta}(A,\theta) = p_{\mathcal{R},\mathcal{I}}(A\cdot\cos\phi, A\cdot\sin\theta)\,|det(J)| \tag{2.37}$$

where J is the Jacobian matrix relating the two variable sets, defined as

$$J = \begin{pmatrix} \frac{\partial\mathcal{R}}{\partial A} & \frac{\partial\mathcal{R}}{\partial\phi} \\ \frac{\partial\mathcal{I}}{\partial A} & \frac{\partial\mathcal{I}}{\partial\phi} \end{pmatrix} \quad \text{and} \quad |det(J)| = A. \tag{2.38}$$

Applying the transformation theorems, we yield

$$p_{A,\theta}(A,\theta) = \frac{A}{2\pi\sigma^2}\exp\left(-\frac{A^2}{2\sigma^2}\right). \tag{2.39}$$

From the joint distribution $p_{A,\theta}(A,\theta)$ for amplitude and phase one can derive the corresponding probability density function by marginalizing w.r.t. the other random variable. The resulting probability density function for the amplitude A is therefore

$$p_A(A) = \int_{-\pi}^{+\pi} p_{A,\theta}(A,\theta)\,\mathrm{d}\theta = \frac{A}{\sigma^2}\exp\left(-\frac{A^2}{2\sigma^2}\right) \quad \text{s.t.} \quad A\in\mathbb{R}^+, \tag{2.40}$$

which corresponds to the *Rayleigh* distribution.

Large Number of Scatterers with Constant Phasor

If the scattering scenario from the previous section is slightly modified by assuming the presence of a coherent component in the resolution cell, we yield a different model. The coherent component might arise due to strong specular scattering or unresolved structures such as periodically located scatterers [Dutt, 1995]. This slight modification of the previous model implicates a different model, however, with very similar derivation. The coherent component can be formalized by extending real part of the resulting phasor from Eq. 2.26 with an additional constant amplitude part A_0, yielding

$$\mathcal{R} = A_0 + \frac{1}{\sqrt{N}}\sum_{i=1}^{N} a_i\cos\phi_i, \tag{2.41}$$

with the following associated joint distributions

$$p_{\mathcal{R},\mathcal{I}}(\mathcal{R},\mathcal{I}) = \frac{1}{2\pi\sigma^2}\exp\left(-\frac{(\mathcal{R}-A_0)^2+\mathcal{I}^2}{2\sigma^2}\right) \tag{2.42}$$

$$p_{A,\theta}(A,\theta) = \frac{A}{2\pi\sigma^2}\exp\left(-\frac{A^2+A_0^2-2AA_0\cos\theta}{2\sigma^2}\right) \tag{2.43}$$

Marginalization over the phase θ yields the probability density distribution for the envelope

$$p_A(A) = \frac{A}{\sigma^2}\exp\left(-\frac{A^2+A_0^2}{2\sigma^2}\right)\cdot I_0\left(\frac{AA_0}{\sigma^2}\right) \quad \text{s.t.} \quad A\in\mathbb{R}^+, \tag{2.44}$$

where $I_0(.)$ represents the modified Bessel function of first kind and order zero. The resulting probability density function is also known as the *Rice* distribution, which is also referred to as the *Rician* or *Ricean* distribution. Unsurprisingly, for very small constant phase A_0 ($A_0 \to 0$) we yield again a *Rayleigh* distribution, therefore the *Rice* distribution is a generalization of the *Rayleigh* distribution. Parameter estimation can be performed in various ways such as closed form based on moments and iterative maximum-likelihood [Talukdar and Lawing, 1991, Sijbers et al., 1998, Benedict and Soong, 1967].

Limited Number of Scatterers

Previous models assumed a large number of scatters per resolution cell. Shedding this assumption, which is only satisfied in very limited situations, by reducing the number of scatters to around 10 per resolution cell, we yield an unlike more realistic but also much more complicated model - for more detail and the derivation see [Jakeman and Pusey, 1976, Jakeman, 1980, Jakeman and Tough, 1987b]. Making again use of the random walk phasor sum representation endowed with the negative binomial distribution to model the number of N scatterers, one yields the *K-Distribution* modeling the echo envelope A. The probability density function defined as [Jakeman and Tough, 1987a, Dutt, 1995]

$$p_A(A) = 2\left(\frac{A}{2}\right)^{\alpha} \frac{b^{\alpha+1}}{\Gamma(\alpha)} K_{\alpha-1}(bA) \quad \text{s.t.} \quad A \in \mathbb{R}^+, \tag{2.45}$$

where K_α represents the modified Bessel function of second kind of order α and $b = \sqrt{\frac{4\alpha}{E[A^2]}}$. It is noteworthy that the parameter α can be directly related to the physical property underlying ultrasound speckle generation. It corresponds to the number of effective scatterers in the resolution cell influencing the echo envelope. Unlike it counterpart for a large number of scatterers, it can also model variations in scattering cross-sections [Shankar, 1995], which are more likely to change in regions of abnormality and are therefore of interest for tissue characterization [Joynt, 1979]. Furthermore, for large α ($\alpha \to \infty$) the *K*-distribution converges to the *Rayleigh* distribution, which further highlights that the *K*-distribution is a generalization of the *Rayleigh* distribution. In spite of its inherent complexity, the parameters of the *K*-distribution can be estimated in closed from directly from the sample moments, due to the closed form expressions of its moments, or iteratively following a maximum-likelihood approach [Dutt, 1995].

Limited Number of Scatterers with Constant Phasor

Similarly to the scenario with a *large number of scatterers*, where switching from *Rayleigh* to *Rice* distribution is due to added coherence, the previous case of *limited number of scatterers* can be extended to cope with coherence. This is achieved by incorporating a constant phasor component, which represents a coherent component in the backscatter echo. Model extension leads from the *K* distribution to the *homodyned-K* distribution, defined as [Jakeman and Tough, 1987b, Dutt and Greenleaf, 1994]

$$p_A(A) = A \int_{u=0}^{+\infty} u \frac{J_0(us) J_0(uA)}{\left(1 + \frac{u^2\sigma^2}{2\alpha}\right)^{\alpha}} \mathrm{d}u \tag{2.46}$$

which evaluates to the rather complex term

$$\begin{aligned} p_A(A) = \frac{1}{\Gamma(\alpha)} \left(\frac{A}{s}\right)^{\frac{1}{2}} \sqrt{\frac{2\alpha}{\pi\sigma^2}} \sum_{m=0}^{\infty} \left[(-1)^m \frac{\Gamma(m+\frac{1}{2})}{m!\Gamma(\frac{1}{2}-m)} \left(\frac{\sigma^2}{sA\alpha}\right)^m \right. \\ \left. \left(\frac{\alpha}{2\sigma^2}\right)^{\frac{\alpha+m-1/2}{2}} |s-A|^{\alpha+m-1/2} K_{\alpha+m-1/2}\left(\sqrt{\frac{2\alpha}{\sigma^2}}|s-A|\right) \right]. \end{aligned} \tag{2.47}$$

Here J_p denotes the Bessel function of the first kind of order p, K_α represents the modified Bessel function of second kind of order α and s^2 corresponds to the coherent and σ^2 to the diffuse signal energy. Similarly to the *K* distribution, the parameter α corresponds to the effective number of scatterers. Furthermore, it should be noted that the *K* distribution is a special case of the *homodyned-K* and the latter converging for large α ($\alpha \to \infty$) to a *Rice* distribution. As a result it can model a wide range of scattering scenarios such as the ones described by Rice, Rayleigh and K distributions.
Although computationally adverse, parameters of the *homodyned-K* distribution can be computed based on fractional order moments [Hruska, 2009, Prager et al., 2002].
Beside the *homodyned-K* distribution, similar speckle conditions can be modeled by the so called *Generalized K* distribution [Barakat, 1986, Jakeman and Tough, 1987a], however, not always modeling weakly scattering media properties reasonably. A case in point is the that predicted mean intensity of the scattered field is subject to fluctuation due to variations in the number of scatterers, even for the case of the mean intensity of the scatterers approaching zero [Destrempes and Cloutier, 2010, Jakeman and Tough, 1987a]. Additionally, it is stated in [Destrempes and Cloutier, 2010] that strictly speaking only the parameters of the *homodyned K* distribution can be reasonably associated with a physical meaning. It is exactly this property that makes this distribution very attractive for application domains such as tissue characterization.

2.4.2 Further Statistical Models for RF Data

Beside the established models for modeling RF data as presented in the previous subsections, there also exist several relatively new models. These models typically capture various speckle scenarios. At first we present the Nakagami distribution and then the Rician Inverse Gaussian distribution.

Nakagami Distribution

Due to their inherent complexity, the practical applicability of distributions such as the *Generalized-K* or *homodyned-K* is often limited. Simpler but nonetheless extremely versatile, the *Nakagami* distribution [Nakagami, 1960] (also

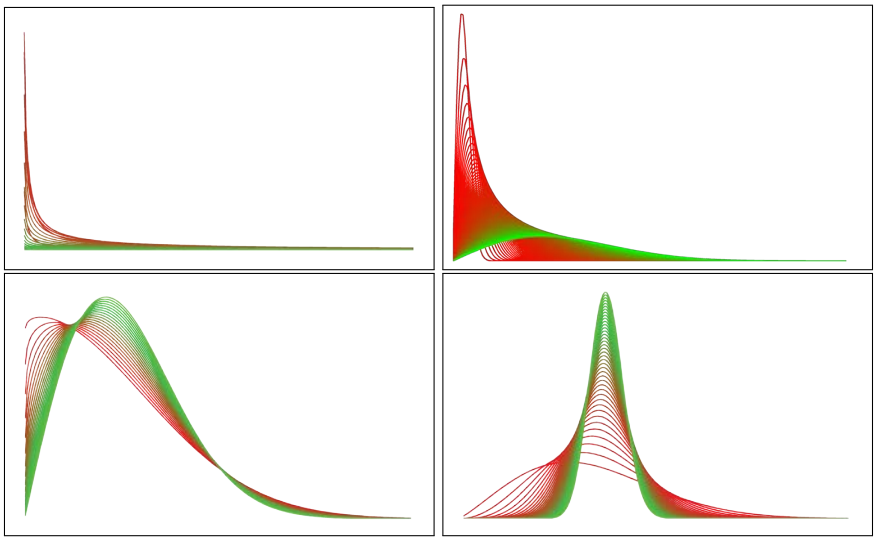


Figure 2.9: Nakagami distribution for various shape parameter ranges. Top-left: *Pre-Rician* with $0 < \mu < 1/2$, Bottom-left: *Generalized Rician* with $1/2 < \mu < 1$, Top-right: *Rayleigh* with $\mu = 1$, Bottom-right: *Rician* with $1 < \mu$.

known as the *Nakagami-m* distribution) was proposed originally for modeling the fading of radio signals, admitting an explicit analytical expression of various speckle scenarios. In particular it allows modeling of large number of scatterers with and without constant phasor as well as small number of scatterers with varying scatterer cross-sections. Merely the scenario of limited number of scatterers in combination with a constant phasor cannot be modeled fully. In this respect, it should be noted that by varying the shape parameter of the *Nakagami* distribution, it is possible to emulate other distributions such as *Rayleigh*, *Rician* or *homodyned-K* [Shankar et al., 2001, Destrempes and Cloutier, 2010]. See Fig. 2.9 for an illustration of the shape parameter affecting the look of the Nakagami distribution. It was shown that the *Nakagami* distribution accurately models backscatter characteristics of ultrasound envelope data [Shankar, 2000] and hence is used in various applications such as segmentation and classification, see [Destrempes et al., 2009, Shankar et al., 2002] and references therein. Furthermore, it was also shown that its parameters can also be directly used for applications such as tissue classification [Shankar et al., 2002]. The *Nakagami* distribution $\mathcal{N}(x \mid \mu, \omega)$ belongs to the exponential family and requires the specification of two parameters, μ and ω, determining shape and scale, respectively,

$$\mathcal{N}(x \mid \mu, \omega) = \frac{2\mu^{\mu} x^{2\mu-1}}{\Gamma(\mu)\omega^{\mu}} \exp\left(-\frac{\mu}{\omega}x^2\right), \forall x \in \mathbb{R}_+. \tag{2.48}$$

Its parameters can be estimated among other options using the method of moments (MOM) [Cheng and Beaulieu, 2002] and iteratively, following a maximum-likelihood (MLE) approach [Cheng and Beaulieu, 2001]. For the

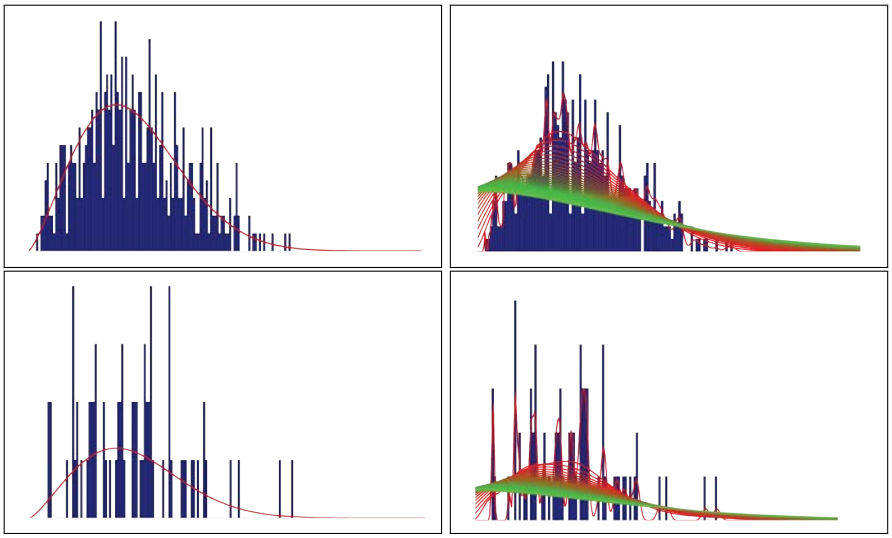


Figure 2.10: Nakagami parameter estimation. Left: Maximum-likelihood estimation, Right: Small-window estimation with different kernel sizes (Kernel size [red to green]: 0.001 - 2.5). Top: Large number of samples, Bottom: Small number of samples.

method of moments approach one considers the 2^{nd} and 4^{th} moments,

$$\omega = \mathrm{E}[x^2] \tag{2.49}$$

and

$$\mu = \frac{[\mathrm{E}(x^2)]^2}{\mathrm{E}[x^2 - \mathrm{E}(x^2)]^2} \tag{2.50}$$

with the latter being referred to as inverse normalized variance estimator [Abdi and Kaveh, 2000] . Again, $E[.]$ denotes the expected value. In order to increase the locality of the parameters, a small window kernel regression based on the non-symmetric Gamma kernel was proposed in [Larrue and Noble, 2011,Chen, 2000]. It allows the estimation of the Nakagami parameters given only a small amount of sample data. See Fig. 2.10 for an illustration of parameter estimation. Assuming data follows Nakagami distribution, the Gamma distribution may be used instead as they are related by a simple transformation given by

$$Y \sim \mathcal{GA}(x \mid \mu_{gam}, \omega_{gam}), X \sim \mathcal{N}(x \mid \mu_{nak}, \omega_{nak}) \tag{2.51}$$

$$\Rightarrow \sqrt{X} = Y(\mu_{nak}, \omega_{nak}/\mu_{nak}). \tag{2.52}$$

This transformation will employed throughout this thesis. It can be derived by applying the transformation of random variables theorems [Papoulis and Pillai, 2002],

$$\int_0^\infty \frac{2\mu_{nak}^{\mu_{nak}}}{\Gamma(\mu_{nak})\omega_{nak}^{\mu_{nak}}} y^{2\mu_n ak-1} \exp\left(-\frac{\mu_{nak}}{\omega_{nak}} y^2\right) \mathrm{d}y =$$

$$\int_0^\infty \frac{2\mu_{nak}^{\mu_{nak}}}{\Gamma(\mu_{nak})\omega_{nak}^{\mu_{nak}}} \sqrt{x}^{2\mu_n ak-1} \exp\left(-\frac{\mu_{nak}}{\omega_{nak}} (\sqrt{x})^2\right) \frac{\mathrm{d}x}{2\sqrt{x}} =$$

$$\int_0^\infty \frac{2\mu_{nak}^{\mu_{nak}}}{\Gamma(\mu_{nak})\omega_{nak}^{\mu_{nak}}} \frac{x^{\mu_{nak}-1/2}}{x^{1/2}} \exp\left(-\frac{\mu_{nak}}{\omega_{nak}} x\right) \mathrm{d}x = \int_0^\infty \frac{\mu_{nak}^{\mu_{nak}}}{\Gamma(\mu_{nak})\omega_{nak}^{\mu_{nak}}} x^{\mu_{nak}-1} \exp\left(-\frac{\mu_{nak}}{\omega_{nak}} x\right) \mathrm{d}x =$$

$$\int_0^\infty \frac{\mu_{gam}^{\mu_{gam}}}{\Gamma(\mu_{gam})\omega_{gam}^{\mu_{gam}} \mu_{gam}^{\mu_{gam}}} x^{\mu_{gam}-1} \exp\left(-\frac{x}{\omega_{gam}}\right) \mathrm{d}x = \int_0^\infty \frac{1}{\Gamma(\mu_{gam})\omega_{gam}^{\mu_{gam}}} x^{\mu_{gam}-1} \exp\left(-\frac{x}{\omega_{gam}}\right) \mathrm{d}x$$

This representation is often case more convenient and permits various simplifications. Furthermore, the Gamma distribution is more of a standard and therefore typically contained in all popular mathematical software programs. For the sake of completeness it should also be mentioned that there exists also the so-called *Generalized Nakagami* distribution [Shankar, 2001], which is defined as

$$\mathcal{N}(x \mid \mu, \omega, s) = \frac{2s\mu^\mu x^{2s\mu-1}}{\Gamma(\mu)\omega^\mu} \exp\left(-\frac{\mu}{\omega} x^{2s}\right), \forall x \in \mathbb{R}_+. \tag{2.53}$$

Compared to the standard *Nakagami* distribution, it has an additional parameter s referred to as shape adjustment parameter. By applying the transformation $x^{1/s}$ on the data one yields again *Nakagami* distributed values. It should be noted that the mean intensity of the *Generalized Nakagami* does not have an analytic expression unless the parameter s is an integer [Destrempes and Cloutier, 2010].

Rician Inverse Gaussian Distribution

In [Eltoft, 2003, Eltoft, 2005] a three-parameter distribution referred to as Rician Inverse Gaussian (RIG) is proposed, which is constructed as a mixture of the Rician probability density function with the Inverse Gamma (IG), defined as

$$p_A(A|\alpha, \beta, \gamma) = \int p_{A|Z}(a|z, \beta) p_Z(z|\delta, \gamma) dz, \tag{2.54}$$

where the conditional distribution of A given Z is Rician

$$p_{A|Z}(a|z, \beta) = \frac{a}{z} \exp\left(-\frac{1}{2z}(a^2 + \beta^2 z^2)\right) I_0(\beta a), \tag{2.55}$$

with β denoting a skewness parameter and $I_0(.)$ denoting the modified Bessel function of first kind and zero order. Furthermore, the Inverse Gamma distribution is defined as [Barndorff-Nielsen, 1997]

$$p_Z(z|\delta, \gamma) = \frac{\delta}{\sqrt{2\pi}} \exp(\delta\gamma) z^{-3/2} \exp\left(-\frac{1}{2}\left[\frac{\delta^2}{z} + \gamma^2 z\right]\right). \tag{2.56}$$

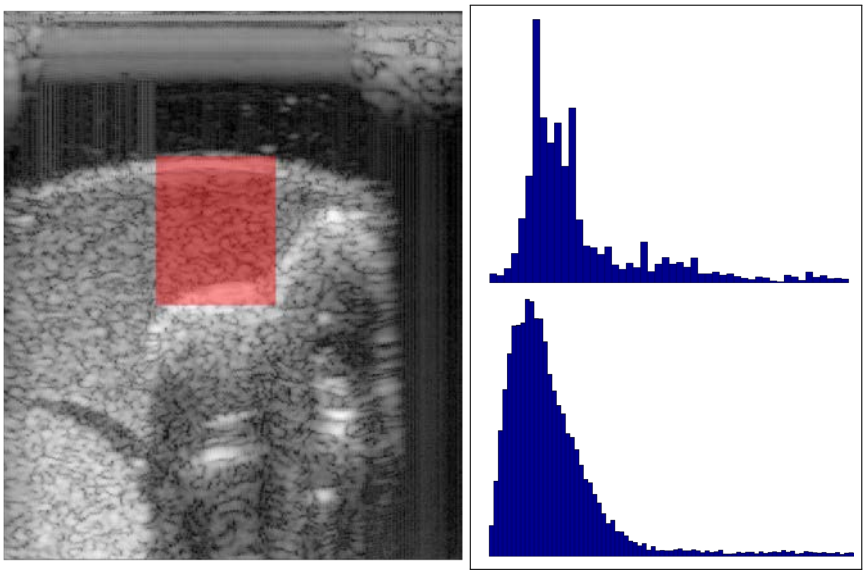


Figure 2.11: Left: Image with region highlighted in red, from which histograms are computed. Right-top: Histogram from B-mode. Right-bottom: Histogram from envelope detected RF data. Clearly visible the comparably increased sparsity and roughness in the B-mode histogram due to quantization and reduced resolution.

Beside the analytical expression given in Eq. 2.54, the RIG distribution can also be represented in closed form defined as

$$p_A(a|\alpha,\beta,\gamma) = \frac{2}{\pi}\alpha^{3/2}\gamma\exp(\delta\gamma)\frac{a}{(\delta^2+a^2)^{3/4}}K_{3/4}\left(\alpha\sqrt{(\delta^2+a^2)}\right)I_o(\beta a), \tag{2.57}$$

where $K_x(.)$ denotes the modified Bessel function of second kind and order x. Using the RIG distribution, non-Rayleigh amplitude statistics can be modeled, which are all scenarios except fully developed speckle. Parameter estimation of the RIG distribution is possible in an iterative fashion, either based on maximum likelihood or based on moments, the latter less computationally demanding [Eltoft, 2005].

2.4.3 Statistical Models for B-Mode

All the above mentioned models basically only apply to envelope detected RF ultrasound data. However, typical ultrasound machines provide access only to B-Mode images, although lately, raw US radio frequency (RF) data has become more readily available to the research community. Compared to RF data, B-Mode imagery is affected by non-linear post-processing steps, such as log-compression and other proprietary filter algorithms that change the speckle statistics for reasons of improved visual appeal as was discussed at length in Sec. 2.3 - see Fig. 2.11 for an illustration of changed speckle statistics in B-mode

vs. RF data. However, there exist several approaches that employ statistical models for B-mode imagery. One possibility is to decompress the B-mode image in order to approximate the original enveloped RF data, whereby the challenge is to recover the parameter D from Eq. 2.22, which represents the input dynamic range to be shown in the output. More specifically, the envelope x can be obtained approximately by inverting the log-compression given the B-mode intensity y by

$$x = \exp\left(\frac{y}{D}\right), \tag{2.58}$$

where according to [Wagner et al., 1983], in regions of fully developed speckle the recovered intensities follow exponential distribution

$$p(y|\sigma) = \frac{1}{2\sigma^2}\exp\left(\frac{-y}{2\sigma^2}\right), \tag{2.59}$$

which is exploited in the decompression approach proposed in [Prager et al., 2003]. However, this process is further complicated by the proprietary filtering, interpolation and image quantization steps specific for different US machine manufacturers. In the approach proposed in [Prager et al., 2003], the image is decompressed whereby the dynamic range parameter is computed iteratively based on fractional moments, which is followed by speckle detection employing a *homodyned-K* distributional model. In [Dutt and Greenleaf, 1996] a distribution for the log-compressed envelope is derived, assuming envelope statistics to follow *K*-distribution. In [Kaplan and Ma;, 1993] the log-compressed B-Mode under the assumption of Rayleigh scattering is modeled employing a *Fisher-Tippet* distribution, which is defined as [Abramowitz and Stegun, 1964]

$$p(y|a,b) = \frac{2}{a}\exp\left(-\log(2\sigma^2) - 2\frac{y-b}{a} - \exp\left(-\log(2\sigma^2) - 2\frac{y-b}{a}\right)\right). \tag{2.60}$$

However, their approach requires access to the ultrasound data before scan conversion. Finally, in [Crawford et al., 1993] the processing pipeline characteristics are analyzed assuming an underlying Rayleigh model and using a phantom. Beside decompression, the log-compression can also be directly integrated into the noise models as was done for similarity metrics in the works of [Cohen and Dinstein, 2002a] and [Myronenko et al., 2009a]. Similarly, in [Sanches and Marques, 2003] non-linear compression was integrated in a 3D freehand reconstruction system in order to improve the quality of volumes, under the assumption that the envelope detected RF data follows *Rayleigh* distribution.

Chapter 3

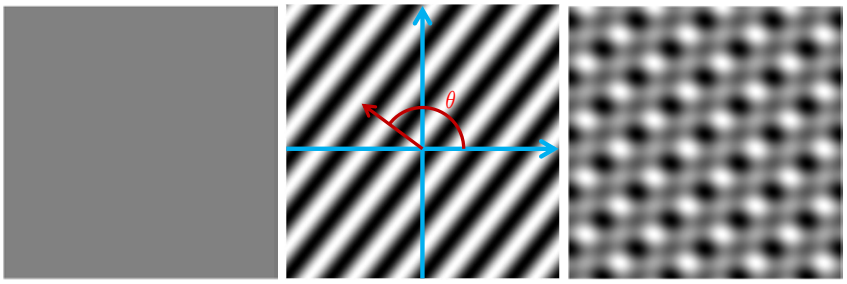
2D Analytical Signal

Underlying the statistical image processing of radio frequency data is the analytic signal. Whereas conventionally, the analytic signal is computed in 1D, on a scanline per scanline basis, this chapter illustrates the advantages of an improved concept for RF demodulation, which permits modeling of the analytic signal in 2D and herein intrinsic structures. Furthermore, the 2D analytical signal suggests superior statistical modeling properties, as is shown by subsequent goodness-of-fit tests to a Nakagami distributional model, which indicate a clear advantages and improved applicability. This chapter is based on the papers 'The 2D Analytic Signal for Envelope Detection and Feature Extraction on ultrasound images' [Wachinger et al., 2011] and 'The 2D Analytic Signal on RF and B-mode Ultrasound Images' [Wachinger et al., 2012a].

3.1 Introduction

The analytic signal (AS) facilitates the extraction of local, low-level features from images, whereby featuring the fundamental property of split of identity - separation of qualitative and quantitative information of a signal in form of the *local phase* and the *local amplitude*, respectively. These quantities further fulfill invariance and equivariance properties [Felsberg and Sommer, 2001]. This in turn allows for an extraction of structural information that is invariant to brightness or contrast changes. It is exactly those favorable properties in image analysis that have led to a multitude of applications in computer vision and medical imaging. Notably among them is employment in registration [Carneiro and Jepson, 2002, Grau et al., 2007, Mellor and Brady, 2005, Zang et al., 2007, Zhang et al., 2007], detection [Estepar et al., 2006, Mulet-Parada and Noble, 2000, Szilágyi and Brady, 2009, Xiaoxun and Yunde, 2006], segmentation [Ali et al., 2008, Hacihaliloglu et al., 2008, Wang et al., 2009], and stereo [Fleet et al., 1991]. In this respect, phase-based processing is particularly interesting for ultrasound imagery due to the significant brightness variations [Grau et al., 2007, Hacihaliloglu et al., 2008, Mellor and Brady, 2005, Mulet-Parada and Noble, 2000].

For 1D domain, the local phase is calculated with the 1D analytic signal. However, for 2D several extensions of the analytic signal are proposed. In



(a) i0D (b) i1D (c) i2D

Figure 3.1: Illustration of 2D signals with different intrinsic dimensionality. For i1D, we show the local orientation

this respect, the *monogenic signal* [Felsberg and Sommer, 2001] represents an isotropic extension. The description of the signal's structural information (phase and amplitude) is extended by a geometric component, the *local orientation*. As the name suggests it indicates the orientation of intrinsic 1D (i1D) structures in 2D images - see Fig. 3.1 for an illustration of various intrinsic structures. Herein, exactly lies the limitation of the monogenic signal - it is restricted to the subclass of i1D signals. Recently, however, [Wietzke et al., 2009] proposed the *2D analytic signal*, which is a natural extension to the monogenic signal permitting the analysis of intrinsic two dimensional (i2D) signals. Therefore, the 2D signal analysis is embedded into 3D projective space and a new geometric quantity, the *apex angle*, is introduced. Furthermore, the 2D analytic signal also has the favorable property of more accurately estimating local features from i1D signals [Wietzke et al., 2009].

In the remainder of this chapter, we show the advantages of the calculation of the 2D analytic signal for in particular radio frequency (RF) as well as B-mode ultrasound images. Notably, the first obvious difference compared to standard RF demodulation approaches is that it is not performed for each scan line separately. Rather, the demodulation is performed in its natural 2D context with 2D Hilbert filters of first- and second-order. Therefore the structural integrity of structures in the 2D domain such as speckle is preserved. Since all further processing steps of the creation of the B-mode image are based on the envelope detected RF data, improvement of this critical step propagates through the entire ultrasound pipeline and directly affects the quality of the ultrasound imagery. Moreover, beside better and more coherent visual appearance, the result from the 2D envelope detection bears better statistical properties, as we illustrate with goodness-of-fit tests w.r.t. the Nakagami distributional model. This in turn has advantageous implications in diverse fields such as classification, registration and segmentation, which is of particular interest for statistical modeling approaches as presented in this thesis. Finally, we show the advantages of the 2D analytic signal for estimating local features on B-mode images, whereby all experiments are performed on clinical ultrasound images.

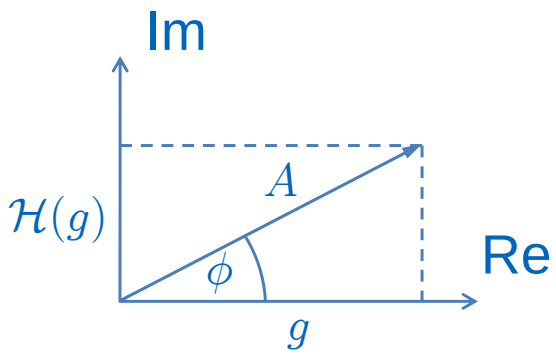


Figure 3.2: Illustration of the analytic signal in the complex plane. Real signal g, Hilbert transformed signal $\mathcal{H}(g)$, phase ϕ, and amplitude A.

3.2 2D Analytic Signal

There are various concepts to analyze the phase of signals, such as Fourier phase, instantaneous phase, and local phase [Granlund and Knutsson, 1995]. However, we are primarily interested in the last two. For 1D signals, $g \in L^2(\mathbb{R})$, the instantaneous phase is defined as the argument of the analytic signal

$$\phi = \arg(g + i \cdot \mathcal{H}\{g\}),$$

with $\mathcal{H}$ being the Hilbert transform and $i = \sqrt{-1}$. See Fig. 3.2 for an illustration of the analytic signal with phase and amplitude in the complex plane. In this respect, the instantaneous amplitude A is the absolute value of the analytic signal

$$A = \sqrt{g^2 + \mathcal{H}\{g\}^2}.$$

Since real signals consist of a superposition of multiple signals of different frequency components, the instantaneous phase, although local, can lead to wrong estimates. Therefore, the signal has to be split up into multiple frequency bands in order to achieve meaningful results. This can be achieved by means of bandpass filters as is further described in Sec. 3.2.2.

Considering 2D signals, $f \in L^2(\mathbb{R}^2)$, the intrinsic dimension represents the number of degrees of freedom describing local structures [Zetsche and Barth, 1990]. Intrinsic zero dimensional (i0D) signals are constant signals, i1D signals are structures such lines and edges, and i2D are all other patterns in 2D - see Fig. 3.1 for an illustration of intrinsic structures. As already mentioned above, the monogenic signal is restricted to i1D signals. It is calculated with the two-dimensional Hilbert transform, which is also known as the Riesz transform. In the frequency domain, the first-order 2D Hilbert transform is obtained with the multiplication of

$$H_x^1(\mathbf{u}) = i \cdot \frac{x}{||\mathbf{u}||}, \qquad H_y^1(\mathbf{u}) = i \cdot \frac{y}{||\mathbf{u}||}, \tag{3.1}$$

with $\mathbf{u} = (x, y) \in \mathbb{C}\backslash\{(0,0)\}$.

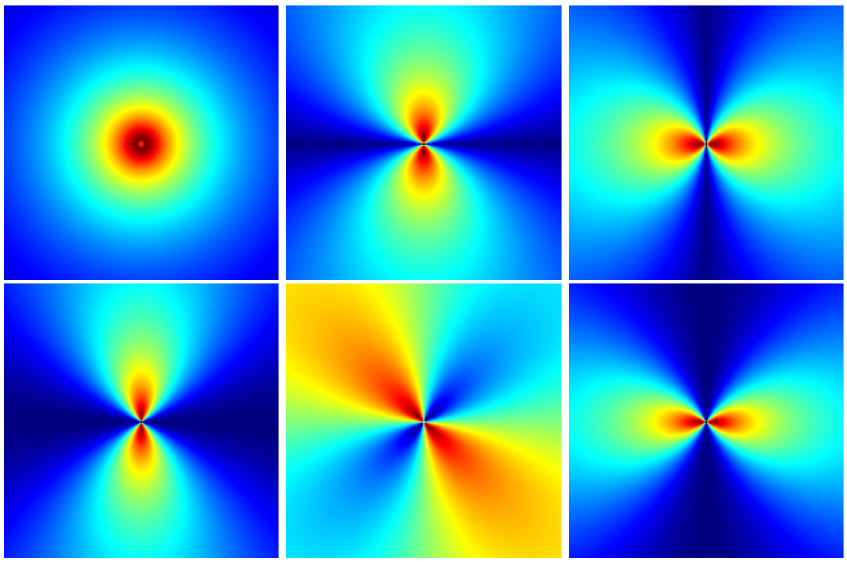


Figure 3.3: Magnitude of 2D Hilbert transforms with log-Gabor kernels in frequency domain. From left to right: B, $B \odot H^1_x$, $B \odot H^1_y$, $B \odot H^2_{xx}$, $B \odot H^2_{xy}$, $B \odot H^2_{yy}$.

For the calculation of the 2D analytic signal, higher order Hilbert transforms are used [Wietzke et al., 2009]. The Fourier multipliers of the second-order Hilbert transform are defined as

$$H^2_{xx}(\mathbf{u}) = -\frac{x \cdot x}{||\mathbf{u}||^2}, \quad H^2_{xy}(\mathbf{u}) = -\frac{x \cdot y}{||\mathbf{u}||^2}, \quad H^2_{yy}(\mathbf{u}) = -\frac{y \cdot y}{||\mathbf{u}||^2}, \tag{3.2}$$

again with $\mathbf{u} = (x, y) \in \mathbb{C}\backslash\{(0,0)\}$. In contrast to [Wietzke et al., 2009], we do not present the formulas of the Hilbert transforms in spatial domain. Rather we employ representation in frequency domain, which is more versatile and therefore favorable for applications such as filtering - see Sec. 3.2.2. Throughout this chapter we use upper case letters for filters and signals in frequency domain and lower case ones for their representation in spatial domain.

3.2.1 Structural and Geometrical Features

The proposed extension of the 2D analytic signal is obtained by an embedding in 3D projective space. This allows for a differentiation of geometrical features (local orientation, local apex) and structural features (local phase, local amplitude). The filtered signal F_p, the first-order Hilbert transformed signals F_x, F_y, and the second-order Hilbert transformed signals F_{xx}, F_{xy}, F_{yy} are calculated with the bandpass filter B and the point wise multiplication $\odot$ in frequency domain as

$$\begin{bmatrix} F_p \\ F_x \\ F_y \end{bmatrix} = \begin{bmatrix} B \odot F \\ H^1_x \odot B \odot F \\ H^1_y \odot B \odot F \end{bmatrix} \tag{3.3}$$

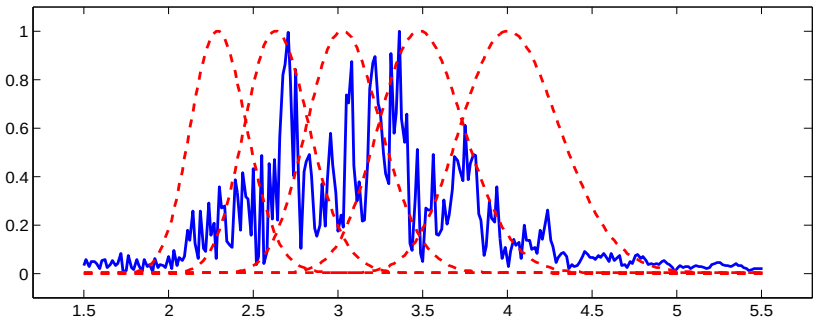


Figure 3.4: Log-Gabor filter bank consisting of 5 filters (red) and ultrasound signal spectrum (x-axis: frequency in MHz). Ultrasound acquisition frequency: 3.3 MHz.

and

$$\begin{bmatrix} F_{xx} \\ F_{xy} \\ F_{yy} \end{bmatrix} = \begin{bmatrix} H_{xx}^2 \odot B \odot F \\ H_{xy}^2 \odot B \odot F \\ H_{yy}^2 \odot B \odot F \end{bmatrix}. \tag{3.4}$$

We illustrate the Hilbert transforms in frequency domain multiplied with log-Gabor bandpass filters in Fig. 3.3. In order to enable an interpretation of second-order Hilbert transformed signals in projective space, an isomorphism between the Hesse matrix and a vector valued representation is used [Wietzke et al., 2009], leading to $f_s = \frac{1}{2}[f_{xx} + f_{yy}]$, $f_+ = f_{xy}$, and $f_{+-} = \frac{1}{2}[f_{xx} - f_{yy}]$.

Finally, the local features are calculated as follows. The apex angle α permits a differentiation between features of different intrinsic dimensionality, is defined as

$$\alpha = \arccos \frac{\sqrt{f_+^2 + f_{+-}^2}}{||f_x||}. \tag{3.5}$$

With the apex angle, the homogeneous signal component f_h of the signal f_p in projective space is defined as

$$f_h = \sqrt{\frac{1 + \cos\alpha}{2}}. \tag{3.6}$$

Furthermore, the features such as local orientation θ, local phase ϕ, and local amplitude A are calculated with

$$\theta = \frac{1}{2} \arctan \frac{f_+}{f_{+-}}, \tag{3.7}$$

$$\phi = \text{atan2}\left(\sqrt{[f_h^{-1} f_x]^2 + [f_h^{-1} f_y]^2}, f_p\right), \tag{3.8}$$

$$A = \frac{1}{2}\sqrt{f_p^2 + [f_h^{-1} f_x]^2 + [f_h^{-1} f_y]^2}. \tag{3.9}$$

In case of i1D signals, the homogeneous component simplifies to $f_h = 1$ and the formulas above reduce to the ones known from the monogenic signal.

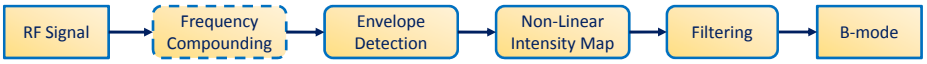


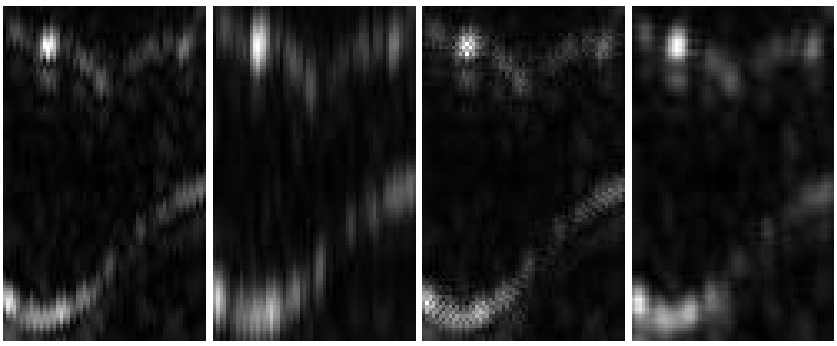
Figure 3.5: Exemplary ultrasound processing pipeline for RF to B-mode conversion.

3.2.2 Frequency Selection

Each signal f can be described with the Fourier series, decomposing the signal into components of different frequency components, each having its own phase and amplitude. The direct application of the Hilbert transform on the entire original signal at once, which represents an accumulation of local signals from different frequencies, does therefore not adequately permit an extraction of local features. Theoretically, we would need to calculate the analytic signal for infinitely narrow bandwidths, i.e., Dirac deltas in the frequency domain. However, following the uncertainty principle this results in filters with global support. Here bandpass filters can be employed as they facilitate an appropriate approximation for localization in spatial and frequency domain. [Felsberg and Sommer, 2001] apply the difference of Poisson kernels for frequency selection. An interesting property of the Poisson filter is that it creates a linear scale-space [Felsberg and Sommer, 2004]. Another filter that is commonly applied, especially in ultrasound, is the log-Gabor filter [Boukerroui et al., 2004, Grau et al., 2007, Hacihaliloglu et al., 2008, Mulet-Parada and Noble, 2000].

Also in our analysis on ultrasound images, we achieve better results with the log-Gabor filter, which is therefore employed in the following. A drawback of the log-Gabor filter is, however, that it does not permit an analytic expression in the spatial domain, which is also the reason why we present the Hilbert transforms in Eq. 3.1 and 3.2 in frequency and not in spatial domain, as it is done in [Wietzke et al., 2009].

See Fig. 3.4 for an illustration of a filter bank with five log-Gabor filters. A study of alternative bandpass filters is presented by [Boukerroui et al., 2004]. For the further analysis, it is either possible to focus on the signal at one specific scale, or accumulate all responses from various scales, as it is e.g. done for the phase congruency [Kovesi, 1999].



(a) 1D AS (b) 1D ASF (c) 2D AS (d) 2D ASF

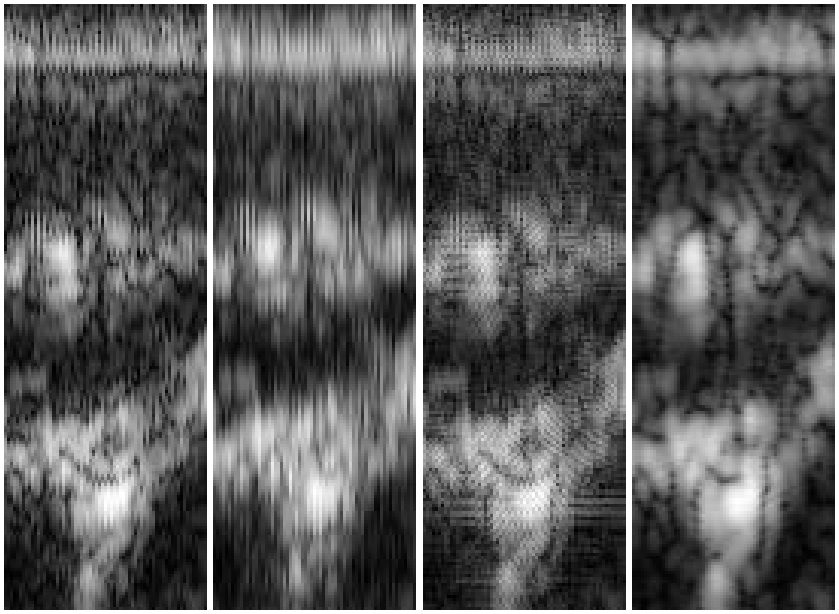
Figure 3.6: Magnified region of envelope detected 2D image for various envelopes.

3.3 2D Analytic Signal on RF Data

In the last section the concepts of the 2D analytic signal were introduced. Now these concepts are applied for estimating the envelope of ultrasound RF data. An overview of the transmission and reception principles in conventional ultrasound machines is presented in Sec. 2.3 and herein in particular Fig. 2.3 and 2.5 illustrate the processes, respectively. Important for the analytic signal generation in the RF processing pipeline (see Fig. 2.6) is the so called signal demodulation.

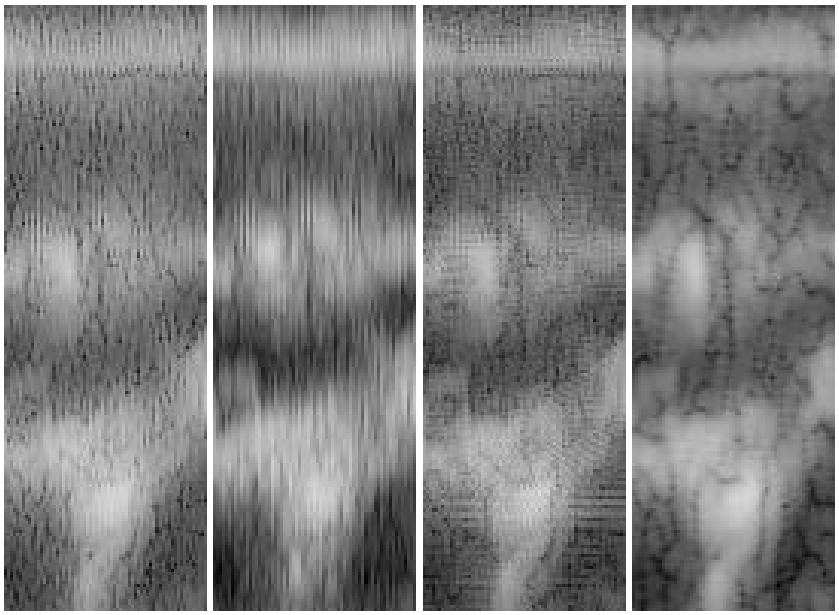
During this step the information bearing part of the signal is extracted by removal of the modulated carrier wave that is required to convey the information. In ultrasound processing, the step demodulation is commonly performed by an *envelope detection*. Hereby, the amplitude of the analytic signal is calculated for each of the 1D scan lines separately. Interestingly, calculating the amplitude of the 1D analytic signal is equivalent to the *instantaneous amplitude*. In the literature of ultrasound imaging, it is noted that the quality of ultrasound images can be increased by multi-frequency decomposition and compounding of the received signal, simply referred to as *frequency compounding* [Cincotti et al., 2001]. This is in fact equivalent to the *local amplitude* estimation. At a closer look, this constitutes an interesting analogy, between the advantages of the frequency compounded signal to the normal one. Furthermore, it bears the advantage of the local amplitude in the 2D case in comparison to the instantaneous amplitude for 1D case. We have neither seen this analogy noted in the literature before, nor the application of local amplitude and local phase techniques to RF data.

It is particularly noteworthy that in contrast to the usual approach where each scan line is processed separately, in the 2D approach all scan lines are considered at once with the resulting 2D analytic signal to estimate the local amplitude. This in fact facilitates an improved envelope detection because the signal is analyzed in its natural 2D context. In particular the information from neighboring scanlines in lateral direction is considered. As a result speckle



(a) 1D AS (b) 1D ASF (c) 2D AS (d) 2D ASF

Figure 3.7: Magnified regions of images after log-compression.



(a) 1D AS (b) 1D ASF (c) 2D AS (d) 2D ASF

Figure 3.8: Magnified regions of images after log-compression and MUCRO.

patterns are not artificially broken up and distorted between neighboring scanlines, leading to more consistent structures in the images featuring a reduced proneness in the formation of artifacts. The balance between influence from lateral and axial direction can be adjusted by the bandwidth in each direction of the bandpass filter. Parameter adaptation should consider the smaller spacing in axial direction accordingly.

3.3.1 Envelope and B-mode Results

We acquired 9 ultrasound RF datasets from different patients with an ultrasound system from Ultrasonix (Vancouver, Canada) using a linear transducer. Each dataset consists of three acquisitions. We experimented with three different acquisition frequencies, 3.3MHz, 6MHz, and 10MHz, and two different depth settings, 4 cm and 6 cm, in order to evaluate the universality of our method. The sampling frequency of the RF data is 40 MHz. We compare the envelope detection for: (i) 1D analytic signal (1D AS), (ii) 1D analytic signal with filter bank (1D ASF), (iii) monogenic signal (MS), and (iv) monogenic signal with filter bank (MSF), (v) 2D AS, and (vi) 2D ASF. Exemplarily, we show the frequency spectrum of one dataset acquired at 3.3Mhz together with the log-Gabor filter bank in Fig. 3.4. We present magnified regions of the various envelope images in Fig. 3.6. Note that we do not show the results of the MS, because the more interesting improvement is for 2D AS. However, we include them into the analysis of noise statistics in Sec. 3.3.2. We can clearly observe that the 2D analytic signal leads to a more accurate and consistent extraction of structures. This becomes particularly clear on the circular structure on the top left, which appears rather ellipsoidal on the estimates from the 1D analytic signals. We also note the positive influence of the filter bank for the estimation of the 2D analytic signal.

We perform an RF to B-mode conversion of local amplitude images A with a log-compression including a translation of 25, $\log(A + 25)$. The results for 1D ASF and 2D ASF are shown in Fig. 3.7. The B-mode image resulting from the 2D analytic signal clearly shows more consistent structures and less noise. Typically, further filtering steps are applied to the log-compressed image to improve its visual appearance. These processing steps are proprietary to the manufacturer and generally not publicly accessible. Ultrasonix, however, distributes a particular research system with a specific SDK including their post-processing filter, called MUCRO. We apply MUCRO to the log-compressed images, with the results shown in Fig. 3.8. Even after the application of MUCRO, the advantages of the images from the 2D analytic signal are clearly visible. This is not self-evident because the post-processing methods are designed to be applied to 1D envelope detected images, still leaving room for improvement by adapting the post-processing to 2D envelope estimation. Finally, one of the reasons for applying the post-processing filtering is to establish consistency between scan lines, which we already achieve by the 2D envelope detection.

3.3.2 Analysis of Envelope Statistics

Next to the visual assessment of the 2D envelope detection, we are interested in supporting the improvement by quantitative measurements. This is challenging because of the non existing ground truth envelope signal that would permit comparison with our results. As presented in Sec. 2.4, various noise distributions were derived from a theoretical analysis of speckle, which appears as if the signal is composed of a multitude of independently phased additive complex component, whose accumulation creates random walks [Dutt, 1995, Goodman, 2007]. In the following, we analyze the effects of the 2D envelope detection on the speckle statistics choosing the Nakagami model - see Sec. 2.4.2. For assessing the impact of the 2D analytic signal quantitative goodness-of-fit tests are performed. This allows, on the one hand, to evaluate the correspondence between the theoretical speckle model and the calculated envelope, and on the other hand, to show the potential for the aforementioned applications based on Nakagami modeling.

The Nakagami distribution with shape μ and scale ω parameters is defined as

$$p(x \mid \mu, \omega) = \frac{2\mu^{\mu} x^{2\mu-1}}{\Gamma(\mu)\omega^{\mu}} \exp\left(-\frac{\mu}{\omega}x^2\right), \forall x \in \mathbb{R}_+. \tag{3.10}$$

Goodness-of-Fit Test

A goodness-of-fit (GOF) test evaluates if the data $d_1, \ldots, d_n$, under the assumption of i.i.d. samples, comes from the given theoretical probability distribution p [D'Agostino and Stephens, 1986]. Note that conventional GOF tests are restricted to the case of single distributions. For inhomogeneous regions in the image, however, a mixture of Nakagami is more appropriate. Figure 3.9 illustrates a misfit of a single Nakagami to mixture Nakagami data as well as a perfect mixture fit. Consequently, we can only achieve good results with the GOF test on homogeneous image regions. The mixture case has to be further evaluated, with similar results to be expected. In order to determine to what kind of (mixture) model is most appropriate, e.g. reversible-jump Markov chain Monte Carlo can be used to infer the parameters [Green, 1995]. For the particular case of a Gamma/Nakagami mixture, a specific adaptation was proposed in [Wiper et al., 2001].

For the GOF test, the range of the data is partitioned into M bins β_i, $i = 1, \ldots, M$, with N_i and the number of samples per bin. Moore suggests to divide the data into $M = 2n^{\frac{2}{5}}$ bins [D'Agostino and Stephens, 1986]. Furthermore, we assume the bins to be equiprobable as suggested in [Bock and Krischer, 1998]. In this regard, we let p_i be the integral of the distribution in the range β_i given the parameters of the distribution $\theta = \{\mu, \omega\}$

$$p_i = \int_{\beta_i} p(x \mid \theta)\, \mathrm{d}x. \tag{3.11}$$

Hence, p_i expresses the likelihood of a sample to be in the bin β_i (identical for all bins). The test statistics underlying the GOF test is the sum of differences

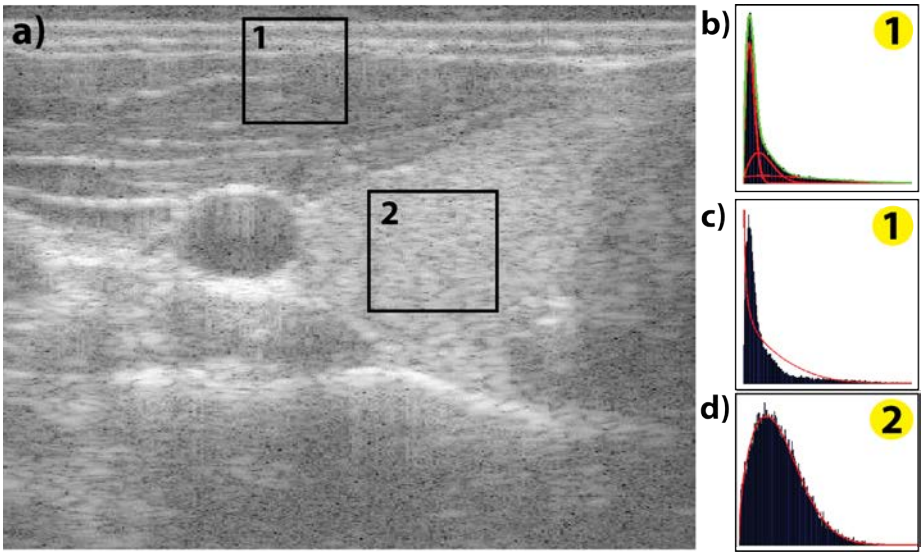


Figure 3.9: RF image with sample distributions estimated for two areas. Data in region 1 is mixture of Nakagami distributed and data of region 2 is single Nakagami distributed. Whereas MLE can fit nicely in region 2 (d) it expectedly performs poorly in region 1 (c), that can only be represented properly by mixture (b).

between observed and expected outcome frequencies

$$X^2 = \sum_{i=1}^{M} \frac{(N_i - np_i)^2}{np_i}. \tag{3.12}$$

This yields a quadratic form in N_i that has approximately a χ distribution with $M - N - 1$ degrees of freedom and $N = 2$ the number of parameters of the distribution, where [Rachev et al., 2010]

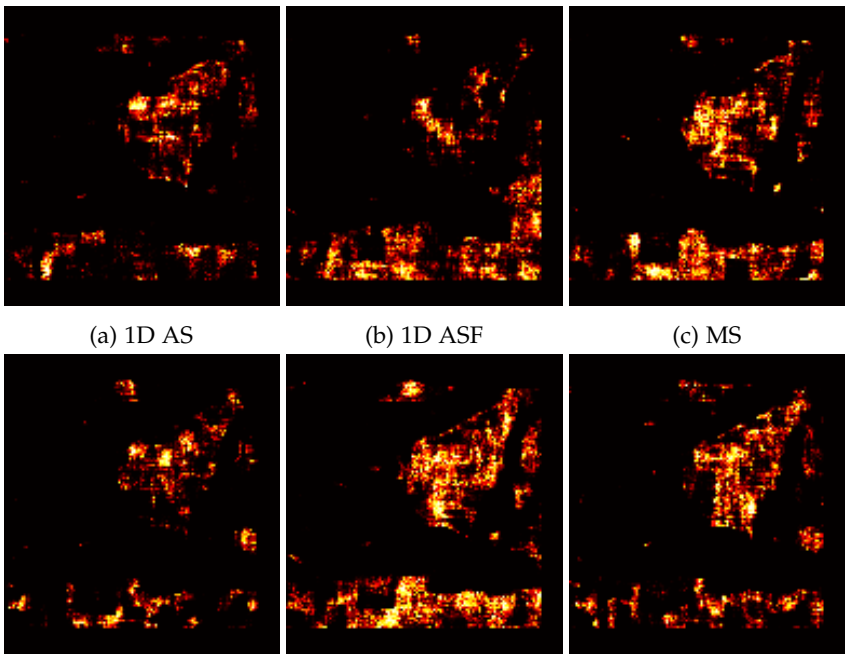
$$\chi(x|k) = \begin{cases} \frac{x^{(k/2)-1}e^{-x/2}}{2^{k/2}\Gamma\left(\frac{k}{2}\right)}, & x \leq 0 \\ 0, & \text{otherwise,} \end{cases} \tag{3.13}$$

denotes the χ distribution with k degrees of freedom and $\Gamma(.)$ corresponds to the Gamma-function.

In order to assess the GOF quantitatively, we employ the P-value based hypothesis test. The P-value serves as an indicator of how likely the null hypothesis H_0 is true. In our case, H_0 is the hypothesis that the observations are Nakagami distributed, leading to the following calculation of the P-value

$$P = \int_{X^2}^{\infty} \chi^2(M - N - 1)\, \mathrm{d}x, \tag{3.14}$$

employing Eq. 3.12 as the lower bound of integration.



(a) 1D AS (b) 1D ASF (c) MS

(d) MSF (e) 2D AS (f) 2D ASF

Figure 3.10: The P-values are calculated for all patches of an envelope image. Pixel brightness indicates P-value. We perform the calculation for various envelope detection techniques. Comparing the P-value images to the B-mode image in Fig. 3.9, we see that the bright regions correspond to homogeneous regions in the ultrasound image. The results shown correspond to patient 1, acquisition 1, and a large window size.

Rao-Robson Statistic

Given the data, we first have to estimate the parameters μ, ω of the Nakagami distribution before the GOF test is performed. This is, however, opposing the general assumption that the parameters of the distribution are a-priori given before the test is performed. Therefore, another quadratic form in N_i has to be used, with the Rao-Robson statistic being one possibility [D'Agostino and Stephens, 1986, Lin et al., 2005, Tao et al., 2006]. Considering the parametric form of the distribution $p(x \mid \theta)$ and the maximum likelihood estimate $\hat{\theta}$, the Rao-Robson statistic is

$$RR = V^{\top}(\hat{\theta})Q(\hat{\theta})V(\hat{\theta}) \tag{3.15}$$

with

$$V(\theta) = \frac{N_i - np_i}{(np_i)^{1/2}} \tag{3.16}$$

$$Q(\theta) = I + D(\theta)[J(\theta) - D^\top(\theta)D(\theta)]^{-1}D^\top(\theta) \tag{3.17}$$

$$D_{ij}(\theta) = p_i(\theta)^{\frac{1}{2}} \frac{\partial p_i(\theta)}{\partial \theta_j}. \tag{3.18}$$

Here $J(\theta)$ denotes the $N \times N$ Fisher information matrix and I is an $M \times M$ identity matrix. The partial derivative with respect to the distribution parameters $\frac{\partial p_i(\theta)}{\partial \theta_j}$ is involved, which is given below with Eq. 3.20 and Eq. 3.21. Making use of the Nakagami-Gamma relationship

$$Y \sim \mathcal{N}(x \mid \mu, \omega) \quad \Rightarrow \quad Y^2 \sim \Gamma(x \mid \mu, \frac{\omega}{\mu}), \tag{3.19}$$

the data can be transformed to follow the Gamma distribution. Hence, the required derivatives w.r.t. to the parameters can be performed on the Gamma distribution. In this regard, the Gamma distribution is defined as

$$\Gamma\left(x \mid \alpha, \beta\right) = \frac{\beta^\alpha x^{\alpha-1} \mathrm{e}^{-x\beta}}{\Gamma\left(\alpha\right)}. \tag{3.20}$$

Then the partial derivative of the binned distribution in the interval $[a_i, b_i]$ with respect to both parameters $\{\alpha, \beta\}$ is required. The derivation with respect to the first parameter α yields

$$\frac{\partial \int_{a_i}^{b_i} \Gamma\left(x \mid \alpha, \beta\right) \, \mathrm{d}x}{\partial \alpha} = \frac{(b_i\beta)^\alpha \, \mathrm{e}^{-b_i\beta} - (a_i\beta)^\alpha \, \mathrm{e}^{-a_i\beta}}{\beta \, \Gamma\left(\alpha\right)}. \tag{3.21}$$

Denoting the generalized hypergeometric function by ${}_2\mathrm{F}_2$, we obtain equation

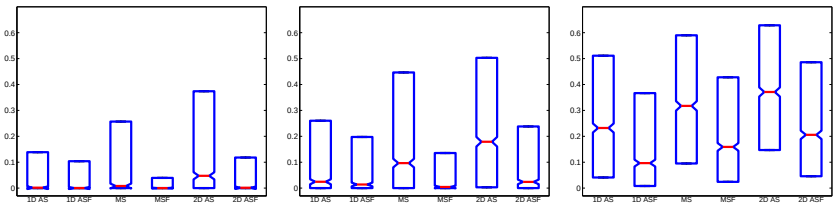
$$\begin{aligned}
\frac{\partial \int_{a_i}^{b_i} \Gamma\left(x \mid \alpha, \beta\right) \, \mathrm{d}x}{\partial \beta} &= -(\alpha^2 \Gamma\left(\alpha\right))^{-1} \left[- (a_i\beta)^\alpha \, {}_2\mathrm{F}_2(\alpha, \alpha;\, 1+\alpha, 1+\alpha;\, -a_i\beta) \right. \\
&+ (b_i\beta)^\alpha \, {}_2\mathrm{F}_2(\alpha, \alpha;\, 1+\alpha, 1+\alpha;\, -b_i\beta) + \alpha^2 \Gamma\left(\alpha, b_i\beta\right) \ln\left(b_i\beta\right) - \alpha^2 \Gamma\left(\alpha\right) \ln\left(b_i\beta\right) \\
&\left. - \alpha^2 \Gamma\left(\alpha, a_i\beta\right) \ln\left(a_i\beta\right) + \alpha^2 \Gamma(\alpha) \ln(a_i\beta) - \alpha^2 \Gamma(\alpha, b_i\beta) \Psi(\alpha) + \alpha^2 \Gamma(\alpha, a_i\beta) \Psi(\alpha) \right].
\end{aligned}$$

for the derivation with respect to β.

The Rao-Robson statistic is χ^2 distributed with $M - N - 1$ degrees of freedom, leading to P-values computed by

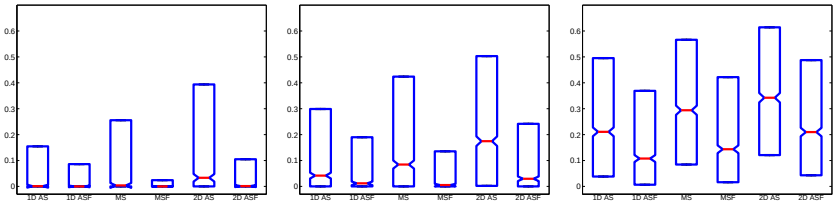
$$P = \int_{RR}^{\infty} \chi^2(M - N - 1) \, \mathrm{d}x, \tag{3.22}$$

with the Rao-Robson statistic RR as lower bound of integration.



(a) Large Window (b) Mid Window (c) Small Window

Figure 3.11: Box plot of P-values for different envelope detections and window sizes for patient 1, acquisition 1.



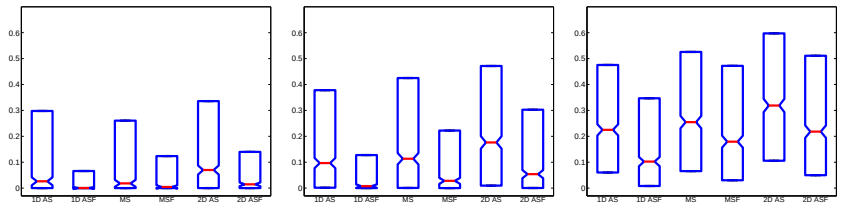
(a) Large Window (b) Mid Window (c) Small Window

Figure 3.12: Box plot of P-values for different envelope detections and window sizes for patient 1, acquisition 3.

Distribution Parameter Estimation

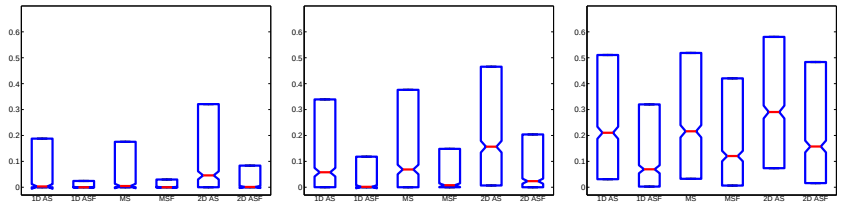
An essential step before the Rao-Robson statistics can be calculated is the fitting of the Nakgami distribution to the data, resulting in the distribution parameters θ. As mentioned in Sec. 2.4.2 there exist several methods to calculate the parameters of the Nakagami distribution. Notably, the maximum likelihood and the method of moments approach.

The distribution parameters are estimated for local windows in the images. Thereby constitutes the size of the window a trade-off between sufficient statistics for the estimation and homogeneous structures in the patch. As presented in detail previously, a mixture model fit is more appropriate for certain regions (see Fig. 3.9), whereas it is more likely to have homogeneous structures within smaller windows. Moreover, for applications such as tissue characterization and segmentation, the interest lies in working with small windows to achieve a high spatial resolution of the parameter map [Noble, 2010]. The moment-based estimation depends on local histograms, which are not good approximations to the actual distribution for small windows, and further, the estimation of the parameters is unstable. The same holds for MLE. [Larrue and Noble, 2011] therefore propose the Gamma kernel density estimation (GKDE) for Nakagami imaging in order to achieve a more stable fit for a very limited number of data points. See Fig. 2.10 for a visualization of the parameter estimation approaches and their corresponding results. The



(a) Large Window (b) Mid Window (c) Small Window

Figure 3.13: Box plot of P-values for different envelope detections and window sizes for patient 2, acquisition 3.



(a) Large Window (b) Mid Window (c) Small Window

Figure 3.14: Box plot of P-values for different envelope detections and window sizes for patient 3, acquisition 3.

usage of non-symmetric Gamma kernels such as

$$K_{x/b+1,b}(t) = \frac{t^{x/b} \cdot e^{-t/b}}{b^{x/b+1} \cdot \Gamma(x/b+1)}, \tag{3.23}$$

was shown to lead to better results for the estimation of highly asymmetric distributions such as Nakagami or Gamma [Chen, 2000]. The resulting Gamma kernel estimator for a window W containing the data $d_1, \ldots, d_l$ is

$$p(x) = \frac{1}{l} \sum_{j=1}^{l} K_{x/b+1,b}(d_j). \tag{3.24}$$

The smoothness is controlled by the parameter b, behaving similarly to the variance in Gaussian-based density estimation. [Larrue and Noble, 2011] achieved good results by setting $b = 0.05$, which performed favorably in our experiments as well.

In our analysis, we work with three different window sizes, 80×10, 60×6, and 20×3, with more pixels being considered along the axial direction. For the two larger patch sizes, we achieve similar results for MLE and GKDE with moments, so that we work with the MLE estimates. For the smallest patch size, we use GKDE with moments because it leads to more robust estimates than MLE.

Statistical Results

We perform the statistical evaluation on all 9 datasets, each one consisting of 3 different acquisitions, resulting in 27 different images. Further, we use three different window sizes in order to evaluate the dependency on the window size. This leads to 81 different configurations for the estimation, where the Rao-Robson GOF test is performed densely throughout the image for all 6 types of envelope detections. By performing the test densely throughout the image, we can create a new image with the intensity values being the P-values. We show these images in Fig. 3.10 for the various envelope detection schemes for one configuration. Brighter images indicate higher P-values, which consequently results in better statistical usability due to better fits. We note that the bright regions are corresponding to the homogeneous areas in the ultrasound image, because only these areas are appropriately modeled with a single distribution, as discussed previously.

Additionally, we calculate the statistics of the P-values. We show the box plot for four different configurations in Fig. 3.11-3.14. The red line indicates the median and the box is constructed from the interquartile range. Our results therefore show that the envelope detection without the filter bank produces better fits, which makes sense, because log-Gabor filters influence the distribution. More importantly, however, we note the improvement from 1D AS to MS, and further from MS to 2D AS. This shows on the one hand, the advantage of applying 2D Hilbert transforms in contrast to 1D ones, and on the other hand, the advantage of the 2D analytic signal in contrast to the monogenic signal. This confirms the visually improved results for 2D envelope detection from the previous section.

While the images give a good overview about the spatial distribution of the P-values and the boxplots nicely illustrates their statistics, it is difficult to show the results for all 81 configurations compactly. Consequently, we quantify the percentage of P-values that are above 0.85 for each image, corresponding to appropriate fits. The results are shown in Tab. 4.1 and Tab. 3.2. Generally, better fits are achieved for smaller window sizes. This is comprehensible, because smaller windows are more likely to contain homogeneous tissue. In the tables, we highlight the percentage of the envelope detection scheme in bold that performed best. Furthermore, the pie diagram in Fig. 3.15 illustrates the percentages of envelope detection schemes performing best. The envelope detection with the 1D AS is best in 3.7%, the MS in 2.5%, the 2D AS in 91.4%, and 2D ASF in 2.5% of the cases. This clearly shows that the application of the 2D analytic signal leads to advantages in the statistical analysis in most of the cases. Just regarding these numbers, one is tempted to assume that 1D AS performs better than MS. This is, however, not the case because the MS outperforms the 1D AS in most of the cases that 2D AS performs best.

Table 3.1: The tables shows the percentage of P-Values above 0.85 for the various images and envelope detection schemes. The highest percentage across the envelope detection techniques is marked in bold.

Patient	Acqu	Frequency	Depth	Window	1D AS	1D ASF	MS	MSF	2D AS	2D ASF
				Large	1.61	1.83	4.10	0.66	**4.76**	1.83
1	1	3.3Mhz	4cm	Medium	2.77	2.38	5.04	1.83	**6.59**	2.60
				Small	5.70	3.32	6.98	4.60	**7.25**	4.49
				Large	1.83	1.50	2.16	0.78	**4.71**	1.27
1	2	3.3Mhz	4cm	Medium	2.60	1.77	4.60	1.38	**6.37**	2.60
				Small	4.60	3.38	6.48	4.37	**7.81**	5.76
				Large	1.38	2.38	2.99	0.83	**4.37**	1.38
1	3	3.3Mhz	4cm	Medium	2.60	1.77	4.60	1.38	**6.37**	2.60
				Small	4.76	2.38	6.92	4.26	**7.14**	4.49
				Large	2.77	0.78	3.32	1.11	**4.32**	2.55
2	1	6Mhz	5cm	Medium	3.10	1.11	4.43	1.66	**5.43**	2.21
				Small	**6.98**	4.32	5.43	4.76	5.32	5.54
				Large	3.10	1.66	2.66	1.55	**3.88**	1.66
2	2	6Mhz	5cm	Medium	3.43	1.00	4.54	1.55	**4.65**	3.21
				Small	6.31	3.65	3.99	4.87	**7.75**	4.43
				Large	**4.21**	0.78	2.44	1.22	3.54	1.33
2	3	6Mhz	5cm	Medium	2.88	1.77	3.77	2.88	**6.31**	3.10
				Small	4.98	3.54	4.65	5.43	**6.53**	5.98
				Large	3.29	0.68	3.00	0.78	**3.59**	1.36
3	1	10Mhz	4cm	Medium	5.04	1.55	4.65	0.58	**6.40**	1.94
				Small	3.29	3.39	5.52	3.88	**5.91**	3.68
				Large	3.00	0.39	2.33	0.87	**3.39**	0.58
3	2	10Mhz	4cm	Medium	3.59	1.84	3.88	1.74	**5.91**	1.16
				Small	6.30	2.81	4.46	3.59	**6.69**	3.88
				Large	2.33	0.87	1.74	0.97	**3.29**	0.87
3	3	10Mhz	4cm	Medium	3.78	1.16	4.26	2.03	**6.30**	2.52
				Small	4.17	2.91	5.04	4.17	**7.17**	3.78
				Large	1.74	0.10	1.26	0.68	**1.94**	0.58
4	1	6Mhz	5cm	Medium	2.62	0.78	2.62	1.36	**4.26**	1.65
				Small	3.88	3.00	4.46	4.65	**5.43**	4.26
				Large	1.26	0.29	1.55	0.78	**3.00**	1.45
4	2	6Mhz	5cm	Medium	2.81	1.74	2.91	2.13	**3.59**	1.84
				Small	4.26	3.10	5.23	3.59	**5.91**	3.10
				Large	1.55	0.39	1.45	0.58	**2.62**	0.68
4	3	6Mhz	5cm	Medium	2.52	0.48	4.07	1.55	**5.33**	1.65
				Small	4.94	2.62	4.55	4.84	4.65	**5.23**
				Large	0.97	0.19	1.55	0.48	**2.23**	0.87
5	1	10Mhz	4cm	Medium	2.23	1.16	1.94	1.26	**4.94**	1.65
				Small	4.46	2.23	4.65	3.49	**5.91**	3.59
				Large	1.55	0.39	2.13	0.87	**2.23**	0.87
5	2	10Mhz	4cm	Medium	2.03	0.87	2.33	0.29	**5.23**	1.55
				Small	5.04	3.59	**6.10**	4.55	4.07	3.68

Table 3.2: The tables shows the percentage of P-Values above 0.85 for the various images and envelope detection schemes. The highest percentage across the envelope detection techniques is marked in bold.

Patient	Acqu	Frequency	Depth	Window	1D AS	1D ASF	MS	MSF	2D AS	2D ASF
				Large	0.97	0.19	1.07	0.58	**2.42**	0.48
5	3	10Mhz	4cm	Medium	2.62	0.97	3.00	1.07	**3.68**	1.36
				Small	3.88	2.23	4.07	3.29	**5.62**	3.88
				Large	2.13	0.68	2.62	1.55	**3.68**	1.84
6	1	6Mhz	4cm	Medium	4.55	1.74	4.46	1.65	**5.72**	3.39
				Small	5.43	2.71	5.23	4.17	4.75	**5.52**
				Large	3.49	0.78	2.81	2.13	**4.65**	1.65
6	2	6Mhz	4cm	Medium	4.84	1.36	4.84	1.84	**6.30**	2.23
				Small	5.81	3.49	5.23	4.17	**6.69**	4.65
				Large	2.91	0.78	3.20	1.45	**3.78**	1.74
6	3	6Mhz	4cm	Medium	4.26	1.26	4.07	2.42	**5.72**	2.71
				Small	**6.20**	3.10	5.91	4.75	6.10	5.04
				Large	2.62	0.58	3.59	0.78	**4.36**	1.94
7	1	10Mhz	4cm	Medium	3.88	1.84	4.36	1.55	**5.04**	2.91
				Small	5.04	3.10	5.62	4.46	**6.10**	5.04
				Large	3.20	0.87	3.49	1.45	**3.88**	1.07
7	2	10Mhz	4cm	Medium	4.55	1.84	4.07	1.45	**5.33**	2.42
				Small	5.43	3.00	5.43	4.26	**6.59**	3.88
				Large	2.62	0.78	3.00	0.87	**4.46**	2.33
7	3	10Mhz	4cm	Medium	3.49	1.45	5.04	1.55	**5.62**	2.13
				Small	3.88	2.91	**5.91**	4.75	5.72	5.14
				Large	2.23	0.68	2.52	0.78	**3.97**	3.00
8	1	6Mhz	5cm	Medium	3.00	1.36	3.59	2.42	**5.72**	3.39
				Small	4.94	2.33	4.84	5.14	**6.10**	4.17
				Large	2.13	0.68	2.71	1.16	**4.65**	1.16
8	2	6Mhz	5cm	Medium	3.59	0.87	4.36	2.52	**5.43**	3.59
				Small	5.72	3.20	5.33	5.04	**6.20**	3.78
				Large	2.81	0.48	3.29	1.74	**3.97**	0.97
8	3	6Mhz	5cm	Medium	3.29	1.36	3.97	2.33	**5.81**	2.81
				Small	4.36	3.10	5.72	4.75	**6.01**	3.29
				Large	1.45	0.48	2.23	0.78	**2.52**	0.78
9	1	10Mhz	5cm	Medium	2.03	0.97	2.91	0.78	**3.29**	1.84
				Small	3.39	2.71	5.33	4.55	**6.59**	4.84
				Large	0.78	0.97	2.23	0.58	**3.29**	0.78
9	2	10Mhz	5cm	Medium	3.10	1.65	3.59	1.07	**4.75**	1.36
				Small	3.49	3.00	5.14	3.20	**6.20**	3.97
				Large	1.16	0.48	1.36	0.29	**2.62**	1.26
9	3	10Mhz	5cm	Medium	2.03	0.68	2.42	1.55	**4.75**	2.13
				Small	3.78	3.20	4.75	4.36	**6.20**	4.36

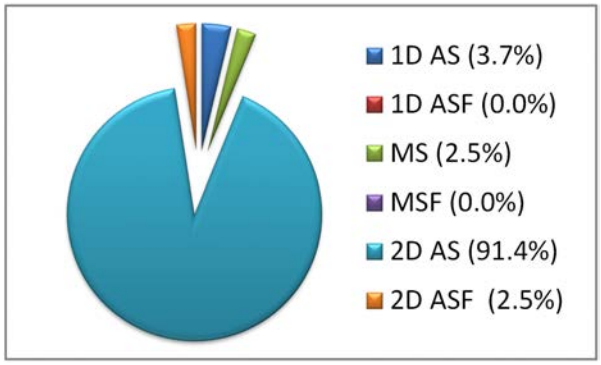


Figure 3.15: Pie diagram showing the percentages that envelope detection techniques performed best in tables 4.1 and 3.2. The envelope detection with the 2D analytic signal performed best in 91.4% of the cases.

3.4 2D Analytic Signal on B-mode Images

Next to the benefits of the 2D analytic signal for the demodulation of RF data, it also allows for a more accurate estimation of local features on B-mode images [Wietzke et al., 2009]. This has the potential to increase the quality of follow-up applications such as registration [Grau et al., 2007, Mellor and Brady, 2005, Zhang et al., 2007], segmentation [Hacihaliloglu et al., 2008], and detection [Mulet-Parada and Noble, 2000], which use the local features as input. To demonstrate this, we calculate the local orientation on B-mode images showing a biopsy needle. In Fig. 3.16, we illustrate the local orientation that is estimated from the monogenic signal and the 2D analytic signal, both with filtering. The estimation from the monogenic signal shows no consistent orientation information in the region of the needle. In contrast, the improved concept of the 2D analytic signal indicates a consistent result.

3.5 Conclusion

We demonstrated that the application of the 2D analytic signal leads to multiple advantageous properties in RF as well as B-mode data. The demodulation of RF signals with the 2D analytic signal enables a more consistent extraction of structures. This can be attributed to the fact that the signal is analyzed in its natural 2D context. We further showed that the improved envelope detection enables the creation of B-mode images of enhanced quality. To validate this, we applied a proprietary post-processing filtering for ultrasound on the log-compressed images and compared the result of 1D and 2D analytic signal. Additionally, we illustrated the improved statistical properties of envelope data resulting from the 2D analytic signal by performing goodness-of-fit tests w.r.t. to the Nakagami distributional model.

Finally, the advanced signal model of the 2D analytic signal leads to benefits in the estimation of local features in B-mode images, as we have illustrated for the case of needle detection.

For the demodulation, we focused on scans from a linear transducer. For curved linear transducers, the application of 2D Hilbert transforms without a previous scan conversion can be achieved with the polar Fourier transform [Averbuch et al., 2006]. This, together with the incorporation of mixture models in the statistical analysis, remains as future work.

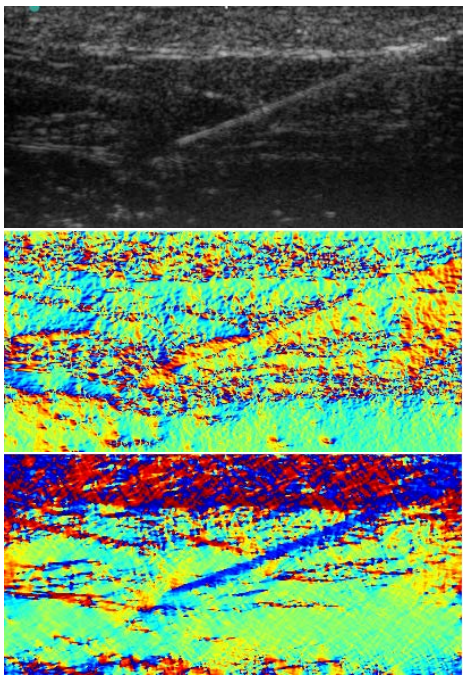


Figure 3.16: Ultrasound image with biopsy needle (top). Calculated local orientation for monogenic signal (middle) and 2D analytic signal (bottom).

Chapter 4

2D Texture Modeling

Modeling of texture seems to be attractive way of modeling objects in images as this beside shape is the main property that characterizes objects and therefore seems most natural to us. In this chapter we give an overview about methods of texture modeling with a particular focus on Markov Random Field texture models and Local Binary Patterns. General models are then presented in the specific context of ultrasound data together with sample applications such as texture learning and shadow detection. Within this chapter we also explain the Expectation-Maximization (EM) algorithm, which is used in various adaptations in this thesis. However, in context of this chapter it is used for the detection of shadow regions. This chapter is based on the papers 'Spatial Statistics Based Feature Descriptor For RF Ultrasound Data' [Klein et al., 2011] and 'Shadow Detection in Ultrasound RF Data' [Klein et al.,].

4.1 Introduction

Texture is typically referred to the small scale detail of objects of interest, although there seems to be no generally agreed upon definition [Petrou and Sevilla, 2006]. More formally, it can also be considered as a function, which exhibits spatial variation in pixel intensities or color. Despite this somehow vague definition, textures can be divided into two classes. Broadly speaking, those classes are based on the inherent texture regularity: statistical textures and regular structured textures [Paget, 2009]. See Fig. 4.1 for an illustration of this two types of textures.

Statistical texture modeling tries to model patterns by means of a probability function, which can be either parametric or non-parametric. Whereas structured textures are more or less composed of repetitive structures or texture elements. Those structured elements are often referred to as texels. Modeling of textures plays an important role in computer vision and image processing applications. Particular for natural scene images, where objects often exhibit a certain structure or pattern, approaches modeling textures are widely applicable. Among the major application in texture processing are classification, segmentation and synthesis. In the domain of ultrasound imagery, textures are a particular interesting source of information given the general lack of rich

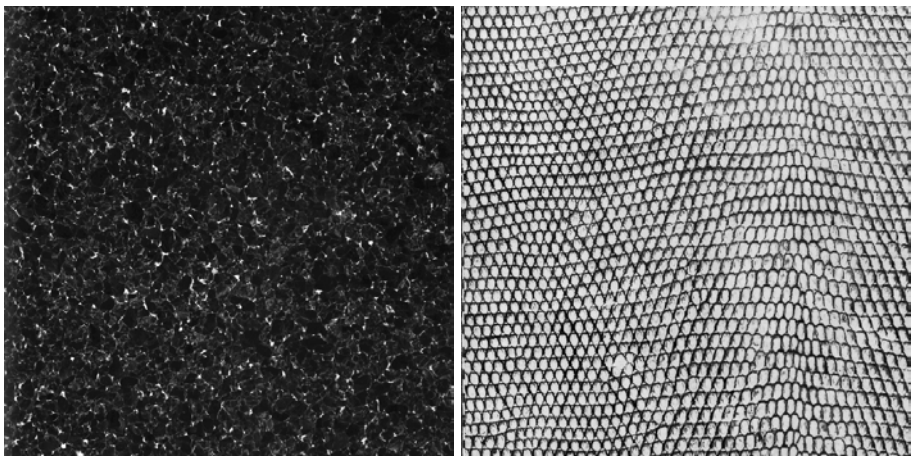


Figure 4.1: Different textures from the Brodatz database. Left: a stochastic texture. Right: a regular and structured texture

features in this noisy modality. As mentioned in detail in Sec. 2.2, different tissues in the body exhibit specific backscatter patterns due to characteristic micro-scale inhomogeneities. Modeling those patterns is in turn employed for applications such as tissue classification and segmentation. In this respect, it is often interesting to discriminate healthy from non-healthy tissue, as backscatter properties are often subject to change induced by an illness [Joynt, 1979].
Focus of this chapter is on statistical methods as those facilitate to describe both the image and disturbances data has been subject to [Lindgren, 2002]. This is in particular important for ultrasound imagery with is noise-like image acquisition artifacts. In this chapter we provide at first an overview about Markov Random Field textures models in Sec. 4.3 with an example machine learning approach, in which pattern (dis-)similarity is learned in Sec. 4.4. This is followed by a first introduction of the Local Binary Pattern (LBP) texture descriptor in Sec. 4.5. Introduction of ultrasound specific LBPs, however, is deferred to Ch. 5, as there they are used in context of a similarity metric. Finally, in Sec. 4.6 we present a shadow detection application, in which shadow regions due to total reflectors containing mirroring and reverberations artifacts are modeled, employing among others a Nakagami MRF texture model.

4.2 Related Work

Due to for ultrasound images comparably high feature richness and highly pronounced distinctive patterns, liver has been studied extensively in the domain of ultrasound texture analysis. However, most of the research activity addresses texture analysis in B-Mode rather than RF data. This can partly be attributed to the added computational cost that comes when working with large volume RF data as well as the fact that only relatively few vendors so far have granted unlimited access to RF signal data [Noble, 2010]. For the sake of readability, in the following, we assume ultrasound data to be B-mode if not stated otherwise. In this respect, in [Wu et al., 1992] classification between normal and diffuse liver diseases is performed based on multi resolution fractal features in ultrasound data. This ties together the concept of multiple resolution with fractional Brownian motion model. The former describing the relationships among texture and contextual tone information, and the latter serving as a measure of roughness or granularity. In this respect, the Brownian motion model, which was originally proposed in [Mandelbrot, 1983], regards naturally occurring rough surfaces as product of random walks, thus resembling the ultrasound speckle model. For the purpose of discrimination between normal and fatty liver tissue, in [Huang et al., 2008] ultrasound textures are analyzed using discrete wavelet transformation (DWT) with classification based on probabilistic neural networks. Similarly, in [Ahmadian et al., 2004, Ahmadian et al., 2005] liver tissue classification employing Gabor Wavelets is proposed. Also making use of Gabor filters, in [Zhan and Shen, 2006] a set of Gabor-support vector machines is used to discriminate voxels around the surface of deformable model into prostate or non-prostate in 3D ultrasound data. Thereby each Gabor-support vector machine consists of a Gabor filter bank for description of rotation-invariant texture features and a kernel support vector machine for texture classification. In [Bleck et al., 1996] autoregressive periodic random field models are used in order to discriminate between liver with and without microfocal lesions. Using a fuzzy variant of local binary patterns as feature descriptor, in [Iakovidis et al., 2010] areas in ultrasound images of the thyroid are used to detect nodules of high malignancy risk employing a support vector machine.

Making use of RF data, in [Bouhlel and Sevestre-Ghalila, 2009] a Markov Random Field texture model is used to discriminate the area inside and outside of choroidal malignant melanoma. In [Moradi et al., 2010] tissue typing on a multitude of animal tissues is performed making us of spectral and fractal features of ultrasound RF time series.

4.3 Markov Random Field Model

A popular technique of texture modeling is the representation as a *Markov Random Field* (MRF), especially in the domain of synthetic aperture radar (SAR) image analysis they enjoy wide-spread use. The underlying assumption for those models thereby is that the value of a single pixel is conditionally dependent on the pixels in its neighborhood. Thereby is the inherent complexity of the system regulated by the size of neighborhood. Theoretically each pixel could depend on all other pixels in the image domain resulting in maximum interaction or with dependence on no other pixel leading to no interaction. Due to its structure MRFs are optimal for modeling homogeneous textures. However, beside for modeling of texture [Hassner and Sklansky, 1980, Cross and Jain, 1983, Derin and Elliott, 1987], MRFs have been used for a large spectrum of applications such as image restoration [Geman and Geman, 1984], texture segmentation [Derin and Elliott, 1987, Hu and Fahmy, 1992] as well as for texture classification [Schroder et al., 1998, Serpico and Moser, 2006, Dellepiane, ,Farag et al., 2005, Bouhlel and Sevestre-Ghalila, 2009, Chellappa and Chatterjee, 1985]. For MRF texture modeling as well as synthesizing there exist numerous approaches for formalizing the underlying distributions, ranging from parametric [Bouhlel et al., 2004, Bouhlel and Sevestre-Ghalila, 2009, Chellappa and Chatterjee, 1985] to non-parametric [Paget, 2004]. Inherent to all of them is the idea of combining the notion of distribution and neighborhood dependency.

4.3.1 Random Field Preliminaries

Following [Ancona et al., 1990] an image is represented as a lattice $\mathcal{S} = \{s_1, ..., s_N\}$ of N sites s_i (= pixel position). Each site $s \in \mathcal{S}$ is associated with a random variable X_s, which gets assigned a value (=pixel intensity) x_s from the state space at the site Λ_s. For simplicity the joint label association $\{X_1 = x_1, ..., X_N = x_n\}$ is abbreviated as $X = x$. The family of random variables $\{X_1, ..., X_N\}$ is referred to as *Random Field*. For the entire lattice we obtain the configuration space Ω

$$\Omega = \prod_{s \in \mathcal{S}} \Lambda_s \quad \text{s.t.} \quad \Lambda_s \subset \mathbb{R}, \tag{4.1}$$

representing all possible realization of the random field.

However, in image processing one commonly assumes a simplified state space that is identical at all sites with elements corresponding to the available gray-values e.g. $\Lambda = \{0, ..., 255\}$ for 8-bit intensity range. As a consequence we yield the state space $\Omega = \Lambda^N$. Furthermore, we define a probability measure Π on Ω with $\Pi(X = x) > 0, \forall x \in \Omega$, which measures the probability associated with a specific configuration x. The corresponding conditional probability density function incorporating dependency on the neighbors is thereby defined as

$$\Pi_i(x_i | x_{\mathcal{S}-\{i\}}) = \Pi(X_i = x_i | X_s = x_s, s \neq i), \quad s \in \mathcal{S}, x \in \Omega, \tag{4.2}$$

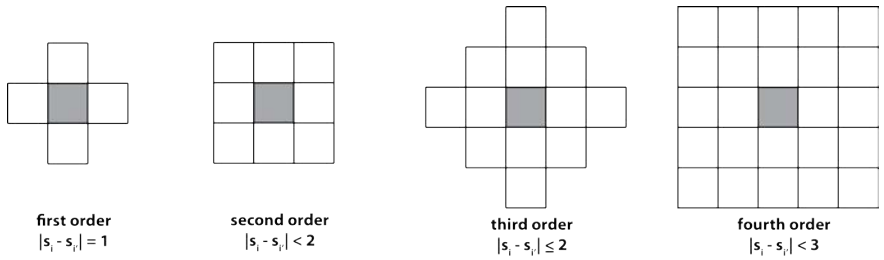


Figure 4.2: Neighborhood illustration of various orders ranging from first to fourth order.

where $\mathcal{S} - \{i\}$ denotes the set difference. Put in other words, $x_{\mathcal{S}-\{i\}}$ represents the set of all labels at sites $\mathcal{S} - \{i\}$, which is all sites in $\mathcal{S}$ except i.

The sites in $\mathcal{S}$ are all related to each other w.r.t. a neighborhood system $\mathcal{V}$. In this regard, the neighborhood system for the entire lattice is defined as

$$\mathcal{V} = \{\mathcal{V}_i | i \in \mathcal{S}\}, \tag{4.3}$$

where $\mathcal{V}_i$ represents the neighboring sites of i and for a regular lattice is defined as

$$\mathcal{V}^r = \left\{i' \in \mathcal{S} | 0 \leq |s_i - s_{i'}| \leq r, i \neq i'\right\}. \tag{4.4}$$

Here $|.|$ denotes the Euclidean distance and r some integer-valued distance threshold, defining the order of the neighborhood. The notion of neighboring induces two properties: (1) a site cannot be a neighbor to itself: $i \notin \mathcal{V}$. (2) the neighboring relationship is always mutual: $i \in \mathcal{V}_j \Leftrightarrow j \in \mathcal{V}_i$. Two common neighborhood system are the *first-order neighborhood*, consisting of 4 nearest neighboring sites (2 vertical, 2 horizontal), and the *second-order neighborhood*, consisting of 8 neighboring sites (2 vertical, 2 horizontal, 4 vertical). See Fig. 4.2 and 4.3 for a an illustration of various neighborhood systems.

5	4	3	4	5	β_{11}	β_7	β_6	β_8	β_{12}
4	2	1	2	4	β_9	β_3	β_2	β_4	β_{10}
3	1	s	1	3	β_5	β_1	α	β_1	β_5
4	2	1	2	4	β_{10}	β_4	β_2	β_3	β_9
5	4	3	4	5	β_{12}	β_8	β_6	β_7	β_{11}

Figure 4.3: Neighborhood structures. Left: the hierarchical neighborhood. Right: Parametrical neighborhood for auto-models.

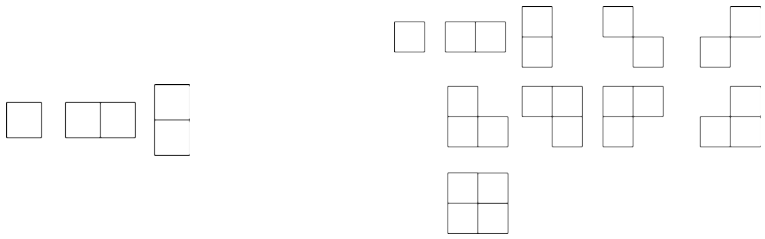


Figure 4.4: Cliques for different neighborhood systems. Left: first order neighborhood cliques. Right: second order neighborhood cliques.

Given the graph-like structure of $(\mathcal{S}, \mathcal{V})$, the neighborhood system can be broken down in subcomponents referred to as *cliques* $\mathcal{C}$, such that $\mathcal{C} \subseteq \mathcal{S}$ with distinct points in the cliques being neighbors. Put another way, a clique defines a subset of points that stands in relationship via a neighborhood system - see Fig 4.4. for an illustration of cliques. More formally, given $i, i' \in \mathcal{C}, i \neq i' \Longrightarrow i' \in \mathcal{V}_i$. In this respect, one distinguishes the cliques by the number of points that simultaneously share the neighborhood relation. That is a single site clique $\mathcal{C}_1$ and pair-site $\mathcal{C}_2$ are defined as

$$\mathcal{C}_1 = \{i | i \in \mathcal{S}\} \tag{4.5}$$

$$\mathcal{C}_2 = \left\{\left\{i, i'\right\} | i' \in \mathcal{V}, i \in \mathcal{S}\right\}. \tag{4.6}$$

Triple-site $\mathcal{C}_3$ cliques and high orders are defined analogously by extending the pair-site case. It should be noted that the number of possible cliques grows exponentially with the order of the neighborhood system.

4.3.2 Markov Random Fields

Given the definition of random fields and neighborhood system in the previous subsection, we are now able to define a Markov Random Field (MRF). Let $X = \{X_1, ..., X_N\}$ be a random field. Then X is said to be a Markov Random Field on $\mathcal{S}$ w.r.t. a neighborhood system $\mathcal{V}$, if the following two conditions are satisfied:

$$\Pi(y) > 0, \quad \forall y \in \Omega \quad \textit{(positivity condition)} \tag{4.7}$$

$$\Pi(y_i | y_{\mathcal{S}-\{i\}}) = \Pi(y_i | y_{\mathcal{V}_i}) \quad \textit{(Markovianity)} \tag{4.8}$$

The positivity condition arises from technical reasons and requires that any configuration must have a non-zero probability. If this condition is fulfilled it was shown in [Besag, 1974] that any random field is uniquely determined by local conditional probability distribution. The Markov property describes the locality of conditional dependence that is the interaction range is limited by the neighborhood.

4.3.3 Gibbs Random Fields

Let $X = \{X_1, ..., X_N\}$ be a random field. Then X is said to be a Gibbs Random Field (GRF) on $\mathcal{S}$ w.r.t. a neighborhood system $\mathcal{V}$, if and only if its configurations follow a *Gibbs* distribution, which has the following form

$$\Pi(x) = Z^{-1} \cdot \exp\left(-\frac{1}{T}U(x)\right) \tag{4.9}$$

$$Z = \sum_{x \in \Omega} \exp\left(-\frac{1}{T}U(x)\right) \tag{4.10}$$

with $\{X = x\}$ being a particular realization, T a constant called temperature, Z normalization constant/particion function. Furthermore, the energy term U is defined as

$$U(x) = -\sum_{c \in \mathcal{C}} V_c(x), \tag{4.11}$$

where V_c is referred to as the clique potential w.r.t. to clique $c \in \mathcal{C}$ and thereby depends on the local configuration. Often case a homogeneous GRF is assumed, then the $V_c(.)$ are independent of the relative position of the clique c in $\mathcal{S}$, such that the energy term is split a up w.r.t. to the size of the cliques e.g.

$$U(x) = \sum_{\{i\} \in \mathcal{C}_1} V_1(x_i) + \sum_{\{i,i'\} \in \mathcal{C}_2} V_2(x_i, x_{i'}) + \sum_{\{i,i',i''\} \in \mathcal{C}_3} V_2(x_i, x_{i'}, x_{i''}) + ... \tag{4.12}$$

It should be noted that due to the large configuration space, even for small random fields, the computation of normalization factor Z is typically intractable both analytically and numerically. However, there exist several approach that allow an approximation of Z.

The equivalence between MRFs and GRFs, also known as the Hammersley-Clifford theorem, was originally established by Hammersley and Clifford although their proof was never published. However, there exists a multitude of alternative proofs that can be found in [Besag, 1974, Grimmett, 1973, Ancona et al., 1990, Moussouris, 1974]. The Hammersley-Clifford theorem states that the family of random variables X is a MRF on the lattice $\mathcal{S}$ w.r.t. to the neighborhood system $\mathcal{V}$, if and only if X is a GRF on the lattice $\mathcal{S}$ w.r.t. to the neighborhood system $\mathcal{V}$ [Li, 2009]. Practical consequence of the theorem is that one can specify the joint probability $\Pi(X = x)$ of a random field with an appropriate energy function and potential functions $V_i(.)$, thereby encoding desired behavior [Li, 2009].

4.3.4 Auto-Models

A particular MRF model arises when assuming energy formulation of a configuration, where dependency is limited to unary and pairwise potentials. This reduces the general formulation of Eq. 4.12 to

$$U(x) = \sum_{\{i\} \in \mathcal{C}_1} V_1(x_i) + \sum_{\{i,i'\} \in \mathcal{C}_2} V_2(x_i, x_{i'}) = \sum_{i \in \mathcal{S}} V_1(x_i) + \sum_{i \in \mathcal{S}} \sum_{i' \in \mathcal{V}_i} V_2(x_i, x_{i'}). \tag{4.13}$$

Further, assuming that the conditional distribution of the MRF $\Pi(y_i|y_{\mathcal{S}-\{i\}})$ now belongs to the exponential family, then the corresponding log-likelihood can be decomposed as [Besag, 1974],

$$\Pi(y_i|y_{\mathcal{V}_i}) = A_i(y_{\mathcal{V}_i})B_i(y_i) + C_i(y_i) + D_i(y_{\mathcal{V}_i}) \tag{4.14}$$

Moreover, then it is possible to write $A_i(y_{\mathcal{V}_i})$ as weighted sum of neighboor intensities defined as

$$A_i(y_{\mathcal{V}_i}) = \alpha_i + \sum_{i \neq i'} \beta_{i,i'} B_{i'}(y_{i'}) \quad \text{s.t.} \quad \forall i, i' \in \mathcal{S}, i \neq i'. \tag{4.15}$$

Furthermore, the unary $\phi_i(.)$ and pairwise $\phi_{i,i'}(.,.)$ dependency functions can be written as

$$\phi_i(y_i) = \alpha_i B_i(y_i) + C_i(y_i) \tag{4.16}$$

$$\phi_{i,i'}(y_i, y_{i'}) = \beta_{i,i'} B_i(y_i) B_{i'}(y_{i'}), \tag{4.17}$$

where α_i and $\beta_{i,i'}$ are specific constants reflecting the single site influence and the interaction relationship between neighboring pixels, respectively. See Fig. 4.3 for the distribution of the parameters on a grid reflecting the different spatial relationships. Intuitively one can think of every pixel in the texture as a random draw from a distribution whereby the distribution changes spatially governed by the neighboring pixels. In particular it is the case that pair-wise interaction parameters have positive values when the respective neighbors and the central site tend to have similar intensities, otherwise they are negative. Additionally, given a texture pattern one can instantiate the MRF texture model on it and use obtained parameters for classification of the texture. Given the interaction coefficients, the energy formulation of Eq. 4.13 can be written as

$$U(x) = \sum_{\{i\} \in \mathcal{C}_1} y_i G_i(y_i) + \sum_{\{i,i'\} \in \mathcal{C}_2} G_{i,i'}(y_i, y_{i'}) y_i y_{i'} = \tag{4.18}$$

$$\sum_{i \in \mathcal{S}} y_i G_i(y_i) + \sum_{i \in \mathcal{S}} \sum_{i' \in \mathcal{V}_i} G_{i,i'}(y_i, y_{i'}) y_i y_{i'}, \tag{4.19}$$

where G_i and $G_{i,i'}$ are functions for the unary and pairwise energy term that are specific for the underlying distributional model. This altogether is referred to as the *Auto-Model* as porposed in [Besag, 1974]. Several auto-models have been developed, among the prominent ones is the *auto-binomial* [Cross and Jain, 1983] defined as

$$\Pi(y_i|y_{\mathcal{V}_i}) = \binom{M-1}{y_i} q^{y_i} (1-q)^{M-1-y_i} \quad \text{s.t.} \quad q = \frac{e^{\alpha_i + \sum_{i' \in \mathcal{V}_i} \beta_{i,i'} y_{i'}}}{1 + e^{\alpha_i + \sum_{i' \in \mathcal{V}_i} \beta_{i,i'} y_{i'}}}, \tag{4.20}$$

with corresponding energy term

$$U(y) = \sum_{\{i\} \in \mathcal{C}_1} \ln \binom{M-1}{y_i} - \sum_{\{i\} \in \mathcal{C}_1} \alpha_i y_i - \sum_{\{i,i'\} \in \mathcal{C}_2} \beta_{i,i'} y_i y_{i'} \tag{4.21}$$

assuming that $y_i \in \{0, 1, ..., M-1\}$. A further frequently used MRF is the *Gaussian auto-model* also known as *auto-normal model* [Chellappa and Chatterjee, 1985, Cohen and Cooper, 1987] with

$$\Pi(y_i|y_{\mathcal{V}_i}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}\left[y_i - \alpha_i - \sum_{i' \in \mathcal{V}_i} \beta_{i,i'}(y_{i'} - \alpha_{i'})\right]^2\right) \tag{4.22}$$

and the associated conditional expectation

$$E\left[y_i|y_{\mathcal{V}_i}\right] = y_i - \alpha_i - \sum_{i' \in \mathcal{V}_i} \beta_{i,i'}(y_{i'} - \alpha_{i'}). \tag{4.23}$$

There exist a multitude of further auto-models such as *Auto–Poisson*, *Auto–Exponential* and so on. Whereas the above mentioned models apply to general imaging, an ultrasound specific auto-model was devised in [Bouhlel et al., 2004, Bouhlel and Sevestre-Ghalila, 2009], namely the *Auto-K* and the *Auto-Nakagami* model. In the following we will elaborate more on the Auto-Nakagami model.

4.3.5 Auto-Nakagami

As was described in Sec. 2.2 and 2.4 backscatter intensities within an organ are of stochastic nature, however, follow a certain distribution. The resulting speckle pattern texture give rise to modeling those properties. In particular the analysis of spatial interaction relationship within a neighborhood underlying the MRF models seems particularly appealing for ultrasound. Assuming the enveloped RF data Y to be Nakagami distributed $Y \sim \mathcal{N}(x \mid m_{nak}, \omega_{nak})$, it follows from the application of the random variable transformation that $Y^2 = X \sim \mathcal{GA}(x \mid \mu_{gam}, \omega_{gam})$ is Gamma distributed (see Eq. 2.53 for the Nakagami-Gamma relationship). For sake of convenience we will assume the Gamma model, knowing that it is equivalent to Nakagami by transformation. This allows for a series of simplifications not least the availability of Gamma related functionality in most mathematical software programs. Given the transformed data, then the log-likelihood of the Gamma is defined as

$$\begin{aligned}\log \mathcal{GA}(x \mid \mu_{gam}, \omega_{gam}) = \mu_{gam} \log(\omega_{gam}) + \\ (\mu_{gam} - 1)\log(x) - \log(\Gamma(\mu_{gam})) - (\omega_{gam} x).\end{aligned} \tag{4.24}$$

Applying the log-decomposition theorem (see Eq. 4.14) we yield the following dependencies

$$A_i(x_{\mathcal{V}_i}) = \mu_{gam} - 1 \tag{4.25}$$

$$B_i(x_i) = \log(x_i) \tag{4.26}$$

$$C_i(x_i) = -\omega_{gam} x_i + \omega_{gam}. \tag{4.27}$$

As a consequence, we obtain the following energy term for a configuration $\{X = x\}$

$$U(Y = y) = \sum_{s \in \mathcal{S}} \left(\alpha \log x - \omega_{gam} x_s + \omega_{gam}\right) + \sum_{s,r \in \mathcal{C}_2} \beta_{s,r} \log x_s \log x_r. \tag{4.28}$$

Then combining the Gamma property of $E[x] = \frac{\mu_{gam}}{\omega_{gam}}$ with the auto-model property of Eq. 4.15, we arrive at the following conditional expectation

$$E[x_i|x_{\mathcal{V}_i}] = \frac{\mu_{gam}}{\omega_{gam}} = \frac{A_i(x_{\mathcal{V}_i}) + 1}{\omega_{gam}} = \frac{1}{\omega_{gam}} \left(\alpha_i + 1 + \sum_{i' \in \mathcal{V}_i} \beta_{i,i'} \log y_{i'} \right). \quad (4.29)$$

The *Auto-Nakagami* model was first derived in [Bouhlel and Sevestre-Ghalila, 2009] and used for tissue classification of ultrasound images. According to the author the α parameter can be related to the density and amplitudes of scatterers and thereby directly reflecting spatial properties. In contrast to that, the β parameters represent the interaction of the diametrically opposite neighbor pixels influencing the center pixel and therefore are closely related to scatterer spacing. This is similar to the case of the generalized and homodyned K-distribution, whose parameters can be given a physical interpretation [Destrempes and Cloutier, 2010]. However, these distributions are not part of the exponential family, due to the Bessel function in their formulation, and thus do not fit into the auto-model framework.

4.3.6 Parameter Estimation

Having set up a MRF texture model with appropriate neighborhood and sites instantiated, the final step is to compute the interaction parameters, that is

$$\psi = \left(\alpha \ , \ \beta_1 \ , \ \dots \ , \ \beta_n \right) \in \mathbb{R}^{n+1} \ , \quad (4.30)$$

employing a neighborhood system with n cliques. In the following we assume that a parameter set is computed for homogeneous texture patches. If homogeneity is not fulfilled in texture patterns, pixels require classification prior to estimation, as otherwise strictly speaking the model does not hold. However, this constitutes a non-trivial problem and becomes even more complex when the number of underlying texture models is not known a priori. In the following we relax this assumption and always assume a single and homogeneous texture model. Basically, this is controlled by the patch size on which the parameters are estimated. The smaller the patch, the more likely a homogeneous texture is present. In contrast, larger patches are more likely to have heterogenous texture but otherwise permit a more stable estimation.

Maximum-Likelihood-Estimation

Given a MRF with configuration $\{X = x\}$ a MLE estimator seeks to maximize the conditional probability $\Pi(x|\theta)$ w.r.t. to the parameters of the auto-model yielding $\theta = \{\alpha, \beta_{i,i'}\}$, that is

$$\hat{\theta} = \arg\max_{\theta} \Pi(x|\theta). \quad (4.31)$$

Generally, in order to compute the parameters via the MLE estimator the first step is to define the log-likelihood of the configuration in Gibbs representation using Eq. 4.9 and 4.10

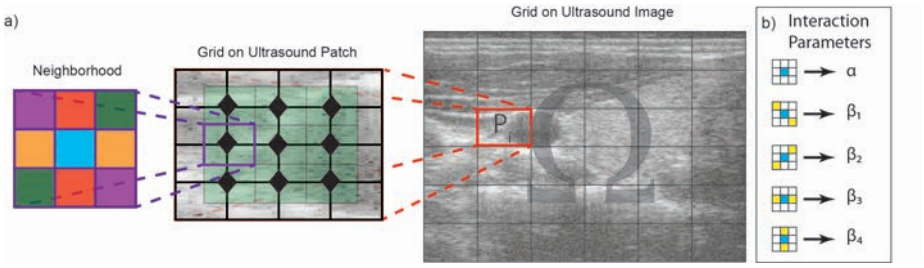


Figure 4.5: Scheme of ultrasound MRF model - from right to left (a) Sparse grid placed on ultrasound image; Example patch $\mathcal{P}_i$ with sub-grid and grid points $\mathcal{S}$ (filled $\diamondsuit$ symbols); second-order neighborhood of a sub-grid point (b) Interaction parameters and the associated neighborhood relation.

$$\log \Pi(x|\theta) = -U(x|\theta) - \log Z(\theta). \tag{4.32}$$

Consequently, optimum values can be found by setting $\frac{\partial \log \Pi(x|\theta)}{\partial \theta} = 0$ and solving for θ. However, this requires the evaluation of the partition function $Z(\theta)$. Due to the general computational intractability of the partition function, this method is not widely applicable, and is already prohibitive even for MRFs of moderate size. Specifically, the computational complexity of the partition function arises due to the large combinatorial number of elements in the configuration space Ω.

Conditional-Least-Squares

The conditional-least squares (CLS) approach provides a tractable solution to compute the MRF parameters. The idea is to find a set of auto-model parameters $\theta = \{\alpha, \beta_{i,i'}\}$ that minimize the variance between the expected value from the model and the observed intensities, that is [Schroder et al., 1998]

$$\hat{\psi} = \underset{\theta}{\operatorname{argmin}} \sum_i (x_i - E[X_i])^2, \theta_i = \{\alpha_i, \beta_{i,i'}\} \tag{4.33}$$

$$E[X_i^2] = \frac{1}{2\omega_i}(\alpha_i + \sum_{i \neq i'} \beta_{i,i'} \log x_i') \,, \tag{4.34}$$

with $E[.]$ being the expected value. It should be noted that CLS implicitly assumes a Gaussian variation around x_i. Each pixel location x_s supplies an equation w.r.t. to its neighbors $x_\mathcal{V}$ and the parameters θ. This can be written as a linear system of equations such as [Petrou and Sevilla, 2006]

$$b = Ax, \tag{4.35}$$

where x contains the unknown parameters, i.e. the parameters of the auto-model,

$$x = \begin{pmatrix} 1+\alpha & \beta_1 \ldots & \beta_M \end{pmatrix}. \tag{4.36}$$

In the following the method is show-cased for a Nakagami/Gamma Auto-Model making use of the parameter relationship provided by Eq. 4.29. At first step the the N intensities from all lattice position scaled by the scale parameter ω_{gam} of the Gamma distribution are stacked up in a vector

$$b = \begin{pmatrix} x_1 \cdot \omega_{gam} & ... & x_N \cdot \omega_{gam} \end{pmatrix}. \tag{4.37}$$

In the next step the summed intensities from each of the M cliques are stacked up for each corresponding pixel lattice position yielding the following matrix

$$A = \begin{bmatrix} 1 & \sum_{i \in \mathcal{C}_1^1} \log x_i & ... & \sum_{i \in \mathcal{C}_M^1} \log x_i \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ 1 & \sum_{i \in \mathcal{C}_1^N} \log x_i & ... & \sum_{i \in \mathcal{C}_M^N} \log x_i \end{bmatrix}, \tag{4.38}$$

where $\mathcal{C}_Y^X$ denotes the clique type Y at location X.

Having defined the vector b in Eq. 4.37 and the matrix A in Eq. 4.38, a linear system of equations as given by Eq. 4.35 can be set up. This can be solved for vector x by simply computing the pseudo-inverse defined as

$$\hat{x} = \left(A^T A \right)^{-1} A^T b. \tag{4.39}$$

However, when the lattice is large building the corresponding large vectors and vectors might be cumbersome. A more memory efficient method avoiding the stacking up process is described in [Björck, 1996, Chapter 2.2], yielding

$$\begin{aligned} \vartheta_s &= \left(1 \ , \ \sum_{i \in \mathcal{C}_1} \log x_i \ , \ ... \ , \ \sum_{i \in \mathcal{C}_n} \log x_i \right) \\ \hat{x} &= \left[\sum_{s \in \mathcal{S}} \vartheta_s^T \vartheta_s \right]^{-1} \sum_{s \in \mathcal{S}} \vartheta_s^T \left(\omega_{gam} x_s \right). \end{aligned} \tag{4.40}$$

The result supplied by the least squares estimator basically assumes that each pixel takes the most probable value permitted in consistence by all other pixels and therefore represents an averaged compromise solution.

Maximum-Pseudo-Likelihood

Given that the MLE parameter estimation is in general intractable an alternative method was proposed in [Besag, 1975, Besag, 1986], which is referred to as the Maximum-Pseudo-Likelihood (MPL). In order to avoid the computation of the computational prohibitive partition function $Z(\theta)$ (see Eq. 4.9 and 4.10), the pseudo-likelihood is simply the product of the conditional probabilities defined as

$$\Pi(x) \sim PL(x) = \prod_{s \in \mathcal{S}} \Pi(x_i \mid x_{\mathbf{N}_i}) = \prod_{s \in \mathcal{S}} \frac{\exp\left[-U_i(x_i, x_{\mathcal{V}_i}) \right]}{\sum_{x_s \in \Lambda} \exp\left[-U_i(x_s, x_{\mathcal{V}_s}) \right]}. \tag{4.41}$$

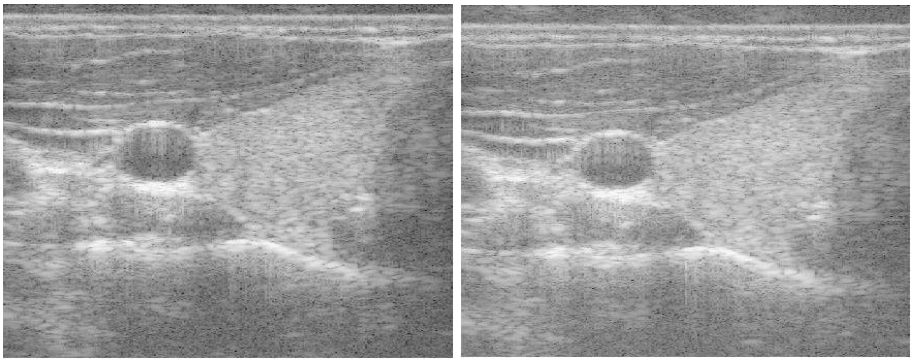


Figure 4.6: Ultrasound image of human neck; left training image, right test image.

Algorithm 4.1 Metropolis Sampler

random initialization $\{X = x\}$ s.t. $\forall i \in \mathcal{S} : x_i \in \Lambda$
repeat
  $i \leftarrow i + 1$
  for all $i \in \mathcal{S}$ **do**
    choose random label: $x_i \in \Lambda$
    update label at site i obtaining a test configuration x'
    Compute probability of new configuration $p = min\{1, \Pi(x')/\Pi(x)\}$
    $x \leftarrow x'$ with probability p
  end for
until $i = N$

Similarly to the MLE approach the MPL seeks to maximize the corresponding likelihood that is

$$\hat{\psi} = \arg\max_{\theta} \Pi(x_s), \quad \text{s.t.} \quad \theta_i = \{\alpha_i, \beta_{i,i'}\}, \tag{4.42}$$

where the optimum is found by zero-setting of the corresponding derivatives $\frac{\partial PL(x)}{\partial \alpha} = 0$ and $\frac{\partial PL(x)}{\partial \beta} = 0$. However, the pseudo-likelihood is just an approximation of the real likelihood as of course x_i and $x_{\mathcal{V}_i}, \forall i \in \mathcal{S}$ are not independent as suggested by the product. Basically, this inherently assumes textural locality such that implicitly long-range dependencies will get lost. However, in [Geman and Graffigne, 1986] it was shown that for large lattices $\mathcal{S}$ the MPL solution converges to the true optimum. Given the independent product terms in the MPL, optimization is quite attractive for GPU-processing such that efficient and fast solving is tractable.

4.3.7 Texture Synthesis

Beside estimation the parameters and analyzing existing textures, it is also possible to synthetically generate textures by random sampling, whereby common approaches perform stochastic relaxation (SR) of a Markov chain

Algorithm 4.2 Gibbs Sampler

random initialization $\{X = x\}$ s.t. $\forall i \in \mathcal{S} : x_i \in \Lambda$
repeat
  $i \leftarrow i + 1$
  for all $i \in \mathcal{S}$ **do**
    for $j = 1 \rightarrow |\Lambda|$ **do**
      $p_j = \Pi(x_i = j | x_{\mathcal{V}_i})$
    end for
    $x_i \leftarrow k$ with probability p_k
  end for
until $i = N$

(MC) [Cross and Jain, 1983, Geman et al., 1990, Geman and Geman, 1984]. Following [Bharucha-Reid, 2010, Kemeny and Snell, 1960], SR is a process of generating a MC, that is a sequence of random variables $(s_t : t = 0, 1, 2, ...)$ with Markov property, which in the equilibrium state yield a distribution $P(s)$. In this respect, two common approaches for generation are the Metropolis sampler [Metropolis et al., 1953] and the Gibbs sampler [Geman and Geman, 1984] - see Fig. 4.1 and 4.2 for corresponding pseudo-code of the samplers. Both algorithms replicate all but one label of the random field at one step. The non-replicated site, however, is updated according to the local conditional likelihood depending on the neighboring replicated intensities. The difference between the algorithm lies in the assignment of new samples. Whereas the Metropolis sampler randomly chooses a label from the set Λ and computes the probability for the new configuration, acceptance for this new configuration is proportional to the change in equilibrium. In contrast to that, the Gibbs Sampler computes the conditional probability $\Pi(x_i = j | x_{\mathcal{V}_i}), \forall j \in \Lambda$ and then accepts a new label by random draw with proportional label probability. As the Gibbs sampler needs to compute $|\Lambda|$ probabilities in each iteration, it can be computationally quite expensive. Furthermore, when the number of labels is quite large, the probabilities tend to become very small and the algorithm inaccurate due to numerical inaccuracies. In such cases the Metropolis Sampler is typically the favored approach. However, the Gibbs sampler has the tendency to converge faster to the equilibrium state than the Metropolis sampler [Petrou and Sevilla, 2006]. See Fig. 4.7 for some example synthesized textures generated using a Metropolis sampler. It should be noted that not every set of parameters $\theta = \{\alpha, \beta_{i,i'}\}$ yields a corresponding texture pattern. Rather the inherent constraints suggest that the parameters need to be compliant e.g. lie on a valid manifold. In particular this requires the MRF to be self-consistent and equivalent to a Gibbs distribution such that its neighborhood structure can be expressed in terms of cliques in order to be realizable [Petrou and Sevilla, 2006].

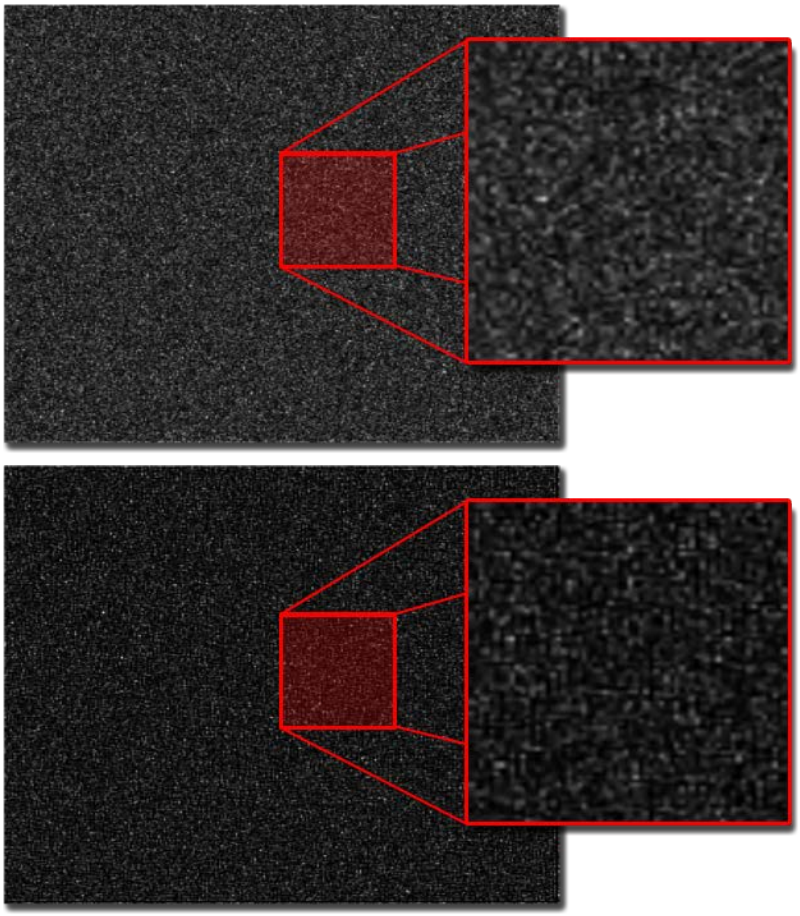


Figure 4.7: Textures generated with a Metropolis sampler with 300 iterations and a second-order neighborhood system. Top: $\mu_{gamma} = 1.4, \omega_{gamma} = 1.5E3, \alpha = 2, \beta_1 = 0.5, \beta_2 = 0.5, \beta_3 = -0.5, \beta_4 = -0.5$. Bottom: $\mu_{gamma} = 3.4, \omega_{gamma} = 2.3E3, \alpha 5 =, \beta_1 = 0.9, \beta_2 = 0.9, \beta_3 = -0.3, \beta_4 = -0.3$. Whereas the texture at the top resembles more or less like random structures, the texture at the bottom shows more slightly more pronounced characteristics in the horizontal and vertical direction, as induced by higher magnitude in interaction features β_1 and β_2 related to those directions.

4.4 Markov Random Field Descriptor for RF Ultrasound

In the following we formulate and analyze a MRF-based feature descriptor for RF ultrasound data. Basis for the method is the assumption of an underlying Nakagami model for the data, which is fulfilled in many clinical scenarios. The closest work, which shares analogy with our method but, however, does not formulate a global feature descriptor, see [Bouhlel and Sevestre-Ghalila, 2009]. There have been attempts at using texture for segmentation of ultrasound, see [Noble, 2009] for an overview. However, these models do not use Besag's auto-model, essential to this work; instead they employ simpler statistics like local mean and variance as feature descriptors. Feature descriptors, such as the one discussed here have many potential uses, e.g. classification, segmentation and most important of all registration.

4.4.1 General Background

As thoroughly described in Sec. 2.2 speckle in ultrasound is the result of the wave-front interference characterized by the spacing and organization of scatterers in tissue [Burckhardt, 1978, Wagner et al., 1983]. The resulting stochastic patterns are distinctive for organs. Backscatter intensities within an organ are of stochastic nature, however, follow a certain distribution and therefore are subject to a spatial interaction relationship within a neighborhood. This spatial interaction within a statistical framework can be nicely modeled with a MRF texture model described in Sec. 4.3.5. Following the approach in [Bouhlel and Sevestre-Ghalila, 2009], the feature descriptor to be discussed employs the descriptive nature of the interaction parameters of auto-models. This contrasts with the conventional MRF based SAR image analysis, where one set of interaction parameters is sought providing the best classification results.
In the sequel we assume all intensities $x \in \Omega$ in the RF envelope image Ω to follow a Nakagami distribution

$$\mathcal{N}(x \mid \mu, \omega) = \frac{2\mu^{\mu} x^{2\mu-1}}{\Gamma(\mu)\omega^{\mu}} \exp\left(-\left(\frac{\mu}{\omega}\right) x^2\right), \forall x \in \mathbb{R}_+, \tag{4.43}$$

with μ, ω the shape and scale parameters, respectively. When the data follows the Nakagami distribution, it can be readily embedded into an auto-model, since the Nakagami belongs to the exponential family. See Sec.4.3.5 for details on the *Auto-Nakagami* model.

4.4.2 Method

The goal of this approach is to derive a descriptor for a small region of the image referred to as patch. As a first step, in order to decompose the image, a sparse grid is placed on the entire image Ω subdividing the data into patches $\mathcal{P}_i \in \Omega$, s.t. $\bigcup_i \mathcal{P}_i = \Omega$. Subsequently, a Nakagami MLE is operated on a patch containing area (or Gamma MLE, making use of the

close Nakagami-Gamma relationship $Y \sim \mathcal{G}(x, m, \theta), X \sim \mathcal{N}(x, \mu, \omega), \sqrt{X} = Y(m = \mu, \theta = \frac{\omega}{\mu})$. Having obtained the MLE for the Nakagami distribution, another more dense grid is placed on the patch, with $\mathcal{S} \subset \mathcal{P}_i$ denoting the set of all grid points on a specific patch. This is followed by an instantiation of the MRF interaction model centered on each sub-grid point $s \in \mathcal{S}$. We define the cliques to be spatially opposite segments of an annulus of varying radius (depending on the degree of neighborhood) centered on sub-grid points, similar to work proposed in [Bouhlel and Sevestre-Ghalila, 2009]. In the following we assume a second-order neighborhood on which the MRF is instantiated - see Fig.4.5 for an schematic illustration of the model. After having computed the MRF parameters $\psi = (\alpha\ ,\ \beta_1\ ,\ \ldots\ ,\ \beta_n) \in \mathbb{R}^{n+1}$ using CLS (see Sec. 4.3.6) and MPL (see Sec. 4.3.6) a distance function has to be defined on the parameter space. In order to cope with the different variation statistic among the individual components, a learning based approach was selected. Concretely, we employed the boosting based learned metric DistBoost proposed in [Hertz et al., 2006, Hertz et al., 2004, Hertz, 2006]. For training the metric, out of a sequence of ultrasound images one was selected and used for supervised-learning. Therefore features $X = \{x_i\}_{i=1..N}$ with $x_i \in \mathbb{R}^{w+h+n+1}$, assuming texture patches of size $[w, h]$ and each feature vector with $dim(\psi) = n + 1$, from various texture areas were manually segmented and supplied with labels $Y = \{y_i\}_{i=1..N}$ where $y_i = \{-1, +1, *\}$ (denoting negative, positive and unknown equivalence relation). Consequently, one yields two equivalence sets that is pairs of index classes that are positively $\left\{(p_j^1, p_j^2)\right\}_{j=1..N_p}$ or negatively $\left\{(n_j^1, n_j^2)\right\}_{j=1..N_n}$ correlated, where N_p, N_n denote the number of corresponding features.

4.4.3 DistBoost

In the following we give a brief summary of the functionality of the learning based distance approach DistBoost as proposed in [Hertz et al., 2006, Hertz et al., 2004, Hertz, 2006]. Looking at its core structure, DistBoost is basically an augmentation of the AdaBoost algorithm [Freund and Schapire, 1995, Schapire and Singer, 1999]. In this regard, the idea underlying AdaBoost is to create a series of weak classifiers, which can have substantial error as long as their overall performance is better than random guess. Furthermore, training is performed in a series of iterations. In each round of the training, the importance of the samples is modified, giving misclassified training points higher weight in order to shift focus in the next training iteration to the wrongly classified points. In the end the classifiers is composed of a weighted sum of weak classifiers with the weight related to the error rate. Similarly, the distance function built by DistBoost is based on the weighted majority vote of weak partition functions. Within the training process the weak learner tries to find reasonable partitions of the input space that comply with the equivalence constraints provided for training. The training process is performed in a loop, whereby each iteration $t = 1..T$ produces a weak classifier with associated weight. At the end, the distance metric is represented as a weighted sum of

the feedback of the individual classifiers. After having described the rough functionality of DistBoost, we next explain the individual steps in more detail. Given the data associated with equivalence constraints as well as unlabeled data, the first step is to obtain a weak classifier. Therefore a Gaussian-Mixture-Model (GMM) is computed using the so called constrained Expectation-Maximization (cEM) [Shental et al., 2003] algorithm, which enforces the given equivalence constraints during parameter inference (see Sec. 4.6.6 and 4.6.7 for a detailed overview about the EM algorithm). The joined distribution on a set of n points is estimated by cEM is defined as

$$p(X, H|\Theta, \Omega) = \frac{1}{Z} \prod_{i=1}^{n} \alpha_{h_i} p(x_i|\theta_{h_i}) \prod_{j=1}^{N_p} \delta_{h_{p_j^1} h_{p_j^2}} \prod_{k=1}^{N_n} (1 - \delta_{h_{n_k^1} h_{n_k^2}}), \tag{4.44}$$

where X is the set of data points with corresponding data constraints Ω (see above, Sec.4.4.2) and $\Theta = \{\alpha_i, \theta_{h_i}\}_{i=1..M}$ the mixture parameter set with overall M components. The mixture is computed from a sample of size N_v, whereby each point x_i is associated with a weight w_i. As a result, during training each point x_i appears $w_i \cdot N_v$ times in the sample set. By this mechanism importance of mis-classified points can be reinforced. Having obtained a mixture model by means of cEM fitting, the next step is to compute a weak distance function. This is achieved by maximum a posteriori (MAP) class assignment, $h_t : X \to [-1, +1]$, returning positive distance when the points $(x_i, x_{i'}) \in X$ are from the same Gaussian source and negative, otherwise. The MAP based classifier gives

$$p_{MAP}(x_i) = \arg\max_j + p(h_i = j|x_i, \Theta), \tag{4.45}$$

then the weak classifier is composed of

$$\tilde{h}_t(x_1, x_2) = \begin{cases} +p_{MAP}(x_1) \cdot p_{MAP}(x_2), & h_1 = h_2 \\ -p_{MAP}(x_1) \cdot p_{MAP}(x_2), & h_1 \neq h_2 \end{cases}, \tag{4.46}$$

where h_1, h_2 are the hidden parameters associated with the data points x_1, x_2 obtained from GMM. By transformation of the weak classifier $\tilde{h}_t(.,.)$, we yield the weak distance function $h_t(.,.)$ defined as

$$h_t(x_i, x_{i'}) = \frac{1}{2} \left(1 - \tilde{h}_t(x_i, x_{i'})\right) \in [0, 1]. \tag{4.47}$$

Having obtained the weak distance function h_t, the next step is to compute a corresponding weight parameter α_t (not to be confused with the mixture weight) that scales the output of the weak distance function. It is defined as

$$\alpha_t = \frac{1}{2} \log\left(\frac{1+r}{1-r}\right) \quad \text{with} \quad r = \sum_{x_i, x_{i'}, y_i = \pm 1}^{N} W_t(i) h(i) > 0 \tag{4.48}$$

$$W^t_{x_i, x_{i'}} = 1/(n^2) \quad , \quad w^1_k = 1/n. \tag{4.49}$$

Next the weighting function $W(.)$ is updated,

$$W^{t+1}_{x_i, x_{i'}}(i) = \begin{cases} W^t_{x_i, x_{i'}}(i) \exp(-\alpha_t y_i h_t(x_i, x_{i'})), & y_i \in \{-1, 1\} \\ W^t_{x_i, x_{i'}}(i) \exp(-\alpha_t), & y_i = * \end{cases}, \tag{4.50}$$

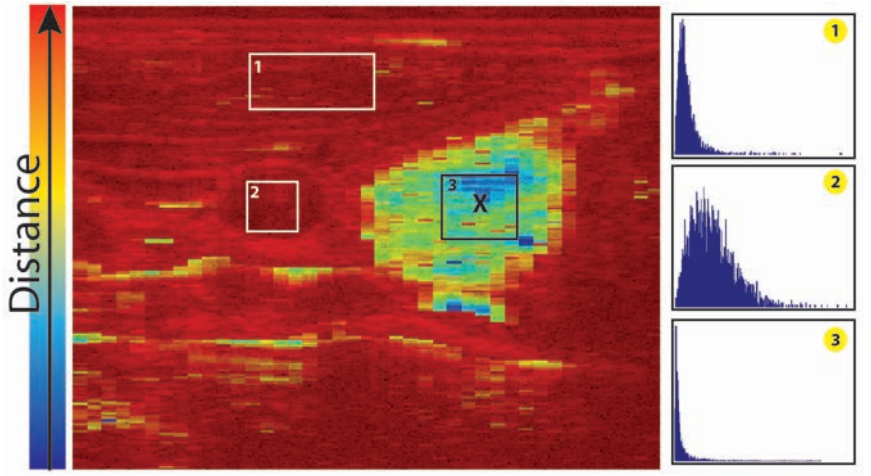


Figure 4.8: Distance map of the feature descriptor. The feature highlighted with a cross is compared with all other feature vectors in the image. Shades from blue to red indicate increasing dissimilarity. Histograms show the distribution in various areas and their similarity as can be seen from the feature vector distance.

where y_i again the denote the manual labels provided for the training. Having computed the weight function, the next step is to compute a normalization factor Z. As a result the weight function output is adapted to provide normalized weighting as defined by

$$W^{t+1}_{x_i,x_{i'}}(i) = W^{t+1}_{x_i,x_{i'}}(i)/Z^{t+1} \quad \text{s.t.} \quad Z_{t+1} = \sum_{i}^{N} W^{t+1}_{x_i,x_{i'}}. \tag{4.51}$$

As the boosting process computes weights $W(.)$ for pairs of points, one needs to deduce weights for points in order to adapt the sample set which is presented to the cEM algorithm. This is achieved by simple marginalization of all point pairs,

$$w^{t+1}_k = \sum_j W^t_{j,k} \tag{4.52}$$

After T iterations the final distance function is the weighted sum of weak classifiers given by

$$f(x_i, x_{i'}) = \sum_{t=1}^{T} \alpha_t h_t(x_i, x_{i'}). \tag{4.53}$$

4.4.4 Results

In order to analyze the feature descriptor we computed a feature image (size: 206 x 40 x 5) for an ultrasound image of the human neck[1](size: 2080 x 256) - see Fig. 4.5 and 4.6. The image was recorded using a linear transducer,

[1]courtesy of Ultrasonix Medical Corporation, Richmond, Canada

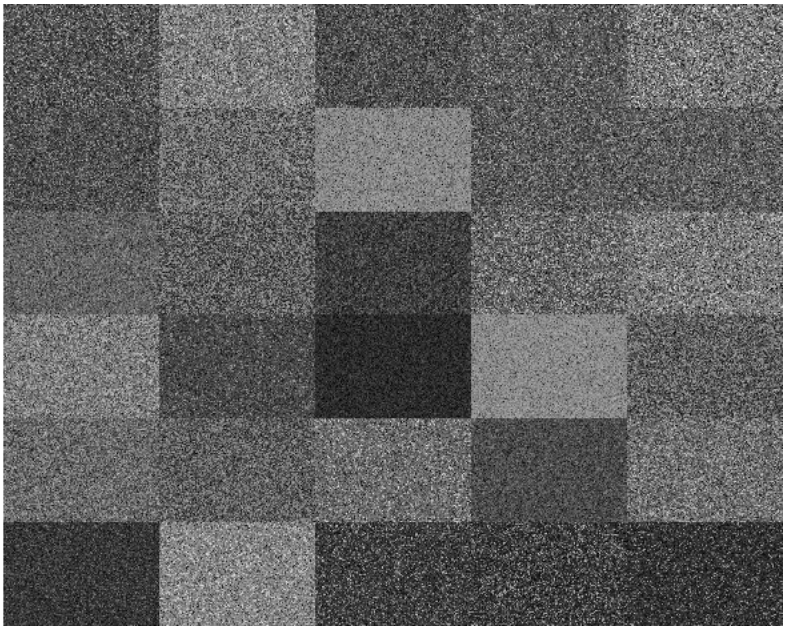


Figure 4.9: Texture data for training DistBoost - consisting of synthesized pattern from real ultrasound data.

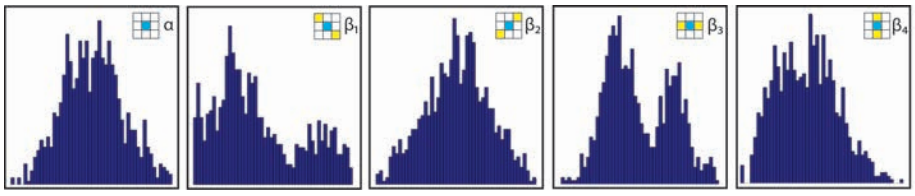


Figure 4.10: Distribution of the interaction parameters for the second-order neighborhood within a homogenous texture.

avoiding the issue of locally varying neighborhood, e.g. prominent in convex or phased array transducers. The descriptive power of the feature description is visualized using distance maps. Thereby a designated feature is chosen in a characteristic area and compared to all other feature vectors - see Fig. 4.8 and 4.6.

We learn the metric on synthesized patches with distributions acquired from the image data, see Fig. 4.9. Although sufficient for our purposes, it would be desirable to train the metric directly on the image. This will require a refinement of our MRF model, to take into account the highly heterogeneous nature of real RF-data. From the comparably low computational complexity, calculating the interaction parameters using the CLS method seems attractive. The underlying local Gaussian variation assumption of CLS is generally not fulfilled due to e.g. heavy tailed distributions. In our case, however, estimates provided by CLS proved to be sufficiently regular in homogeneous

areas. Parameters derived using MPL tend to show even more regularity in homogeneous areas. It is this property that facilitates the learning of distance metric and the computation of distance maps. However, MPL optimization is quite time consuming. Employing other metrics such as Mahalanobis was considered, but deemed inappropriate due to the non-Gaussian distribution of the parameters - see Fig. 4.10. An example distance map is shown in Fig. 4.8. It indicates high similarity in the proximity in to the neighborhood of the comparison feature (marked with X). With increasing distance and growing pattern dissimilarity the feature distance grows. On the border regions, one can also find regions of similar distributions, which is correctly expressed by the feature descriptor with low distance. Furthermore, dissimilarity to the feature of interest is captured correctly as can be seen in regions of high intensity streaks, resulting in high distance peaks in the map.
It should be noted that we are *not* assuming that the proposed procedure is independent of specific machine settings (gain, frequency etc). In fact, it is quite evident that there is a strong inherent dependency between machine settings and features. Therefore we have formulated a procedure for computing a feature descriptor by which one can learn tissue specific similarities for images obtained using a *specific* machine setting. For future work, improvement of the generality of the model is envisaged.

4.4.5 Conclusion

We presented MRF-based feature descriptor for RF ultrasound data. Modeling the ultrasound speckle patterns promises to capture the underlying nature of the ultrasound data. Further research will be devoted on improving the reliability of the descriptor as well as investigating the influence of different neighborhood systems on the descriptive power. Potential fields of applications of this feature descriptor could be registration and segmentation as will be studied in future.

4.5 Local Binary Pattern

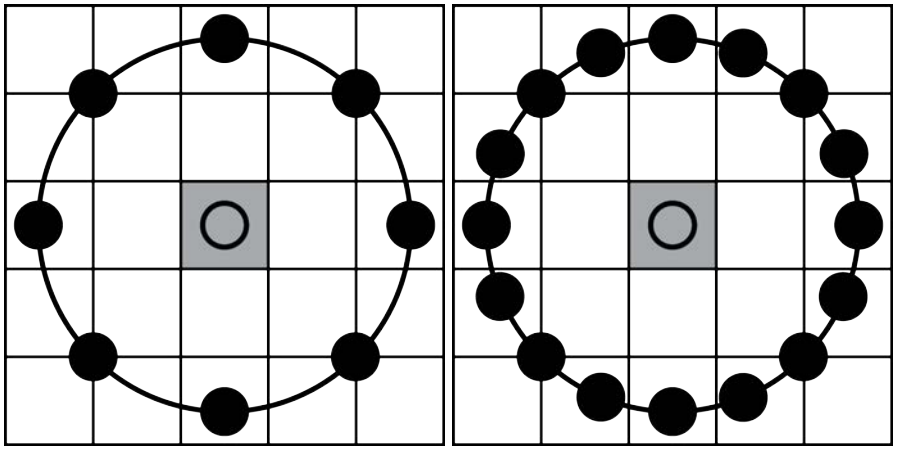


Figure 4.11: Illustration of circular LBP neighborhood system. The gray box denotes the central site g_c. Black dots indicatores neighbors g_i that are obtained using interpolation. Left: 8-neighbor system. Right: 16-neighbor system.

Texture classification is a wide domain in which numerous approaches exist. A relatively simple, but powerful and popular technique in computer vision, are the so called Local Binary Patterns (LBP), which were originally introduced in [Ojala et al., 1994, Ojala et al., 1996]. Later this work was extended to a multi-resolution and rotation-invariant feature descriptor [Ojala et al., 2002]. Since then it has found a wide range of application in diverse computer vision domains such texture segmentation [Ojala and Pietikäinen, 1999] or face recognition [Ahonen et al., 2004, Zhang et al., 2005]. Beside the analysis of static textures, recently application in employing of LBPs for spatio temporal texture modeling has come up for motion and activity analysis [Zhao and Pietikainen, 2007], facial dynamics [Zhao and Pietikainen, 2007] and human activity recognition [Kellokumpu et al., 2008, Kellokumpu et al., 2009, Kellokumpu et al., 2011].
In its standard formulation a LBP encodes a second order neighborhood of a pixel into a 2^8 bits code. Encoding is based on the inequality relationship between the central site g_c intensity and its N neighbors $\{g_i\}_{i=0..N-1}$, that is

$$s(x - g_c) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} . \tag{4.54}$$

A pixel position with lower intensity than the central one x_c is attributed the binary code 0, otherwise 1. Concatenation of those binary digits then yields an integer-valued descriptor of the neighborhood, defined as [Ojala et al., 2002]

$$LBP = \sum_{i=0}^{N-1} s(g_i - g_c)2^i. \tag{4.55}$$

The nature of the descriptor results in a certain invariance of the codes towards gray scale transformations. As long as the order of the gray values remains unchanged, the generated LBP code will always be identical. Instead of using a single code built from binary digits for describing a region, rather a histogram of codes is employed serving as regional descriptor. Given that the origin of LBPs lies in computer vision, their application is mainly tailored to natural scene images. However, there also exist adaptations to cope with images that are inherently subject to noise. Apart from traditional hierarchical systems, any type of neighborhood can be employed such as circular ones where the intensities are obtained by interpolation, such as

$$\begin{aligned} x_p &= x + R \cdot \cos(2\pi p/P) \\ y_p &= y - R \cdot \sin(2\pi p/P), \end{aligned} \tag{4.56}$$

with $g_i = (x_i, y_i)$, where P denotes the number of sampling points and R is the radius - see Fig. 4.11 for an illustration of circular neighborhood system with different numbers of elements. In this formulation the LBP associates with each distinct neighbor position a binary digit. Rotating the image or patch would lead to the same identical distribution of binary digits but in different order. To account for that, one can also generate rotation invariant codes. A very simple method is to rotate the bits until the most significant bits in the bit string have the maximum number of leading zero series. As a result, this produces the smallest integer-valued descriptor for a patch [Ojala et al., 2002],

$$LBP^{ri} = min\{ROR(LBP, i) \quad | \quad i = 0, 1, ..., P-1\}, \tag{4.57}$$

where $ROR(., i)$ denotes a circular bit-wise right shift operator, performing i shifts on a P bits wide code. This, however, also reduces the number of unique codes drastically. For the case of 8 neighbors, the numbers of codes is reduced from non-rotation invariant number of codes being 2^8 to merely 36 for rotation invariant codes. An alternative to the bit-rotation, providing better performance are the so-called *uniform* patterns. They achieve rotation invariance by simply counting the number of (0/1) bit transitions [Ojala et al., 2002]

$$LBP^{uniform} = \begin{cases} \sum_{p=0}^{(P-1)} s(g_p - g_c), \text{if } U(LBP) \leq 2 \\ P+1, \text{otherwise,} \end{cases} \tag{4.58}$$

with

$$U(LBP) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P} |s(g_p - g_c) - s(g_{p-1} - g_c)|. \tag{4.59}$$

The improved discriminative power compared to the rotational operator can be attributed to the fact that such codes are not so prone to crude quantization and high variance in the occurrence frequency of the individual patterns as LBP^{ri} codes are [Pietikäinen et al., 2000, Ojala et al., 2002]. Using uniform patterns will result in $P + 2$ unique codes for 2^P wide binary input values. Whereas the traditional LBPs are made for general computer vision applications making use of natural scene imaginary, a fuzzy LBP was proposed [Iakovidis

et al., 2010] specifically for ultrasound and classification of patterns in thyroid images. Due to the noisy nature of the images, regions may not only be encoded by a single LBP, but various codes can represent a region depending on the uncertainty. Therefore less histogram variation is achieved, providing a more robust descriptor for noisy image domains. In this respect, a threshold F specifies an intensity range in which fuzziness is assumed and a ramp (membership function) associates the probability or confidence for the binary class encoding. This leads to the following membership function for class 0 [Iakovidis et al., 2010] ,

$$\mu_0(g_i) = \begin{cases} 0, \text{ if } (g_i - g_c) \geq F \\ \frac{F-(g_i-g_c)}{2 \cdot F}, \text{ if } -F < g_i - g_c < F \\ 1, \text{ if } (g_i - g_c) \leq -F, \end{cases} \tag{4.60}$$

and similarly for class 1, defined as

$$\mu_1(g_i) = \begin{cases} 1, \text{ if } (g_i - g_c) \geq F \\ \frac{F+(g_i-g_c)}{2 \cdot F}, \text{ if } -F < g_i - g_c < F \\ 0, \text{ if } (g_i - g_c) \leq -F. \end{cases} \tag{4.61}$$

Given the membership values for each digit $\{d_i\}_{i=1..N} = \{s(g_i)\}_{i=1..N}$, individual LBPs are generated according to the confidence in their appearance proportionally in the histogram according to

$$c_{LBP} = \prod_{i=1}^{N} \mu_{d_i}. \tag{4.62}$$

Common to all the approaches, once a histogram is built from the individual codes, similarity has to estimated. For measuring similarity between LBP histograms, several methods have been proposed such as Histogram intersection

$$D(\mathbf{S}, \mathbf{M}) = \sum_i min(s_i, m_i), \tag{4.63}$$

Log-likelihood statistic

$$L(\mathbf{S}, \mathbf{M}) = -\sum_i s_i \log m_i, \tag{4.64}$$

and Chi-square statistic

$$\chi^2(\mathbf{S}, \mathbf{M}) = \sum_i \frac{(s_i - m_i)^2}{s_i + m_i} \tag{4.65}$$

where $\mathbf{S}, \mathbf{M}$ denote the LBP histograms, containing the individual bins $\mathbf{S} = \{s_i\}_{i=1..N}$ and $\mathbf{M} = \{m_i\}_{i=1..N}$. A comparative study of these measures w.r.t. face recognition was performed in [Ahonen et al., 2004].

4.6 Shadow Detection in Ultrasound RF Data

Acoustic shadowing is an ultrasound specific artifact caused by an abrupt and large change in acoustic impedance at transmission medium interfaces, e.g. tissue and bone. A large difference causes strong reflection and limited or even zero energy transmission of sound into underlying tissue and results in shadow [Hedrick et al., 2004]. Although largely considered an artifact, it can also be of use in diagnosis. In particular it has shown useful for detecting gallstones, calcifications, and bone structures [Hedrick et al., 2004, Hellier et al., 2010]. However, on the downside it influences the efficiency of image processing algorithms negatively, including registration and 3D reconstruction. In particular multi-view 3D reconstruction of ultrasound in the presence of shadow artifacts requires processing in order to yield useful results. Consequently, it is important for many applications to accurately and reliably detect these artifacts for both diagnostic and image processing applications [Hellier et al., 2010].

Automatic ultrasound shadow detection has received relatively limited attention. In [Penney et al., 2004] each ultrasound scanline is sampled and marked as artifact/shadow region until an empirical threshold is reached. Afterwards, the resulting mask is utilized for multi-modal US-MR registration. In [Leroy et al., 2004] a heuristic exponential function is correlated to the ultrasound scanline profile and shadow is detected when correlation reaches an empirical threshold. Subsequently, the detection is used to reduce the impact of shadow regions in multi-modal US-CT registration. Furthermore, in [Madabhushi et al., 2006] breast lesions are discriminated on the basis of posterior acoustic shadowing in a machine learning approach by previously extracting multi-scale texture features from a training dataset. In [Hellier et al., 2010] possible shadow starting points are detected along each scanline and tested for shadowing by evaluating local robust patch statistics. Examples are presented that illustrate the benefit of integrating shadow detection into mono-modal ultrasound registration and US reconstruction. There acoustic shadow areas are assumed to have low signal and noise. Although, this assumption is generally valid for intraoperative brain images [Hellier et al., 2010], it does not hold for shadow regions that occur after highly reflective interfaces. The ultrasound image area after such interfaces, like bone-tissue and lung-liver interface, typically contains a broad range of texture patterns and noise (see Fig. 4.16). Among the prominent patterns are reverberation, mirroring, and backscattered echo [Hedrick et al., 2004]. As these effects lead to non-homogeneous shadow regions, which in turn complicate shadow detection and necessitate specialized approaches to handle them.

In the following we present a fully automatic shadow detection approach for both, regions with low and high noise (we will refer to reverberation and mirror artifacts also as noise). The approach is based on envelope detected RF data, which has the advantage of not being subject to post-processing steps such as log-compression and proprietary filters that change the noise statistics. The higher information content compared to conventional B-mode makes it attractive for statistical analysis.

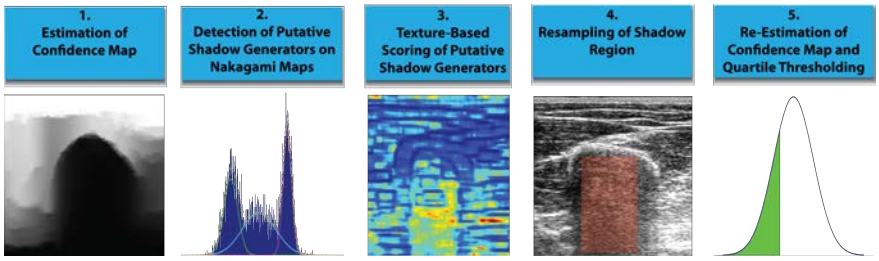


Figure 4.12: Illustration of the shadow detection pipeline consisting of five steps.

For robust detection a series of components is required. A fundamental step is the estimation of confidence in the acquired signal and to pre-determine possible shadow candidates. This information is further utilized to suppress the influence of noise in shadow areas, in order to produce accurate and robust shadow detection results. Coupled together with statistical pattern analysis, this yields a robust shadow detection. In our evaluation we demonstrate the applicability of this method on datasets with different imaging settings and levels of noise in shadow regions.

4.6.1 Method

In order to detect the shadow region(s) the image data undergoes several processing steps. The pipeline for shadow detection consists of five steps ranging from confidence estimation, putative shadow generator detection to the final classification of shadow regions followed by a thresholding operation on re-estimated confidence maps - see Fig. 4.12 for an schematic illustration of the individual steps. The first step is the computation of a confidence image C_1 of the data Ω, taking into account intensity gradients and the associated attenuation effects - see Sec. 4.6.2. Varying on the degree of noise and artifacts, this yields already a reasonable map of potential shadow regions only disturbed by major reverberations.
In the next stage of the image processing chain, strong reflection causing interfaces, potentially generating shadow, are localized and segmented, yielding a list L of putative shadow generating elements. The localization/segmentation step is performed on the Nakagami image corresponding to the RF data by means of a mixture modeling - see Sec. 4.6.3 for details on segmentation of reflectors. Furthermore, see Sec. 4.6.6 and 4.6.7 for an introduction to the EM algorithm that is used to obtain mixture estimation. Additionally, within the context of mixture-based segmentation see Sec.4.6.4 for the maximum-a-posteriori MRF formalism and an outline of the alpha-expansion algorithm in Sec. 4.6.5 that is used for obtaining the segmentation.
After having obtained the putative shadow generating interface segments, inconsistent elements are discarded from L. In particular, implausible potential reflectors that are located in high confidence regions as well as isolated and blocky elements are removed. The remaining interface segments are more

in-depth analyzed based on associated spatial statistics. Thereby scoring is performed of the region stretching from the interface to the distal end denoted as S_i with $\{S_i\}_{i=1..N}$ s.t. $S_i \subset \Omega$ and $N = |L|$, quantifying the potential of containing shadow. This is achieved by analyzing texture and statistical features for shadow specific patterns. Regions that exceed a certain score-threshold τ are considered as shadow. In this respect see Sec. 4.6.8 for details on the computation of the shadow texture descriptor and Sec. 4.6.9 for information about the scoring function.
After classification of shadow regions, reverberations and mirroring artifacts are removed in the designated shadow regions. This is achieved by random re-sampling intensities from the fourth quartile - $Q_4 \colon H(0.75 \leq X \leq 1)$ - of the cumulative intensity histogram H from up to the third quartile, $Q_{i\leq 3} \colon H(0 \leq X < 0.75)$ in the designated areas. This virtually removes the high intensity artifacts by artificial attenuation and eventually generates consistently distributed data. On the resulting image data that consists of replicated non-shadow regions and resampled shadow regions, a new confidence map C_2 is estimated providing more reliable maps as in the first step of the pipeline. Finally, shadow classification can be directly performed on the new confidence map. More precisely, pixels with confidence from the first confidence quartile, $Q_1 \colon H(0 \leq X \leq 0.25)$ are assumed to be shadow and simple thresholding is sufficient for segmentation. See Fig. 4.3 for a pseudo code listing of the processing pipeline.

Algorithm 4.3 Shadow Detection Processing Chaing

Given image Ω
Compute confidence map $C_1 = Confidence(\Omega)$
for all $\bigcup_i \mathcal{P}_i = \Omega$ **do**
  Compute Nakagami image $[\mu_i, \omega_i] = \mathcal{N}(\mathcal{P}_i)$
  Compute MRF texture parameters $\psi = CLS(\mathcal{P}_i)$
end for
Compute putative bone segments L based on Nakagami image $[\mu, \omega]$
for all $l_i \in L$ **do**
  if l_i implausible **then**
    $L \leftarrow L - \{l_i\}$
  end if
end for
for all $l_i \in L$ **do**
  Compute score s_i based on $\{\psi\}$ parameter set in $S_i \subset \Omega$ and L
  if $l_i \geq \tau_1$ **then**
    Shadow detected in region S_i
    Resample associated region: $S_i = resample(S_i)$
  end if
end for
Compute confidence map $C_2 = Confidence(\Omega)$
$Shadow = H(0 \leq C_2 \leq 0.25)$

4.6.2 Confidence Maps

Shadow detection and reduction of artifact influence in shadow area is performed with ultrasound confidence maps [Karamalis et al., 2012a], which were previously applied for emphasizing uncertainty in Intravascular Ultrasound (IVUS) acquisitions. The confidence maps are based on the Random-Walk algorithm [Grady, 2006] in its original formulation proposed for K-way image segmentation. Later on, the algorithm was applied for several other applications in computer vision such as mesh segmentation [Zhang et al., 2010], mesh denoising [Sun et al., 2008] and stereo matching [Shen et al., 2008]. The random walk algorithm formulates the segmentation task as a graph-based problem, based on initial user defined K seed points, objects are then segmented in the image or volume domain. In this respect, the image is modeled as an undirected weighted graph $G = (V, E)$ with pixels/voxels represented as nodes $v \in V$ and associated edges $e_{ij} \in E$ connecting neighboring pixels v_i and v_j. Based on the local image content each edge $e_{ij} \in E$ is assigned a weight $w_{ij} > 0$, which represents the likelihood of a random walker crossing the edge. In the original work [Grady, 2006], a Gaussian weighting function was employed

$$w_{ij} = \exp\left[-\beta(g_i - g_j)^2\right], \tag{4.66}$$

where g_x denotes the pixel intensity at node x and β a free parameter. In the segmentation process, for each unlabeled node, i.e. not a seed node, the probability is computed that a random walker starting at the node will reach each of the K-seed nodes. As a result, each node will be associated with a K-tuple vector representing the probabilistic associativity to the respective seed points. Finally, a segmentation is obtained by assigning each node the most likely seed point class. As outlined in [Grady, 2006], the solution to the random walk problem is equivalent to the combinatorial Dirichlet problem, which can be solved analytically. The underlying framework is quite general belonging to the class of equilibrium state modeling of discrete systems, applicable in a wide range of domains such as electrical networks, structural mechanics or in mass-spring systems [Gao, 2000, Strang, 2007]. This requires the computation of the so called combinatorial Laplacian matrix [Dodziuk, 1984] or weighted Graph Laplacian [Strang, 2007], which is used to represent connectivity in a graph which is defined as

$$L_{ij} = \begin{cases} d_i & \text{if } i = j \\ -w_{ij} & \text{if } v_i \text{ and } v_j \text{ adjacent nodes} \\ 0 & \text{otherwise} \end{cases}, \tag{4.67}$$

where $d_i = \sum_j w_{ij}$. Noteworthy properties of the Laplacian matrix are sparsity, symmetry and positive definiteness. As can be seen from the definition, the elements on the diagonal correspond to the accumulated weights of all connected neighbors. Adjacent nodes are associated with negative weight, therefore giving affinity to the neighbor. Non-connected nodes are represented with a zero. An alternative construction of the Laplacian matrix is possible as $L = A^T C A$, via the common graphic incidence matrix A in combination

with diagonal matrix C, consisting of the edge weights, with the elements of A defined as

$$A_{e_{ij}v_k} = \begin{cases} +1 & \text{if } i = k \\ +1 & \text{if } i = j \\ 0 & \text{otherwise,} \end{cases} \tag{4.68}$$

for every node v_k and edge e_{ij}. In the process of solving for the K-tuple probabilities, the matrix L is decomposed in blocks, followed by an rearrangement, yielding

$$L = \begin{bmatrix} L_M & B \\ B^T & L_U \end{bmatrix}, \tag{4.69}$$

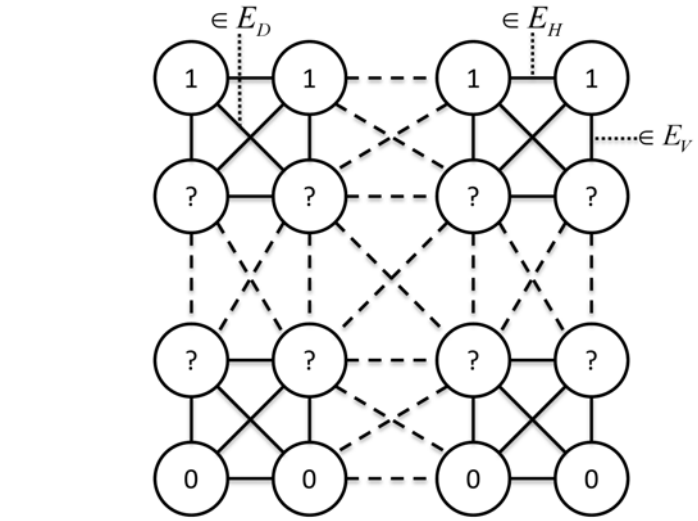
where M and U correspond to marked (seed points) and unmarked nodes. Finally, the desired probabilities are obtained by solving the system of linear equations given as

$$L_U x_U = -B^T x_M, \tag{4.70}$$

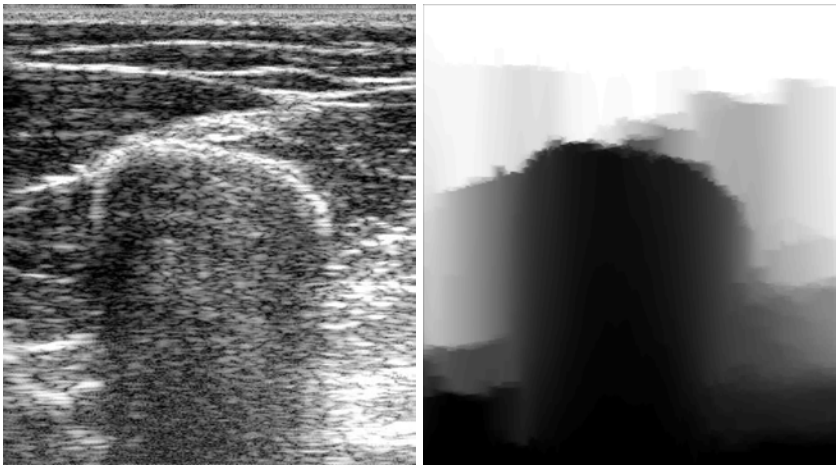
where x_U denotes the unknown probabilities for the unmarked nodes and x_M the known unit probabilities of the seeds.

For the confidence/uncertainty estimation of ultrasound imagery, the random walk problem is modified in order to incorporate ultrasound specific constraints. Thereby the notion of seed points as virtual transducer elements at the beginning of each scanline is employed. Consequently, the random walk formulation models the problem of obtaining the probability of a random walker starting from a pixel arriving at the transducer. However, the weights of the Graph Laplacian L are fundamentally different than in the original approach due to ultrasound specific properties. First and foremost image intensity cannot assumed to be constant across the image domain. This is mainly due to the attenuation effects - see Eq. 2.16 in Sec. 2.2 for more detail. This necessitates encoding the notion of distance into the model. Practically, the likelihood of a pixel reaching a virtual transducer should decay with increasing distance to accommodate the physical processes underlying attenuation. Another important property that has to be considered is that ultrasound is not propagating in a straight line fashion through the transmission medium. Rather it is subject to reflection, refraction and scattering. Although this random behavior is already partly modeled within the random walk, further modification is required. Given that most of the ultrasound energy is concentrated in a narrow beam, contribution from outside this area is unlikely. Therefore the random walker should be given the freedom to move on the beam, while accepting limited deviation in horizontal/perpendicular direction. Thereby the probability of horizontal deviation decreases with distance from the scanline, thus representing the width of the narrow beam. All those ultrasound specific constraints taken together lead to a new and specific formulation of the edge weights that are defined as,

$$w_{ij} = \begin{cases} w_{ij}^H & \text{if } i,j \text{ adjacent and } e_{ij} \in E_H \\ w_{ij}^V & \text{if } i,j \text{ adjacent and } e_{ij} \in E_V \\ 0 & \text{otherwise} \end{cases} \tag{4.71}$$



(a)



(b) (c)

Figure 4.13: (a) shows an illustration of graph setup for the confidence estimation using 8-connected lattice is shown. Virtual transducer elements nodes are marked with unity and absorption 'no signal' nodes with zero potential. (b) shows the envelope detected image from raw RF and (c) shows the corresponding confidence map.

$$w_{ij}^{H} = \exp(-\beta((g_i \exp(-\alpha x_i) - g_j \exp(-\alpha x_j))^2 + \gamma)) \tag{4.72}$$

$$w_{ij}^{V} = \exp(-\beta((g_i \exp(-\alpha x_i) - g_j \exp(-\alpha x_j))^2)), \tag{4.73}$$

where E_H, E_V are the edges along the horizontal and vertical graph direction, respectively. In this respect, see Fig. 4.13 for an illustration of the underlying graph, with the symmetric difference $E_H \ominus E_V = E$, and x_i the normalized closest distance from node v_i to the virtual transducer elements (seeds). The

adaptation of the model with respect to ultrasound physics becomes more intuitive when discussed in terms of potential theory (as described in [Grady, 2006]). In this respect, the beam-width is represented by the γ parameter that simultaneously acts as a penalty on the random walks, reducing the probability of a random walk moving in scanline perpendicular direction. Specifically, as mentioned above the horizontal move probability decays with distance to scanline is implicitly encoded into this formulation due to the inherent accumulative increase in circuit resistance, whereby the additive term acts as resistor. Sharing similarity to the original random walks approach, high intensity gradients represented by $(g_i \exp(-\alpha x_i) - g_j \exp(-\alpha x_j))^2)$, result in proportionally reduced energy transmission depending on the echo amplitude. Regarded from a potential theory viewpoint, the potential along possible circuit paths is reduced due to high resistances. Taken all together, the random walks solution [Grady, 2006, Karamalis et al., 2012a] to the graph Laplacian with the previously defined weights yields the desired ultrasound specific confidence estimate.

4.6.3 High Reflection Interface Segmentation

In order to find interfaces that are potentially highly reflective in the image domain we perform a rough segmentation of the corresponding Nakagami image (see step 2 in 4.6.8) following the observation that the Nakagami scale parameters can be represented by a log-Normal $\mathcal{LN}$ finite mixture model

$$p(x|\theta) = \sum_i^K \pi_i \mathcal{LN}(x|\mu_i, \sigma_i) = \sum_i^K \frac{w_i}{\sqrt{2\pi}\sigma_i x} \exp\left(-\frac{(\ln x - \mu_i)^2}{2\sigma_i^2}\right) \quad \text{s.t.} \quad \sum_{k=1}^K \pi_k = 1, \tag{4.74}$$

which is estimated using the Expectation-Maximization (EM) algorithm [Dempster et al., 1977, McLachlan and Krishnan, 2007]. Specifically, by applying the logarithm on the data, $\log(x)$, a common Gaussian mixture model (GMM) can be employed to derive the distributional parameters $\theta = \{\pi_i, \mu_i, \sigma_i\}_{i=1..K}$. Almost identical to the log-Normal $\mathcal{LN}$ finite mixture model, the Gaussian $\mathcal{G}$ mixture model is defined as

$$p(x|\theta) = \sum_i^K \pi_i \mathcal{G}(x|\mu_i, \sigma_i) = \sum_i^K \frac{w_i}{\sqrt{2\pi}\sigma_i x} \exp\left(-\frac{(x - \mu_i)^2}{2\sigma_i^2}\right) \quad \text{s.t.} \quad \sum_{k=1..K} \pi_k = 1, \tag{4.75}$$

by omitting the $\log(.)$ on the data. For the number of classes within the mixture model, we assume a $K = 3$ model facilitating the discrimination between *high-reflection* H_0, *intermediate* H_1 and *no-reflection* H_2. The intermediate class serves to avoid the merger of potential reflectors and thus allows for a more precise localization of shadow and reverberations. The actual segmentation is performed using a Markov-Random-Field (MRF) that provides regional coherence employing the MAP MRF formalism - see Sec. 4.6.4. Given the image

Ω consisting of M pixels, we have to determine the labelset $\{l_1, ..., l_M\}$ with $l_i \in \{H_0, H_1, H_2\}$, which can be solved with any multi-label MRF optimizer, i.e. graph-cuts based algorithm [Boykov et al., 2001, Kolmogorov and Zabih, 2004, Boykov and Kolmogorov, 2004] - see Sec. 4.6.5 for a brief outline of the algorithm. Following the Hammersley-Clifford theorem, the segmentation can therefore be stated as an energy minimization problem given the label posterior [Dubes and Jain, 1989]

$$P(l_j = H_i | x_j, \mathcal{V}_j) = \frac{\exp\left[-U(H_i|x_j, \mathcal{V}_j)\right]}{\sum_{t=0}^{K-1} \exp\left[-U(H_t|x_j, \mathcal{V}_j)\right]}. \tag{4.76}$$

Formulated as a second-order isotropic Potts model [Kindermann et al., 1980, Dubes and Jain, 1989] we yield the corresponding energy term

$$U(H_i|x_j, \mathcal{V}_j) = -\kappa \ln \mathcal{LN}(x_j|\mu_i, \sigma_i) + \chi \sum_{j\prime \in \mathcal{V}_j} \delta(l_j, l_{j\prime}) \quad \text{s.t.} \quad \kappa, \chi \in \mathbb{R}^+, \tag{4.77}$$

consisting of unary and pairwise potentials. Thereby $\mathcal{V}_j$ denotes the neighborhood of point x_j, whereas δ is the Kronecker delta and l_j represents the label at site j. The unary potential controls the data consistency scaled by parameter κ, whereas the pairwise potentials associated term weights the label consistency in the neighborhood. As a consequence, a high parameter χ enforces larger and homogeneous regions.

4.6.4 MAP-MRF

The idea of maximum-a-posteriori MRF (MAP-MRF) is strongly related to the Bayes risk estimation [Li, 2009]. In this respect, for a particular label configuration $\{X = x\}$ the Bayes risk can be written as

$$R(\tilde{x}) = \int_{x\in\Omega} C(\tilde{x}, x) p(x|d) \mathrm{d}x = \int_{x\in\Omega} C(\tilde{x}, x) p(x|d) \mathrm{d}x, \tag{4.78}$$

where d corresponds to the observed data and $C(\tilde{x}, x)$ is the cost of obtaining a labeling x when the ground truth is $\tilde{x}$. Additionally, $p(x|d)$ corresponds to the probability of a certain labeling given the data. A popular choice for the cost function is

$$C(\tilde{x}, x) = \begin{cases} 0, \text{ if } \|\tilde{x} - x\| \leq \delta \\ 1, \text{ otherwise }, \end{cases} \tag{4.79}$$

such that the associated Bayes risk can be formulated as

$$R(\tilde{x}) = \int_{x:\|\tilde{x}-x\|>\delta} p(x|d) \mathrm{d}x = 1 - \int_{x:\|\tilde{x}-x\|\leq\delta} p(x|d) \mathrm{d}x \underset{\delta\to 0}{\approx} 1 - \tau p(x|d), \tag{4.80}$$

where τ denotes the volume containing points for which the condition $\|\tilde{x} - x\| \leq \delta$ holds. As a result, minimizing the risk is equal to maximization of the posterior probability $p(x|d)$. Further, making use of the Bayes theorem

$p(x|d) = \frac{p(d|x)p(x)}{p(d)}$, this can be rewritten as

$$\hat{x} = \arg\max_{x \in \Omega} \underbrace{p(d|x)}_{\text{Likelihood}} \underbrace{p(x)}_{\text{Prior}} \tag{4.81}$$

Combining that with the findings of the Hammersley-Cliffort theorem

$$p(x|d) \propto e^{-U(x|d)}, \tag{4.82}$$

whereby $U(x|d)$ decomposes as

$$U(x|d) = U(d|x) + U(x), \tag{4.83}$$

Eq. 4.81 can be rewritten such that it is formulated w.r.t to energy terms $U(.)$ as maximization of posterior energy,

$$\hat{x} = \arg\max_{x \in \Omega} U(x|d). \tag{4.84}$$

For discrete labels this corresponds to the Generalized Potts model

$$U(x|d) = -\sum_{\{i,i'\} \in \mathcal{C}_2} \underbrace{J_{ii'}\delta(x_i, x_{i'})}_{\text{Likelihood: } \log p(d|x)} + \sum_{i \in \mathcal{S}} \underbrace{H_i x_i}_{\text{Prior: } \log p(x)}, \tag{4.85}$$

where H_i and $J_{ii'}$ correspond to unary and pairwise potential scalars. In this respect, the unary term corresponds to the data fidelity whereas the pairwise term assigns a cost with respect to the neighborhood labeling. For homogeneous fields, where parameters are independent of spatial location, we have $H_i \equiv H$ and $J_{ii'} \equiv J$, yielding

$$U(x|d) = -\sum_{\{i,i'\} \in \mathcal{C}_2} J\delta(x_i, x_{i'}) + \sum_{i \in \mathcal{S}} H x_i, \tag{4.86}$$

which represents an isotropic Potts model.

4.6.5 Energy Minimization via Graph Cuts

The underlying idea of the graph cuts based energy minimization algorithm, as proposed in [Boykov et al., 2001, Kolmogorov and Zabih, 2004, Boykov and Kolmogorov, 2004], is that an optimal solution i.e. a specific binary labeling configuration, can be obtained by solving the minimum-cut/maximum flow problem. In this respect, minimum cut solving refers to the problem of finding a binary partition in a graph, which has the minimum cost among all possible cuts. The flow across the minimum-cut thereby corresponds to the maximum possible flow, which is stated in the famous maximum-flow/minimum-cut theorem. Specifically, this corresponds to finding solutions for energy terms of type

$$E(L) = \sum_{p \in \Omega} D_p(L_p) + \sum_{(p,q) \in \mathcal{V}} V_{p,q}(L_p, L_q), \tag{4.87}$$

where $L = \{L_p | p \in \mathcal{S}\}$ is a particular label configuration of image $\mathcal{S}$. Here $D_p(.)$ corresponds to the unary data term and $V_{p,q}$ to the pairwise interaction

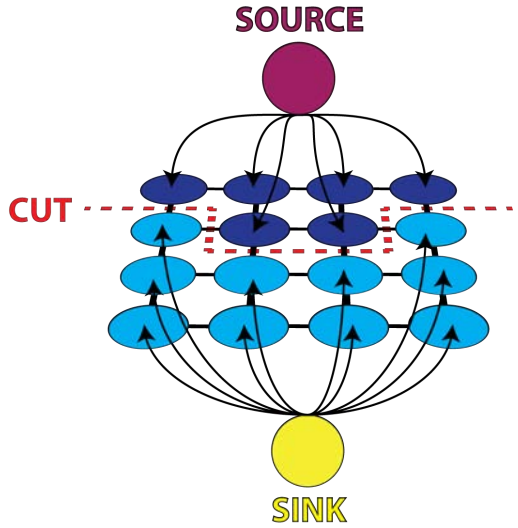


Figure 4.14: Illustration of graph-cuts. The cut through the graph induces a particular label configuration. The binary labeling is achieved by connecting exclusively to one of the terminal nodes.

potentials, where $\mathcal{V}$ indicates the neighboring elements. Similarly to binary label problems, graph cuts methods can be generalized and applied to multi-labeling problems, however, requiring the transformation of the problem. In case of the *alpha-expansion* algorithm, multi-label problems require solving a series of iterative binary maximizations. In this respect, it is noteworthy that in case of *alpha-expansion*, the binary labeling problem can be solved globally optimal, whereas the multi-label problem can only be solved locally optimal. Basic idea behind the algorithm is that cuts through the graph give rise to a particular labeling. Thus the inherent binary nature of the problem assumes that the image is represented as a graph, where each pixel/node is initially connected to a sink and a source, whereby both represent a particular class association. Any cut through the graph will then sever for any node/pixel in the graph the connection to a terminal node - either to sink or source. This link severing or terminal node disconnection then in turn induces a particular label configuration, as nodes are associated with the label class with retained link. The min-cut/max-flow label then gives rise to an optimal solution. Specifically, in each iteration pixels decide whether to change labels or keep the old label configuration, depending on a data term and the neighborhood configuration. The neighborhood configuration, however, is subject to change from move to move due to the relabeling operations. See Fig.4.14 for a schematic illustration of a binary partition of an image with corresponding connections to source and sink terminal nodes.

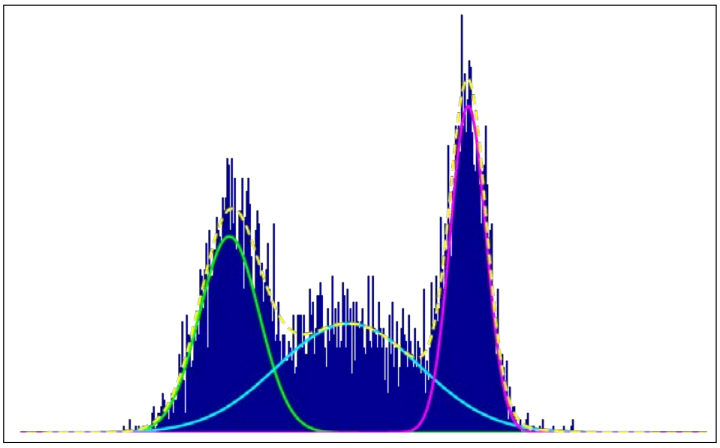


Figure 4.15: Illustration of Gaussian mixture model fitted to data. The mixture model consists of three components, represented with solid lines. The dashed line represents the weighted sum of all components, whereas the multi-modal data is shown in dark blue.

4.6.6 Expectation Maximization for Gaussian Mixture

In the following we will give a brief outline the Expectation-Maximization (EM) algorithm specifically showcased on a Gaussian mixture model (see Eq. 4.75). The EM algorithm is an iterative method for finding maximum likelihood (ML) (or in a specific adaptation the maximum a posteriori) estimates of distributional parameters in statistical models. Specifically, the algorithm tries to determine the distributional components modeling the data as well associate the data to the respective source distributions. See Fig. 4.15 for an illustration of data arising from a Gaussian mixture model as well as the visualization of the components and the distributional model. The underlying model thereby implicitly assumes unobserved or latent variables that basically function as indicator variables defining the data source association. In the iterative estimation process of the algorithm, the latent variables are modeled probabilistically.
The EM algorithm basically consists of two fundamental steps, whose execution is alternated during the iterative estimation process: (1) an expectation (E) step, which computes the expectation of the log-likelihood evaluated using the estimated parameters of the previous iteration (or the initial values), and (2) a maximization (M)-step, which computes parameters by maximizing the expected log-likelihood found in the preceding E-step. In the next iteration, the resulting parameter-estimates from the M-step are then again used to determine the distribution of the latent variables in the E-step, and so on until convergence is reached.
It should be noted that for the derivation of the EM algorithm there exist two different views, where we largely follow the presentation given in [Bishop, 2009]. At first we will focus on the more intuitive view, whereas the alternative

general view will take more in account the role of the latent variable.

Input to the EM algorithm is a data set $X = \{x_1, ..., x_N\}$ as well as the specification of the number of components K, from which in case of the Gaussian distribution the following parameters (see Eq. 4.75) $\theta = \{\pi_i, \mu_i, \sigma_i\}_{i=1..K}$ are recovered, where μ_i and σ_i^2 indicate mean and variance, respectively. Implicitly, the data will then be augmented by an unknown K-dimensional binary latent vector $z_k \in \{0, 1\}$, under the constraint $\sum_{k=1}^{K} z_k = 1$, which for each single data point indicates component association. For simplicity and sake of presentation we assume a single point for the moment. At first we can define the marginal distribution over z

$$p(z_k = 1) = \prod_{k=1}^{K} \pi_k^{z_k} = \pi_k, \tag{4.88}$$

which corresponds to the component weight of the mixture model. Furthermore, we can define the conditional distribution of x given the latent variable z, which is defined as

$$p(x|z_k = 1) = \prod_{k=1}^{K} \mathcal{N}(x|\mu_k, \sigma_k)^{z_k} = \mathcal{N}(x|\mu_k, \sigma_k). \tag{4.89}$$

From this we yield a joint distribution $p(z)p(x|z)$ of the data. Furthermore, the latent variables can be combined for the mixture representation to

$$p(x) = \sum_{z} p(z)p(x|z) = \sum_{k=1}^{K} \pi_k \mathcal{N}(x|\mu_k, \sigma_k). \tag{4.90}$$

Additionally, we can define the conditional probability of z given the data x as

$$p(z_k = 1|x) = \frac{p(z_k = 1)p(x|z_k = 1)}{\sum_{j=1}^{K} p(z_k = j)p(x|z_k = j)} = \frac{\pi_k \mathcal{N}(x\ |\mu_k, \sigma_k)}{\sum_{j=1}^{K} \pi_k \mathcal{N}(x\ |\mu_k, \sigma_k)} = \gamma(z_k), \tag{4.91}$$

which denotes the probability of a point x belonging to class k, such that $z_k = 1$ and $\forall j \neq k, z_j = 0$. Having specified the distributional models with the underlying latent variable, next we focus on the mixture estimation process itself. Goal of the EM algorithm is to find the distributional parameters that represent the most probable solution. Therefore the EM algorithm specifically seeks the maximization of the log-likelihood

$$p(X|\theta) = \sum_{i}^{N} \left[\log \sum_{j}^{K} \pi_i \mathcal{N}(x_i|\mu_j, \sigma_j) \right] \tag{4.92}$$

w.r.t. the parameters θ given data $X = \{x_1, ..., x_N\}$. This requires zero-setting of the derivatives w.r.t. to the parameters to be optimized. The zero-set derivative of the log-likelihood and the parameter of the mean μ_k yields

$$0 = -\sum_{i=1}^{N} \underbrace{\frac{\pi_k \mathcal{N}(x_i|\mu_k, \sigma_k)}{\sum_{j=1}^{K} \pi_j \mathcal{N}(x_i|\mu_j, \sigma_j)}}_{\gamma(z_{ik})} \sigma_k (x_i - \mu_k), \tag{4.93}$$

where the parameters $\gamma(z_{ik})$ can be seen as the proportion that component k takes in explaining the observation x_i w.r.t. to all other components [Bishop, 2009]. In other words it equals the class association probability of point i w.r.t. to class k. Strictly speaking, $\gamma(z_{ik})$ should be written as $\gamma(z_{ik}^{(t)})$, to highlight the fact that the computation depends on parameters computed at time-step (t), such that $\gamma(z_{ik})$ is treated as a constant for the steps to follow. Re-arranging of the variables leads to

$$\mu_k = \frac{1}{N_k} \sum_{i=1}^{N} \gamma(z_{ik}) x_i \tag{4.94}$$

$$N_k = \sum_{i=1}^{N} \gamma(z_{ik}), \tag{4.95}$$

where N_k can be seen as the effective number of points associated with class k. In the same line, zero-setting of the derivative w.r.t. to σ_k results in

$$\sigma_k = \frac{1}{N_k} \sum_{i=1}^{N} \gamma(z_{ik})(x_i - \mu_k)^2. \tag{4.96}$$

Finally, we need the zero-set derivative with respect to the mixing weights π_k. However, here the constraint of $\sum_{i=1}^{K} \pi_k = 1$ must be incorporated, which can be achieved using the Lagrange multipliers λ. Altogether, we need to calculate derivative of the following augmented term

$$\log p(X|\theta) + \lambda \left(\sum_{k=1}^{K} \pi_k \right), \tag{4.97}$$

yielding

$$0 = \sum_{i=1}^{N} \frac{\mathcal{N}(x_i|\mu_k, \sigma_k)}{\sum_{j=1}^{K} \pi_j \mathcal{N}(x_i|\mu_j, \sigma_j)} + \lambda. \tag{4.98}$$

Multiplying both sides by π_k and summing over k, making use of the summed weight constraint, one finds $\lambda = -N$. This relation can be used to remove the Lagrange multipliers such that after re-arrangement of variables we yield

$$\pi_k = \frac{N_k}{N}, \tag{4.99}$$

which states that the mixing coefficient corresponds to the average responsibility the corresponding component takes in explaining the data points. Although it might appear is if Eq. 4.94 - 4.96 and 4.99 might permit a closed-form solution that is clearly not the case. This is due to the dependency of those terms on $\gamma(.)$, which in turn again depend on those parameters. Therefore, as mentioned above, introducing the time-step indicator (t) on the $\gamma(z_{ik}^{(t)})$ variables makes this dependency much clearer. However, this intrinsic dependency gives rise to an iterative solution process, which is exactly what the EM algorithm is. Based on some initial parameters θ_{init}, the distribution of hidden parameters are computed, corresponding to the E-step. Given the current class association

probability of each point, represented by $\gamma(z_{ik})$, the maximum-likelihood or maximum-a-posteriori estimate of the mixture parameters is performed, corresponding to the M-step. This is repeated until convergence. See Fig. 4.4 for an outline of the complete EM algorithm.

Algorithm 4.4 Expectation Maximization for Gaussian Mixture Model

Input: Data $X = \{x_1, ..., x_N\}$
Definite initial values: $\theta^{(0)} = \left\{\pi_i^{(0)}, \mu_i^{(0)}, \sigma_i^{(0)}\right\}_{i=1..K}$
repeat
$t \leftarrow t + 1$
E-Step: Compute class association probabilities/responsibilities

for all $i \in N, \quad k \in K$ **do**

$$\gamma(z_{ik}^{(t)}) = \frac{\pi_k^{(t-1)} \mathcal{N}(x_i | \mu_k^{(t-1)}, \sigma_k^{(t-1)})}{\sum_{j=1}^{K} \pi_j^{(t-1)} \mathcal{N}(x_i | \mu_j^{(t-1)}, \sigma_j^{(t-1)})}$$

end for

M-Step: Estimate the the parameters given the current class associations/responsibilities
for all $k \in K$ **do**

$$\mu_k^{(t)} = \frac{1}{N_k} \sum_{i=1}^{N} \gamma(z_{ik}^{(t)}) x_i$$

$$\sigma_k^{(t)} = \frac{1}{N_k} \sum_{i=1}^{N} \gamma(z_{ik})^t (x_i - \mu_k^{(t)})^2$$

$$N_k^{(t)} = \sum_{i=1}^{N} \gamma(z_{ik}^{(t)})$$

$$\pi_k^{(t)} = \frac{N_k^{(t)}}{N},$$

end for
until Convergence of parameters or log-likelihood

4.6.7 General View of Expectation Maximization

In the previous section we already introduced the latent variables z_k, which is a binary K-vector indicating class association with constraint $\sum_{k=1}^{K} z_k = 1$. Now suppose we are given a latent variable for each data point $X = \{x_1, ..., x_N\}$ then we have $Z = \{z_1, ..., z_N\}$. In this respect, we refer to the data set $\{X, Z\}$ as the *complete data*. However, typically we are not given the the latent variables, therefore we refer to $\{X\}$ as the *incomplete data*. As a result, the log-likelihood for the *complete data* is $p(X, Z|\theta)$ and for the incomplete data $p(X|\theta)$. Since

we are not given the latent variable, the only information we are left with is its posterior $p(Z|X,\theta)$. As as consequence we cannot compute the complete data log-likelihood directly. However, what we can do is approximate it by computing its expected value, namely $E\left[\log p(X,Z|\theta)\right]$, where $E\left[.\right]$ denotes the expectation.
In case of a Gaussian mixture model, we let $\theta = \{\pi_i, \mu_i, \sigma_i\}_{i=1..K}$ and $\psi = \{\mu_i, \sigma_i\}_{i=1..K}$, where again μ_i and σ_i^2 denote the mean and variance, respectively. Then we can write the expected log-likelihood as [Lindgren, 2002]

$$\mathcal{Q}(\theta,\theta^t) = E\left[\log p(X,Z|\theta)|X,\theta^{(t)}\right] = \sum_Z p(Z|X,\theta^{(t)}) \log p(X,Z|\theta). \quad (4.100)$$

Furthermore, we have the incomplete data likelihood defined as,

$$p(X|\theta) = \prod_{i=1}^{N}\sum_{k=1}^{K} p(y_i|z_{ik}=1,\psi_k)\pi_k \quad (4.101)$$

or in case of known class association the complete data likelihood

$$p(X,Z|\theta) = \prod_{i=1}^{N} p(x_i|z_i,\psi_{x_i})\pi_{x_i}. \quad (4.102)$$

Then it follows that the complete data likelihood is defined as,

$$\log p(X,Z|\theta) = \sum_{i=1}^{N} \left(p(x_i|z_i,\psi_{x_i}) + \log \pi_{x_i}\right) \quad (4.103)$$

and the class posterior distribution posterior for $Z|X,\theta^{(t)}$

$$\gamma(z_{ik}) = p(z_{ik}=1|X,\theta^{(t)}) = \frac{p(x_i|z_{ik}=1,\psi_k^{(t)})\pi_k^{(t)}}{p(x_i|\theta^{(t)})} = \frac{p(x_i|z_{ik}=1,\psi_k^{(t)})\pi_k^{(t)}}{\sum_{j=1}^{K} p(x_i|z_{ij}=1,\psi_j^{(t)})\pi_j^{(t)}}, \quad (4.104)$$

where we again use (t) to highlight the parameters that are estimated in previous time step t. This all taken together corresponds exactly to the E-step of the EM algorithm. Following the computation of the expectation of the log-likelihood, we maximize the log-likelihood w.r.t. to the distributional parameters θ, which corresponds to the M step given by

$$\hat{\theta} = \arg\max_{\theta} \mathcal{Q}(\theta,\theta^{(t)}). \quad (4.105)$$

Applying the general deduction from above to the specific case of a Gaussian mixture we have the following likelihood term

$$p(X,Z|\theta) = \prod_{i=1}^{N}\prod_{k=1}^{K} \pi_k^{z_{ik}} \mathcal{N}(x_i|\psi_k)^{z_{ik}} = \prod_{i=1}^{N}\prod_{k=1}^{K} \pi_k^{z_{ik}} \mathcal{N}(x_i|\mu_k,\sigma_k)^{z_{ik}} \quad (4.107)$$

Algorithm 4.5 General Expectation Maximization

Input: Data $X = \{x_1, ..., x_N\}$
Definite initial values: $\theta^{(0)} = \left\{\pi_i^{(0)}, \mu_i^{(0)}, \sigma_i^{(0)}\right\}_{i=1..K}$
repeat
$t \leftarrow t + 1$
E-Step: Compute $p(Z|X, \theta^{(t-1)})$, in particular $\forall i \in N, k \in K$ compute $\gamma(z_{ik}^{(t)})$.
M-Step: Evaluate θ^t

$$\theta^t = \arg\max_{\theta} \mathcal{Q}(\theta, \theta^{(t-1)})$$
$$\mathcal{Q}(\theta, \theta^{(t-1)}) = E\left[\log p(X, Y|\theta)|X, \theta^{(t-1)}\right] = \sum_Z p(Z|X, \theta^{(t-1)}) \log p(X, Z|\theta). \tag{4.106}$$

until Convergence of parameters or log-likelihood

with the corresponding log-likelihood

$$\log p(X, Z|\theta) = \sum_{i=1}^{N} \sum_{k=1}^{K} z_{ik} \left[\log \pi_k + \log \mathcal{N}(x_i|\psi_k)\right]. \tag{4.108}$$

The maximum likelihood parameter estimates are obtained analogously to previous Sec. 4.6.6 from the derivative of the log-likelihood.

The class posterior distribution posterior for $Z|X, \theta^{(t)}$ is obtained from

$$p(Z|X, \theta) \propto p(Z) \cdot p(X|Z) = \prod_{i=1}^{N} \pi_k^{z_k} \prod_{k=1}^{K} \mathcal{N}(x_i|\psi_k)^{z_k} = \prod_{i=1}^{N} \prod_{k=1}^{K} \left[\pi_k \mathcal{N}(x_i|\psi_k)\right]^{z_{ik}}. \tag{4.109}$$

Furthermore, it follows that the expected value of the class indication variable z_{ik} is defined as

$$E\left[z_{ik}\right] = \frac{\sum_{z_{ik}} z_{ik} \left[\pi_k \mathcal{N}(x_i|\psi_k)\right]^{z_{ik}}}{\sum_{k=1}^{K} \left[\pi_k \mathcal{N}(x_i|\psi_k)\right]^{z_{ik}}} = \frac{\pi_k \mathcal{N}(x_i|\psi_k)}{\sum_{k=1}^{K} \left(\pi_k \mathcal{N}(x_i|\psi_k)\right)} = \gamma(z_{ik}), \tag{4.110}$$

which is exactly what is computed $\forall i \in N, k \in K$ during the E-step. Finally, we yield the expected value of the complete-data log-likelihood as

$$E\left[\log p(X, Z|\theta)\right] = \sum_{i=1}^{N} \sum_{k=1}^{K} \gamma(z_{ik}) \cdot \left[\log \pi_k + \log \mathcal{N}(x_i|\psi_k)\right], \tag{4.111}$$

which is maximized during the M-step.

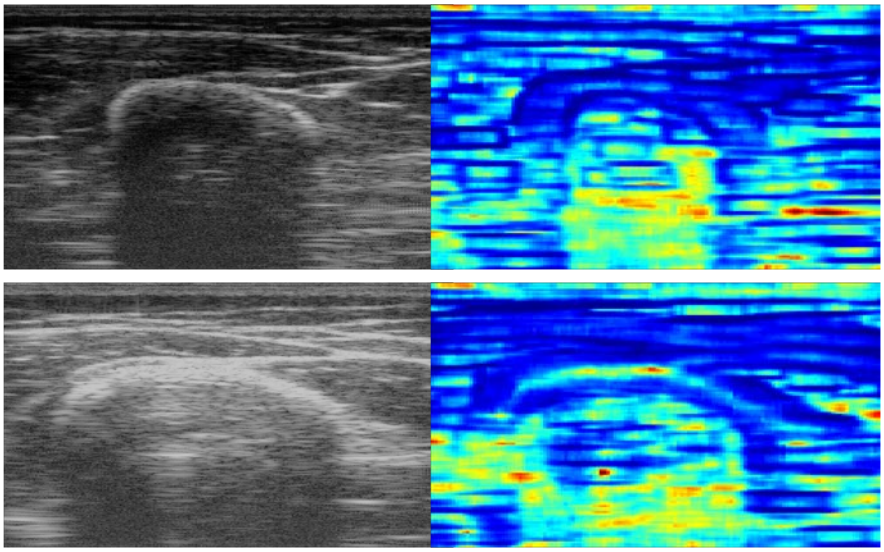


Figure 4.16: Feature descriptor component for lateral direction $\xi_2 \in \psi$. Shadow regions are highly expressed within the feature descriptor domain

4.6.8 RF Shadow Descriptor

The distribution of the envelope of the RF signal, resulting from backscattered tissue echo, has been shown to be modeled by the Nakagami distribution [Shankar, 2000]. Thus we assume all intensities in the RF envelope 2D image Ω to follow a Nakagami distribution

$$\mathcal{N}(x|\mu,\omega) = \frac{2\mu^{\mu}x^{2\mu-1}}{\Gamma(\mu)\omega^{\mu}} \exp\left(-\frac{\mu}{\omega}x^2\right), \forall x \in \mathbb{R}_+ \tag{4.112}$$

with μ, ω the shape and scale parameters, respectively. The Nakagami distribution is part of the exponential family of distribution, and as such we can model the spatial statistics of the RF envelope image, by an auto-model as described in Sec. 4.3.4 and 4.3.5.

$$\log \mathcal{N}(x_i|\mu,\omega) = A(x_{\mathcal{V}_i})B(x_i) + C(x_i) + D(x_{\mathcal{V}_i}) \tag{4.113}$$

$$A(x_{\mathcal{V}_i}) = \lambda + \sum_{i \neq i'} \xi_{i,i'} B_i(x'_i) \; . \tag{4.114}$$

Parameters $A(.), D(.)$ depend on the intensities $x_{\mathcal{V}_i}$ in the neighborhood $\mathcal{V}_i$ of the central pixel x_i, whereas $B(.), C(.)$ depend only on x_i. Furthermore, λ is the weight of the central site x_i and $\xi_{i,i'}$ are the interaction weights between neighboring pixels. Combining the eqns. 4.114 results in

$$\mu = \frac{1}{2}(\lambda + 1 + \sum_{i \neq i'} \xi_{i,i'} \log x'_i) \; . \tag{4.115}$$

Inspired by [Bouhlel and Sevestre-Ghalila, 2009], we use the interaction parameters of the auto-model as a feature-descriptor for shadow. This model applies

particularly well for shadow, as those regions have a specific and homogeneous texture pattern. The RF descriptor characterizes small subpatches of the image domain, as follows. A sparse 2D grid is placed in the image Ω subdividing the data into subpatches $\mathcal{P}_i \in \Omega$, s.t. $\bigcup_i \mathcal{P}_i = \Omega$. For *each* subpatch $\mathcal{P}_i$ the following steps are performed: **1)** distribution parameters are obtained by a Nakagami MLE (as whole representing a Nakagami image) **2)** a dense 2D grid is placed on the subpatch, with $\mathcal{S} \subset \mathcal{P}_i$ denoting the set of all grid points **3)** centered on each grid point $s \in \mathcal{S}$, an MRF interaction model given by Eq. 4.115 is instantiated. In this MRF model, cliques are taken to be spatially opposite segments of a second-order neighborhood centered on sub-grid points, similar to the work proposed in [Bouhlel and Sevestre-Ghalila, 2009].

Finally interaction parameters for all subpatches $\mathcal{P}_i$ are computed using the conditional-least-squares (CLS) approach [Bouhlel and Sevestre-Ghalila, 2009], as presented in Sec. 4.3.6. For each subpatch with a neighborhood system of n cliques, an interaction feature vector

$$\psi = \left(\lambda \ , \ \xi_1 \ , \ \ldots \ , \ \xi_n\right) \in \mathbb{R}^{n+1} \ , \tag{4.116}$$

is generated. The λ parameters can be related to the density and amplitudes of scatterers, while the ξ parameters represent the interaction of the diametrically opposite neighbor pixels influencing the center pixel and therefore are closely related to scatterer spacing [Bouhlel and Sevestre-Ghalila, 2009]. Noteworthy within the context of shadow detection is thereby that only the feature corresponding to lateral interaction ξ_2 is of interest. As can be seen in Fig. 4.16, this feature becomes pronounced in homogeneous shadow areas.

Following CLS, the interaction weights are obtained by minimizing the variance between observed and expected intensities

$$\hat{\psi} = \underset{\{\lambda,\xi_{i,i'}\}}{\operatorname{argmin}} \sum_i (x_i^2 - E[X_i^2])^2 \quad \text{s.t.} \quad E[X_i^2] = \frac{1}{2\omega_i}(\lambda + 1 + \sum_{i \neq i'} \xi_{i,i'} \log x_i') \tag{4.117}$$

with $E[.]$ being the expected value. The solution for Eq. 4.117, yielding the interaction feature vector ψ (see Eq. 4.116), is obtained by least-squares.

4.6.9 Shadow Classification

In order to classify shadow predominant regions, we make use of the observation that in those areas the MRF-feature corresponding to the speckle interaction in lateral direction (see Eq. 4.116) becomes highly expressed - see Fig. 4.16. At first the feature map is classified using a $K = 2$ Gaussian finite mixture model (GMM)

$$\begin{aligned} p(x|\theta) = \sum_i^K \pi_i \mathcal{N}(x|\mu_i, \sigma_i) = \sum_i^K \frac{w_i}{\sqrt{2\pi}\sigma_i x} \exp\left(-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right) \\ \text{s.t.} \quad \sum_{k=1}^K \pi_k = 1, \end{aligned} \tag{4.118}$$

which analogously to Sec. 4.6.3 is estimated using the EM algorithm [Dempster et al., 1977]. See Sec. 4.6.6 for details on the EM algorithm for Gaussian

Expert	**P1** 5MHz	**P1** 10MHz	**P2** 5MHz	**P2** 10MHz	**P3** 5MHz	**P3** 10MHz
# 1	0.84 ± 0.05	0.87 ± 0.02	0.84 ± 0.04	0.9 ± 0.01	0.89 ± 0.03	0.91 ± 0.03
# 2	0.78 ± 0.04	0.82 ± 0.02	0.83 ± 0.05	0.88 ± 0.01	0.87 ± 0.03	0.86 ± 0.03
GT Dice	0.92 ± 0.03	0.93 ± 0.01	0.96 ± 0.02	0.95 ± 0.01	0.94 ± 0.01	0.93 ± 0.01

Table 4.1: Dice coefficients for shadow regions on basis of manual segmentations by two experts for six patient datasets (PX = patient X) - Avg.: 0.86 ± 0.03. Last row provides inter-user variability in terms of Dice coefficient - Avg.: 0.94 ± 0.02.

mixture estimation. Performing this Gaussian mixture model yields a binary clustering of feature activity, consisting of the two classes *inactive* C_0 and *active* C_1. Looking at the shadow regions one can see that within those areas the ratio of active features is higher compared to inactive ones. Only reverberations and mirroring next to the reflection causing interface interfere that characteristic feature activity pattern. As a result, next to the reflector the feature ratio is not higher, as those regions from their patterns resemble regular image texture. See Fig. 4.16 for example images highlighting the feature activity in shadow regions. Given the characteristic activity patterns and the interference, we propose a distance weighted score function S accounting for those conditions, which is defined as

$$S = \frac{\log(w)}{h} \sum_{x}^{h} \left[\frac{C_N(x)}{w} e^{(\Delta_N \cdot x + D_{N,1})} + \frac{C_R(x)}{w} (\Delta_R \cdot x + D_{R,1}) \right], \Delta_i = \frac{D_{i,2} - D_{i,1}}{h}.$$

Summation is performed over the row space of each segment of size $[0, w] \times [0, h]$. Furthermore, $C_1(x)$ and $C_R(x)$ denote the number of features in row x from the specific class of shadow noise or from reflector, respectively. Whereas C_1 is obtained from the Gaussian mixture model as mentioned above to classify into active and inactive elements, $C_R(x)$ is obtained from the initial reflector segmentation - see Sec. 4.6.3 for details. At last, the parameters $D_{i,1}, D_{i,2}$ denote distance weighting scalars specific for shadow noise $i = N$ and reflectors $i = R$. Thereby we follow the notion that more distant from the reflector reverberations are less likely and do not disturb the shadow specific distribution patterns and vice versa.

4.6.10 Results

For testing the performance of the shadow detection in RF ultrasound, numerous data sequences of the Humerus were acquired from three patients using an Ultrasonix MDP machine (Ultrasonix Medical Corporation, Richmond, Canada) in combination with a linear transducer L14-5/38 GPS. Data was recorded with depth 4 cm yielding RF images of resolution 1304×256 pixels. For validation, manual segmentation of the shadow regions was performed in 50 frames for each patient dataset by two ultrasound experts. Image sequences for segmentation were selected to contain various mirroring/reverberation artifact

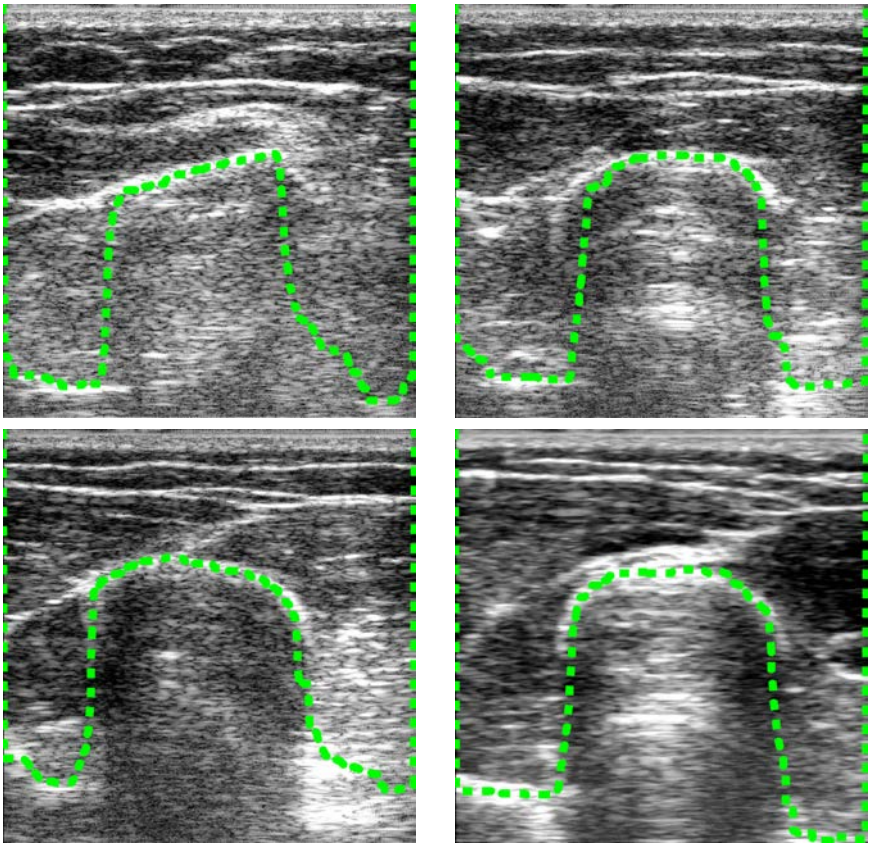


Figure 4.17: Segmentation results for different frames.

scenarios ranging from low to high. It should be noted that these artifacts naturally give rise to ambiguities, which can be seen in the consistent difference between expert segmentations - see Tab. 4.1. Nevertheless, comparison between ground-truth and automatic segmentation shows overall high Dice coefficients, which suggests accurate and robust shadow detection. Qualitative examples are shown in Fig. 4.17. Occasional slight drop in Dice coefficient is due the over-segmentation of reflectors caused by highly attenuating ligaments/muscular tissue attached to the bone that is barely distinguishable from bone itself, but not causing full reflection. The evaluation dataset was recorded at 5 MHz and 10 MHz to study the effect of frequency dependency, with the result indicating that frequency does not noticeably affect it. This further demonstrates the robustness of the method. Finally, comparison with the method in [Hellier et al., 2010] was conducted. However, the results obtained were inconsistent; producing strong over-segmentation of shadow areas. We omitted the results as a direct comparison seems not fair given that the other method was proposed for brain data only.

4.6.11 Conclusion

We presented a novel technique for detection of shadow in RF envelope detected ultrasound images. The combination of confidence estimation and statistical pattern analysis methods allows accurate detection, even in regions containing severe reverberation and mirroring artifacts. Evaluation on ultrasound image series of the Humerus demonstrates good results that are consistent with manual segmentation. Future work will focus on further evaluation of the method with different datasets, including femur and tibia acquisitions, and improved classification of strong reflectors. Furthermore, it is of interest to study the effect of integrating the proposed shadow detection in registration and 3D reconstruction approaches.

Chapter 5

Ultrasound Similarity Measures

In this chapter we present two similarity measures for 2D ultrasound. At first a hybrid method combining global statistics with local texture information referred to as Hybrid Local Binary Patterns (HLBP) is discussed. In this respect, for local statistics a textural descriptor based on an adaptation of Local Binary Patterns (LBP) is used. Furthermore, for global statistics a Hellinger distance for Gamma distributions is employed. The presentation of the hybrid approach is followed by an overview of ultrasound likelihood functions for registration. Finally, we present the locally adaptive Nakagami-based similarity measure based on its bivariate distribution. This chapter is based on the papers 'Locally Adaptive Nakagami-based Ultrasound Similarity Measures' [Wachinger et al., 2012b] and 'Registration of RF Ultrasound Data Using Hybrid Local Binary Patterns' [Klein et al., 2012a].

5.1 Introduction

Registration of images is a fundamental problem in several fields, particularly in the domain of medical imaging. However, issues such as noise and artifacts complicate this process and often make automatic registration processes intractable. In this respect, the alignment of ultrasound images is considered to be especially challenging due to the inherent omnipresence of speckle noise in the images. A further issue complicating image processing that is particularly prevalent in ultrasound imaging is view dependency and non-linear behavior due to its complex physical nature. Nevertheless, there exist a multitude of applications in ultrasound registration that require registration. Among the clinical applications for ultrasound registration are panorama imaging [Henrich et al., 2003, Wachinger et al., 2008], elastography [Salcudean et al., 2006, Basarab et al., 2007], tracking [Gee et al., 2006], and motion recovery [Cohen and Dinstein, 2002b]. Hence, the development of similarity metrics in ultrasound registration is an active field of research. A common approach in modeling this problem is to employ statistical methods. Unlike in the conventional vision domain, where the Gaussian intensity distribution has been successfully applied

in a variety scenarios, in ultrasound other models are more appropriate. In the domain of speckle tracking, the Rayleigh distribution is applied [Strintzis and Kokkinidis, 1997, Cohen and Dinstein, 2002b, Boukerroui et al., 2003, Revell et al., 2004]. For segmentation and classification applications, research on various distribution models for ultrasound scattering has been performed over the last years [Destrempes and Cloutier, 2010]. Numerous distributions were introduced that deal with varying numbers of scatterers per resolution cell and the presence of coherent structures [Goodman, 2007] - see also Sec. 2.4.

In the following we will focus on processing of envelope detected ultrasound data. The statistical properties of the echo envelope of ultrasound data depend on numerous factors. Among them is, in particular, the density and spatial distribution of scatterers in the medium. As different types of biological tissue exhibit various characteristics w.r.t. density and scatterer distribution, this can be utilized for registration purposes. Given enveloped RF data, the following models have been proposed as was discussed in Sec. 2.4. For fully developed speckle the Rayleigh distribution [Wagner et al., 1983] applies. The Rice distribution has been suggested [Wagner et al., 1983, Insana and et al., 1986] for the case of coherent, or structured, high density scatterers, while for partially developed speckle, the K-distribution [Jakeman, 1999, Shankar, 1993] has been shown to be appropriate. Moreover, general models have been proposed in the literature such as the generalized K-distribution and the homodyned K-distribution (see e.g., [Jakeman, 1999] and [Dutt and Greenleaf, 1994]) and most recently the Rician Inverse of Gaussian distribution [Eltoft, 2003, Eltoft, 2005] - see Sec. 2.4 for details on various distribution models. These models can account for a multitude of scattering conditions, however, at the price of comparatively high analytical complexity. As already mentioned at length in Sec. 2.4.2 because of its simplicity and versatility we focus on the Nakagami distribution [Shankar, 2000, Nakagami, 1960], which can account for varying scattering conditions. Furthermore, it is computationally efficient [Shankar, 2000] as compared to many other models that exhibit a tremendous complexity. As mentioned above, the advantage of the Nakagami distribution is to model various scattering scenarios, however, this requires to specify the parameters adequately. In this regard, recently a Nakagami-based similarity measure has proposed in [Myronenko et al., 2009b]. Since usually multiple scattering scenarios are present within an ultrasound image, depending on the depicted structures, we argue that a local adaptation increases the descriptiveness and better exploits the benefits of the Nakagami distribution.

In this chapter we present two ultrasound specific similarity measures. At first we present in Sec. 5.3 the Hybrid Local Binary Patterns (HLBP), which are a variant of the Local Binary patterns (LBP) presented in Sec. 4.5, employing confidence maps as introduced in Sec. 4.6.2. This is followed by a locally adaptive similarity measure based on the bivariate Nakagami distribution in Sec. 5.4, which further makes use of the 2D analytical signal as presented in Ch. 3.

5.2 Related Work

Registration of ultrasound images is addressed in a large number of articles [Strintzis and Kokkinidis, 1997, Krucker et al., 2002, Cohen and Dinstein, 2002b, Xiao et al., 2002, Boukerroui et al., 2003, Revell et al., 2004, Zikic et al., 2006, Poon and Rohling, 2005, Wachinger et al., 2007, Grau et al., 2007, Basarab et al., 2007, Elen et al., 2008, Esther Leung et al., 2008, Myronenko et al., 2009b]. Among the typical application domains are (i) motion measurements in echocardiography for detecting and characterizing abnormalities, (ii) breast deformation analysis to assess the elastic properties of tissues, (iii) assessment of tissue strain with elastography, and (iv) multi-view compounding. In this respect, statistics based registration is a common approach. Assuming multiplicative Rayleigh noise for ultrasound images, likelihood functions are presented in [Strintzis and Kokkinidis, 1997]. This notion is further extended in [Cohen and Dinstein, 2002b], assuming that both - the moving and the fixed image - are affected by multiplicative speckle noise. Futhermore, for improved performance on B-mode imagery, the log-compression is incorporated in the imaging model. In [Boukerroui et al., 2003, Revell et al., 2004] these similarity metrics are successfully applied for motion estimation employing a block matching approach. Similarly, in [Basarab et al., 2007], block matching is used in combination with normalized cross-correlation (NCC) for flow estimation and elasticity imaging. Furthermore, in [Krucker et al., 2002, Poon and Rohling, 2005], a block matching approach is applied in context of image quality improvement by combining various compounded images. In this respect, several similarity measures are evaluated in [Krucker et al., 2002], with the conclusion that sum of squared differences (SSD) is suitable only in imagery that is subject to low noise levels. However, in the work [Poon and Rohling, 2005] focus is on NCC for the compounding optimization. Instead of univariate data modeling, [Myronenko et al., 2009b] presented similarity measures based on the bivariate Rayleigh and Nakagami distributions. However, in contrast to previous approaches, [Myronenko et al., 2009b] do not work on B-mode images but on log-compressed envelope data. This is of importance due to the fact that B-mode images have corrupted speckle statistics as a result of a series of applied non-linear transformations and non-linear proprietary filter operations, as was discussed in detail in Sec. 2.3. Common to all the presented methods is that the parameters of the distributions are set heuristically on a global basis, which is at odds with the underlying local data variation. For instance, the variance of the Rayleigh distribution is set to $2/\pi$ in [Cohen and Dinstein, 2002b], or the shape and correlation parameters are set to $m = 0.5, \rho = 0.8$ in [Myronenko et al., 2009b].

Next to ultrasound specific approaches, we also want to mention related work in terms of general registration. A typical situation that challenges the application of mutual information is the registration of images, which contain high intensity non-uniformity due to the bias field. Approaches that address this issue try to estimate the joint density in local regions or do a combination of local and global estimation [Loeckx et al., 2010, Zhuang et al., 2009]. Recently, learning-based similarity measure were proposed. In [Babenko et al., 2009] a

similarity metric is learned using a boosting approach. Similarly, in [Xiong et al., 2012] Random Forest [Breiman, 2001] classification was employed for learning. Metric learning for multi-modal registration was proposed in [Lee et al., 2009, Bronstein et al., 2010]. In this respect, supervised learning is performed on previously registered data to learn an appropriate similarity function. However, these methods require access to correctly aligned data for training, which is a major drawback of these methods.

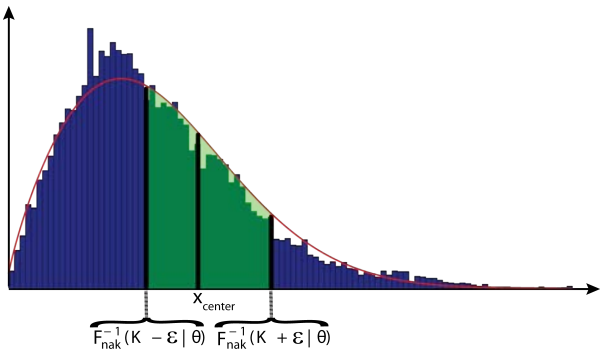


Figure 5.1: The threshold parameter ϵ defining a probability interval around the central site intensity value.

5.3 Hybrid Local Binary Patterns

5.3.1 Method

Given the noisy nature of ultrasound, using statistical measures for the sake of robustness seems evident. In order to improve the registration accuracy a hybrid approach is proposed that couples the global concept of distribution matching with a local one measuring texture patterns. We refer to this in the following as hybrid local binary pattern (HLBP). The two components of HLBP are based on the Gamma Hellinger distance metric, see Sec. 5.3.2, and Local Binary Patterns, see Sec. 4.5, respectively. Additionally, we make use of ultrasound confidence maps for parameter estimation, see Sec. 4.6.2.

5.3.2 Gamma Hellinger distance metric

We use the Nakagami distribution [Shankar, 2000, Nakagami, 1960] for modeling the speckle distribution in envelope detected RF ultrasound data. The Nakagami distribution is closely related to the Gamma distribution, and so may be used instead, by application of a simple transformation of data (see Sec. 2.4.2 and 2.53), specifically

$$\begin{aligned} \text{if} \quad Y \sim f_{gam}(x \mid m,k) \quad &\text{and} \quad X \sim f_{nak}(x \mid \mu,\omega) \\ &\text{then} \quad \sqrt{X} = Y(\mu, \frac{\omega}{\mu}) \,, \end{aligned} \tag{5.1}$$

where f_{gam} and f_{nak} denotes the Gamma and Nakagami PDF, respectively.

For measuring the difference between two densities there exist a multitude of possibilities. One option in distribution matching is the Hellinger distance metric. Basis for this metric is that for any family of densities $\{p(x|\theta)\}$, the set $\left\{\sqrt{p(.|\theta)}, \theta \in \Theta\right\}$ represents a parametric curve on the unit sphere in L_2-space, as can be seen from the basic property of [Korostelev and Korosteleva, 2011]

$$1 = \int_{\mathbb{R}} p(x|\theta)\mathrm{d}x = \int_{\mathbb{R}} \left(\sqrt{p(x|\theta)}\right)^2 \mathrm{d}x = \|\sqrt{p(.|\theta)}\|_2^2, \tag{5.2}$$

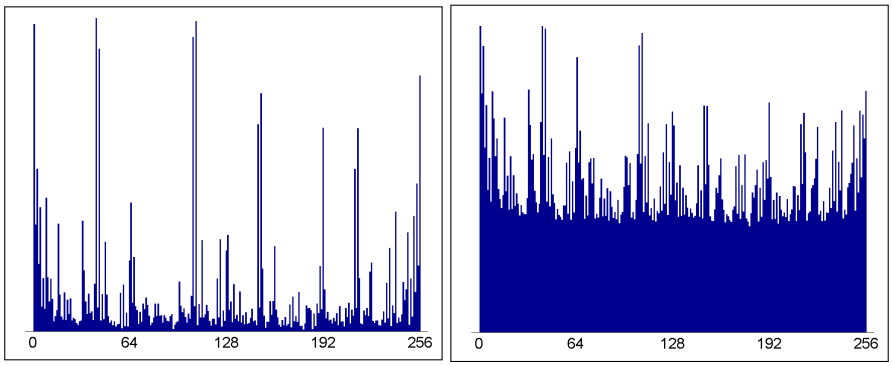


Figure 5.2: Feature histograms for low (*left*) and high (*right*) threshold.

where $\|.\|_2$ denotes the L_2-norm. The squared Hellinger metric $H^2(\theta_0, \theta_1)$ for general distributions $p(.|\theta_0)$ and $p(.|\theta_1)$ is then defined as

$$H(\theta_0, \theta_1) = \| \sqrt{p(x|\theta_0)} - \sqrt{p(x|\theta_1)} \|_2^2 = 2 \left(1 - \int_{\mathbb{R}} \sqrt{p(x|\theta_0) p(x|\theta_1)} \mathrm{d}x \right), \quad \theta_0, \theta_1 \in \Theta, \tag{5.3}$$

from which follows that $0 \leq H(\theta_0, \theta_1) \leq \sqrt{2}$ (sometimes the Hellinger is defined normalized in that case one has $0 \leq H(\theta_0, \theta_1) \leq 1$).

In the following we employ the Hellinger distance H_{gam} between two Gamma probability distributions F and G, as expressed by their associated probability density functions $f \sim f_{\text{gam}}(x|m_1, k_1)$ and $g \sim f_{\text{gam}}(x|m_2, k_2)$ as

$$H_{gam}(f, g) = 1 - \frac{\Gamma\left(\frac{(m_1+m_2)}{2}\right)}{\left[\Gamma(m_1)\Gamma(m_2)\right]^{\frac{1}{2}}} \cdot \frac{\left(\frac{(k_1+k_2)}{2}\right)^{-\left(\frac{m_1+m_2}{2}\right)}}{\left[\frac{1}{k_1^{m_1}} \cdot \frac{1}{k_2^{m_2}}\right]^{\frac{1}{2}}}, \tag{5.4}$$

with $H_{gam}(f, g) \in [0, 1]$. We apply Hellinger distance alone and in conjunction with Local Binary Patterns (see Sec. 5.3.3) in registration of envelope detected ultrasound data.

Statistics-based Membership Function

Unlike the original Fuzzy Local Binary Patterns (FLBP) [Iakovidis et al., 2010], which defines the membership function as a ramp, we propose a non-linear function associated with the underlying statistical properties of data. Let

$$f_{\text{nak}}(x \mid \mu, \omega) = \frac{2\mu^{\mu} x^{2\mu-1}}{\Gamma(\mu)\omega^{\mu}} \exp\left(-\frac{\mu}{\omega} x^2\right), \forall x \in \mathbb{R}_+ \tag{5.5}$$

be the probability density function of the Nakagami distribution with μ, ω the shape and scale parameters, respectively. Its corresponding cumulative density function (CDF) is

$$F_{\text{nak}}(x \mid \theta) = \frac{\gamma(\mu, \frac{\mu}{\omega}x^2)}{\Gamma(\mu)}\,, \tag{5.6}$$

with $\theta = \{\mu, \omega\}$ and γ being the lower incomplete gamma function, defined as

$$\gamma(s, x) = \int_0^x t^{s-1}e^{-t}\,dx. \tag{5.7}$$

Following the idea of FLBP, a membership function $m_j(x)$ is defined, denoting the confidence of a class association $j \in \{0, 1\}$ given intensity x. Membership symmetry is assumed, i.e. $m_0(x) = 1 - m_1(x)$. Here the membership function m_1 is non-linear, given by

$$m_1(x_i) = \begin{cases} 1, & x_i > F^{-1}_{\text{nak}}(K + \epsilon|\theta), \\ \frac{F_{\text{nak}}(x_i|\theta) - K}{2\epsilon} + \frac{1}{2}, & x_i \in (x_{\text{cntr}}, F^{-1}_{\text{nak}}(K + \epsilon|\theta)] \\ \frac{K - F_{\text{nak}}(x_i|\theta)}{2\epsilon}, & x_i \in [F^{-1}_{\text{nak}}(K - \epsilon|\theta)], x_{\text{cntr}}] \\ 0, & x_i < F^{-1}_{\text{nak}}(K - \epsilon|\theta) \end{cases}, \tag{5.8}$$

where $K = F_{\text{nak}}(x_{\text{cntr}}|\theta)$, and x_{cntr} and x_{i} denote the intensity at the center site and site i, correspondingly. This definition follows the notion that the uncertainty in membership should be associated with the underlying data distribution. Thereby the probability adapts to varying distributions, rather than being tied to a fixed parameter. As a result, within the tail of a distribution a wider capture range of membership is allowed, whereas in areas next to the mode, membership association is more concentrated. In this respect, in the fuzzy region $[F^{-1}_{\text{nak}}(K - \epsilon|\theta), F^{-1}_{\text{nak}}(K + \epsilon|\theta)]$, as illustrated in Fig. 5.1, the membership function is the normalized cumulative within class probability. As ϵ increases, more fuzziness is assumed, and thus more noise is compensated for. It is noteworthy that the influence of ϵ is quite data dependent and there is flexibility in integrating it. One possibility is to define it as a combination of a probability threshold T with a confidence value C,

$$\epsilon = T \cdot C^{\gamma}, \tag{5.9}$$

where $\gamma < 1$ allows for some compression, e.g. gamma correction. Here we employ a confidence map (see Sec. 5.3.2) corresponding to the ultrasound image, to approximate C. The underlying idea is that low confidence regions should be compensated in terms of increased fuzziness as there the membership association is not reliable. See Fig. 5.2 for a visualization how the magnitude of the threshold ϵ affects the feature histogram.

Confidence Maps

We formulate the ultrasound signal confidence estimation as a random walk problem. The solution to this is based on the algorithm proposed for image segmentation [Grady, 2006] as described at length in Sec. 4.6.2. There a

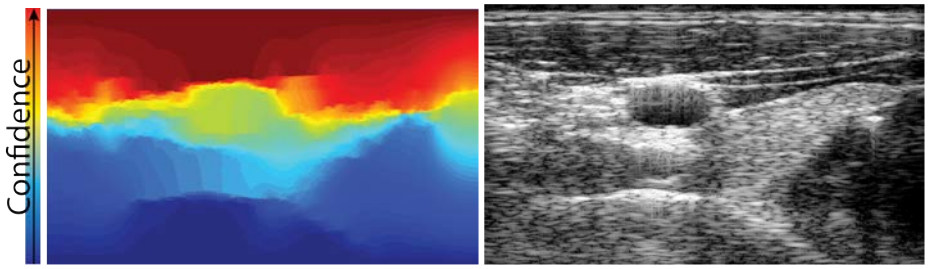


Figure 5.3: Confidence map (left) and the corresponding RF image from which it is computed (right).

multi-label image segmentation is obtained by the analytic computation of the probabilities for random walks reaching user-defined image labels. For the confidence estimation we are interested in the probabilities of random walks reaching the transducer elements under ultrasound specific constraints. More specifically, the random walks behavior is adjusted by modifying the graph Laplacian [Karamalis et al., 2012a] to model ultrasound transmission, beam-width, and depth-dependent attenuation. Subsequently, the analytic solution expresses the probability of ultrasound energy reaching a point in the RF data domain. Regions which are associated with low confidence are prone to contain noise. In turn high noise should be reflected by a wider fuzzy membership association, as data is more subject to noise in those areas. In contrast to that, areas with high confidence are less subject to noise artifacts such that the fuzzy membership function should be more narrow there.

5.3.3 Hybrid Similarity Measure

A common approach to measuring the similarity between LBPs is histogram intersection

$$D_{\mathbf{X},\mathbf{Y}} = \sum_i min(x_i, y_i) \ , \tag{5.10}$$

where $\mathbf{X}, \mathbf{Y}$ denote the histograms made up from N individual bins $\mathbf{X} = \{x_i\}_{i=1..N}$ and $\mathbf{Y} = \{y_i\}_{i=1..N}$. However, standard histogram intersection is prone to yield several local minima. In order to avoid this, we endow the standard histogram intersection with a component measuring the statistical similarity of distributions. This follows the notion that patterns, in the two patches being considered, should be assumed relevant only where the underlying distributions significantly exhibit high similarity. Statistical similarity alone on the other hand is globally precise (*rough scale*), but is locally (*fine scale*) imprecise. However, in combination with the in Sec. 5.3.2 proposed statistics-based FLBP, which has diametrically opposite behavior, i.e. globally imprecise and locally precise (see Fig. 5.6), we yield overall high reliability. This reasoning is validated by our results, see Tab. 5.1, which show that superior performance is achieved by combining Gamma Hellinger measure with statistics-based FLBP, as opposed to applying them individually. HLBP is able to take into account very fine textural features in data, which can discriminate between small local

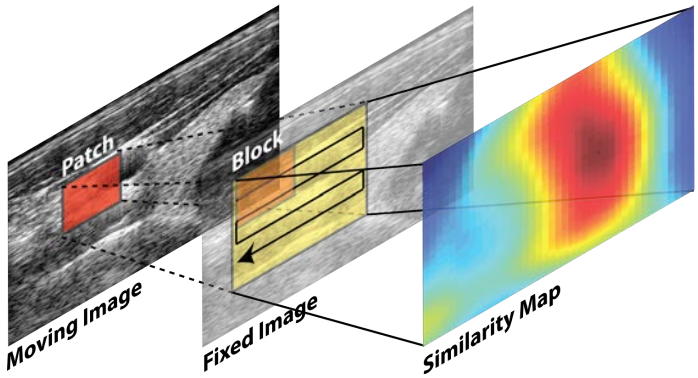


Figure 5.4: Concept of block matching and derivation of the similarity maps. A patch is taken in the moving image with the goal to find its best match within a limited region in the fixed image referred to as block. Moving the patch all over the block and computing the similarity yields a similarity map, which provide an idea about cost function smoothness. Optimizers need to find the optimum within the similarity map.

differences between images. When performing registration, HLBP would then be a useful measure when performing small-scale alignments. However, when differences are large between images, fine textural differences are not relevant. Here a purely statistical (dis-)similarity measure like e.g. Hellinger distance, is a better candidate for detecting global differences between images. This is because the distance (divergence) is defined between estimated distributions on the image. Distributional differences between images is however imprecise, when images are very similar, or in the case of registration, when close to alignment. The hybrid Hellinger weighted histogram intersection is defined as

$$D_{\text{Hybrid}} = D_{X,Y} \cdot \exp(-\Delta \bar{H}_{gam}(f,h)) + \bar{H}_{gam}(f,h), \tag{5.11}$$

where Δ denotes the influence radius of the patch-averaged distribution distance $\bar{H}_{gam}$, which implies multiple Hellinger distance computations on corresponding (homogenous) sub-patches. Furthermore, X, Y contain the LBP feature histograms (see Sec. 5.3.2). This reinforces the notion that LBP is primarily used to enhance the accuracy of Hellinger, which determines a robust rough approximation of the location of the global optima.

5.3.4 Experiments

Experiments are performed on 5 human neck datasets each acquired with a 10 MHz linear array probe yielding RF envelope images of size 2048 x 256 pixels. Images were recorded with an Ultrasonix MDP machine (Ultrasonix Medical Corporation, Richmond, Canada). The registration dataset consists of pairs of images constituting a moving and a fixed image for the registration (translation, 2 degrees of freedom). Consequently, block matching (block size 20 x 40 pixels) is performed at 28 regular spaced points on the image domain, yielding

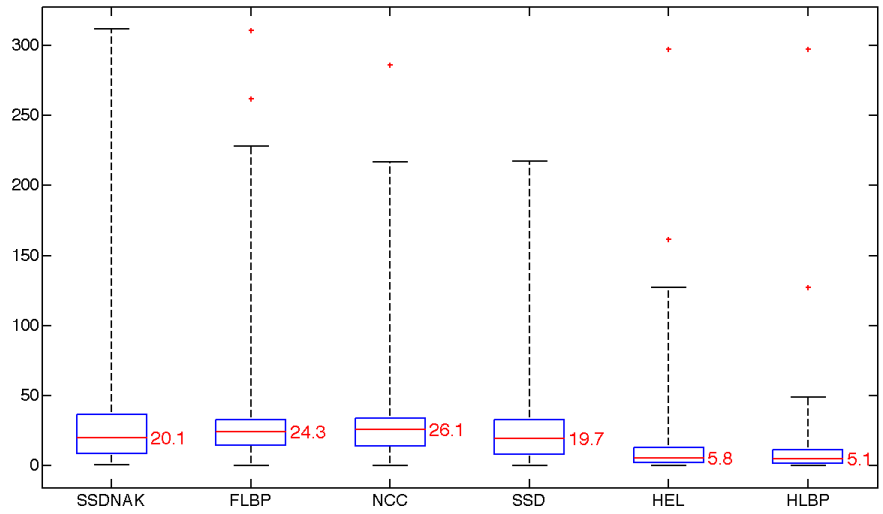


Figure 5.5: Global registration performance for all datasets combined - error measured in pixels.

Dataset	SSD	SSDNAK	NCC	FLBP	HEL	HLBP
#1	20.1	23.0	26.1	28.1	4.7	**3.5**
#2	16.1	19.8	26.3	28.2	7.4	**6.3**
#3	20.1	12.1	26.2	27.2	7.9	**7.4**
#4	17.9	16.5	25.0	24.0	8.5	**7.4**
#5	24.4	19.9	27.2	3.3	1.4	**1.3**

Table 5.1: Median errors (pixels) of random registration study for various datasets and similarity metrics.

similarity maps. See Fig. 5.4 for a schematic visualization of the block matching process. For each block 20 registration runs are performed with random initial start points to estimate the susceptibility of metrics towards building local minima. Manual alignment of each dataset pair served as ground truth. The following methods were used for the registration test: SSD (Sum-Of-Squared Differences), SSDNAK (SSD on the shape parameter of Nakagami images), NCC (Normalized Cross-Correlation), FLBP (threshold $\epsilon = 30$, confidence weighted), Hellinger and HLBP (threshold $\epsilon = 0.35$, confidence weighted). Finally, the global similarity is computed by accumulating estimates from all blocks on each dataset - see Tab. 5.1 for median pixel errors. For registration performance evaluation, comparison results from all datasets are combined - see Fig.5.5. Additionally, in order to assess the different similarity measures we extract a patch in the moving image, and compute corresponding similarity maps - see Fig. 5.6 for an example.

5.3.5 Conclusion

We presented a similarity metrics Nakagami Hellinger and Hybrid Local Binary Patterns. Results of the experiments show that HLBP greatly outperforms standard methods, such SSD and NCC, and also improves upon the result obtained by use of Hellinger distance metric. The robustness of HLBP is due to its hybrid local and global strategy. On the one hand, LBP and its variants are able to take into account very fine textural features in data, which can discriminate small local differences between images, and are thus suitable for small-scale alignments. On the other hand, when differences are large between images, fine textural differences are not relevant. Here a purely statistical (dis-)similarity measure like e.g. Hellinger distance, is a better candidate for detecting global differences between images. For future work, optimization of the neighborhood system of HLBP will be investigated, as this has been shown in initial experiments to be a source of potential major improvement.

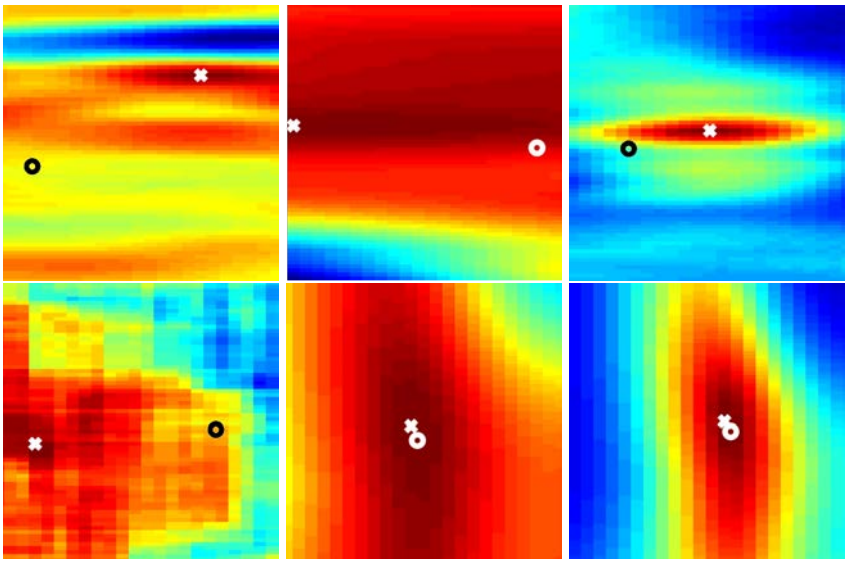
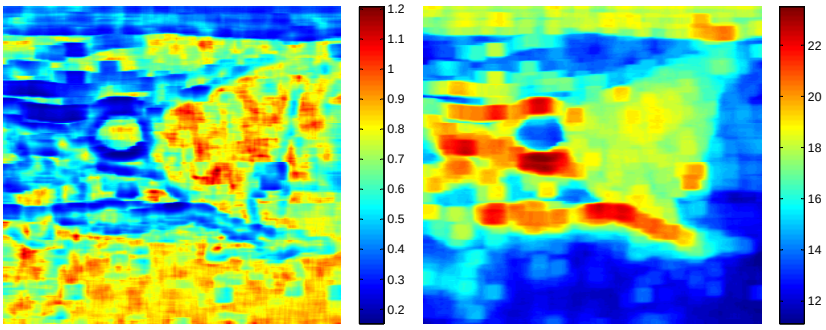


Figure 5.6: Similarity maps from left to right, top to bottom: SSD, SSDNAK, NCC, FLBP, HEL and HLBP. Circle: ground truth optimum; Cross: optimum in similarity map.



(a) m map (b) ω map

Figure 5.7: Illustration of Nakagami MLE parameters shape m and scale ω, calculated densely for one image of a human neck.

5.4 Locally adaptive Nakagami-based ultrasound similarity measures

5.4.1 RF Data

One of the first fundamental processing steps of RF data is envelope detection as described in Sec.2.3. For the envelope detection we apply the approach based on the 2D analytic signal as presented in Ch.3, although the method is also applicable to conventionally enveloped data based on 1D Hilbert transform. However, as was demonstrated the incorporation of lateral information leads to a more robust estimation of the local amplitude from ultrasound RF data, which is beneficial in terms of registration. Moreover, the extracted envelope bears superior statistical properties, as was shown with goodness-of-fit tests to Nakagami distributions. Since we want to accurately model the statistical properties of the envelope data in the derived likelihood functions, the calculation of the 2D analytic signal is therefore advantageous for our application.

5.4.2 Maximum Likelihood Ultrasound Registration

Given ultrasound envelope detected images I, J and the transformation T. Then registration can be formulated as maximum likelihood estimation [Viola, 1995, Roche et al., 2000] by

$$\hat{T} = \arg\max_{T} \log p(I \mid J, T, \varepsilon), \tag{5.12}$$

where $\hat{T}$ corresponds to the estimated transformation of maximum probability and ε to the noise. For the derivation of SSD, correlation ratio, or mutual information, an additive Gaussian noise is incorporated in the deduction [Roche et al., 2000].

5.4.3 Ultrasound Likelihood Functions

Given the general definition of registration as maximum likelihood estimation problem, we present an overview of ultrasound specific similarity measures proposed in the literature. This is followed by a derivation of a new locally adaptive Nakagami-based metric. We list several imaging models and denote them with the initials of the authors, prevailing in the literature.

SK_1: Multiplicative Rayleigh Noise

In [Strintzis and Kokkinidis, 1997], an imaging model based on a multiplicative Rayleigh noise is presented

$$I = J \cdot \varepsilon, \tag{5.13}$$

where the noise ε is assumed to be Rayleigh distributed. It is noteworthy, that in this noise model only one of the images is assumed to be degraded by noise, while the other one is assumed to be void of noise. Knowing the distribution of the noise $p(\varepsilon)$, we have to find the distribution for the likelihood $p(I \mid J, T)$. Assuming that we know the distribution of a random variable x and we want to calculate the distribution of a random variable y, with both being related by the function g, so $y = g(x)$. Then applying the fundamental theorem of random variable transformation [Papoulis and Pillai, 2002, p.93] permits the derivation of distribution of the random variable y according to

$$p(y) = \frac{p(x)}{|g'(x)|}. \tag{5.14}$$

In our case, this implies setting $g(x) = x \cdot J$ and therefore yielding $\frac{dg(x)}{dx} = J$. Altogether, this leads to the likelihood function defined as,

$$p(I \mid J, T) = \frac{1}{J} \cdot p(\varepsilon). \tag{5.15}$$

Furthermore, setting the variance of the Rayleigh distribution to $\frac{2}{\pi}$ leads to the log-likelihood function of SK_1

$$\log p(I \mid J, T) = \log \left[\frac{1}{J} \cdot p\left(\frac{I}{J}\right) \right] \tag{5.16}$$

$$\approx \log \left(\frac{I}{J^2} \right) - \frac{\pi}{4} \frac{I^2}{J^2}. \tag{5.17}$$

In [Strintzis and Kokkinidis, 1997] a second model is derived with the specific property of a signal dependent Gaussian noise, which is not further considered here.

CD_1: Division of Rayleigh Noises

A more realistic and sophisticated model is proposed in [Cohen and Dinstein, 2002b]. Here the underlying and more realistic assumption is made that each image is subject to a multiplicative Rayleigh noise denoted by ε_1 and ε_2,

respectively. Considering the underlying noise-free scene S, the imaging model is formulated as $I = S \cdot \varepsilon_1$ and $J = S \cdot \varepsilon_2$. The relationship can then be written as

$$I = J \cdot \frac{\varepsilon_1}{\varepsilon_2} = J \cdot \eta, \tag{5.18}$$

where $\eta = \frac{\varepsilon_1}{\varepsilon_2}$ denotes the division of probabilities. Following [Papoulis and Pillai, 2002, p.138], the distribution of the divisional noise $p(\eta)$ is defined as

$$p(\eta) = \int_{-\infty}^{\infty} \varepsilon_2 \cdot p(\eta\varepsilon_2, \varepsilon_2)\, \mathrm{d}\varepsilon_2. \tag{5.19}$$

Further, integrating the assumption of the noise in the images being independent $p(\varepsilon_1, \varepsilon_2) = p(\varepsilon_1)p(\varepsilon_2)$ with the property of equal variance and underlying Rayleigh distributional model, the integration results in

$$p(\eta) = \frac{2 \cdot \eta}{(\eta^2 + 1)^2}. \tag{5.20}$$

Given the distribution of the noise $p(\eta)$, next we have to find the distribution for the likelihood $P(I \mid J, T)$. This is again obtained by employing the fundamental theorem of random variable transformation [Papoulis and Pillai, 2002, p.93]. In particular setting $g(x) = x \cdot J$ the derivative part reduces to $\frac{dg(x)}{dx} = J$. Taken together the likelihood is defined as

$$p(I \mid J, T) = \frac{1}{J} p(\eta) = \frac{1}{J} \int_{-\infty}^{\infty} \varepsilon_2 \cdot p(\eta\varepsilon_2, \varepsilon_2)\, \mathrm{d}\varepsilon_2. \tag{5.21}$$

Furthermore, the log-likelihood function of CD_1 is defined

$$\begin{aligned}
\log p(I \mid J, T) &= \log \left[\frac{1}{J} \cdot p\left(\frac{I}{J}\right) \right] && (5.22)\\
&= \log \left[\frac{1}{J} \frac{2 \cdot \frac{I}{J}}{\left(\left(\frac{I}{J}\right)^2 + 1 \right)^2} \right] && (5.23)\\
&\approx \log I - 2 \log J - 2 \log \left[\left(\frac{I}{J}\right)^2 + 1 \right]. && (5.24)
\end{aligned}$$

CD_2: Logarithm of Division of Rayleigh Noises

The second model proposed in [Cohen and Dinstein, 2002b] considers beside the noise presence in both images also the log-compression, which is applied B-mode ultrasound images, yielding

$$\log I = \log(J \cdot \eta) = \log J + \log \eta. \tag{5.25}$$

Denoted the log compressed images as $\tilde{I} = \log I$ and $\tilde{J} = \log J$, respectively, we can define

$$\eta = \exp(\tilde{I} - \tilde{J}). \tag{5.26}$$

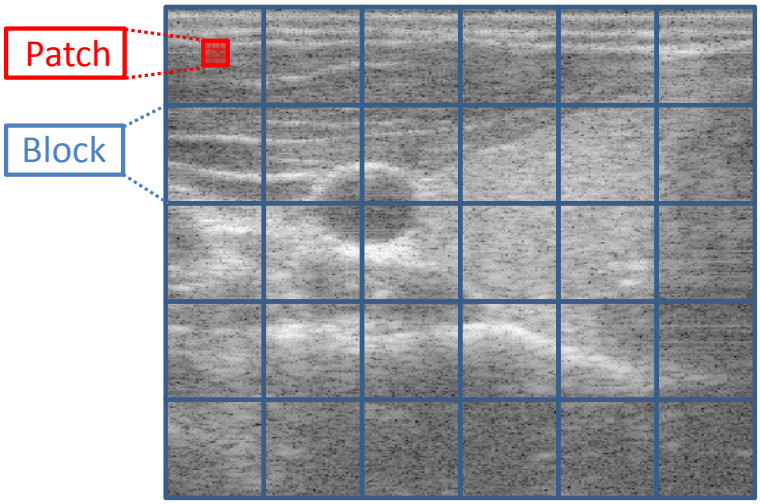


Figure 5.8: Illustration of the subdivision of the image domain into blocks and patches.

The likelihood function can then be obtained by applying the fundamental theorem of random variable transformation. However, this time with respect to the log-compressed images. Therefore the likelihood for the log-compressed images is defined as

$$p(\tilde{I} \mid \tilde{J}, T) = \eta \cdot p(\eta). \tag{5.27}$$

This is obtained by setting $g(\eta) = \log J + \log \eta$ and the corresponding derivative $g'(\eta) = \frac{1}{\eta}$. As a result, the corresponding log-likelihood function for CD_2 is defined as

$$\begin{aligned}
\log p(\tilde{I} \mid \tilde{J}, T) &= \log \left[\frac{\exp(\tilde{I})}{\exp(\tilde{J})} \cdot p(\exp(\tilde{I} - \tilde{J})) \right] && (5.28)\\
&= \log \left[\frac{\exp(\tilde{I})}{\exp(\tilde{J})} \cdot \frac{2 \cdot \exp(\tilde{I} - \tilde{J})}{\left[\exp(\tilde{I} - \tilde{J})^2 + 1\right]^2} \right] && (5.29)\\
&= \log \left[\frac{2 \cdot \exp(2(\tilde{I} - \tilde{J}))}{\left[\exp(2(\tilde{I} - \tilde{J})) + 1\right]^2} \right] && (5.30)\\
&\approx \tilde{I} - \tilde{J} - \log[\exp(2(\tilde{I} - \tilde{J})) + 1]. && (5.31)
\end{aligned}$$

5.4.4 Bivariate Nakagami

As explained in Sec. 2.4, the Nakagami distribution is a versatile distribution well suited for modeling various backscatter scenarios and speckle statistics in ultrasound. Assuming a divisional noise model in combination with noise independence between the images, a bivariate distributional model is required - see Eq. 5.21. Such assumptions are made in [Cohen and Dinstein, 2002b], where a Rayleigh distributional model is employed. Given the physical principles

underlying ultrasound imaging, which is speckle arising due to physical interaction of the beam with the micro-scale inhomogeneities in tissue, the same speckle patterns appear under constant and identical acquisition conditions. Nowadays, modern ultrasound systems provide high frame rates for acquisitions such that for images taken from approximately the same viewing-angle, it is reasonable to assume assume a correlation between the noise related random variables ε_1 and ε_2 [Cobbold, 2007]. In this respect, following [Myronenko et al., 2009b] the bivariate Nakagami distribution is defined as

$$p(\varepsilon_1, \varepsilon_2) = \frac{2^{1-m} m^{m+1} (\varepsilon_1 \varepsilon_2)^m}{\sigma^{2(m+1)} (1-\rho) \rho^{(m-1)/2} \Gamma(m)} e^{-\frac{m(\varepsilon_1^2+\varepsilon_2^2)}{2(1-\rho)\sigma^2}} I_{m-1}\left(-\frac{m\sqrt{\rho}\varepsilon_1\varepsilon_2}{(1-\rho)\sigma^2}\right), \quad (5.32)$$

where ρ denotes a squared correlation coefficient, I_{m-1} is the modified Bessel function of the first kind of order $m-1$, and the distribution widths σ^2 are identical. Next, incorporation of the bivariate Nakagami distribution in the conditional density in Eq. 5.21, yields the following likelihood

$$p(I|J,T) = \frac{1}{J}\int_{-\infty}^{\infty} \varepsilon_2 \cdot p(\eta\varepsilon_2, \varepsilon_2)\, \mathrm{d}\varepsilon_2 \quad (5.33)$$

$$= \frac{1}{J} 2(1-\rho)^m \frac{\Gamma(2m)}{\Gamma(m)^2} \frac{\eta^{2m-1}}{(\eta^2+1)^{2m}} \left(1 - \frac{4\rho\eta^2}{(\eta^2+1)^2}\right)^{-\frac{2m+1}{2}} \quad (5.34)$$

$$= \frac{1}{J} \frac{2(1-\rho)^m}{\beta(m)} \cdot \frac{1}{\eta} \cdot \frac{\eta^{2m}}{(\eta^2+1)^{2m}} \cdot \left(1 - \frac{4\rho\eta^2}{(\eta^2+1)^2}\right)^{-\frac{2m+1}{2}}, \quad (5.35)$$

where $\beta(m)$ is defined as $\beta(m) = \frac{\Gamma(m)^2}{\Gamma(2m)}$. More details on the exact deduction, however, assuming log-compressed envelope data and therefore a different imaging model, are presented in [Myronenko, 2010]. Finally, the Nakagami-based similarity measure is defined as the log-likelihood of the bivariate Nakagami

$$\begin{aligned}
&\log p(I|J,T) \\
&= -\log J + \log \frac{2(1-\rho)^m}{\beta(m)} - \log\eta + 2m\log\left(\frac{\eta}{\eta^2+1}\right) - \frac{2m+1}{2}\log\left(1 - \frac{4\rho\eta^2}{(\eta^2+1)^2}\right) \\
&= -\log J - \log\left(\frac{I}{J}\right) + 2m\log\left(\frac{\left(\frac{I}{J}\right)}{\left(\frac{I}{J}\right)^2+1}\right) - \frac{2m+1}{2}\log\left(1 - \frac{4\rho\left(\frac{I}{J}\right)^2}{(\left(\frac{I}{J}\right)^2+1)^2}\right),
\end{aligned} \quad (5.36)$$

where the parameter m denotes the shape and ρ specifies the correlation. In the next section, an adaptive method is presented that computes the parameters m and ρ locally.

5.4.5 Parameter estimation

During each iteration of the similarity computation, the bivariate Nakagami model has to be instantiated. This requires the specification of a shape parameter m as well as a correlation coefficient ρ - see Eq. 5.36. Considering

the need for high locality in order to obtain reliable registration results, the patch size has to be kept at minimum. However, this data sparsity constraint is detrimental for the computation of the m parameter of the bivariate Nakagami distribution. Additionally, because of the high frequency of this similarity computation in course of the registration procedure, complexity has to be kept at bare minimum to keep tractability. Therefore, we decided to use a fast approximation scheme for the parameters.
In this respect, the distribution parameters are calculated separately for the moving (m_I, ω_I) and the fixed patches (m_J, ω_J) employing a standard maximum likelihood estimation (MLE) procedure - see Sec. 2.4.2. Then, given the individual MLE distribution parameters, the joint shape parameter m of the bivariate Nakagami is approximated by the mean,

$$m = \frac{(m_I + m_J)}{2}. \tag{5.37}$$

For computing the correlation coefficient ρ for the bivariate Nakagami, probably the most obvious way is to directly use the corresponding intensities of the patches. However, due to the noise-susceptibility of ultrasound, a probabilistic correlation score was chosen in order to increase the overall reliability. In this respect, we employ the Bhattacharyya coefficient (BC) [Bhattacharyya, 1943], which is divergence-type measure between distributions and is defined as

$$BC(p,q) = \int \sqrt{p(x)q(x)}\, dx, \tag{5.38}$$

with $0 \leq BC \leq 1$. Noteworthy, the BC is related to the Hellinger distance by $\sqrt{1 - BC}$. The Bhattacharyya coefficient is a measure of the relative overlap between two probability distributions p and q. In our case, the distributions correspond to the Nakagami MLE estimates of the moving and the fixed image, respectively.
In this regard, the Bhattacharyya coefficient has an intuitive geometric interpretation [Comaniciu et al., 2003], as the cosine of the angle between (in the discrete case) N-dimensional vectors unit vectors $\left(\sqrt{p_1}, ..., \sqrt{p_N}\right)^T$ and $\left(\sqrt{q_1}, ..., \sqrt{q_N}\right)^T$,

$$cos(\theta) = \sum_{i=1}^{N} \sqrt{p(i)q(i)}. \tag{5.39}$$

Furthermore, BC adopts maximum value in case of identical distributions, which implies $\forall i \quad : \quad p(i) = q(i)$. Then given p and q representing probability distributions, it naturally follows $\sum_{i=1}^{N} p(i) = 1$. Similarly, following the cosine angle representation we have

$$cos(\theta) = \sum_{i=1}^{N} \sqrt{p(i)q(i)} = \sum_{i=1}^{N} \sqrt{p(i)p(i)} = \sum_{i=1}^{N} p(i) = 1. \tag{5.40}$$

The BC itself is not a proper metric as it does not fulfill all required metric axioms. However, by slight modification it can be turned into the so called Bhattacharyya metric D_B, defined as

$$D_B(p,q) = -\log BC(p,q) = -\log \sqrt{p(x)q(x)}\, dx, \tag{5.41}$$

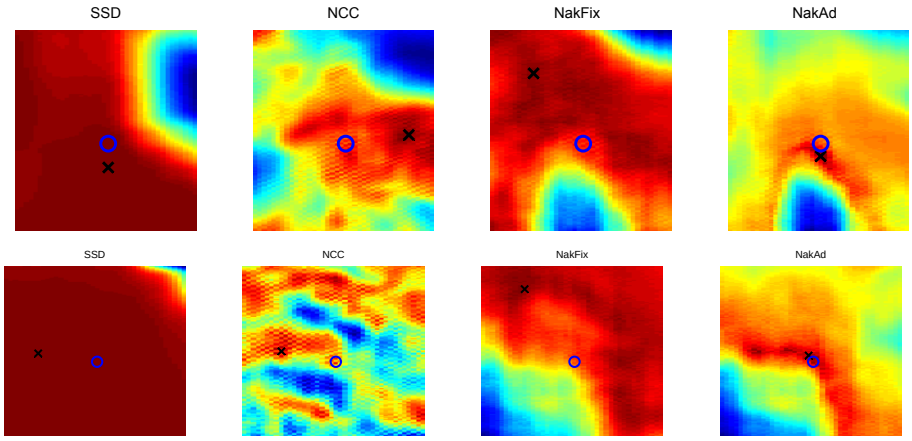


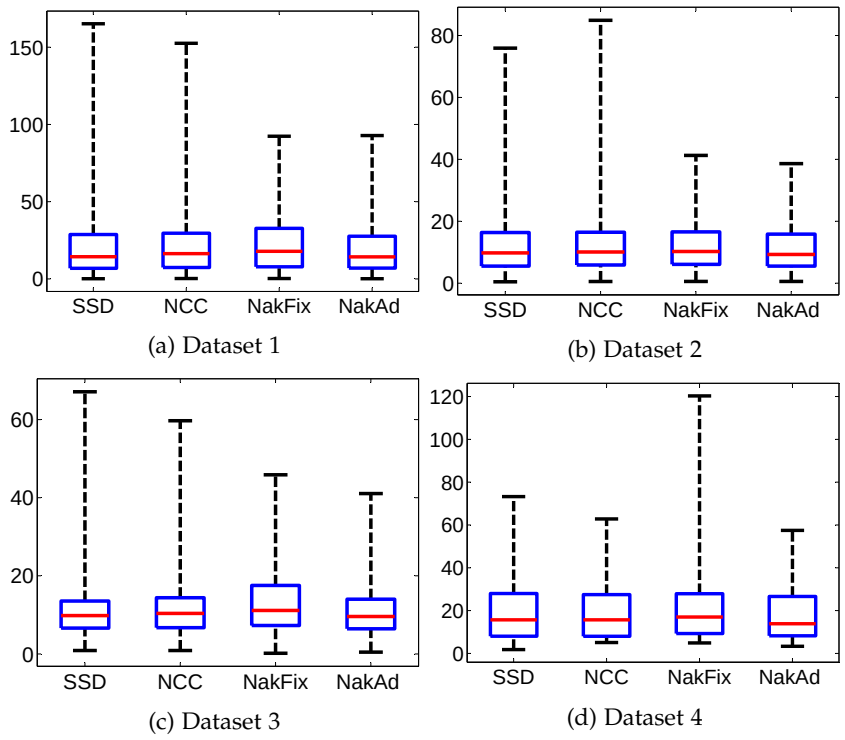
Figure 5.9: Similarity plots for various measures. Blue circle: correct alignment, black cross: similarity maximum.

with $0 \leq D_B \leq \infty$.

5.4.6 Experiments

Performance of the similarity metric is evaluated in experiments that are performed on several RF datasets. Each dataset is acquired with a linear transducer at 3.3 MHz, at a sampling rate of 40 MHz. Depending on the depth transducer setting, the images obtained have a resolution between 1157 and 2080 pixels in axial, and 256 pixels in lateral direction. For registration procedure evaluation each dataset was set up with two images: a moving and a fixed image. On this data a block matching procedure is performed. Specifically, 100 blocks equally distributed across the image are employed for the matching. The process is schematically illustrated in Fig. 5.8. For each block alignment in matching manual alignment serves as ground truth data. Performance is further compared with state of the art ultrasound metrics, SSD and NCC with the similarity measure presented in Eq. 5.36. Additionally, we once use heuristic values for the distribution parameters, referred to as NakFix, and once we estimate them on the images, referred to as NakAd. For the heuristic case, we choose the parameters used in [Myronenko et al., 2009b] $m = 0.5$ and $\rho = 0.8$. NakFix is therefore the analogue of the similarity measure presented in [Myronenko et al., 2009b] for non log-compressed envelope data.

In order to further compare the different similarity measures similarity maps are generated. For that purpose a patch of 91×11 pixels is extracted in the moving image. The extracted image is then in turn shifted it over the block in the fixed image. Each shift position then yields a similarity value. Taken all block matchings over the entire image domain one yields a similarity map - see Fig. 5.4 for an illustration of block matching and similarity maps. For NakAd the distribution parameters are estimated on these patch intensities. It should be noted that the patch size is a trade off between sufficient statistics for param-



(a) Dataset 1 (b) Dataset 2

(c) Dataset 3 (d) Dataset 4

Figure 5.10: Boxplots of errors from random registration study for block matching.

eter estimation and the detection of fine grained deformations. In particular for heterogeneous texture large patches are detrimental for the registration procedure as the averaged distributional parameters. We illustrate similarity plots for two different blocks in Fig. 5.9. As one can observe, in both cases, SSD and NCC are unable to correctly indicate the correct alignment of the patches. Moreover, the similarity plot of NCC shows several local minima. The existence of many local optima impedes the registration procedure. Furthermore, the maximum of NakFix is also far off the correct alignment. However, the local adaptation of the parameters in the similarity metric, as it is performed in NakAd, significantly changed the similarity function. As can be seen in the figures, the local adaptation leads to good registration results in this case.

The similarity plots provide only a first qualitative impression of the performance of similarity metrics. Therefore we further evaluate the performance by block matching. As already discussed in Sec. 5.2, block matching is commonly applied in ultrasound registration. For registration, we extract patches from the moving image and try to find the corresponding or best matching patch in the fixed image. This procedure is performed for one patch per block, all in all with 100 different patches. Since we know the alignment of the blocks, we can calculate the Mahalanobis distance with respect to the ground truth position, serving as error measure. Here we select the Mahalanobis distance

Dataset	SSD	NCC	NakFix	NakAd
Dataset 1	9.3	9.9	10.4	**9.1**
Dataset 2	9.8	10.1	10.2	**9.2**
Dataset 3	9.9	10.4	11.1	**9.5**
Dataset 4	15.7	15.7	17.0	**13.8**

Table 5.2: Median errors of random registration study for various datasets and similarity metrics.

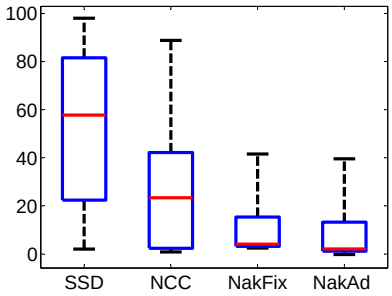


Figure 5.11: Boxplot of errors from rigid registration study.

to compensate for the significantly higher resolution in axial direction. Furthermore, in order to be able to perform significant statistics on the results of the registration, we perform a random registration study. In this respect, we randomly displace the patch 100 times from the ground truth position, with maximal initial deviation of ± 40 pixels in axial and ± 10 pixels in lateral direction. The random displacement serves as initial position of the registration procedure. As a result, the result of theandom registration study provides an idea about the robustness of the similarity metric. The errors over all patches and all runs are shown in the boxplots in Fig. 5.10 for 4 datasets. In Tab. 6.1 we enlist the median errors of the registration experiment. As can be observed the median, the box, and the whiskers are lowest for NakAd similarity metric in all cases. The performance of SSD and NCC is comparable, however, with slight advantages for NCC with respect to the whiskers. Furthermore, the performance of NakFix for datasets 1 - 3 is slightly worse than NakAd. For dataset 4, NakFix is not leading to good results.

In previously presented experiments performance was based individual block matching procedure. Next we also perform experiments for global rigid registration. Therefore the image is again separated into blocks, as already discussed previously. Then we again estimate the parameters for each block and evaluate the similarity measure on the block. However, this time we do not consider the block matchings individually. Rather we accumulate the similarity estimates from all blocks together. This in turn leads to the global similarity estimate. Analogously to the individual block matchings, we also perform a random registration study with 100 runs in order to assess the robustness. In this respect, the initial values had deviations up to ± 240 pixels in axial

and ± 70 pixels in lateral direction. The results are shown in Fig. 5.11. As can be seen, the median errors are SSD: 57.8, NCC: 23.4, NakFix: 4.1, and NackAd: 2.1, respectively. NCC performs better than SSD, but the best results are obtained for the Nakagami-based similarity measures, with the adaptive version outperforming the fixed version.

5.4.7 Conclusion

We introduced a locally-adaptive similarity measure based on the bivariate Nakagami distribution, employing local adaptation of the measure by estimating the distribution parameters on the ultrasound images. As discussed, the adaptation of the similarity measure to various scattering scenarios is necessary to model the noise correctly. Experiments are performed on ultrasound RF data. The results from block matching indicate the improvement of incorporating the Nakagami distribution and the necessity in locally adapting the parameters. Moreover, a clear improvement of the proposed similarity measures is shown for the case of global rigid registration.

Chapter 6

3D Freehand Ultrasound

3D freehand ultrasound allows to easily extend the field of view by means of reconstructed volumes generated from tracked transducer data. Several reconstruction approaches exist, which are all application dependent. This chapter gives an overview about various methods of reconstruction and deals with a related application in the field of neurology - the early diagnosis of Parkinson's disease. In this regard, advanced registration approaches such as based on confidence maps and multi-view mixture models are presented. Furthermore, in context of multi-view mixture modeling, a derivation for the Nakagami Expectation-Maximization algorithm is presented. This chapter is based on the papers 'Towards a More Objective Visualization of the Midbrain and its Surroundings Using 3D Transcranial Ultrasound' [Plate et al., 2010], 'Modeling of Multi-View 3D Freehand Radio Frequency Ultrasound' [Klein et al., 2012b], 'Ultrasound Confidence Maps using Random Walks' [Karamalis et al., 2012b] and '3D Sonographic Examination of the Midbrain for Computer-Aided Diagnosis of Movement Disorders' [Plate et al., 2012].

6.1 Introduction

Nowadays ultrasound examinations are performed mostly with conventional machines providing two dimensional imagery. However, there exist a multitude of applications where doctors could benefit from three dimensional ultrasound providing better judgement, due to the extended spatial view. 3D freehand ultrasound allows acquisition of images by means of a tracking device attached to the ultrasound transducer. Unfortunately, ultrasound specificities make the 3D representation of ultrasound a non-trivial task. This is largely due to the inherent process of ultrasound imaging, which is above all view-dependent, as well as subject to noise and prone to containing various types of artifacts. As already mentioned in Sec. 2.2, the characteristic speckle noise is dependent on factors such as spatial arrangement and size of scatterers. Thereby patterns are formed, which are largely characteristic for various types of tissue. In order to represent data in a meaningful way, those ultrasound specific issues need to be addressed. Because of the same reason, the interpretation of 3D ultrasound imagery is typically not straightforward as well as of quite

subjective nature, and therefore highly dependent on the expertise of its investigator. Furthermore, the view dependency is responsible for the limited repeatability of ultrasound data acquisitions, which is one of the major factors restricting applicability of this modality.

In this chapter we will give at first an overview of conventional freehand systems as found in recent articles - see Sec. 6.2. Thereby different reconstruction approaches are reviewed. Next we elaborate on backward-warping reconstruction in Sec. 6.3. This is followed by a brief introduction into the brain anatomy, which is succeeded by a discussion of an application using of transcranial ultrasound in neurology in Sec. 6.4. Subsequently, in Sec. 6.5 we present an advanced reconstruction method, employing confidence maps to weight intensity information from ultrasound images. In Sec. 6.6 follows a presentation of statistically motivated multi-view reconstruction using raw ultrasound radio frequency (RF) data, in order to take full advantage of speckle statistics of ultrasound. This is advantageous against the background that the underlying data is unaffected by non-linear post-processing steps, such as log-compression and other proprietary filter algorithms that change speckle statistics for reasons of improved visual appeal. Furthermore, when used in compounding, RF data also avoids multiple interpolations of data (compared to ultrasound compounding from B-mode see e.g. [Wein et al., 2006]), which is detrimental to analysis. In this respect, Nakagami finite mixture model is discussed in-depth in Sec. 6.6.3. Employing the Nakagami finite mixture model for 3D freehand provides improved reconstruction and registration capabilities and further allows the embedding of the view-dependency property of ultrasound in a statistical formulation. Performance and use is again show-cased on transcranial ultrasound brain data.

6.2 Related Work

Generation of 3D ultrasound volumes from the pooled data acquired by means of tracking a 2D probe has been performed since the end of the 90s. A good overview over state-of-the art reconstruction techniques is available in the review article [Solberg et al., 2007], whereby we largely follow the exposition presented therein. Following this article reconstruction methods can be divided into three basic types: (i) *Voxel-Based*, (ii) *Pixel-Based* and (iii) *Function-Based* methods.

Voxel-based (also known as backward-warping) methods generate a 3D voxel grid space covering all spatially organized tracked 2D slices. In [McCann et al., 1988, Sherebrin et al., 1996, Rohling et al., 1999] a voxel nearest neighbor scheme is proposed, where each voxel of the grid is traversed and for the nearest pixel(s) on one or several slices is searched. In [Trobaugh et al., 1994] the two nearest surrounding slices of each voxels are determined and then a distance weighted interpolation scheme is employed to compute the voxel intensity value. Instead of using an orthogonal projection of the nearest slices, assuming constant speed a trajectory between the nearest planes is estimated in [Coupé et al., 2007], which in turn is used to find corresponding points for interpolation. An efficient reconstruction approach with different weighting functions is proposed in [Wein et al., 2006]. Instead of interpolating from the scan-converted 2D images, nearest neighbor 1D scanline interpolations can be used as proposed in [Berg et al., 1999, Martens and Gilja, 2005], which avoids multiple interpolation steps. Similarly, 1D scanline interpolation is proposed in [Thune et al., 1996], however, in combination with a trilinear interpolation scheme.

The other major group of reconstruction algorithms is referred to as *Pixel-Based* (also known as forward-warping). In this approach each pixel in each the input domain is traversed and the intensity values distributed to the corresponding associated voxels. In [Hottier and Billon, 1990] a pixel nearest neighbor approach is proposed that associates each pixel to its nearest voxel. Multiple pixel intensities contribution with averaging was proposed in [Nelson and Pretorius, 1997, Gobbi and Peters, 2002]. Beside averaging there are also approaches proposing the most recent [Nelson and Pretorius, 1997] or maximum value [Ohbuchi et al., 1992]. As the pixel-based methods are prone to produce holes in the volume, there exist a multitude of strategies to encounter this problem. Taking the average, maximum value or median in a local neighborhood was proposed in [San Jose Estepar et al., 2003]. Furthermore, kernel based methods have been proposed. In doing so a 3D kernel is placed on each filled voxel and the intensity is applied with a distance-based weight in a local neighborhood (kernel) [José-Estépar et al., 2003]. In contrast to the voxel-based methods, the pixel-based methods do not require all data before reconstruction can be initiated, facilitating real-time reconstruction on CPU [Gobbi and Peters, 2002] and GPU [Karamalis et al., 2009].

Last but not least the *Function-Based* methods employ functional interpolation. This requires the selection of an interpolation function and during the reconstruction the computation of associated coefficients, such that the function

passes through the irregular spaced input data. Having computed the coefficients, the function can be evaluated at regular 3D grid points, yielding a dense and smooth 3D reconstruction. In this respect, Radial Basis Functions (RBF) were proposed in [Rohling et al., 1999]. Moreover, in [Sanches and Marques, 2000] a Rayleigh distribution-based reconstruction is proposed, employing a Bayesian framework, which is further improved in terms of speed by making use of a multi-scale approach [Sanches and Marques, 2002]. Furthermore, multi-angle ultrasound compounding , also referred to as spatial compounding, has demonstrated improvement in image quality, i.e., an increase in signal-to-noise ratio (SNR) and reduction of ultrasound speckle, as described in the works of [Shankar, 1986, Wilhjelm et al., 2000, Behar and Nikolov, 2006].
Having covered related work on general 3D freehand reconstruction, next we provide an overview about applications in the domain of neurosurgery and neurology. Due to the skull and its inhomogeneities ultrasound defocuses in course of transcranial examinations, which results in loss of spatial resolution and increased blur. Using a 2D array probe real-time phase aberration correction for transcranial imaging is proposed in [Ivancevich et al., 2004, Ivancevich et al., 2006, Dahl et al., 2005]. The work was later extended in a helmet with multiple transducer, for the early detection of stroke [Smith et al., 2009]. 3D freehand ultrasound imaging technique has also been suggest and used for neurosurgical interventions. In this respect, in [Unsgaard et al., 2006] an overview of related target application is provided. In [Reinertsen et al., 2007], based on a phantom data a vessel tree is reconstructed from Doppler ultrasound images and registered to the corresponding structures obtained from MRI for the purpose of brain shift analysis.
Beyond the reconstruction of multi-view ultrasound as applied in this thesis, (Nakagami/Gamma) finite mixture models represent a widely used method for data analysis. In this respect, finite mixture modeling has already been applied successfully to ultrasound data, however, only for the case of 2D data. Specifically, a Nakagami mixture model was used for segmentation of the carotid artery in [Destrempes et al., 2009]. Beyond the medical domain, mixture models enjoy wide-spread use. Furthermore, in [Copsey and Webb, 2003, Webb, 2000] Gamma mixture models are used for target detection in naval radar imaging (multi-view). In [Hanson, 2006] a mixture of Gamma was used to model censored lifetime data.
The target applications of models to be presented in the next sections are registration and reconstruction. In particular a parametric statistical approach in conjunction with similarity measures (in our case the J-divergence distance metric) to registration of 3D ultrasound data is shown. Previous works [Ijaz et al., 2011] have applied similarity metrics, e.g Kullback-Leibler (KL), Hellinger and Bhattacharyaa directly, however, without any parametric distributional assumptions on the data.

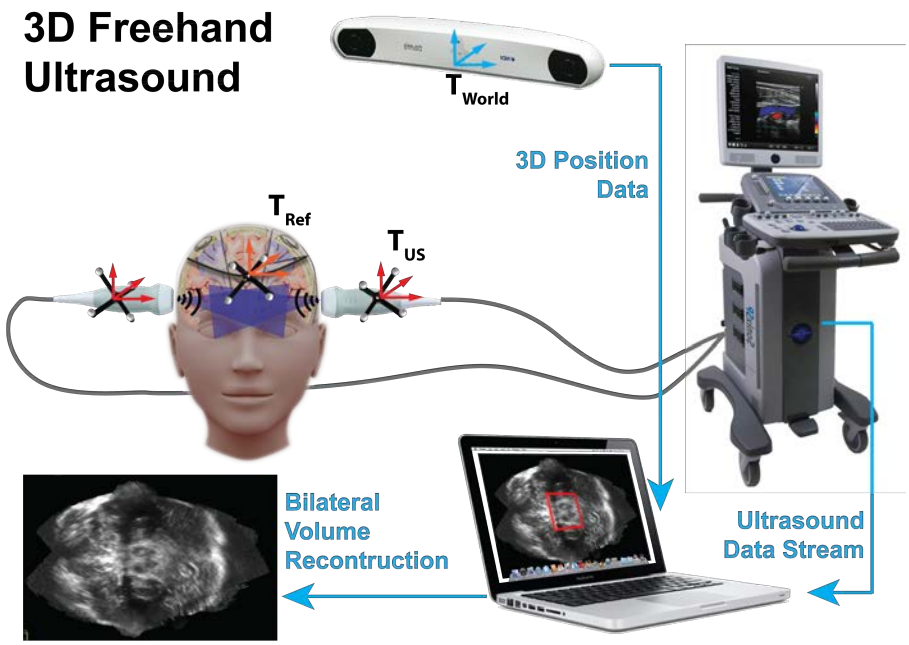


Figure 6.1: System setup for a 3D freehand system using an optical tracker with associated coordinate system T_{world}. Illustrated is the special setup of a transcranial freehand ultrasound system for bilateral reconstructions. Here the patient wears a reference target, allowing the establishment of a reference coordinate system denoted as T^{ref}. Multiple sweeps such as obtained for bilateral scans can be merged into a single volume. Tracking data (transducer T^{US}, reference T^{ref}) is supplied by the camera and together with ultrasound image data sent to a computer, where streams are synchronized. After acquisition, 3D reconstruction is performed on a computer.

6.3 Backward-Warping Reconstruction

In this section we will elaborate on the fundamentals of the voxel-based method for 3D freehand ultrasound reconstruction. Three-dimensional freehand ultrasound reconstruction usually refers to the process of interpolating a series of spatially arbitrarily arranged 2D ultrasound slices into a regular 3D Cartesian grid, i.e. a regular voxel volume. The acquisition of 3D freehand ultrasound follows a general setup - see Fig. 6.1 for an illustration of a freehand system for bilateral reconstructions. In order to obtain the ultrasound data within a spatial context, the ultrasound transducer needs the attachment of a tracking device. Typically this is either a tracking target for an optical system or some sensor such as for electromagnetic tracking devices. With this setup ultrasound images can be acquired in 3D space. The purpose of this acquisition setup is primarily to extend the field-of-view of ultrasound and to image anatomies from different angles to provide complementary imaging information.

In the following we will formalize the 3D reconstruction process. Concerning this matter, we will denote the transformation from coordinate frame A to coordinate frame B as ${}^{B}T_{A}$. Furthermore, we adopt the following scheme of abbreviations to denote coordinate frames for transformations: (T)racker, (P)robe (U)S Plane, (C)alibration phantom. The transformations are supposed to be Euclidean with 6 degrees of freedom (DOF) consisting of 3 rotations (α, β, γ) and 3 translations (x, y, z), such that ${}^{B}T_{A}$ decomposes to

$$
{}^{B}T_{A}(\alpha,\beta,\gamma,x,y,z) = \begin{pmatrix} \cos\alpha\cos\beta & \cos\alpha\sin\beta\sin\gamma - \sin\alpha\cos\gamma & \cos\alpha\sin\beta\cos\gamma + \sin\alpha\sin\gamma & x \\ \sin\alpha\cos\beta & \sin\alpha\sin\beta\sin\gamma + \cos\alpha\cos\gamma & \sin\alpha\sin\beta\cos\gamma - \cos\alpha\sin\gamma & y \\ -\sin\beta & \cos\beta\sin\gamma & \cos\beta\cos\gamma & z \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{6.1}
$$

Rigidly attaching a tracking object to the transducer probe provides the spatial position of the transducer with respect to the tracking device coordinate system, ${}^{T}T_{P}$. However, as one needs to correlate the pixel position in 2D on the ultrasound plane to the 3D space, an ultrasound calibration needs to be performed, providing ${}^{T}T_{U}$. See [Prager et al., 1998, Hsu et al., 2007, Mercier et al., 2005] for a good review over different calibration methods and their expected accuracy. Due to its simplicity and relatively good accuracy the single wall calibration method is a frequently used method. Thereby a planar object such as a nylon membrane immersed in a water tank requires being imaged at different positions and angles. The imaged membrane appears as a (distorted) line in the ultrasound image. Placing the phantom coordinate origin within the plane one yields the following constraints for the k_{th} image

$$
\begin{pmatrix} x_k \\ y_k \\ z_k = 0 \\ 1 \end{pmatrix} = {}^{C}T_{T} \cdot {}^{T}T_{P} \cdot {}^{P}T_{U} \cdot \begin{pmatrix} s_x \cdot u_k \\ s_y \cdot v_k \\ 0 \\ 1 \end{pmatrix}, \tag{6.2}
$$

where x_k, y_k, z_k denote the 3D positions on the image plane, which has the constraint of $z = 0$ and u_k, v_k are 2D image coordinates of the imaged ultrasound plane (obtained from B-mode image) in pixels that are mapped to metric scale (mm) using scale factors s_x, s_y. Each imaged line gives up to two equations (that are not necessarily independent) subject to the constraint $z_k = 0$. For calibration 11 parameters need to be estimated. That is ${}^{C}T_{P}$ (2 rotations, 1 translation), ${}^{T}T_{U}$ (3 rotations, 3 translations) as well as the scale factors s_x and s_y require estimation. Having obtained a series of images such that an overdetermined set of equations can be constructed, the parameters can be obtained using an optimizer such as Levenberg-Marquardt.

Given the calibrated probe and tracking data, each pixel on the ultrasound plane can be spatially located w.r.t to a 3D reference coordinate system. Commonly, the reference coordinate system by the tracking device is used, such that we have the following chain of transformations, ${}^{T}T_{P} \cdot {}^{P}T_{U} = {}^{T}T_{U}$, providing for each pixel the position in the coordinate system of the tracker. Having a acquired a series of images in combination with their spatial location,

a 3D volume can be generated from the measurements. Placing a grid in the set of measurements and subsequently for each grid-position looking for measurements within a certain range corresponds to the backward-warping reconstruction process. This in contrast to forward-warping is less prone in producing holes. In the following Sec. 6.3.1 we will explain several methods that deal with the intensity reconstruction given the measurement collection determined for each voxel.

6.3.1 Intensity Reconstruction Methods

As presented in Sec. 6.2 there exist a multitude of algorithms for determining the voxel intensity from multiple measurements. In the following we will formalize some popular methods as can also be found in [Wein et al., 2006]. For each voxel $x_i \in X$, where X denotes the 3D volume containing all slices collected during an freehand acquisition phase, we have a set of tuples $A_i = \{(d_j, y_j)\}$ denoting the measurements that are in range of x_i, consisting of distance values d_j and intensity values y_j as obtained from the ultrasound slices. Then reconstruction methods are function f that map for each voxel x_i the corresponding tuples A_i to an intensity value $f(A_i)$.

Nearest Neighbor

One of the simplest possibilities of reconstruction is the nearest neighbor approach. Thereby only the intensity value, which comes from the closest sample point to the voxel centroid is considered [Sherebrin et al., 1996]. Therefore the corresponding mapping function is defined as

$$f(A) = y_j | d_j = min\left\{d_j\right\}. \tag{6.3}$$

The absence of smoothing in this approach generally leads to sharp images [Rohling et al., 1999] and due to its simplicity facilitates rapid image reconstruction.

Inverse Distance

Originally proposed for the interpolation of irregular spaced data [Shepard, 1968], in the inverse distance weighting approach, intensities are proportionally weighted according to their relative distance to the voxel centroid with respect to all other associated measurements. In the extreme case of $d_j \to 0$ the reconstruction fairly well approximates the original measured intensities.

$$f(A) = \sum_j y_j \frac{d_j^{-\mu}}{\sum_k d_k^{-\mu}}, \tag{6.4}$$

where $\mu > 1$ is a smoothness parameter.

Gaussian Kernel Smoothing

In order to obtain smooth images, kernel smoothing approaches are appropriate. By placing a 3D Gaussian kernel on each voxel and distance-weighting the intensities accordingly, smooth images can be generated. The corresponding mapping function is defined as

$$f(A) = \frac{\sum_j y_j e^{-d_j^2/\sigma^2}}{\sum_j e^{-d_j^2/\sigma^2}}, \tag{6.5}$$

where σ determines the width of the kernel and thus affects the strength of the smoothing.

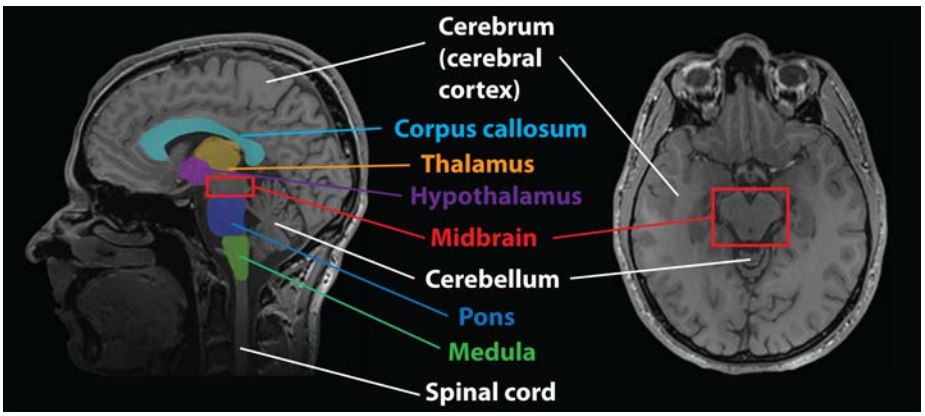


Figure 6.2: MRI slices of the brain. Highlighted are the main structures in the forebrain (cerebral cortex, corpus callosum, thalamus, hypothlamus and brainstem (midbrain, pons, medula).

6.4 Application in Neurology

6.4.1 Brief Overview of Brain Anatomy

In the following we give a brief overview of the brain parts and their functionality - see Fig. 6.2 for an illustration of the main structures of the brain. The brain is the center of the nervous system in human beings, housing an estimated number of 100 billion neurons, which are the biological processing units [Sherwood, 2012]. Neurons are electrically excitable cells with the ability to transmit and process information by means of electric and chemical means. Assembled together into a complex network they build up a sophisticated system that in its entirety provides important functionality such as cognition and emotionality. Protected by the skull bone, the brain is suspended in a liquid referred to as the cerebrospinal fluid (CSF), which fills the space between the skull and the brain surface. Thereby the CSF fulfills a series of functions such mechanical protection as shock absorber and immunological purposes.
Next we will have a closer look at the major structures that build up the brain. At first we start with the brainstem, which broadly speaking connects the brain to the body via the spinal coord. Despite its relatively small size compared to the rest of the brain, the brainstem is an extremely important part as it mainly acts as a relay station. Thereby it is largely responsible for activities such as information transmission of the the motor and sensory signals from the main part of the brain to the rest of the body. In this respect, the brainstem is build up from several structural components with characteristic functions, whereby we largely follow the description in [Saladin, 2010]. Being part of the brainstem, the midbrain is responsible for the control of sensory processes as well as providing signal relay functionality for vision and hearing [Barry E. Stein and Stanford, 2000]. Next the medulla is involved in the control of unconscious, essential functions such as blood circulation and breathing. Finally, the pons is involved in the control of sleep, thereby playing a role in

dream formation. Furthermore, the pons is responsible for the regulation of the change of inspiration to expiration and to relay signals [Saladin, 2010]. Located next to the pons, the cerebellum is responsible for precise and temporally well-coordinated execution of motor processes as well as regulatory control over muscle tone [Baehr and Frotscher, 2012].
Having covered the brain stem, we next discuss the main structures summarized as the forebrain. Within the forebrain, the cerebral cortex constitutes the largest part, which basically is the surface of the cerebrum. Broadly speaking, it is a complex folded neuronal tissue sheet with a thickness of 1-2 mm in human beings, which is often referred to as gray matter [Kandel et al., 2000]. Within this structure important information processing functionality is located such as memory, thought, language and thought. The other tissue in the cerebrum, below the cerebral cortex is referred to as white matter and mainly serves as interconnection between the neurons of the cortex. Another distinguished structure is the hypothalamus. It is responsible for the control of hunger, thirst, body temperature, aggression, fatigue, sleep as well as sexual behavior [Clark et al., 2010]. Last but not least, the thalamus has sensory relaying functions to the ceberal cortex related to vision, touch, hearing and taste, and it is involved in the regulation of consciousness and sleep [Nevid, 2012].

6.4.2 3D Freehand Ultrasound for Neurology

A promising domain for 3D freehand ultrasound technology is transcranial ultrasound for applications such as early diagnosis of Parkinson's disease. Within the corresponding data acquisition procedure image sequences are taken through a narrow bone window at the temporal lobe of the patient. Obtained imagery is capable of providing information about pathological changes within the midbrain and its surroundings - see Fig. 6.2 for images of the brain, highlighting certain major brain structures. Broadly speaking, the midbrain constitutes are part of the brainstem (midbrain, pons, medulla) which connects to the forebrain (cerebral cortex, corpus callosum, thalamus, hypothlamus), which is clearly visible in transcranial ultrasound. This has already been an active domain of research since the mid 90's [Becker et al., 1994a], however, only based on 2D images. Among the diseases that can be diagnosed using this ultrasound-based technique are Parkinson's [Becker et al., 1995b], atypical Parkinson syndromes [Behnke et al., 2005, Walter et al., 2003, Walter et al., 2004], depression [Becker et al., 1994b, Becker et al., 1995a], Dystonia [Naumann et al., 1996] as well as multiple sclerosis [Horowski et al., 2011, Walter et al., 2009]. Concerning Parkinson's disease, it has been established that the in 90% of the cases, its development process is correlated with a hyperechogenic substantia nigra [Berg et al., 2008, Walter et al., 2007]. Recent findings suggest that the process related to the development of Parkinson's disease is associated with the agglomeration of ferrite deposits [Berg et al., 2002], which is responsible for the formation of hyperechogenic areas in the midbrain region of substantia nigra that are visible in transcranial ultrasound. Noteworthy, beside related to serious and progressive movement disorder pathologies, the iron accumulation in midbrain structures as such is associated with the normal aging process of

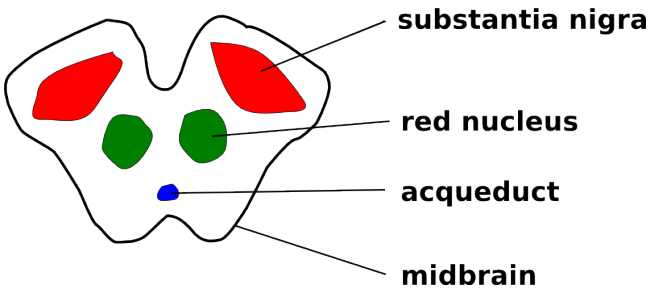
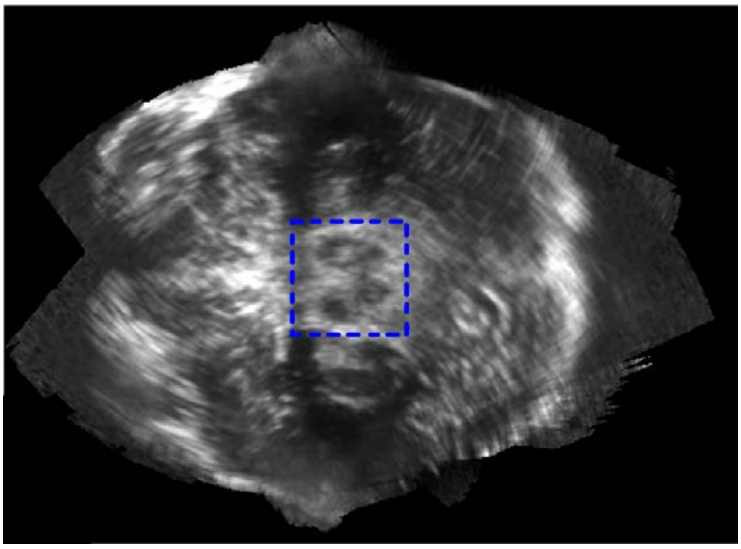


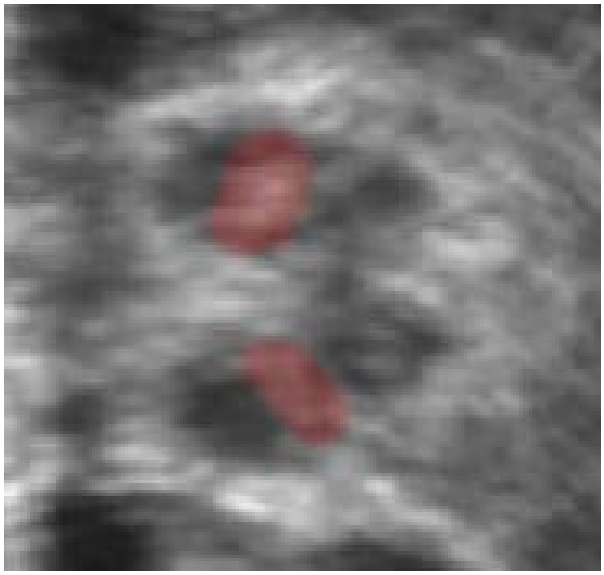
Figure 6.3: Schematic illustration of midbrain and structures within.

the brain [Hof and Mobbs, 2001].
Transcranial ultrasound for the early detection of Parkinson's disease is currently still performed using conventional 2D imagery, which makes the evaluation quite subjective due to the required experience in the method as well on the ultrasound machine. During the diagnosis, a specific ultrasound plane cutting through the midbrain is sought. On this plane the circumference of the hyperechogenetic substantia nigra is measured - see Fig. 6.3 for an illustration of the midbrain and structures herein. Risk is then assessed based on empirical thresholds. In this respect, machine settings and imaging properties clearly affect the differential threshold. Furthermore, user-experience affects the choice of the appropriate ultrasound plane, which is another element of uncertainty in the 2D based approach. The resulting subjectivity is therefore responsible for a limited inter- and intra-observer reliability as was recently shown in [Školoudík et al., 2007, Vlaar, 2011]. Altogether this typically leads to a wide range of sensitivity varying between varying from 48-100% compared to the final clinical diagnosis. Furthermore, the quality of the bone window through which the transcranial ultrasound is performed leads to a fairly high number of inconclusive results as reported in [Vlaar et al., 2009]. In this regard, 3D freehand ultrasound can be helpful to increase the diagnostic yield. See an illustration of the 3D freehand setup for the acquisition of bilateral scans Fig. 6.1. Noteworthy, in this setup a reference target is attached to the patient. This allows for the reconstruction of volumes w.r.t. to this specific target. As a result, multiple ultrasound sweeps can be merged to a single volume. In particular this facilitates the reconstruction of bilateral volumes that have shown useful to physicians. This is due to the fact that in bilateral reconstructions the outline of almost the entire skull is visible, which is helpful in getting an orientation. Furthermore, the skull represents a feature that potentially can be employed for registration tasks. See Fig. 6.4 for a 3D bilateral transcranial scan reconstruction and midbrain structures therein.

Generating 3D data from 2D images by means of a tracked transducer suggests being useful in (1) reducing the subjectivity in diagnosis, as no specific ultrasound plane has to be selected for classification and (2) reducing the drop out rate related to poor imaging of intracranial structures. Furthermore, accurate registration of data is required to perform continuous staging of hyperechogenic substantia nigra regions in patients. This is quite problematic in case of 2D due to the general limited repeatability of ultrasound acquisitions,



(a)



(b)

Figure 6.4: Ultrasound slices obtained from transcranial reconstructions. (a) shows image reconstructed from a bilateral 3D freehand ultrasound transcranial brain scan, highlighted structure in the center corresponds to midbrain. (b) shows corresponding magnified midbrain area with hyperechogenic substantia nigra segments highlighted in red, which are used for the classification of Parkinson's disease.

as this requires the recovery of virtually same plane position.
In order to assess the applicability of 3D ultrasound data for Parkinson's disease a first pilot study was conducted. Therefore data was classified based on

measurements of the midbrain and the substantia nigra echogenicities. The data set consisted of 23 subjects from which 3D freehand transcranial scans were acquired: 11 subjects with Parkison's disease, 11 healthy controls and 1 drop out (due to no or partially visualizable intracranial structures). Evaluation was performed on blinded segmentations of relevant structures in the remaining 22 subjects.
For classification a support vector machine (SVM) was employed using Radial Basis Functions (RBF) with associated variance parameters Γ as well as complexity parameter C (determining the smoothness of the classifying hyper-plane). The optimal parameter set for Γ and C were obtained using an automatic grid search with associated ranges $\Gamma \in \left[10^{-}8, 10^{6}\right]$, $C \in \left[10^{-}2, 10.^{9}\right]$. Given the 3D freehand reconstructed volumes, several features were tested for classification, including volumetric as well as planar properties. However, best results were achieved using a multi-dimensional feature vector consisting of the volumes of substantia nigra echogenicities in the left and right hemispheres, and the midbrain volume. Thereby classification yielded 90.9% sensitivity and 72.7% specificity. Furthermore, we obtained a drop-out rate of around 4%, significantly lower than using the 2D method, which is commonly reported to be in a range of 10-20% [Walter et al., 2007]. However, independent studies based on classification of substantia nigra area measurements directly obtained from B-mode imagery (on more than 100 subjects) have shown similar performance: sensitivity and specificity of around 90% were reported in [Berg, 2006]. Concerning this matter, employing volumetric features instead of only measuring 2D image properties promises powerful in diagnostics. First results in the pilot study suggest comparable classification performance to the B-mode based method, simultaneously being less dependent on the investigator's experience and bone window quality.

6.5 Confidence Weighted Reconstruction

Given that one of the main purposes of 3D freehand ultrasound is to extend the field-of-view of ultrasound and to visualize image anatomies from different angles, the reconstruction process is often confronted with complementary imaging information. This is particular the case in the presence of shadow artifacts, which occur in images due to among others total reflection of ultrasound beams. Shadow areas in image might cause severe reconstruction artifacts, if not treated properly. In the following we present a method incorporating confidence maps as presented in Sec. 4.6.2, to guide the reconstruction process for images that are prone to contain shadow regions.

In our example application we acquired 3D-freehand ultrasound sequences of the human femur from different acquisition angles. The acquisition was performed with an Ultrasonix MDP machine (Ultrasonix Medical Corporation, Richmond, Canada) in combination with a convex C5-2/60 transducer that was tracked with an optical NDI Polaris Spectra system (Northern Digital, Waterloo, Canada). For the purpose of the experiment and the sake of simplicity the leg of the volunteer was immobilized, i.e., tied up to a bed to ensure consistent acquisitions from different angles. Clearly, this is not an advisable procedure for patients, however, clinically viable solutions exist in this direction. In [Rohling et al., 1997] a pre-registration of volumes was proposed for the purpose of spatial compounding. Thereby multiple volumes are merged to a single improved on featuring higher quality.

Within the ultrasound scans of the femur, beside muscle tissue pretty much only the bone surface perpendicular to the ultrasound beam direction is visible because of the view-dependency and shadowing after the tissue-bone interface. Therefore, acquiring the bone from different views allows to obtain the missing bone surface and finally after reconstruction recover the complete bone structure. Such advanced reconstructions could be of use for a series of applications such as computer-aided orthopedic surgery with ultrasound guidance, as described in the works of [Penney et al., 2006, Barratt et al., 2006, Barratt et al., 2008].

As the images with view of the bone from different directions are containing complimentary information, the question of how to weight the contribution of overlapping or close ultrasound slices arises during the reconstruction process [Rohling et al., 1999, Grau et al., 2006]. This is of particular interest considering the presence of artifacts, i.e. shadow areas that contain low-confidence intensity information. Specifically, in the case of overlap between bone and shadow information due to view dependency issues, it is desirable to suppress shadow information in order not to lose important intensity information from relevant structures.

6.5.1 Method

In the following we will describe the process of reconstruction more formally. During the 3D freehand acquisition, a series of 2D images B-mode or RF images is obtained. During the phase of reconstruction, the image intensity is

distributed on a 3D lattice Ω, where at each discrete voxel position $i \in \Omega$ an intensity is computed from a set of measurements denoted as A_i. Thereby A_i is a set of tuples $A_i = \{(y_j, d_j)\}$, where each $a \in A_i$ consists of a distance d_j and associated pixel intensity value y_j. Specifically, d_j denotes the Euclidean distance between the voxel center and the pixel position. Thereby all points in the tuple set fulfill the distance constraint $d_i \leq D$, where D denotes the radius of a sphere limiting the measurements range of each voxel. A common choice for intensity-based reconstruction is to employ a Gaussian weighting function as given by

$$R_G(A) = \frac{\sum_i y_i e^{d_i^2/\sigma^2}}{\sum_j e^{d_i^2/\sigma^2}}, \tag{6.6}$$

where $R_G(A)$ is the reconstructed voxel intensity given the set of tuples A, containing the measurements. As a result, the reconstruction is a combination of a smoothed distance-weighted sum of pixel intensities in proximity to the voxel y. However, this results in pixels from shadow areas and tissue areas to contribute alike in the reconstructed voxel intensity. In the optimal case, one would like to reduce the contribution of shadow areas to areas were complementary anatomical information is available. Here the confidence estimation can be directly used for exactly this purpose. More specifically, for each 2D ultrasound image acquired (either B-mode or RF), we estimate a confidence map, which for each pixel gives the probability of containing reliable intensity information - see Sec. 4.6.2 for details on the confidence maps and their computation. Subsequently, in the reconstruction process we take the confidence into account to weight the intensity information accordingly. More formally, we define a confidence-based weighting function as

$$R_C(A) = \sum_j y_i \exp\left(-\frac{(c_i - \max(C))^2}{(\max(C))^2}\right), \tag{6.7}$$

where $A = \{(y_j, d_j, c_j)\}$ is now a triple, augmented by the confidence values c_j, where $C = \{c_j\}$ denotes the set of all confidence values associated to the voxel. Taking the confidence maps into account, the corresponding weighting function allows to weight the contribution of different slices based on their confidence; reducing significantly the contribution of low confidence regions. In this respect, regions like shadow are typically associated with low confidence values and are therefore down weighted in the confidence based reconstruction approach. This in turn provides predominance to useful structural information.

6.5.2 Results

Integration of confidence-weighting into the reconstruction process leads to images of superior quality with comparatively higher contrast than conventional approaches. Figure 6.5 and 6.6 show transverse and sagittal slices from an ultrasound volume of a human femur, reconstructing with a Gaussian (R_G) and the confidence-based (R_C) weighting function. The improvement in contrast is evident for the reconstruction utilizing the confidence estimate, especially

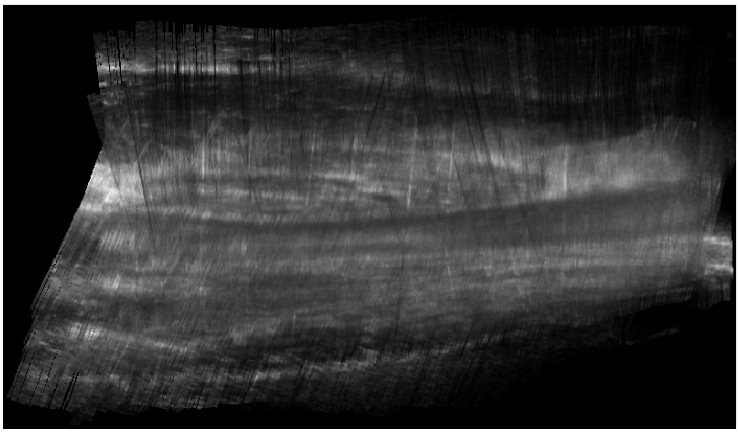
for the bone region. We measure the improvement with the Weber contrast as given by

$$W = \frac{I - I_b}{I_b}, \tag{6.8}$$

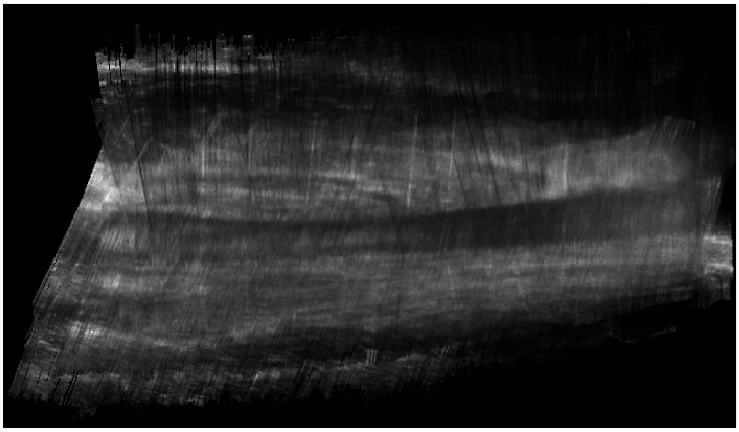
where I is the mean intensity inside the femur, I_b the background intensity outside the femur, and $I, I_b \in [0..1]$. For the reconstruction using the R_G function the contrast is $W = 0.32$ and using the R_C function it improves to $W = 0.45$. Noteworthy, the improvement in contrast is not the result of a linear intensity remapping, but the result of reducing the contribution of low confidence region. This is evident when examining the joint histogram, shown in Fig. 6.5(c), between two volumes reconstructed with R_C and R_G.

6.5.3 Conclusion

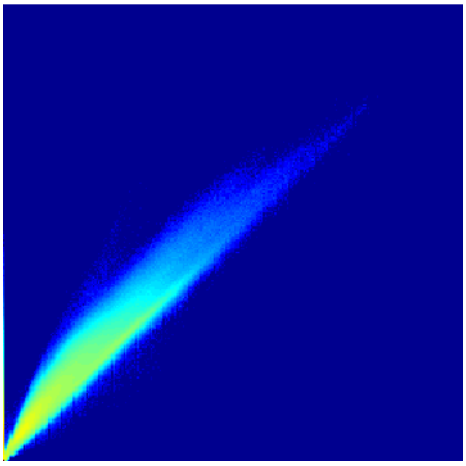
We presented confidence-based weighting approach for the process of 3D ultrasound reconstruction. The weighting aims to mitigate the effect of imaging artifacts and noise. Applicability was demonstrated on reconstructions of the femur due to the large presence of shadowing behind bone surface, which is obviously detrimental for standard reconstruction approaches. However, further fields of application are conceivable such as ultrasound of the brain where signal drop-out is prominent. This and further domains have to be evaluated in future studies.



(a) Gaussian R_G

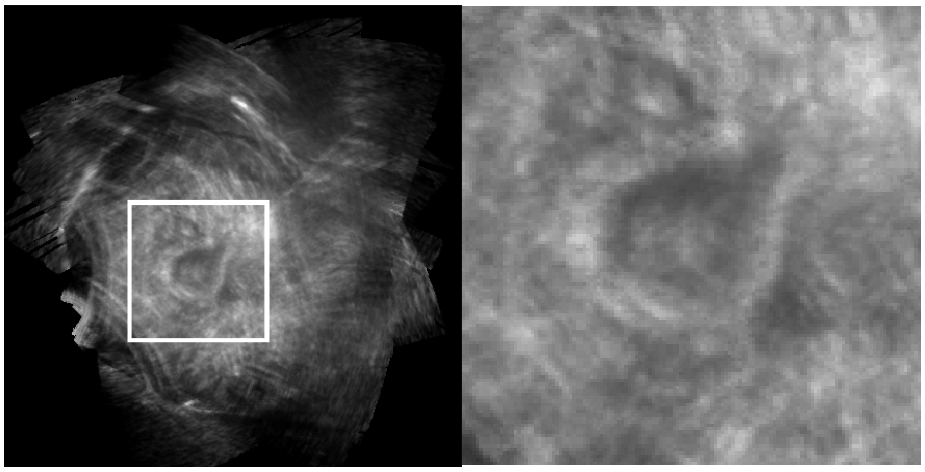


(b) Confidence R_C

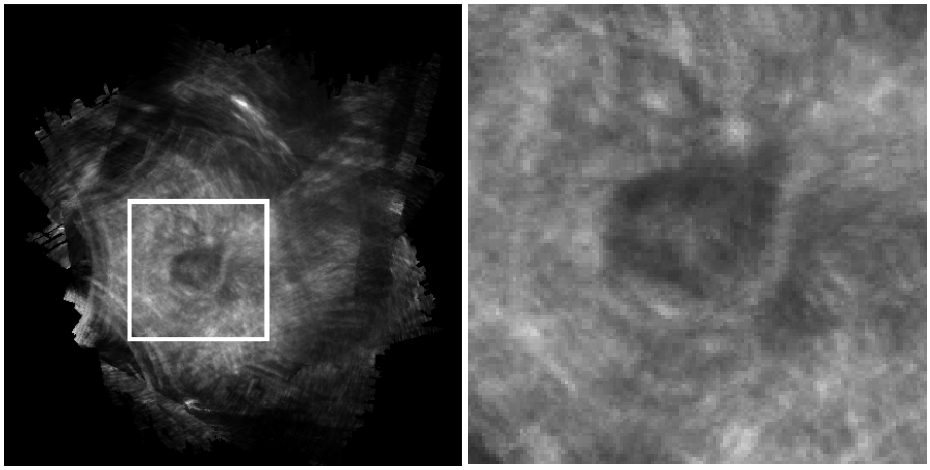


(c) Joint Histogram

Figure 6.5: Image (a) and (b) show a transverse slice of a human femur from a reconstructed volume using the R_G and the confidence-based R_C weighting function, respectively. Image (c) shows the joint histogram of two volumes, one reconstructed with R_G and the other with R_C.



(a) Gaussian R_G (b) Gaussian Enlarged



(c) Confidence R_C (d) Confidence Enlarged

Figure 6.6: Image (a) and (c) show a sagittal slice of a human femur from a reconstructed volume using the R_G and using the confidence-based R_C weighting function, respectively. Image (b) and (d) show an enlarged version of the regions marked with the white box.

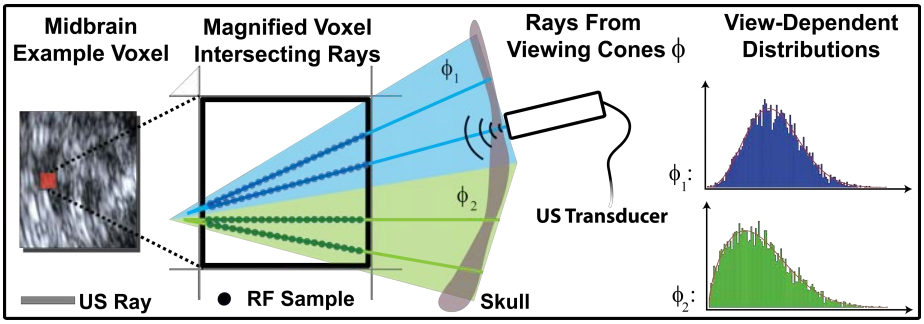


Figure 6.7: Distributional change for cones ϕ_1 and ϕ_2 within a single view of midbrain.

6.6 Mixture Modeling of Multi-View 3D RF Ultrasound

6.6.1 Freehand 3D RF Data

In order to obtain 3D ultrasound RF freehand data we follow the approach of [Wein et al., 2006], with the important and fundamental distinction of using RF data rather than B-mode data. Due to the different spatial relationship of the data, RF requires disintegration into individual scanlines such that the reconstruction process becomes ray-based - see Fig. 6.7. Beside the intensity data we also record geometric information such as view point and direction. This additional data is used in a follow-up processing step. Using RF instead of B-mode is advantageous, since we have access to the unfiltered native signal, uncorrupted by proprietary filters.

The distribution of the envelope of the RF signal, resulting from backscattered tissue echo, has been shown to be modeled, in a simple and versatile way, by the Nakagami distribution [Shankar, 2000]. Thus we assume all intensities in the RF envelope image Ω to follow a Nakagami distribution

$$\mathcal{N}(x|\mu,\omega) = \frac{2\mu^{\mu}x^{2\mu-1}}{\Gamma(\mu)\omega^{\mu}} \exp\left(-\frac{\mu}{\omega}x^2\right) \quad \text{s.t.} \quad \forall x \in \mathbb{R}_+ \,, \tag{6.9}$$

with μ, ω denoting the shape and scale parameters respectively.

6.6.2 Mixture Model Motivation

It is well-known that ultrasound is highly view-dependent [Hedrick et al., 2004], and thus it desirable to incorporate this property when modeling backscatter. However, we are not interested in the individual backscatter intensities, but in the distribution within a small finite volumetric element (voxel) w.r.t views of a data point x, i.e.

$$p(x) = \int_{\mathcal{D}(x)} p(x|\phi)p(\phi)d\phi \,, \tag{6.10}$$

where $\mathcal{D}(x)$ is the set of all possible viewing cones ϕ of x. The distribution $p(x)$ is approximated by a finite mixture model of K Nakagami densities [Destrempes et al., 2009],

$$p(x) \approx \sum_{k=1}^{K} p(\phi_k)p(x|\phi_k) = \sum_{k=1}^{K} w_k \mathcal{N}(x|\mu_k, \omega_k) \quad \text{s.t.} \quad \sum_{k=1}^{K} w_k = 1 \,, \tag{6.11}$$

where the distribution $p(\phi_k)$ of the k^{th} cone is represented by a mixture weight. The K viewing cones are assumed approximately evenly spaced around the object of interest. See Fig. 6.7 for illustration of viewing cones of the midbrain with beams originating at different skull positions. A popular choice for finite mixture model estimation is the Expectation-Maximization (EM) algorithm [Dempster et al., 1977]. However, as the EM algorithm can potentially over-fit the data and is also quite flexible in component modeling, we do not instantiate it on the pooled data from all views. Rather, we instantiate individual mixture estimations within geometrical subspaces obtained from each view (each containing viewing cones) that were recorded during the acquisition process. Altogether this yields robust component estimation, s.t.

$$p(x) = \sum_{i=1}^{N} \sum_{k=1}^{K_i} w_k p(x|\phi_{i,k}) = \sum_{i=1}^{N} \sum_{k=1}^{K_i} w_k \mathcal{N}(x|\mu_{i,k}, \omega_{i,k}) \,. \tag{6.12}$$

Specifically, within each view $i \in N$ we determine the number of components $K \leq K_{max}$, following the approach of Frayley and Raftery [Fraley and Raftery, 2002]. Here the Asymptotic Minimum Description Length (AMDL / BIC) principle selects the finite mixture model with d free parameters, which minimizes the quantity $-2\log\mathcal{L} + d\log n$, where $\mathcal{L}$ is the likelihood of data given model parameters and n the number of observations. As a direct consequence the number of components naturally varies from point to point. This estimation process is performed voxelwise in the entire volume, yielding a mixture model representation for each voxel.

6.6.3 Mixture of Nakagami

In the following section we will elaborate on the gamma mixture estimation for envelope detected RF data as proposed in [Destrempes et al., 2009]. Although the data follows Nakagami distribution, the Gamma distribution may be used instead as they are related by a simple random variable transformation [Papoulis and Pillai, 2002, p.93], given by $Y \sim \mathcal{GA}(x \mid \mu_{gam}, \omega_{gam}), X \sim \mathcal{N}(x \mid \mu_{nak}, \omega_{nak}) \Rightarrow \sqrt{X} = Y(\mu_{nak}, \omega_{nak}/\mu_{nak})$. The mixture of gamma is defined as

$$p(x) \approx \sum_{k=1}^{K} p(\phi_k)p(x|\phi_k) = \sum_{k=1}^{K} w_k \mathcal{GA}(x|\mu_k, \omega_k) \quad \text{s.t.} \quad \sum_{k=1}^{K} w_k = 1 \,, \tag{6.13}$$

where $\theta = \{w_i, \mu_i, \omega_i\}_{i=1..K}$. See Fig. 6.8 for an illustration of a Nakagami mixture model. Following the Bayesian paradigm we need to define a prior distribution for the parameters in θ. In this respect, the mixture weights w_k are modeled using a Dirichlet prior [Escobar, 1994]

$$p(w_1, ..., w_K) = \mathcal{D}(w_1, ..., w_K|A_0, \alpha_1, ..., \alpha_K), \tag{6.14}$$

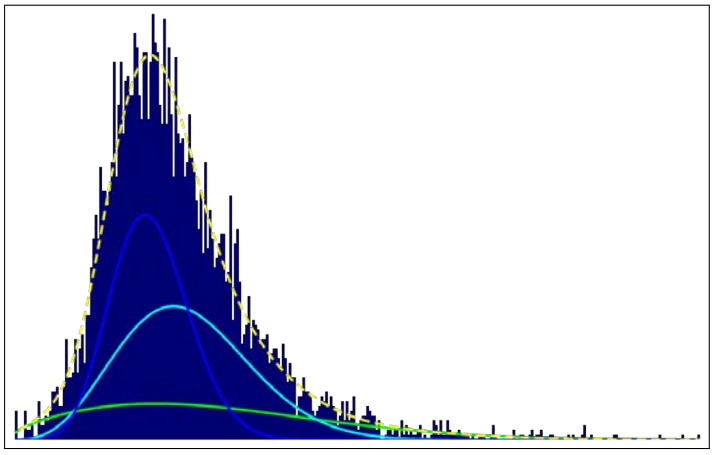


Figure 6.8: Illustration of Nakagami mixture model. The mixture model consists of three components, represented with solid lines. The dashed line represents the weighted sum of all components.

where $\mathcal{D}$ denotes the Dirichlet distribution defined as

$$\mathcal{D}(x_1,...,x_K|A_0,\alpha_1,...,\alpha_K) = \frac{\Gamma(A_0)}{\prod_{i=1}^{K}\Gamma(A_0\alpha_i)}\prod_{i=1}^{K}x_i^{A_0\alpha_i-1} \quad \text{s.t.} \quad \sum_{k=1}^{K}x_k = 1, x_i \geq 0\,, \tag{6.15}$$

where A_0 denotes a confidence level. Given $p_1,...,p_K \sim \mathcal{D}$ it follows that the mean of p_i corresponds to α_i, with the associated variance of $\alpha_i(1-\alpha_i)/(A_0+1)$. Furthermore, we assume an uniform prior for the shape μ_k and the scale ω_k parameters of each Gamma component $\mathcal{GA}$ defined as

$$p(\mu_1,\omega_1,...,\mu_K,\omega_K) \propto \begin{cases} 1, & \text{on} \quad \text{U} \\ 0, & \text{otherwise} \end{cases}\,, \tag{6.16}$$

which is restricted to a compact set $U = ([\mu_{min},\mu_{max}]\times[\omega_{min},\omega_{max}])^K$. Given the priors on the distributional parameters θ, we are now able to estimate parameters using a maximum a posteriori (MAP)

$$\tilde{\theta} = \arg\max \underbrace{p(y|\theta)}_{\text{Likelihood}} \cdot \underbrace{p(\theta)}_{\text{Prior}} = \arg\max \underbrace{p(\theta|y)}_{\text{Posterior}}, \tag{6.17}$$

estimator by employing the EM-algorithm - see Sec. 4.6.7. The maximum-a-posterior EM (MAP-EM) approach was originally proposed in [Green, 1990] for reconstructions of emission tomography. Note, estimation without the prior corresponds to the normal EM-algorithm as proposed in [Dempster et al., 1977]. In particular, computing the MAP-EM instead of the standard EM leaves the E-step unchanged. Only the M-step is affected, as in case of MAP-EM not the expected log-likelihood (see Eq. 4.100) is maximized but the expected

posterior probability,

$$\mathcal{Q}(\theta, \theta^{(t)}) = E\left[\log p(X, Z|\theta) + \log p(\theta)|X, \theta^{(t)}\right] = \sum_{z \in Z} \log\left\{p(X, z|\theta)p(\theta)\right\} p(z|X, \theta^{(t)}), \tag{6.18}$$

which in case of a Gamma mixture is up to a constant equal to the following term

$$\sum_{j=1}^{N}\sum_{i=1}^{K}\left(\log \mathcal{GA}(x_j|\mu_i, \omega_i) + \log w_i\right) p(z_j = i|X, \theta^{(t)}) + \sum_{i=1}^{K}(A_0\alpha_i - 1)\log w_i. \tag{6.19}$$

As outlined in Sec. 4.6.7 the E-step consists in computing the class association probability, which corresponds to $p(z_j = i|X, \theta^{(t)})$ and in the case of a Gamma mixture is defined as

$$p(z_j = i|X, \theta^{(t)}) = \frac{w_i \mathcal{GA}(x_j|\mu_i, \omega_i)}{\sum_{i=1}^{K} w_i \mathcal{GA}(x_j|\mu_i, \omega_i)}. \tag{6.20}$$

For the M-step we need to augment Eq. 6.19 with Lagrange-multipliers λ in order to accommodate for the constraints $\sum_{i=1}^{K} w_i = 1$, such that we yield

$$\sum_{j=1}^{N}\sum_{i=1}^{K}\left(\log \mathcal{GA}(x_j|\mu_i, \omega_i) + \log w_i\right) p(z_j = i|X, \theta^{(t)}) + \sum_{i=1}^{K}(A_0\alpha_i - 1)\log w_i + \lambda\left(1 - \sum_{i=1}^{K} w_i\right). \tag{6.21}$$

Next we have to find the parameter maxima by zero-setting the derivatives of $\mathcal{Q}(\theta, \theta^{(t)})$ - Eq. 6.21 - w.r.t. the parameters in θ. First we start with the derivatives w.r.t. the mixing weights w_i.

$$0 = \frac{\partial \mathcal{Q}(\theta, \theta^{(t)})}{\partial w_i} = \frac{1}{w_i}\left(\sum_{j=1}^{N} p(z_j = i|X, \theta^{(t)}) + A_0\alpha_i - 1\right) - \lambda \tag{6.22}$$

Setting $P_i = \sum_{j=1}^{N} p(z_j = i|X, \theta^{(t)})$ we yield $P_i + A_0\alpha_i - 1 = w_i\lambda$. With $\lambda = N + A_0 - K$ that we obtained by reformulation,

$$\sum_{j=1}^{N} p(z_j = i|X, \theta^{(t)}) + A_0\alpha_i - 1 = w_i\lambda \quad \Leftrightarrow \tag{6.23}$$

$$\sum_{i=1}^{K}\left(\sum_{j=1}^{N} p(z_j = i|X, \theta^{(t)}) + A_0\alpha_i - 1\right) = \sum_{i=1}^{K} w_i\lambda \quad \Leftrightarrow \tag{6.24}$$

$$\sum_{j=1}^{N}\sum_{i}^{l} p(z_j = i|X, \theta^{(t)}) + A_0\sum_{i=1}^{K}\alpha_i - \sum_{i=1}^{K} = \sum_{i=1}^{K} w_i\lambda \quad \Leftrightarrow \tag{6.25}$$

$$N + A_0 - K = \lambda \tag{6.26}$$

(since $\sum_{i=1}^{l} p(z_j = i|X, \theta^{(t)}) = 1$ and $\sum_{i=1}^{l} \alpha_i = 1$) we yield the following expression

$$w_i = \frac{P_i + A_0\alpha_i - 1}{N + A_0 + K}. \tag{6.27}$$

Next the derivative w.r.t to the shape parameter μ_i is defined as

$$0 = \frac{\partial \mathcal{Q}(\theta, \theta^{(t)})}{\partial \mu_i} = \sum_{j=1}^{N} \left(-\log \omega_i - \psi(\mu_i) + \log x_j \right) p(z_j = i|X, \theta^{(t)}). \tag{6.28}$$

Finaly, the derivative w.r.t. to the scale parameter ω_i is given by

$$0 = \frac{\partial \mathcal{Q}(\theta, \theta^{(t)})}{\partial \omega_i} = \sum_{j=1}^{N} \left(-\frac{\mu_i}{\omega_i} + \frac{x_j}{\omega_i^2} \right) p(z_j = i|X, \theta^{(t)}). \tag{6.29}$$

Setting $Q_i = \sum_{j=1}^{N} x_j p(z_j = i|X, \theta^{(t)})$ in combination with Eq. 6.29 implies $\omega_i = \frac{Q_i}{\mu_i P_i}$. Furthermore, setting $R_i = \sum_{j=1}^{N} \log x_j p(z_j = i|X, \theta^{(t)})$ and substituting this together with Q_i in Eq. 6.28, one yields the updating equation for the shape parameter μ_i defined as

$$\log \mu_i - \psi(\mu_i) = \log(Q_i/P_i) - (R_i/P_i), \tag{6.30}$$

with $\psi(x)$ denoting the digamma function. Given the concavity property of the log function on $(0, \infty)$ and Jensen's inequality, it follows that $\log(Q_i/P_i) - (R_i/P_i) \geq 0$, where equality holds if and only if all x_j are equal, which has probability 0. Furthermore, since $\log x - \psi(x)$ is a decreasing function on $(0, \infty)$ with the properties $\lim_{x\to 0} \log x - \psi(x) = \infty$ and $\lim_{x\to\infty} \log x - \psi(x) = 0$, the scale parameters with $\mu_i > 0$ can be found using binary search.

6.6.4 Registration

For registration we perform a voxel-wise distribution matching employing the J-divergence in conjunction with a data fidelity term. This provides higher robustness compared to a pure intensity-based model, as was also observed in [Ijaz et al., 2011]. Considering a fixed volume (A) and moving volume (B), we seek the rigid transformation $\hat{T}$ that yields optimal spatial alignment between the two, s.t. $\hat{T} = \arg\min_T D_{PJD}(A, T(B))$ for the pseudo-distance

$$D_{PJD}(C, D) = \arg\min_{i,j} \sum_{k=1}^{Z} \mathcal{J}(f^i_{C_k}, f^j_{D_k}) \cdot \exp\left(\lambda \cdot \left((1 - w^i_{C_k}) + (1 - w^j_{D_k}) \right) \right), \tag{6.31}$$

where $w_i, w_j \geq \tau$. We refer to this pseudo-distance as *Pseudo-J-Divergence* (PJD). For each voxel $k \in Z$ it takes the mixture components pair $f^i_{C_k}$ and $f^j_{D_k}$ from the two volumes C and D, respectively, to be registered with least J-divergence [1] [Jeffreys, 1946] times the exponentially scaled sum of corresponding mixture weights $w^i_{C_k}, w^j_{D_k}$ (with minimum weight τ, and a scaling parameter λ). The

[1] Jeffreys (J) divergence is also known as symmetric KL distance.

Algorithm 6.1 Expectation Maximization for Gamma Mixture Model

Input: Data $X = \{x_1, ..., x_N\}$
Define initial values: $\theta^{(0)} = \left\{w_i, \mu_i^{(0)}, \sigma_i^{(0)}, \sigma_i^{(0)}\right\}_{i=1..K}$
for j=1,..,N **do**
Random draw of $c_j = i$ with uniform probability $w_i = 1/K$
end for
for i=1,...,K **do**
Set $w_i = \frac{N_i + A_0\alpha_i - 1}{N + A_0 - K}$, with N_i denoting the number of elements in X associated with class i
Compute the MLE estimates μ_i, θ_i for class i based on the associated data
end for
repeat
$t \leftarrow t + 1$
E-Step: Compute class association probabilities/responsibilities

for all $j \in N, \quad k \in K$ **do**

$$p(z_j = k|X, \theta^{(t)}) = \frac{w_i \mathcal{GA}(x_j|\mu_i, \omega_i)}{\sum_{i=1}^{K} w_i \mathcal{GA}(x_j|\mu_i, \omega_i)}$$

end for

M-Step: Estimate the the parameters given the current class associations/responsibilities
for all $k \in K$ **do**
Update class probabilities:

$$w_k^{(t)} = \frac{P_i + A_0\alpha_i - 1}{N + A_0 - K} \quad \text{with} \quad P_k = \sum_{j=1}^{N} p(z_j = k|X, \theta^{(t)})$$

Compute the shape parameter μ_k using binary search

$$\text{Solve: } \log x - \psi(x) = \log \frac{Q_k}{P_k} - \frac{R_k}{P_k}$$

$$Q_k = \sum_{j=1}^{N} x_j p(z_j = k|X, \theta^{(t)})$$

$$R_k = \sum_{j=1}^{N} \log x_j p(z_j = k|X, \theta^{(t)})$$

$x \rightarrow \mu_k$
Update the scale parameter ω_k: $\theta_k = \frac{Q_k}{\mu_k P_k}$
end for
until Convergence of parameters or log-posterior

pseudo-distance D_{PJD} does not satisfy the triangle inequality, but inherits symmetry and uniqueness from the J-divergence $\mathcal{J}(f,g)$. The exponential weight in Eq. 6.31 is a measure of confidence for the pseudo-distance, punishing distances formed from components with low mixture weights, since these components are assumed to be less descriptive of the underlying distribution. The J-divergence between two Gamma distributions $f, g \in \mathcal{GA}$, is the sum of two non-symmetric KL distances with switched arguments, s.t.

$$\begin{aligned}\mathcal{J}(f,g) = KL(f,g) + KL(g,f) = \int \log\frac{f(x)}{g(x)}(f(x)-g(x))dx = \\ (\mu_a - 1)\theta(\mu_a) - \log\omega_a - \mu_a - \log\frac{\Gamma(\mu_a)}{\Gamma(\mu_b)} \\ + \quad \mu_b\log\omega_b - (\mu_b - 1)(\theta(\mu_a) + \log\omega_a) + \frac{\omega_a\mu_a}{\omega_b}\,.\end{aligned} \tag{6.32}$$

Although the data follows Nakagami distribution, the Gamma distribution may be used instead as they are related by a simple transformation, given by $Y \sim \mathcal{GA}(x \mid \mu_{gam}, \omega_{gam}), X \sim \mathcal{N}(x \mid \mu_{nak}, \omega_{nak}) \Rightarrow \sqrt{X} = Y(\mu_{nak}, \omega_{nak}/\mu_{nak})$.

6.6.5 Reconstruction

Given the voxel finite mixture model representation we can perform a novel type of reconstruction. Therefore a reference component is chosen from the mixture model for each voxel. Here we opted for the component associated with the maximum mean intensity, although other approaches are conceivable. However, this criterion guarantees that no high intensity backscatter is missed during reconstruction. Note that artifacts, such as shadows, might require specific treatment. By optimizing the reference component parameters, s.t. the sum of geodesic distances to neighbors component parameters, on the manifold $\mathcal{G}$ of Gamma model parameters, is minimized, smoothness of the reconstructed volume is achieved. Thus, given a point $\theta_a = (\mu_a, \omega_a) \in \mathcal{G}$, the geodesic distance $D_{geodesic}$ to a locally neighboring point $\theta_b = (\mu_b, \omega_b) \in \mathcal{G}$ is bounded s.t.

$$D_{geodesic}(\mathcal{GA}(x|\theta_a), \mathcal{GA}(x|\theta_b)) \leq \left|\frac{d^2\log\Gamma}{d\mu^2}(\mu_b) - \frac{d^2\log\Gamma}{d\mu^2}(\mu_a)\right| + \left|\omega_a \cdot \log\frac{\omega_a}{\mu_b}\right|\,, \tag{6.33}$$

[Arwini et al., 2008, Ch. 7] and [Dodson and Matsuzoe., 2002]. Applying Eq. 6.33, the reference distribution θ_{ref} is optimized by minimizing the sum of geodesic distances to all of its neighbors s.t.

$$\hat{\theta}_{\text{ref}} = \underset{\theta_{\text{ref}}}{\operatorname{argmin}} \sum_{k \in N_{\text{ref}}} D_{geodesic}(\theta_{\text{ref}}, \theta_k)\,, \tag{6.34}$$

keeping the neighbours $\theta_k \in N_{\text{ref}}$ fixed, where N_{ref} defines the neighborhood of a reference voxel. This yields a spatially consistent image without over-smoothing or loss of detail in terms of highlights. Typically a few optimization steps are sufficient and allow for a fast reconstruction.

For data reconstruction within a voxel we apply a Gaussian-weighted (GW) [Wein et al., 2006] reconstruction scheme in order to increase homogeneity,

$$y_j = \frac{1}{Z}\sum_{i=1}^{N} x_i e^{-d_i^2/\sigma^2} \quad \text{s.t.} \quad Z = \sum_{i=1}^{N} e^{-d_i^2/\sigma^2}\,, \tag{6.35}$$

yielding the reconstructed intensity y_j at voxel position j. Here the intensities x_i are sampled from the reference distribution, where d_i are the distances from the voxel centroid that can be obtained by regression from the measured data.

6.6.6 Results

Registration

For testing the registration performance multiple transcranial 3D volumes for several patients were acquired. An acquisition consisted of bilateral scans using an Ultrasonix MDP (Ultrasonix Medical Corporation, Richmond, Canada) ultrasound machine in combination with an optical tracking system NDI Polaris Spectra (Northern Digital, Waterloo, Canada); in doing so various sweeps from numerous possible views were obtained. Furthermore, each patient was recorded with a reference target rigidly attached to the head, which allows to establish ground truth position between numerous volumes.

The three-dimensional RF datasets were acquired with a phased-array probe with a frequency of 3.3 MHz and depth 14 cm. RF data is sampled with 40 MHz. Each 3D RF data set is built from approximately 4000 RF images (approx. 2000 images from each side of the skull), each having a resolution of 3648 x 96 pixels. Volumes were reconstructed with isotropic voxel size of 0.65 mm. For the mixture modeling we assumed $N = 2$ views (bilateral) as well as $K_{max} = 2$. This yielded a maximum of four mixture components per voxel and sufficiently modeling the data while avoiding over-fits.

For evaluating the quality of the proposed approach, we performed registration by block matching, which is commonly used for ultrasound [Poon and Rohling, 2005, Krucker et al., 2002]. For each patient two multi-view volumes are constructed for distinct ultrasound data. These two distinct volumes are then registered by taking 27 equally distributed blocks (each of size $6 \times 6 \times 6$ voxels) for matching, within each multi-view volume. Parameters for our distance metric were set to $\lambda = 2$ and $\tau = 0.3$. In order to be able to perform statistics on the quality of the registration approach, we perform a random rigid-registration study with 10 runs per individual block. For each run we randomly displace the moving block with initial deviation of ± 6 cm in each spatial direction from its ground-truth position. This is followed by registration runs at each position using state of the art similarity metrics for ultrasound, aligning the moving and the fixed volume. In spite of rigidity, registration of transcranial brain ultrasound brain data is quite challenging due to the relative low SNR as a result of variable transmission through skull bone and the low transducer frequency required. Nevertheless, the proposed pseudo-J-Divergence (PJD) yields up to 15% better registration results compared to NCC and SSD [Poon and Rohling, 2005, Krucker et al.,

Dataset	SSD	NCC	PJD
#1	3.2 ± 1.3	3.0 ± 1.3	$\mathbf{2.7 \pm 1.1}$
#2	2.7 ± 1.3	2.7 ± 1.2	$\mathbf{2.4 \pm 1.2}$
#3	3.2 ± 1.2	3.3 ± 1.1	$\mathbf{2.9 \pm 1.1}$
#4	3.7 ± 1.1	3.4 ± 1.1	$\mathbf{3.0 \pm 1.1}$
#5	3.4 ± 1.1	3.2 ± 1.1	$\mathbf{2.9 \pm 1.1}$
#6	3.3 ± 1.2	3.4 ± 1.2	$\mathbf{2.9 \pm 1.1}$

Table 6.1: Median errors (mm) and standard deviation of registration study

2002] (intensity volumes created using GW [Wein et al., 2006]), as can be seen in Tab. 6.1. In order to assess the significance of the results statistical tests were performed. The Hodges-Lehmann (HL) confidence interval for difference in median error between PJD and NCC is (-0.25, -0.08), at confidence level 0.95. The corresponding confidence interval for PJD/SDD is (-0.3, -0.12). The P-value for a Mann-Whitney U-test for PJD/NCC is 0.14E-3, and for PJD/SSD the P-value is 0.2E-6. Thus it is clear that the median error of PJD is statistical significantly lower than both SSD and NCC, at level 0.05. Obtained results are quite close to technical possible limits due to the errors that are propagated through the processing chain. In this respect, ultrasound calibration affects the accuracy significantly, depending largely on the approach used. In our setup, calibration was performed using a single wall phantom, yielding an estimated error between 1-2 mm [Hsu et al., 2007, Mercier et al., 2005].

Reconstruction

As reconstruction methods are difficult to compare due to the highly subjective nature of ultrasound image analysis, we will elaborate on the obvious differences between a state of the art method (Gaussian-weighted, increased homogeneity favorable for segmentation and registration [Wein et al., 2006]) and our approach as well as discuss the potential implications. Reconstructions following the geodesic approach yield more coherent and homogeneous images - see Fig. 6.9. Furthermore, checker-board and radial artifacts disappear, which are visible in the Gaussian-weighted approach due to the multi-view mixture nature of the signal. Additionally, areas like the midbrain exhibit sharp edges. Regions that are clinically relevant for classification such as the substantia nigra, where the level of hyperechogenicity is a risk-assessment criterion for Parkinson's disease [Walter et al., 2007], are more pronounced and less blurred (see Fig. 6.3 for a schematic illustration of the midbrain and structures within). The clinical benefit for applications such as classification, however, remains to be evaluated in further clinical trials.

6.6.7 Conclusion

We have presented a finite mixture model representation of 3D RF data exploiting its view-dependent statistical properties. Making use of view-dependency has potential to facilitate several applications; specifically, results from a block-

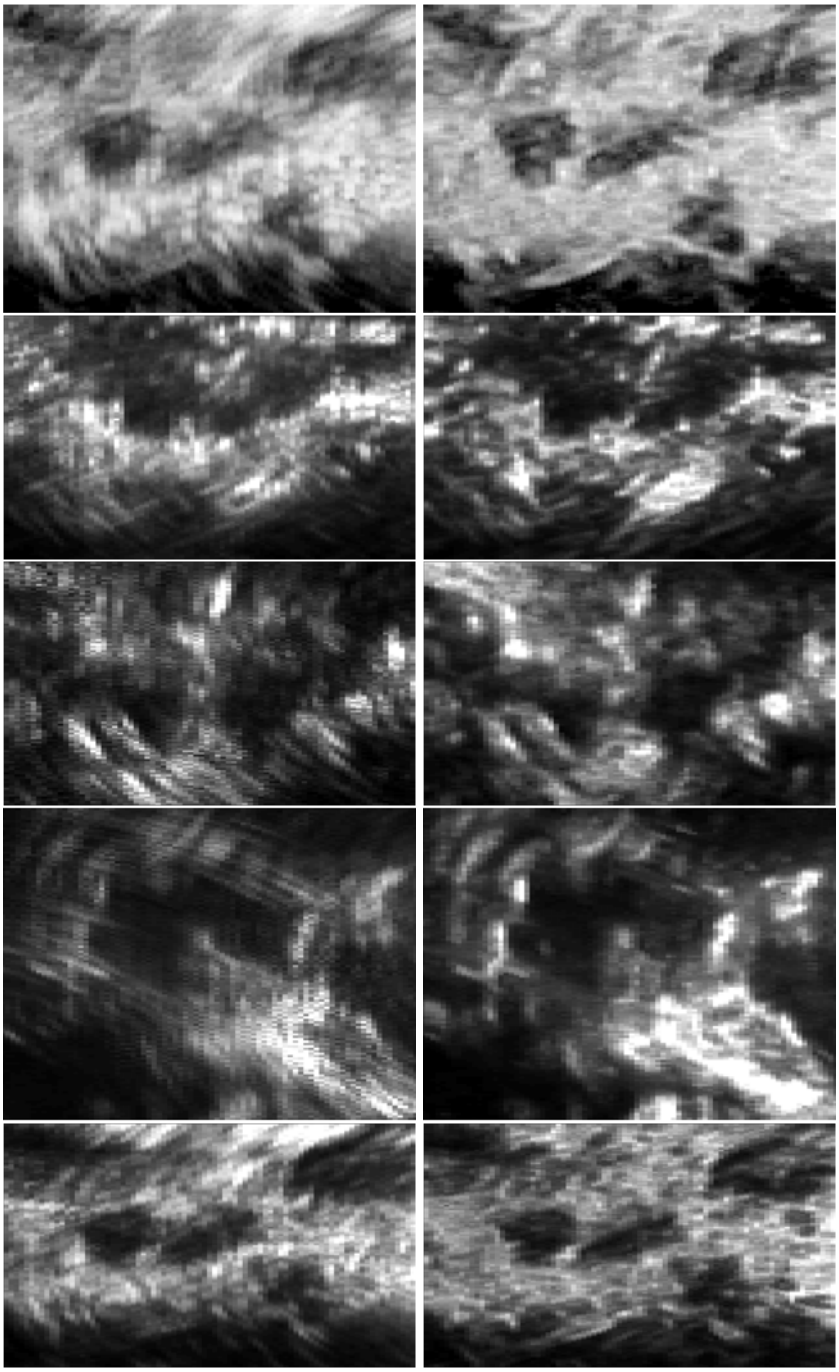


Figure 6.9: Transcranial ultrasound reconstructions. Butterfly-shaped region in the image center corresponds to the midbrain. Left: Gaussian-weighted. Right: Geodesic.

matching based rigid registration study suggest improvements in terms of accuracy compared to conventional similarity metrics. Moreover, image reconstruction promises to be an interesting domain of application. In this respect further application scenarios will be studied. A possible domain could be ultrasound for the bone, where 3D could prove beneficial for fracture reduction images taken from multiple views.

Chapter 7

Conclusion

In this dissertation, we discussed several issues in the field of ultrasound imaging. The main focus was reconstruction of 3D ultrasound from 2D images and the image registration by incorporating ultrasound specific statistical properties. In this respect, issues of ultrasound specific texture modeling incorporating modality specific statistic properties. Furthermore, we also presented the 2D analytic signal and its positive effect on statistical modeling of ultrasound data.

Ultrasound Data Modeling

We presented an overview of the ultrasound processing pipeline in conventional machines. The underlying concepts served as basis for the following adaptation of methods to ultrasound, in particular the usage of RF data instead of B-mode. This was followed by an overview over the various existing distributional models for RF data, from which due to its versatility and simplicity the Nakagami was chosen for follow-up methods. Next the application of the 2D analytic signal was proposed for computing the envelope of RF data. The resulting improved statistical properties were illustrated by means of goodness-of-fit tests. Incorporation of the 2D analytic signal ultrasound specific similarity measures then allowed us to directly benefit from the advanced statistical properties.

Ultrasound Texture Modeling

Next we presented an overview about texture modeling in ultrasound imaging. Particular focus was devoted to statistical approaches. We showed how texture patterns can be tied together with a learning based metric to model the notion of textural similarity. This was followed by an application of shadow detection in ultrasound images containing mirroring and reverberation artifacts due to total reflection. The proposed method made use of statistical texture analysis, mixture modeling of Nakagami images as well as confidence maps.

Ultrasound Similarity Metrics

In this chapter we presented two similarity metrics for 2D ultrasound images. At first we presented a hybrid similarity metric, combining local and global properties. In this respect, distribution matching was performed on global scale due to its robustness towards noise. On the local scale, a statistics-based variant of the Fuzzy Local Binary Patterns was used to capture small textural details. This was followed by a similarity metric based on the bivariate Nakagami distribution, for which we proposed a locally adaptive parameter estimation scheme, implicitly making use of the 2D analytic signal. All proposed similarity metrics yielded superior results in random registration studies as compared to conventional metrics such as SSD and NCC.

3D Freehand Ultrasound

We discussed the use of 3D freehand ultrasound systems and how its use can lead to a qualitative better and more subjective early diagnosis of Parkinson's disease. This was followed by the presentation of several advanced techniques for improved 3D freehand ultrasound reconstruction. In this respect, we considered the creation of bilateral ultrasound volumes for neurological diagnosis. Next the incorporation of confidence maps as an intensity weight function for the reconstruction process. Finally, we concluded the chapter with the multi-view 3D RF ultrasound data representation employing Nakagami finite mixture model formalism. This allowed for more coherent images with reduced artifacts, while increasing the accuracy in registration procedures.

Chapter 8

Perspectives

In this dissertation, we discussed several topics in the domain of ultrasound image analysis and presented extensions or alternatives to state of the art approaches. In the following we discuss further possible steps on how to continue this line of research.

3D Multi-View Ultrasound

We presented an approach for multi-view data representation of 3D freehand ultrasound employing a finite mixture model under Nakagami distribution assumption. We believe that his modeling approach has many potential applications that should be investigated. At first, the influence of different reconstruction methods based on the mixture components should be evaluated. Different applications require reconstructions methods that are tailored to the specific needs. For multi-view transcranial imaging we chose the maximum-mean component approach, which seemed attractive for the envisaged diagnostic purpose of early detection of Parkinson's disease. Given that the target structure e.g. substantia nigra, becomes more echogenic with an increased risk of Parkinson's disease, high intensities should be preserved, while not degrading the overall image contrast. In this respect, it has to be investigated in a clinical study how different reconstructions method affect sensitivity and specificity, in particular for the case of early detection of Parkinson's disease.
Based on the mixture model assumption we proposed the Pseudo-J-Divergence (PJD) similarity measure making use of the distributional J-Divergence measure and the mixture weights. However, PJD does not fulfill the triangle inequality and therefore is only a pseudo metric. Future work could consider the derivation of a real metric, which has the potential to further improve the registration accuracy and thus even more exploit the robustness of the statistical data representation.
The mixture distributions in this thesis have all been estimated using the Expectation-Maximization (EM) algorithm. However, the computational complexity is not negligible in particular when the reconstruction is performed on large scale 3D volumes. In this respect, it might be interesting to develop strategies to port estimation to GPU.
In near future, 2D arrays will probably, at least partially, replace conventional

transducers and make the acquisition of 3D data less cumbersome. The issue of multi-view data representation will prevail, however, under slightly modified conditions. In this respect, mixture modeling of 3D RF data acquired from 2D arrays will have to be investigated.
Moreover, the mixture representation of 3D RF data might be relevant for multi-modal image registration. Instead of using purely intensity-based approaches, employing mixture representation as 3D feature descriptor might prove beneficial.

Ultrasound Data and Texture Modeling

Beside being an application for the modeling of multi-view data, finite mixture representations seem also very attractive in the domain of texture modeling. That is the case for e.g. Markov Random Field auto-models, as stationarity is assumed in the texture. However, unless in very homogeneous regions this assumption is usually violated. Unfortunately, Markov Random Field texture auto-models defy a simple integration of the finite mixture modeling, as this type of distributional representation is not part of the exponential family, which is a necessary requirement. Therefore alternative means and ways have to be found to employ a mixture model as part of texture representation. Continuing this line of thought, mixture models could be the basis for superior similarity measures, similarly as was proposed for the 3D multi-view data.

Ultrasound Similarity Metric

Apart from the already mentioned possibility of integrating mixture model assumptions in ultrasound similarity metrics, there exist several possibilities of extension. On the one hand, it would be reasonable to incorporate the ultrasound machine acquisition parameters into a similarity measure, as those directly affect the image generation. Besides from very obvious influences, such as amplification, speckle and other phenomenons are directly affected by parameters such as frequency. On the other hand, it is obvious that the measures presented in the thesis could be extended from the 2D to 3D domain. This again might require the incorporation of view-information.

'Properly speaking, such work is never finished; one must declare it so when, according to time and circumstances, one has done one's best.'

[Johann Wolfgang von Goethe, 1749–1832]

9 List of Figures

List of Publications

Journal Publications

1. Christian Wachinger, Tassilo Klein, Nassir Navab. Locally Adaptive Nakagami-based Ultrasound Similarity Measures. *Ultrasonics*, 52(4): pages 547 – 554, 2012.
2. Christian Wachinger, Tassilo Klein, Nassir Navab. The 2D Analytic Signal for Envelope Detection and Feature Extraction on ultrasound images. *Medical Image Analysis*, 16(6), pages 1073–1084, 2012.
3. Athanasios Karamalis, Wolfgang Wein, Tassilo Klein, Nassir Navab. Ultrasound Confidence Maps using Random Walks. *Medical Image Analysis*, 16(6), pages 1101–1112, 2012.
4. Annika Plate, Seyed-Ahmad Ahmadi, Olivier Pauly, Tassilo Klein, Nassir Navab, Kai Boetzel. 3D Sonographic Examination of the Midbrain for Computer-Aided Diagnosis of Movement Disorders. *Ultrasound in Medicine & Biology*, 38(12), pages 2041–2050, 2012.

Conference Publications

1. Seyed-Ahmad Ahmadi, Maximilian Baust, Athanasios Karamalis, Annika Plate, Kai Boetzel, Tassilo Klein, Nassir Navab. Midbrain Segmentation in Transcranial 3D Ultrasound for Parkinson Diagnosis. *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 362–369, 2011.
2. Christian Wachinger, Tassilo Klein, Nassir Navab. The 2D Analytic Signal on RF and B-mode Ultrasound Images. In *International Conference on Information Processing in Medical Imaging (IPMI)*, volume 6801 of *Lecture Notes in Computer Science*, pages 359–370. Springer, 2011.
3. Tassilo Klein, Mattias Hansson, Nassir Navab. Spatial Statistics Based Feature Descriptor For RF Ultrasound Data. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pages 33–36, 2011.

4. Tassilo Klein, Mattias Hansson, Athanasios Karamalis, Nassir Navab. Registration of RF Ultrasound Data Using Hybrid Local Binary Patterns. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pages 1072–1075, 2012.

5. Tassilo Klein, Mattias Hansson, Nassir Navab. Modeling of Multi-View 3D Freehand Radio Frequency Ultrasound. *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 422–429, 2012.

6. Annika Plate, Seyed-Ahmad Ahmadi, Tassilo Klein, Nassir Navab, Jeanette Weisse, Jan Mehrkens, Kai Boetzel. Towards a More Objective Visualization of the Midbrain and its Surroundings Using 3D Transcranial Ultrasound. In *54. Jahrestagung der Deutschen Gesellschaft für Klinische Neurophysiologie und Funktionelle Bildgebung (DGKN)*, 2010.

Under Review

1. Tassilo Klein, Athanasios Karamalis, Nassir Navab. Shadow Detection in Ultrasound RF Data. *Ultrasound in Medicine & Biology*.

Bibliography

[Abdi and Kaveh, 2000] Abdi, A. and Kaveh, M. (2000). Performance comparison of three different estimators for the nakagami m parameter using monte carlo simulation. *IEEE Communications Letters*, 4(4):119 –121.

[Abramowitz and Stegun, 1964] Abramowitz, M. and Stegun, I. (1964). *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Applied mathematics series. Dover Publications.

[Ahmadi et al., 2011] Ahmadi, S.-A., Baust, M., Karamalis, A., Plate, A., Boetzel, K., Klein, T., and Navab, N. (2011). Midbrain segmentation in transcranial 3d ultrasound for parkinson diagnosis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 362–369.

[Ahmadian et al., 2004] Ahmadian, A., Mostafa, A., Abolhassani, M., and Alam, N. (2004). A method for texture classification of ultrasonic liver images based on gabor wavelet. In *International Conference on Signal Processing (ICSP)*, volume 2, pages 971 – 974.

[Ahmadian et al., 2005] Ahmadian, A., Mostafa, A., Abolhassani, M., and Salimpour, Y. (2005). A texture classification method for diffused liver diseases using gabor wavelets. In *IEEE Annual International Conference of the Engineering in Medicine and Biology Society (EMBS)*, pages 1567 –1570.

[Ahonen et al., 2004] Ahonen, T., Hadid, A., and Pietikainen, M. (2004). Face Recognition with Local Binary Patterns. In *European Conference on Computer Vision (ECCV)*, pages 469–481.

[Ali et al., 2008] Ali, R., Gooding, M., Christlieb, M., and Brady, M. (2008). Advanced phase-based segmentation of multiple cells from brightfield microscopy images. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pages 181—184.

[Ancona et al., 1990] Ancona, A., Geman, D., Ikeda, N., and Geman, D. (1990). Random fields and inverse problems in imaging. In *École d'Été de Probabilités de Saint-Flour XVIII - 1988*, volume 1427 of *Lecture Notes in Mathematics*, pages 113–193. Springer Berlin / Heidelberg.

[Arwini et al., 2008] Arwini, K., Dodson, C., Doig, A., Sampson, W., Scharcanski, J., and Felipussi, S. (2008). *Information geometry: near randomness and near independence*, volume 1953 of *Lecture Notes in Mathematics*. Springer Verlag.

[Averbuch et al., 2006] Averbuch, A., Coifman, R., Donoho, D., Elad, M., and Israeli, M. (2006). Fast and accurate polar Fourier transform. *Applied and Computational Harmonic Analysis*, 21(2):145–167.

[Babenko et al., 2009] Babenko, B., Branson, S., and Belongie, S. (2009). Similarity metrics for categorization: from monolithic to category specific. *IEEE International Conference on Computer Vision*, pages 293–300.

[Baehr and Frotscher, 2012] Baehr, M. and Frotscher, M. (2012). *Duus' Topical Diagnosis in Neurology: Anatomy, Physiology, Signs, Symptoms*. Georg Thieme Verlag.

[Bakhvalov et al., 1987] Bakhvalov, N., Zhileĭkin, I., and Zabolotskaya, E. (1987). *Nonlinear theory of sound beams*. American Institute of Physics translation series. American Institute of Physics.

[Barakat, 1986] Barakat, R. (1986). Weak-scatterer generalization of the k-density function with application to laser scattering in atmospheric turbulence. *Journal of the Optical Society of America A*, 3(4):401–409.

[Barndorff-Nielsen, 1997] Barndorff-Nielsen, O. E. (1997). Normal inverse gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, 24(1):1–13.

[Barratt et al., 2008] Barratt, D., Chan, C., Edwards, P., Penney, G., Slomczykowski, M., Carter, T., and Hawkes, D. (2008). Instantiation and registration of statistical shape models of the femur and pelvis using 3D ultrasound imaging. *Medical Image Analysis*, 12(3):358–374.

[Barratt et al., 2006] Barratt, D., Penney, G., Chan, C., Slomczykowski, M., Carter, T., Edwards, P., and Hawkes, D. (2006). Self-calibrating 3D-ultrasound-based bone registration for minimally invasive orthopedic surgery. *IEEE Transactions on Medical Imaging*, 25(3):312–323.

[Barry E. Stein and Stanford, 2000] Barry E. Stein, M. T. W. and Stanford, T. R. (2000). *Merging sensory signals in the brain: the development of multisensory integration in the superior colliculus*, chapter 5, pages 55–71. The New Cognitive Neurosciences.

[Basarab et al., 2007] Basarab, A., Aoudi, W., Liebgott, H., Vray, D., and Delachartre, P. (2007). Parametric deformable block matching for ultrasound imaging. In *IEEE International Conference on Image Processing (ICIP)*, volume 2, pages 429 –432.

[Becker et al., 1995a] Becker, G., Becker, T., Struck, M., Lindner, A., Burzer, K., Retz, W., Bogdahn, U., and Beckmann, H. (1995a). Reduced echogenicity of brainstem raphe specific to unipolar depression: a transcranial color-coded real-time sonography study. *Biol Psychiatry*, 38(3):180–184.

[Becker et al., 1995b] Becker, G., Seufert, J., Bogdahn, U., Reichmann, H., and Reiners, K. (1995b). Degeneration of substantia nigra in chronic parkinson's disease visualized by transcranial color-coded real-time sonography. *Neurology*, 45(1):182–184.

[Becker et al., 1994a] Becker, G., Struck, M., Bogdahn, U., and Becker, T. (1994a). Echogenicity of the brainstem raphe in patients with major depression. *Psychiatry Research: Neuroimaging*, 55(2):75 – 84.

[Becker et al., 1994b] Becker, G., Struck, M., Bogdahn, U., and Becker, T. (1994b). Echogenicity of the brainstem raphe in patients with major depression. *Psychiatry Res*, 55(2):75–84.

[Behar and Nikolov, 2006] Behar, V. and Nikolov, M. (2006). Statistical analysis of image quality in multi-angle compound imaging. In *IEEE International Symposium on Modern Computing*, pages 197–201.

[Behnke et al., 2005] Behnke, S., Berg, D., Naumann, M., and Becker, G. (2005). Differentiation of parkinson's disease and atypical parkinsonian syndromes by transcranial ultrasound. *Journal of Neurology, Neurosurgery and Psychiatry*, 76(3):423–425.

[Benedict and Soong, 1967] Benedict, T. and Soong, T. (1967). The joint estimation of signal and noise from the sum envelope. *IEEE Transactions on Information Theory*, 13(3):447 – 454.

[Berg, 2006] Berg, D. (2006). Marker for a preclinical diagnosis of parkinson's disease as a basis for neuroprotection. In Parvez, H. and Riederer, P., editors, *Oxidative Stress and Neuroprotection*, volume 71 of *Journal of Neural Transmission. Supplementa*, pages 123–132.

[Berg et al., 2008] Berg, D., Godau, J., and Walter, U. (2008). Transcranial sonography in movement disorders. *The Lancet Neurology*, 7(11):1044 – 1055.

[Berg et al., 2002] Berg, D., Roggendorf, W., Schroder, U., Klein, R., Tatschner, T., Benz, P., Tucha, O., Preier, M., Lange, K. W., Reiners, K., Gerlach, M., and Becker, G. (2002). Echogenicity of the substantia nigra: Association with increased iron content and marker for susceptibility to nigrostriatal injury. *Archives of Neurology*, 59(6):999–1005.

[Berg et al., 1999] Berg, S., Torp, H., Martens, D., Steen, E., Samstad, S., Høivik, I., and Olstad, B. (1999). Dynamic three-dimensional freehand echocardiography using raw digital ultrasound data. *Ultrasound in Medicine & Biology*, 25(5):745–753.

[Besag, 1974] Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 36(2):192–236.

[Besag, 1975] Besag, J. (1975). Statistical analysis of non-lattice data. *The statistician*, 24(3):179–195.

[Besag, 1986] Besag, J. (1986). On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 259–302.

[Bharucha-Reid, 2010] Bharucha-Reid, A. (2010). *Elements of the Theory of Markov Processes and Their Applications*. Dover books on mathematics. Dover Publications.

[Bhattacharyya, 1943] Bhattacharyya, A. (1943). On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of the Calcutta Mathematical Society*, 35:99–109.

[Bishop, 2009] Bishop, C. (2009). *Pattern recognition and machine learning*. Springer.

[Björck, 1996] Björck, Å. (1996). *Numerical methods for least squares problems*. Miscellaneous Bks. SIAM.

[Bleck et al., 1996] Bleck, J., Ranft, U., Gebel, M., Hecker, H., Westhoff-Bleck, M., Thiesemann, C., Wagner, S., and Manns, M. (1996). Random field models in the textural analysis of ultrasonic images of the liver. *IEEE Transactions on Medical Imaging*, 15(6):796 –801.

[Bock and Krischer, 1998] Bock, R. K. and Krischer, W. (1998). *The Data Analysis Briefbook*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1st edition.

[Boon, 2011] Boon, J. (2011). *Veterinary Echocardiography*. John Wiley & Sons.

[Bouhlel and Sevestre-Ghalila, 2009] Bouhlel, N. and Sevestre-Ghalila, S. (2009). Nakagami markov random field as texture model for ultrasound rf envelope image. *Ultrasound in Medicine & Biology*, 39(6):535–544.

[Bouhlel et al., 2004] Bouhlel, N., Sevestre-Ghalila, S., Rajhi, H., and Hamza, R. (2004). New markov random field model based on k-distribution for textured ultrasound image. *Ultrasonic Imaging and Signal Processing. SPIE International Symposium*, 5373:363–372.

[Boukerroui et al., 2003] Boukerroui, D., Noble, J., and Brady, M. (2003). Velocity estimation in ultrasound images: a block matching approach. *Inf Process Med Imaging*, 18:586–598.

[Boukerroui et al., 2004] Boukerroui, D., Noble, J., and Brady, M. (2004). On the Choice of Band-Pass Quadrature Filters. *Journal of Mathematical Imaging and Vision*, 21(1):53–80.

[Boykov and Kolmogorov, 2004] Boykov, Y. and Kolmogorov, V. (2004). An experimental comparison of min-cut/max- flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9):1124–1137.

[Boykov et al., 2001] Boykov, Y., Veksler, O., and Zabih, R. (2001). Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1222–1239.

[Breiman, 2001] Breiman, L. (2001). Random forests. *Machine Learning*, 45:5–32.

[Bronstein et al., 2010] Bronstein, M., Bronstein, A., Michel, F., and Paragios, N. (2010). Data fusion through cross-modality metric learning using similarity-sensitive hashing. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*, pages 3594–3601.

[Brooks and Chan, 2002] Brooks, A. and Chan, J. C. O. (2002). *Ultrasound in Emergency Care*. BMJ Books.

[Burckhardt, 1978] Burckhardt, C. B. (1978). Speckle in Ultrasound B-Mode Scans. *IEEE Transactions on Sonics and Ultrasonics*, 25:1–6.

[Burgers, 1948] Burgers, J. (1948). A mathematical model illustrating the theory of turbulence. volume 1 of *Advances in Applied Mechanics*, pages 171 – 199. Elsevier.

[Carneiro and Jepson, 2002] Carneiro, G. and Jepson, A. D. (2002). Phase-based local features. In *European Conference on Computer Vision (ECCV)*, pages 282–296.

[Chellappa and Chatterjee, 1985] Chellappa, R. and Chatterjee, S. (1985). Classification of textures using gaussian markov random fields. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 33(4):959– 963.

[Chen, 2000] Chen, S. (2000). Probability density function estimation using gamma kernels. *Annals of the Institute of Statistical Mathematics*, 52(3):471–480.

[Cheng and Beaulieu, 2001] Cheng, J. and Beaulieu, N. (2001). Maximum-likelihood based estimation of the nakagami m parameter. *IEEE Communications Letters*, 5(3):101 –103.

[Cheng and Beaulieu, 2002] Cheng, J. and Beaulieu, N. (2002). Generalized moment estimators for the nakagami fading parameter. *IEEE Communications Letters*, 6(4):144 –146.

[Cincotti et al., 2001] Cincotti, G., Loi, G., and Pappalardo, M. (2001). Frequency decomposition and compounding of ultrasound medical images with wavelet packets. *IEEE Transactions on Medical Imaging*, 20(8):764–771.

[Clark et al., 2010] Clark, D., Boutros, N., and Mendez, M. (2010). *The Brain and Behavior: An Introduction to Behavioral Neuroanatomy*. Cambridge Medicine. Cambridge University Press.

[Cobbold, 2007] Cobbold, R. (2007). *Foundations of biomedical ultrasound*. Biomedical engineering series. Oxford University Press.

[Cohen and Dinstein, 2002a] Cohen, B. and Dinstein, I. (2002a). New maximum likelihood motion estimation schemes for noisy ultrasound images. *Pattern Recognition*, 35(2):455 – 463.

[Cohen and Dinstein, 2002b] Cohen, B. and Dinstein, I. (2002b). New maximum likelihood motion estimation schemes for noisy ultrasound images. *Pattern Recognition*, 35(2):455–463.

[Cohen and Cooper, 1987] Cohen, F. S. and Cooper, D. B. (1987). Simple parallel hierarchical and relaxation algorithms for segmenting noncausal markovian random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(2):195 –219.

[Comaniciu et al., 2003] Comaniciu, D., Ramesh, V., and Meer, P. (2003). Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(5):564–577.

[Copsey and Webb, 2003] Copsey, K. and Webb, A. (2003). Bayesian gamma mixture model approach to radar target recognition. *IEEE Transactions on Aerospace Electronic Systems*, 39:1201–1217.

[Coupé et al., 2007] Coupé, P., Hellier, P., Morandi, X., and Barillot, C. (2007). Probe trajectory interpolation for 3d reconstruction of freehand ultrasound. *Medical Image Analysis*, 11(6):604 – 615.

[Crawford et al., 1993] Crawford, D. C., Bell, D. S., and Bamber, J. C. (1993). Compensation for the signal processing characteristics of ultrasound b-mode scanners in adaptive speckle reduction. *Ultrasound in Medicine & Biology*, 19(6):469–485.

[Cross and Jain, 1983] Cross, G. and Jain, A. (1983). Markov random field texture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5(1):25–39.

[D'Agostino and Stephens, 1986] D'Agostino, R. and Stephens, M. (1986). Goodness-of-fit techniques. *Marcel Dekker, Inc. New York, NY, USA*, page 560.

[Dahl et al., 2005] Dahl, J., Ivancevich, N., Keen, C., Trahey, G., and Smith, S. (2005). Phase correction of skull aberration with 1.75-d and 2-d arrays using speckle targets. *IEEE Ultrasonics Symposium*, 2:1323–1326.

[Dellepiane,] Dellepiane, S. Image segmentation: Errors, sensitivity, and uncertainty. *Engineering in Medicine and Biology Society*.

[Dempster et al., 1977] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood for incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B*, 39:1–38.

[Derin and Elliott, 1987] Derin, H. and Elliott, H. (1987). Modeling and segmentation of noisy and textured images using gibbs random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 9(1):39 – 55.

[Destrempes and Cloutier, 2010] Destrempes, F. and Cloutier, G. (2010). A critical review and uniformized representation of statistical distributions modeling the ultrasound echo envelope. *Ultrasound in Medicine & Biology*, 36(7):1037–1051.

[Destrempes et al., 2009] Destrempes, F., Meunier, J., Giroux, M.-F., Soulez, G., and Cloutier, G. (2009). Segmentation in ultrasonic b-mode images of healthy carotid arteries using mixtures of nakagami distributions and stochastic optimization. *IEEE Transactions on Medical Imaging*, 28(2):215–229.

[Dodson and Matsuzoe., 2002] Dodson, C. and Matsuzoe., H. (2002). An affine embedding of the gamma manifold. *InterStat*.

[Dodziuk, 1984] Dodziuk, J. (1984). Difference equations, isoperimetric inequality and transience of certain random walks. *Transactions of the American Mathematical Society*, 284(2):787–794.

[Dubes and Jain, 1989] Dubes, R. C. and Jain, A. K. (1989). Random field models in image analysis. *Journal of Applied Statistics*, 16(2):131–164.

[Dussik, 1952] Dussik, K. T. (1952). Weitere ergebnisse der ultraschalluntersuchung bei gehirnerkrankungen. *Acta Neurochirurgica*, 2:379–401.

[Dutt, 1995] Dutt, V. (1995). *Statistical analysis of ultrasound echo envelope*. PhD thesis, Mayo Graduate School, Rochester, MN, USA.

[Dutt and Greenleaf, 1994] Dutt, V. and Greenleaf, J. (1994). Ultrasound echo envelope analysis using a homodyned k distribution signal model. *Ultrasonic Imaging*, 16(4):265 – 287.

[Dutt and Greenleaf, 1996] Dutt, V. and Greenleaf, J. F. (1996). Statistics of the log-compressed echo envelope. *The Journal of the Acoustical Society of America*, 99(6):3817–3825.

[Elen et al., 2008] Elen, A., Choi, H. F., Loeckx, D., Gao, H., Claus, P., Suetens, P., Maes, F., and D'hooge, J. (2008). Three-dimensional cardiac strain estimation using spatio temporal elastic registration of ultrasound images: A feasibility study. *IEEE Transactions on Medical Imaging*, 27(11):1580 –1591.

[Eltoft, 2003] Eltoft, T. (2003). Speckle: Modeling and filtering. In *Norwegian Signal Processing Symposium*.

[Eltoft, 2005] Eltoft, T. (2005). The rician inverse gaussian distribution: a new model for non-rayleigh signal amplitude statistics. *IEEE Transactions on Image Processing*, 14(11):1722–1735.

[Escobar, 1994] Escobar, M. D. (1994). Estimating normal means with a dirichlet process prior. *Journal of the American Statistical Association*, 89(425):268–277.

[Estepar et al., 2006] Estepar, R., Washko, G., Silverman, E., Reilly, J., Kikinis, R., and Westin, C. (2006). Accurate airway wall estimation using phase congruency. *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 9:125–134.

[Esther Leung et al., 2008] Esther Leung, K., van Stralen, M., Nemes, A., Voormolen, M., van Burken, G., Geleijnse, M., ten Cate, F., Reiber, J., de Jong, N., van der Steen, A., and Bosch, J. (2008). Sparse registration for three-dimensional stress echocardiography. *IEEE Transactions on Medical Imaging*, 27(11):1568–1579.

[Farag et al., 2005] Farag, A., Mohamed, R., and El-Baz, A. (2005). A unified framework for map estimation in remote sensing image segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, 43(7):1617 – 1634.

[Felsberg and Sommer, 2001] Felsberg, M. and Sommer, G. (2001). The monogenic signal. *IEEE Transactions on Signal Processing*, 49(12):3136–3144.

[Felsberg and Sommer, 2004] Felsberg, M. and Sommer, G. (2004). The monogenic scale-space: A unifying approach to phase-based image processing in scale-space. *Journal of Mathematical Imaging and vision*, 21(1):5–26.

[Fleet et al., 1991] Fleet, D., Jepson, A., and Jenkin, M. (1991). Phase-based disparity measurement. *CVGIP: Image Understanding*, 53(2):198–210.

[Flury, 1997] Flury, B. (1997). *A first course in multivariate statistics*. Springer texts in statistics. Springer.

[Fraley and Raftery, 2002] Fraley, C. and Raftery, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, 97(458):611–631.

[Freund and Schapire, 1995] Freund, Y. and Schapire, R. E. (1995). A decision-theoretic generalization of on-line learning and an application to boosting. In *European Conference on Computational Learning Theory (EuroCOLT)*, pages 23–37.

[Gao, 2000] Gao, D. (2000). *Duality principles in nonconvex systems: theory, methods, and applications*. Nonconvex optimization and its applications. Kluwer Academic Publishers.

[Gee et al., 2006] Gee, A. H., Housden, R. J., Hassenpflug, P., Treece, G. M., and Prager, R. W. (2006). Sensorless freehand 3d ultrasound in real tissue: Speckle decorrelation without fully developed speckle. *Medical Image Analysis*, 10(2):137 – 149.

[Geman et al., 1990] Geman, D., Geman, S., Graffigne, C., and Dong, P. (1990). Boundary detection by constrained optimization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):609 – 628.

[Geman and Geman, 1984] Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6):721 – 741.

[Geman and Graffigne, 1986] Geman, S. and Graffigne, C. (1986). Markov random field image models and their applications to computer vision. In *International Congress of Mathematicians*, volume 1, pages 1496–1517.

[Gobbi and Peters, 2002] Gobbi, D. G. and Peters, T. M. (2002). Interactive intra-operative 3d ultrasound reconstruction and visualization. *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 156–163.

[Goodman, 2007] Goodman, J. (2007). *Speckle phenomena in optics: theory and applications*. Roberts & Co.

[Grady, 2006] Grady, L. (2006). Random walks for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(11):1768–1783.

[Granlund and Knutsson, 1995] Granlund, G. H. and Knutsson, H. (1995). *Signal Processing for Computer Vision*. Kluwer Academic Publishers.

[Grau et al., 2006] Grau, V., Becher, H., and Noble, J. (2006). Phase-based registration of multi-view real-time three-dimensional echocardiographic sequences. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 612–619.

[Grau et al., 2007] Grau, V., Becher, H., and Noble, J. (Sept. 2007). Registration of multiview real-time 3-d echocardiographic sequences. *IEEE Transactions on Medical Imaging*, 26(9):1154–1165.

[Green, 1990] Green, P. J. (1990). Bayesian reconstructions from emission tomography data using a modified em algorithm. *IEEE Transactions on Medical Imaging*, pages 84–93.

[Green, 1995] Green, P. J. (1995). Reversible jump markov chain monte carlo computation and bayesian model determination. *Biometrika*, 82(4):711–732.

[Grimmett, 1973] Grimmett, G. R. (1973). A theorem about random fields. *Bulletin of the London Mathematical Society*, 5(1):81–84.

[Hacihaliloglu et al., 2008] Hacihaliloglu, I., Abugharbieh, R., Hodgson, A., and Rohling, R. (2008). Bone segmentation and fracture detection in ultrasound using 3d local phase features. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 287–295.

[Hamilton, 1998] Hamilton, M. F. (1998). *Nonlinear Acoustics*. Academic Press.

[Hanson, 2006] Hanson, T. (2006). Modeling censored lifetime data using a mixture of gammas baseline. *Bayesian Analysis*, 1:575–594.

[Hassner and Sklansky, 1980] Hassner, M. and Sklansky, J. (1980). The use of markov random fields as models of texture. *Computer Graphics and Image Processing*, 12(4):357 – 370.

[Hedrick et al., 2004] Hedrick, W. R., Hykes, D. L., and Starchman, D. E. (2004). *Ultrasound Physics and Instrumentation*. Mosby.

[Hellier et al., 2010] Hellier, P., Coupé, P., Morandi, X., and Collins, D. (2010). An automatic geometrical and statistical method to detect acoustic shadows in intraoperative ultrasound brain images. *Medical Image Analysis*, 14(2):195–204.

[Henrich et al., 2003] Henrich, W., Schmider, A., Kjos, S., Tutschek, B., and Dudenhausen, J. W. (2003). Advantages of and applications for extended field-of-view ultrasound in obstetrics. *Archives of Gynecology and Obstetrics*, 268(2):121–127.

[Hertz, 2006] Hertz, T. (2006). *Learning Distance Functions: Algorithms and Applications*. PhD thesis, The Hebrew University of Jerusalem.

[Hertz et al., 2004] Hertz, T., Bar-Hillel, A., and Weinshall, D. (2004). Boosting margin based distance functions for clustering. In *International Conference on Machine learning*, pages 393–400.

[Hertz et al., 2006] Hertz, T., Hillel, A., and Weinshall, D. (2006). Learning a kernel function for classification with small training samples. *International Conference on Machine learning*, pages 401–408.

[Hill et al., 2004] Hill, C., Bamber, J., and Haar, G. (2004). *Physical principles of medical ultrasonics*. John Wiley & Sons.

[Hof and Mobbs, 2001] Hof, P. and Mobbs, C. (2001). *Functional Neurobiology of Aging*. Academic Press.

[Horowski et al., 2011] Horowski, S., Zettl, U. K., Benecke, R., and Walter, U. (2011). Sonographic basal ganglia alterations are related to non-motor symptoms in multiple sclerosis. *Journal of Neurology*, 258(2):195–202.

[Hottier and Billon, 1990] Hottier, F. and Billon, A. C. (1990). *3D imaging in medicine*, chapter 3D echography: Status and perspective. Springer Verlag.

[Hruska, 2009] Hruska, D. (2009). Improved parameter estimates based on the homodyned k distribution. In *IEEE Transactions on Ultrasonics, Ferroelectris and Frequency Control*, volume 56, pages 2471–2481.

[Hsu et al., 2007] Hsu, P.-W., Prager, R. W., Gee, A. H., Treece, G. M., wei Hsu, P., Prager, R. W., Gee, A. H., and Treece, G. M. (2007). Freehand 3d ultrasound calibration: A review. Cued/f-infend/tr 579, Cambridge University.

[Hu and Fahmy, 1992] Hu, R. and Fahmy, M. M. (1992). Texture segmentation based on a hierarchical markov random field model. *Signal Processing*, 26(3):285 – 305.

[Huang et al., 2008] Huang, Y., Wang, L., and Li, C. (2008). Texture analysis of ultrasonic liver image based on wavelet transform and probabilistic neural network. In *International Conference on BioMedical Engineering and Informatics (BMEI)*, volume 2, pages 248 –252.

[Iakovidis et al., 2010] Iakovidis, D. K., Keramidas, E. G., and Maroulis, D. (2010). Fusion of fuzzy statistical distributions for classification of thyroid ultrasound patterns. *Artificial intelligence in medicine*, 50:33–41.

[Ijaz et al., 2011] Ijaz, U. Z., Prager, R. W., Gee, A. H., and Treece, G. M. (2011). A Study of Similarity Measures for In Vivo 3D Ultrasound Volume Registration Acoustical Imaging. In André, M. P., Jones, J. P., and Lee, H., editors, *Acoustical Imaging*, volume 30 of *Acoustical Imaging*, chapter 36, pages 315–323. Springer Netherlands, Dordrecht.

[Insana and et al., 1986] Insana, M. and et al. (1986). Analysis of ultrasound image texture via generalized rician statistics. *Optical Engineering*, 6(4-5):743–748.

[Ivancevich et al., 2004] Ivancevich, N., Chu, K., Dahl, J., and Light, E. (2004). Real time 3d ultrasound imaging of the brain. *IEEE Ultrasonics Symposium*, 1:110–113.

[Ivancevich et al., 2006] Ivancevich, N., Dahl, J., Trahey, G., and Smith, S. (2006). Phase-aberration correction with a 3-d ultrasound scanner: Feasibility study. *IEEE Transactions on Ultrasonics Ferroelectrics and Frequency Control*, 53(8):1432–1439.

[Jakeman, 1980] Jakeman, E. (1980). On the statistics of k-distributed noise. *Journal of Physics A: Mathematical and General*, 13(1):31–48.

[Jakeman, 1999] Jakeman, E. (1999). K-distributed noise. *Journal of Optics A: Pure and Applied Optics*, 1(4-5).

[Jakeman and Pusey, 1976] Jakeman, E. and Pusey, P. (1976). A model for non-rayleigh sea echo. *IEEE Transactions on Antennas and Propagation*, 24(6):806 – 814.

[Jakeman and Tough, 1987a] Jakeman, E. and Tough, R. (1987a). Generalized k distribution: a statistical model for weak scattering. *Journal of the Optical Society of America A*, 4(9):1764–1772.

[Jakeman and Tough, 1987b] Jakeman, E. and Tough, R. J. A. (1987b). Generalized k distribution: a statistical model for weak scattering. *Journal of the Optical Society of America A*, 4(9):1764–1772.

[Jeffreys, 1946] Jeffreys, H. (1946). An Invariant Form for the Prior Probability in Estimation Problems. *Proc. R. Soc. A. Series A, Mathematical and Physical Sciences*, 186:453–461.

[José-Estépar et al., 2003] José-Estépar, R. S., Martín-Fernández, M., Caballero-Martínez, P., Alberola-López, C., and Ruiz-Alzola, J. (2003). A theoretical framework to three-dimensional ultrasound reconstruction from irregularly sampled data. *Ultrasound in Medicine & Biology*, 29(2):255 – 269.

[Joynt, 1979] Joynt, L. (1979). *A stochastic approach to ultrasonic tissue characterization*. PhD thesis, Stanford University.

[Kandel et al., 2000] Kandel, E., Schwartz, J., and Jessell, T. (2000). *Principles of Neural Science*. McGraw-Hill, Health Professions Division.

[Kaplan and Ma;, 1993] Kaplan, D. and Ma;, Q. (1993). On the statistical characteristics of log-compressed rayleigh signals: theoretical formulation and experimental results. *IEEE Ultrasonics Symposium*, 2:961 – 964.

[Karamalis et al., 2012a] Karamalis, A., Carlier, S., Katouzian, A., and Navab, N. (2012a). Confidence estimation in IVUS radio-frequency data with random walks. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*.

[Karamalis et al., 2012b] Karamalis, A., Wein, W., Klein, T., and Navab, N. (2012b). Ultrasound confidence maps using random walks. *Medical Image Analysis*, 16(6):1101–1112.

[Karamalis et al., 2009] Karamalis, A., Wein, W., Kutter, O., and Navab, N. (2009). Fast hybrid freehand ultrasound volume reconstruction. In *SPIE Medical Imaging*, Orlando, Florida, USA.

[Kellokumpu et al., 2009] Kellokumpu, V., Zhao, G., Li, S. Z., and Pietikäinen, M. (2009). Dynamic texture based gait recognition. In *International Conference on Advances in Biometrics*, pages 1000–1009.

[Kellokumpu et al., 2008] Kellokumpu, V., Zhao, G., and Pietikäinen, M. (2008). Human activity recognition using a dynamic texture based method. In *British Machine Vision Conference (BMVC)*, pages 88.1–88.10.

[Kellokumpu et al., 2011] Kellokumpu, V., Zhao, G., and Pietikäinen, M. (2011). Recognition of human actions using texture descriptors. *Machine Vision and Applications*, 22(5):767–780.

[Kemeny and Snell, 1960] Kemeny, J. and Snell, J. (1960). *Finite Markov chains*. Undergraduate texts in mathematics. Springer-Verlag.

[Kindermann et al., 1980] Kindermann, R., Snell, J., and Society, A. M. (1980). *Markov random fields and their applications*. Contemporary mathematics. American Mathematical Society.

[Klein et al., 2012a] Klein, T., Hansson, M., Karamalis, A., and Navab, N. (2012a). Registration of rf ultrasound data using hybrid local binary patterns. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pages 1072–1075.

[Klein et al., 2011] Klein, T., Hansson, M., and Navab, N. (2011). Spatial statistics based feature descriptor for rf ultrasound data. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pages 33–36.

[Klein et al., 2012b] Klein, T., Hansson, M., and Navab, N. (2012b). Modeling of multi-view 3d freehand radio frequency ultrasound. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 422–429.

[Klein et al.,] Klein, T., Karamalis, A., and Navab, N. Shadow detection in ultrasound rf data.

[Kolmogorov and Zabih, 2004] Kolmogorov, V. and Zabih, R. (2004). What Energy Functions Can Be Minimizedvia Graph Cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2):147–159.

[Korostelev and Korosteleva, 2011] Korostelev, A. and Korosteleva, O. (2011). *Mathematical statistics: asymptotic minimax theory*. Graduate studies in mathematics. American Mathematical Society.

[Kovesi, 1999] Kovesi, P. (1999). Image features from phase congruency. *Journal of Computer Vision Research*, 1(3).

[Krucker et al., 2002] Krucker, J., LeCarpentier, G., Fowlkes, J., and Carson, P. (2002). Rapid elastic image registration for 3-d ultrasound. *IEEE Transactions on Medical Imaging*, 21(11):1384–1394.

[Larrue and Noble, 2011] Larrue, A. and Noble, J. A. (2011). Nakagami imaging with small windows. *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pages 887–890.

[Lee et al., 2009] Lee, D., Hofmann, M., Steinke, F., Altun, Y., Cahill, N., and Scholkopf, B. (2009). Learning similarity measure for multi-modal 3d image registration. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*, pages 186 –193.

[Leroy et al., 2004] Leroy, A., Mozer, P., Payan, Y., and Troccaz, J. (2004). Rigid registration of freehand 3D ultrasound and CT-scan kidney images. *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 837–844.

[Li, 2009] Li, S. (2009). *Markov random field modeling in image analysis*. Advances in pattern recognition. Springer.

[Lin et al., 2005] Lin, N., Yu, W., and Duncan, J. (2005). Left ventricular boundary segmentation from echocardiography. *Medical Imaging Systems Technology*, 3:89–108.

[Lindgren, 2002] Lindgren, F. (2002). Image modelling and estimation: A statistical approach.

[Loeckx et al., 2010] Loeckx, D., Slagmolen, P., Maes, F., Vandermeulen, D., and Suetens, P. (2010). Nonrigid image registration using conditional mutual information. *IEEE Transactions on Medical Imaging*, 29(1):19 –29.

[Long et al., 2005] Long, M., Levy, M., and Stern, R. (2005). *Architectural Acoustics (Applications of Modern Acoustics)*. Academic Press.

[Madabhushi et al., 2006] Madabhushi, A., Yang, P., Rosen, M., and Weinstein, S. (2006). Distinguishing lesions from posterior acoustic shadowing in breast ultrasound via non-linear dimensionality reduction. In *IEEE Annual International Conference of the Engineering in Medicine and Biology Society (EMBS)*, pages 3070–3073.

[Mandelbrot, 1983] Mandelbrot, B. (1983). *The Fractal Geometry of Nature*. W.H. Freeman.

[Martens and Gilja, 2005] Martens, D. and Gilja, O. H. (2005). *Basic and new aspects of gastrointestinal ultrasonography*, chapter The EchoPAC-3D software for image analysis. World Scientific Publishing.

[McCann et al., 1988] McCann, H., Sharp, J., Kinter, T., McEwan, C., Barillot, C., and Greenleaf, J. (1988). Multidimensional ultrasonic imaging for cardiology. *Proceedings of the IEEE*, 76(9):1063 –1073.

[McLachlan and Krishnan, 2007] McLachlan, G. and Krishnan, T. (2007). *The Em Algorithm and Extensions*. Wiley Series in Probability and Statistics. John Wiley & Sons.

[Mellor and Brady, 2005] Mellor, M. and Brady, M. (2005). Phase mutual information as a similarity measure for registration. *Medical Image Analysis*, 9(4):330–343.

[Mercier et al., 2005] Mercier, L., Lango, T., Lindseth, F., and Collins, D. (2005). A review of calibration techniques for freehand 3-d ultrasound systems. *Ultrasound Med. Biol.*, 31(4):449–471.

[Metropolis et al., 1953] Metropolis, N., Rosenbluth, A., Rosenbluth, M., and Teller, E. (1953). Equations of state calculations by fast computational machine. *Journal of Chemical Physics*, 21:1087–1091.

[Misaridis, 2001] Misaridis, T. (2001). *Ultrasound imaging using coded signals*. Ørsted DTU, Electronics & Signal Processing.

[Moradi et al., 2010] Moradi, M., Abolmaesumi, P., and Mousavi, P. (2010). Tissue typing using ultrasound rf time series: Experiments with animal tissue samples. *Medical Physics*, 37(8):4401–4413.

[Moussouris, 1974] Moussouris, J. (1974). Gibbs and markov random systems with constraints. *Journal of Statistical Physics*, 10(1):11–33.

[Mulet-Parada and Noble, 2000] Mulet-Parada, M. and Noble, J. (2000). 2D+ T acoustic boundary detection in echocardiography. *Medical Image Analysis*, 4(1):21–30.

[Myronenko, 2010] Myronenko, A. (2010). *Non-rigid Image Registration: Regularization, Algorithms and Applications*. PhD thesis, Oregon Health and Science University.

[Myronenko et al., 2009a] Myronenko, A., Song, X., and Sahn, D. (2009a). Maximum likelihood motion estimation in 3d echocardiography through non-rigid registration in spherical coordinates. In *Functional Imaging and Modeling of the Heart*, volume 5528 of *Lecture Notes in Computer Science*, pages 427–436.

[Myronenko et al., 2009b] Myronenko, A., Song, X., and Sahn, D. (2009b). Maximum Likelihood Motion Estimation in 3D Echocardiography through Non-rigid Registration in Spherical Coordinates. *Functional Imaging and Modeling of the Heart*, pages 427–436.

[Nakagami, 1960] Nakagami, N. (1960). The m-distribution, a general formula for intensity distribution of rapid fadings. In Hoffman, W. G., editor, *Statistical Methods in Radio Wave Propagation*. Oxford, England: Pergamon.

[Naumann et al., 1996] Naumann, M., Becker, G., Toyka, K., Supprian, T., and Reiners, K. (1996). Lenticular nucleus lesion in idiopathic dystonia detected by transcranial sonography. *Neurology*, 47(5):1284–90.

[Nelson and Pretorius, 1997] Nelson, T. R. and Pretorius, D. H. (1997). Interactive acquisition, analysis, and visualization of sonographic volume data. *International Journal of Imaging Systems and Technology*, 8(1):26–37.

[Nevid, 2012] Nevid, J. (2012). *Psychology: Concepts and Applications*. Cengage Learning.

[Noble, 2010] Noble, J. (2010). Ultrasound image segmentation and tissue characterization. *Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 224(2):307–316.

[Noble, 2009] Noble, J. A. (2009). Ultrasound image segmentation and tissue characterization. *Proc. IMechE*, 224(8):307–216.

[Ohbuchi et al., 1992] Ohbuchi, R., Chen, D., and Fuchs, H. (1992). Incremental volume reconstruction and rendering for 3d ultrasound imaging. In *The International Society for Optical Engineering*, pages 312–323.

[Ojala and Pietikäinen, 1999] Ojala, T. and Pietikäinen, M. (1999). Unsupervised texture segmentation using feature distributions. *Pattern Recognition*, 32(3):477 – 486.

[Ojala et al., 1994] Ojala, T., Pietikainen, M., and Harwood, D. (1994). Performance evaluation of texture measures with classification based on kullback discrimination of distributions. In *International Conference on Pattern Recognition*, volume 1, pages 582–585.

[Ojala et al., 1996] Ojala, T., Pietikäinen, M., and Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1):51 – 59.

[Ojala et al., 2002] Ojala, T., Pietikainen, M., and Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987.

[Oliver and Quegan, 2004] Oliver, C. and Quegan, S. (2004). *Understanding synthetic aperture radar images*. The SciTech radar and defense series. SciTech Publ.

[Oppenheim and Schafer, 2010] Oppenheim, A. and Schafer, R. (2010). *Discrete-time signal processing*. Prentice-Hall signal processing series. Prentice Hall.

[Paget, 2004] Paget, R. (2004). Strong markov random field model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(3):408–413.

[Paget, 2009] Paget, R. (2009). *Handbook of Texture Analysis*, chapter Texture Modelling and Synthesis, pages 33–60. World Scientific Pub Co.

[Papoulis and Pillai, 2002] Papoulis, A. and Pillai, S. (2002). *Probability, random variables, and stochastic processes*. McGraw-Hill electrical and electronic engineering series. McGraw-Hill.

[Penney et al., 2006] Penney, G., Barratt, D., Chan, C., Slomczykowski, M., Carter, T., Edwards, P., and Hawkes, D. (2006). Cadaver validation of intensity-based ultrasound to CT registration. *Medical Image Analysis*, 10(3):385–395.

[Penney et al., 2004] Penney, G., Blackall, J., Hamady, M., Sabharwal, T., Adam, A., and Hawkes, D. (2004). Registration of freehand 3D ultrasound and magnetic resonance liver images. *Medical Image Analysis*, 8(1):81–91.

[Petrou and Sevilla, 2006] Petrou, M. and Sevilla, P. (2006). *Image processing: dealing with texture*. John Wiley & Sons Inc.

[Pietikäinen et al., 2000] Pietikäinen, M., Ojala, T., and Xu, Z. (2000). Rotation-invariant texture classification using feature distributions. *Pattern Recognition*, 33:43–52.

[Plate et al., 2010] Plate, A., Ahmadi, S.-A., Klein, T., Navab, N., Weisse, J., Mehrkens, J., and Boetzel, K. (2010). Towards a more objective visualization of the midbrain and its surroundings using 3d transcranial ultrasound. In *54. Jahrestagung der Deutschen Gesellschaft für Klinische Neurophysiologie und Funktionelle Bildgebung (DGKN)*.

[Plate et al., 2012] Plate, A., Ahmadi, S.-A., Pauly, O., Klein, T., Navab, N., and Boetzel, K. (2012). Three-dimensional sonographic examination of the midbrain for computer-aided diagnosis of movement disorders. *Ultrasound in Medicine & Biology*, 38(12):2041–2050.

[Poon and Rohling, 2005] Poon, T. and Rohling, R. (2005). Three-dimensional extended field-of-view ultrasound. *Ultrasound in Medicine & Biology*, 32(3):357–369.

[Postema, 2011] Postema, M. (2011). *Fundamentals of Medical Ultrasonics*. Taylor & Francis Group.

[Prager et al., 2002] Prager, R., Gee, A., Treece, G., and Berman, L. (2002). Analysis of speckle in ultrasound images using fractional order statistics and the homodyned k-distribution. *Ultrasonics*, 40(1-8):133–137.

[Prager et al., 2003] Prager, R., Gee, A., Treece, G., and Berman, L. (2003). Decompression and speckle detection for ultrasound images using the homodyned k-distribution. *Pattern Recognition Letters*, 24(4-5):705–713.

[Prager et al., 1998] Prager, R. W., Rohling, R. N., Gee, A. H., and Berman, L. (1998). Rapid calibration for 3-d freehand ultrasound. *Ultrasound in Medicine & Biology*, 24(6):855–869.

[Rachev et al., 2010] Rachev, S., Hoechstoetter, M., Frank J. Fabozzi, C., and Focardi, S. (2010). *Probability and Statistics for Finance*. Frank J. Fabozzi Series. John Wiley & Sons.

[Raichel, 2006] Raichel, D. R. (2006). *The Science And Applications of Acoustics*. Springer US.

[Reinertsen et al., 2007] Reinertsen, I., Lindseth, F., Unsgaard, G., and Collins, D. L. (2007). Clinical validation of vessel-based registration for correction of brain-shift. *Medical Image Analysis*, 11(6):673–684.

[Revell et al., 2004] Revell, J., Mirmehdi, M., and McNally, D. (2004). Combined ultrasound speckle pattern similarity measures. In *Medical Image Understanding and Analysis*, pages 149–153. BMVA Press.

[Roche et al., 2000] Roche, A., Malandain, G., and Ayache, N. (2000). Unifying maximum likelihood approaches in medical image registration. *International Journal of Imaging Systems and Technology: Special Issue on 3D Imaging*, 11(1):71–80.

[Rohling et al., 1997] Rohling, R., Gee, A., and Berman, L. (1997). Three-dimensional spatial compounding of ultrasound images. *Medical Image Analysis*, 1(3):177–193.

[Rohling et al., 1999] Rohling, R., Gee, A., and Berman, L. (1999). A comparison of freehand three-dimensional ultrasound reconstruction techniques. *Medical Image Analysis*, 3(4):339 – 359.

[Saladin, 2010] Saladin, K. (2010). *Anatomy & physiology: the unity of form and function*. McGraw-Hill.

[Salcudean et al., 2006] Salcudean, S., French, D., Bachmann, S., Zahiri-Azar, R., Wen, X., and Morris, W. (2006). Viscoelasticity modeling of the prostate region using vibro-elastography. *Medical Image Computing and Computer-Assisted Intervention*, pages 389–396.

[San Jose Estepar et al., 2003] San Jose Estepar, R., Martin-Fernandez, M., Alberola-Lopez, C., Ellsmere, J., Kikinis, R., and Westin, C.-F. (2003). Freehand ultrasound reconstruction based on roi prior modeling and normalized convolution. In Ellis, R. E. and Peters, T. M., editors, *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 382–390.

[Sanches et al., 2011] Sanches, J., Sanches, J., Laine, A., and Suri, J. (2011). *Ultrasound Imaging: Advances and Applications*. Springer.

[Sanches and Marques, 2000] Sanches, J. M. and Marques, J. S. (2000). A rayleigh reconstruction/interpolation algorithm for 3d ultrasound. *Pattern Recognition Letters*, 21(10):917 – 926.

[Sanches and Marques, 2002] Sanches, J. M. and Marques, J. S. (2002). A multiscale algorithm for three-dimensional free-hand ultrasound. *Ultrasound in Medicine & Biology*, 28(8):1029–1040.

[Sanches and Marques, 2003] Sanches, J. M. and Marques, J. S. (2003). Compensation of log-compressed images for 3-d ultrasound. *Ultrasound in Medicine & Biology*, 29(2):239 – 253.

[Schapire and Singer, 1999] Schapire, R. E. and Singer, Y. (1999). Improved Boosting Using Confidence-rated Predictions. *Machine Learning*, 37(3):297–336.

[Schroder et al., 1998] Schroder, M., Rehrauer, H., Seidel, K., and Datcu, M. (1998). Spatial information retrieval from remote-sensing images. ii. gibbs-markov random fields. *IEEE Transactions on Geoscience and Remote Sensing*, 36(5):1446–1455.

[Serpico and Moser, 2006] Serpico, S. and Moser, G. (2006). Weight parameter optimization by the ho–kashyap algorithm in mrf models for supervised image classification. *IEEE Transactions on Geoscience and Remote Sensing*, 44(12):3695 – 3705.

[Shankar, 1986] Shankar, P. (1986). Speckle reduction in ultrasound B-scans using weighted averaging in spatial compounding. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 33(6):754–758.

[Shankar, 1993] Shankar, P. (1993). Use of non-rayleigh statistics for the identification of tumors in ultrasonic b-scans of the breast. *IEEE Transactions on Medical Imaging*, 12(4-5):687–692.

[Shankar, 1995] Shankar, P. (1995). A model for ultrasonic scattering from tissues based on the k distribution. *Physics in medicine and biology*, 40(10):1633–1649.

[Shankar et al., 2001] Shankar, P., Dumane, V., Reid, J., Genis, V., Forsberg, F., Piccoli, C., and Goldberg, B. (2001). Classification of ultrasonic B-mode images of breast masses using Nakagami distribution. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 48(2):569–580.

[Shankar et al., 2002] Shankar, P., Dumane, V., Reid, J., Genis, V., Forsberg, F., Piccoli, C., and Goldberg, B. (2002). Classification of ultrasonic b-mode images of breast masses using nakagami distribution. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 48(2):569–580.

[Shankar, 2000] Shankar, P. M. (2000). A general statistical model for ultrasonic scattering from tissues. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 47(3):339–343.

[Shankar, 2001] Shankar, P. M. (2001). Ultrasonic tissue characterization using a generalized nakagami model. *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, 48(6):1716–20.

[Shen et al., 2008] Shen, R., Cheng, I., Li, X., and Basu, A. (2008). Stereo matching using random walks. In *International Conference on Pattern Recognition (ICPR).*, pages 1–4.

[Shental et al., 2003] Shental, N., Bar-hillel, A., Hertz, T., and Weinshall, D. (2003). Computing gaussian mixture models with em using equivalence constraints. In *In Advances in Neural Information Processing Systems 16*. MIT Press.

[Shepard, 1968] Shepard, D. (1968). A two-dimensional interpolation function for irregularly-spaced data. In *ACM National Conference*, pages 517–524.

[Sherebrin et al., 1996] Sherebrin, S., Fenster, A., Rankin, R. N., and Spence, D. (1996). Freehand three-dimensional ultrasound: implementation and applications. In Metter, R. L. V. and Beutel, J., editors, *SPIE Medical Imaging*, volume 2708, pages 296–303.

[Sherwood, 2012] Sherwood, L. (2012). *Human Physiology: From Cells to Systems*. Cengage Learning.

[Shiavi, 2007] Shiavi, R. (2007). *Introduction to applied statistical signal analysis: guide to biomedical and electrical engineering applications*. Academic Press Series in Biomedical Engineering. Academic.

[Sijbers et al., 1998] Sijbers, J., Dekker, A. J. D., Scheunders, P., and Dyck, D. V. (1998). Maximum likelihood estimation of rician distribution parameters. *IEEE Transactions on Medical Imaging*, 17(3):357–361.

[Školoudík et al., 2007] Školoudík, D., Fadrná, T., Bártová, P., Langová, K., Ressner, P., Zapletalová, O., Hluštík, P., Herzig, R., and Kaňnovský, P. (2007). Reproducibility of sonographic measurement of the substantia nigra. *Ultrasound in Medicine & Biology*, 33(9):1347–1352.

[Smith et al., 2009] Smith, S. W., Ivancevich, N. M., Lindsey, B. D., Whitman, J., Light, E., Fronheiser, M., Nicoletto, H. A., and Laskowitz, D. T. (2009). The ultrasound brain helmet: Feasibility study of multiple simultaneous 3d scans of cerebral vasculature. *Ultrasound in Medicine & Biology*, 35(2):329–338.

[Soergel, 2010] Soergel, U. (2010). *Radar Remote Sensing of Urban Areas*. Remote Sensing and Digital Image Processing. Springer.

[Solberg et al., 2007] Solberg, O. V., Lindseth, F., Torp, H., Blake, R. E., and Hernes, T. A. N. (2007). Freehand 3d ultrasound reconstruction algorithms—a review. *Ultrasound in Medicine & Biology*, 33(7):991 – 1009.

[Strang, 2007] Strang, G. (2007). *Computational science and engineering*. Wellesley-Cambridge Press.

[Strintzis and Kokkinidis, 1997] Strintzis, M. and Kokkinidis, I. (1997). Maximum likelihood motion estimation in ultrasound image sequences. *IEEE Signal Processing Letters*, 4(6):156–157.

[Sun et al., 2008] Sun, X., Rosin, P., Martin, R., and Langbein, F. (2008). Random walks for feature-preserving mesh denoising. *Computer Aided Geometric Design*, 25(7):437–456.

[Szilágyi and Brady, 2009] Szilágyi, T. and Brady, S. M. (2009). Feature extraction from cancer images using local phase congruency: a reliable source of image descriptors. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pages 1219–1222.

[Talukdar and Lawing, 1991] Talukdar, K. K. and Lawing, W. D. (1991). Estimation of the parameters of the rice distribution. *The Journal of the Acoustical Society of America*, 89(3):1193–1197.

[Tao et al., 2006] Tao, Z., Tagare, H., and Beaty, J. (2006). Evaluation of four probability distribution models for speckle in clinical cardiac ultrasound images. *IEEE Transactions on Medical Imaging*, 25(11):1483–1491.

[Taraldsen, 2001] Taraldsen, G. (2001). A generalized westervelt equation for nonlinear medical ultrasound. *The Journal of the Acoustical Society of America*, 109(4):1329–1333.

[Thomas et al., 2005] Thomas, A., Banerjee, A., and Busch, U. (2005). *Classic Papers In Modern Diagnostic Radiology*. Springer.

[Thune et al., 1996] Thune, N., Gilja, O. H., Hausken, T., and Matre, K. (1996). A practical method for estimating enclosed volumes using 3d ultrasound. *European Journal of Ultrasound*, 3(1):83 – 92.

[Trobaugh et al., 1994] Trobaugh, J. W., Trobaugh, D. J., and Richard, W. D. (1994). Three-dimensional imaging with stereotactic ultrasonography. *Computerized Medical Imaging and Graphics*, 18(5):315 – 323.

[Unsgaard et al., 2006] Unsgaard, G., Rygh, O. M., Selbekk, T., Müller, T. B., Kolstad, F., Lindseth, F., and Hernes, T. A. N. (2006). Intra-operative 3d ultrasound in neurosurgery. *Acta Neurochirurgica*, 148:235–253.

[Viola, 1995] Viola, P. A. (1995). *Alignment by Maximization of Mutual Information*. Ph.d. thesis, Massachusetts Institute of Technology.

[Vlaar, 2011] Vlaar, A. (2011). The reliability of transcranial duplex scanning in parkinsonian patients: Comparison of different observers and ultrasound systems. reliabilität der transkraniellen duplex-sonografie bei parkinsonpatienten: Vergleich verschiedener untersucher und ultraschallsysteme. *Ultraschall in der Medizin 1431-4894*, 32(S1):83.

[Vlaar et al., 2009] Vlaar, A., Bouwmans, A., Mess, W., Tromp, S., and Weber, W. (2009). Transcranial duplex in the differential diagnosis of parkinsonian syndromes. *Journal of Neurology*, 256(4):530–538.

[Wachinger et al., 2011] Wachinger, C., Klein, T., and Navab, N. (2011). The 2d analytic signal on rf and b-mode ultrasound images. In Székely, G. and Hahn, H. K., editors, *IPMI*, volume 6801 of *Lecture Notes in Computer Science*, pages 359–370.

[Wachinger et al., 2012a] Wachinger, C., Klein, T., and Navab, N. (2012a). The 2d analytic signal for envelope detection and feature extraction on ultrasound images. *Medical Image Analysis*, 16(6):1073–1084.

[Wachinger et al., 2012b] Wachinger, C., Klein, T., and Navab, N. (2012b). Locally adaptive nakagami-based ultrasound similarity measures. *Ultrasonics*, 52(4):547 – 554.

[Wachinger et al., 2007] Wachinger, C., Wein, W., and Navab, N. (2007). Three-dimensional ultrasound mosaicing. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*.

[Wachinger et al., 2008] Wachinger, C., Wein, W., and Navab, N. (2008). Registration strategies and similarity measures for three-dimensional ultrasound mosaicing. *Academic Radiology*, 15:1404–1415.

[Wagner et al., 1983] Wagner, R. F., Smith, S. W., Sandrik, J. M., and Lopez, H. (1983). Statistics of Speckle in Ultrasound B-Scans. *IEEE Transactions on Sonics and Ultrasonics*, 30(3):156–163.

[Walter et al., 2007] Walter, U., Dressler, D., Probst, T., Wolters, A., Abu-Mugheisib, M., Wittstock, M., and Benecke, R. (2007). Transcranial brain sonography findings in discriminating between parkinsonism and idiopathic parkinson disease. *Archives of Neurology*, 64(11):1635–40.

[Walter et al., 2004] Walter, U., Dressler, D., Wolters, A., Probst, T., Grossmann, A., and Benecke, R. (2004). Sonographic discrimination of corticobasal degeneration vs progressive supranuclear palsy. *Neurology*, 63(3):504–9.

[Walter et al., 2003] Walter, U., Niehaus, L., Probst, T., Benecke, R., Meyer, B., and Dressler, D. (2003). Brain parenchyma sonography discriminates parkinson's disease and atypical parkinsonian syndromes. *Neurology*, 60(1):74–7.

[Walter et al., 2009] Walter, U., Wagner, S., Horowski, S., Benecke, R., and Zettl, U. (2009). Transcranial brain sonography findings predict disease progression in multiple sclerosis. *Neurology*, 73(13):1010–1017.

[Wang et al., 2009] Wang, P., Kelly, C., and Brady, M. (2009). Application of 3d local phase theory in vessel segmentation. In *IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI)*, pages 1174–1177.

[Webb, 2000] Webb, A. R. (2000). Gamma mixture models for target recognition. *Pattern Recognition*, 33(12):2045 – 2054.

[Wein et al., 2006] Wein, W., Pache, F., Roper, B., and Navab, N. (2006). Backward-warping ultrasound reconstruction for improving diagnostic value and registration. *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 4191:750.

[Wietzke et al., 2009] Wietzke, L., Sommer, G., and Fleischmann, O. (2009). The geometry of 2d image signals. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*, pages 1690–1697.

[Wilhjelm et al., 2000] Wilhjelm, J., Jensen, M., Brandt, T., Sahl, B., Martinsen, K., Jespersen, S., and Falk, E. (2000). Some imaging strategies in multi-angle spatial compounding. In *IEEE Ultrasonics Symposium*, volume 2, pages 1615–1618.

[Wiper et al., 2001] Wiper, M., Insua, R. D., and Ruggeri, F. (2001). Mixtures of Gamma Distributions With Applications. *Journal of Computational & Graphical Statistics*, 10,(3):440–454.

[Wu et al., 1992] Wu, C.-M., Chen, Y.-C., and Hsieh, K.-S. (1992). Texture features for classification of ultrasonic liver images. *IEEE Transactions on Medical Imaging*, 11(2):141 –152.

[Xiao et al., 2002] Xiao, G., Brady, J., Noble, J., Burcher, M., and English, R. (2002). Nonrigid registration of 3-D free-hand ultrasound images of the breast. *IEEE Transactions on Medical Imaging*, 21(4):405–412.

[Xiaoxun and Yunde, 2006] Xiaoxun, Z. and Yunde, J. (2006). Local Steerable Phase (LSP) Feature for Face Representation and Recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*, volume 2, pages 1363–1368.

[Xiong et al., 2012] Xiong, C., Johnson, D., Xu, R., and Corso, J. J. (2012). Random forests for metric learning with implicit pairwise position dependence. *CoRR*, abs/1201.0610.

[Zagzebski, 1996] Zagzebski, J. A. (1996). *Essentials of Ultrasound Physics*. Mosby.

[Zang et al., 2007] Zang, D., Wietzke, L., Schmaltz, C., and Sommer, G. (2007). Dense optical flow estimation from the monogenic curvature tensor. *Scale Space and Variational Methods in Computer Vision*, pages 239–250.

[Zetsche and Barth, 1990] Zetsche, C. and Barth, E. (1990). Fundamental limits of linear filters in the visual processing of two dimensional signals. *Vision Research*, (30).

[Zhan and Shen, 2006] Zhan, Y. and Shen, D. (2006). Deformable segmentation of 3-d ultrasound prostate images using statistical texture matching method. *IEEE Transactions on Medical Imaging*, 25(3):256 –272.

[Zhang et al., 2010] Zhang, J., Zheng, J., and Cai, J. (2010). Interactive mesh cutting using constrained random walks. *IEEE Transactions on Visualization and Computer Graphics*, 17:357–367.

[Zhang et al., 2007] Zhang, W., Noble, J. A., and Brady, J. M. (2007). Spatio-temporal registration of real time 3d ultrasound to cardiovascular mr sequences. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 343–350.

[Zhang et al., 2005] Zhang, W., Shan, S., Gao, W., Chen, X., and Zhang, H. (2005). Local gabor binary pattern histogram sequence (lgbphs): A novel non-statistical model for face representation and recognition. *IEEE International Conference on Computer Vision (ICCV)*, 1:786–791.

[Zhao and Pietikainen, 2007] Zhao, G. and Pietikainen, M. (2007). Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6):915 –928.

[Zhuang et al., 2009] Zhuang, X., Hawkes, D., and Ourselin, S. (2009). Unifying encoding of spatial information in mutual information for nonrigid registration. In *Information Processing in Medical Imaging*, volume 5636 of *Lecture Notes in Computer Science*, pages 491–502. Springer Berlin / Heidelberg.

[Zikic et al., 2006] Zikic, D., Wein, W., Khamene, A., Clevert, D.-A., and Navab, N. (2006). Fast deformable registration of 3d-ultrasound using a variational approach. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 9(1):915–923.

This work was fully sponsored by the EU grants:

FP7-ICT-2007-215190 and FP7-ICT-2009-6-270460