TECHNISCHE UNIVERSITÄT MÜNCHEN
Lehrstuhl für Biologische Bildgebung

Towards Real-Time Clinical Imaging with Multi-Spectral Optoacoustic Tomography: Reconstruction Approaches and Initial Experimental Studies

Christian Lutzweiler

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:
Prof. Dr.-Ing. Bernhard U. Seeber
Prüfer der Dissertation:
1. Prof. Dr. Vasilis Ntziachristos
2. Prof. Dr. Hans-Joachim Bungartz
3. Prof. Dr. Oliver Hayden

Die Dissertation wurde am 08.07.2016 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 07.07.2017 angenommen.

Abstract

Optical imaging plays a central role in basic biological research and clinical practice because of its rich and versatile contrast and its low implementation cost. In purely optical imaging, high spatial resolution is limited to shallow depths due to the intense photon scattering in biological tissues. Optoacoustic imaging has shattered the resolution barrier of traditional deep tissue optical imaging by exploiting the optoacoustic effect which is the conversion of transient absorbed photon energy to transient ultrasonic pressure waves. Due to its intrinsically hybrid nature, optoacoustic imaging combines the advantages of both energy forms, i.e. the contrast and the spectroscopic capabilities (known as multi-spectral optoacoustic tomography) of optical and the spatio-temporal resolution of ultrasonic imaging. The unique combination has enabled several novel preclinical applications in the fields of cancer, brain, or cardiovascular imaging. Optoacoustic approaches are also currently translated towards clinical operation for skin or breast cancer imaging.

In optoacoustic tomography, optical absorption images are formed computationally from the detected pressure signals. Advanced iterative model-based image reconstruction methods have been demonstrated to provide superior image quality as compared to analytical approaches. Yet, model-based methods are orders of magnitude more computationally demanding and thus their application is often not feasible for clinical imaging under high-throughput conditions. The first goal of this work was to identify characteristics of the optoacoustic reconstruction problem that allow to drastically decrease the computational cost of model-based reconstructions so that routine application in clinical imaging becomes feasible. The second goal was to abate the need for manual user input in the model-based reconstruction process and develop novel methods that can automatically and yet efficiently provide superior image quality especially in non-ideal clinical imaging scenarios.

For the experimental part, the goal was to contribute to the clinical translation of optoacoustic imaging and demonstrate the possibility to perform accurate imaging of the human finger. The finger is affected by several prevalent peripheral vascular diseases such as rheumatoid arthritis or Raynaud's phenomenon. The intrinsic sensitivity to blood contrast constitutes multi-spectral optoacoustic tomography a highly promising candidate to overcome the limitations of current modalities. The third goal of the thesis was to test multi-spectral optoacoustic tomography for clinically relevant applications in the finger and to extract characteristic anatomical and physiological parameters from the experimental studies.

For the acceleration of the reconstruction process two transformation-based approaches have been developed. The first exploited tomographic symmetries

of the acquisition geometry, enabled splitting and subsequent inversion of the large model matrix, and achieved video-rate high quality reconstructions for the most common optoacoustic system. The novel method offered the possibility of two orders of magnitude faster image reconstructions than established numerical approaches. The second approach was based on a sparse problem formulation rooted on general properties of the optoacoustic reconstruction problem. For typical experimental data-sets a factor of 40 - 700 x in computation time was gained over standard iterative reconstructions of similar image quality.

In order to automatically provide superior image quality, noise artifacts in multi-spectral data-sets were considerably reduced by the development of a state-of-the-art de-noising algorithm. Furthermore, causes of unphysical negative image values were identified, and a computationally efficient constrained inversion method was developed. A residual-based homogeneous auto-focusing method was shown to successfully resolve even fine image structures. In finger imaging however, a heterogeneous speed-of-sound distribution is beneficial and a novel signal domain analysis could retrieve improved acoustic propagation properties and in addition the outline of the bone.

Using a previously developed optoacoustic tomography scanner enriched by a custom finger holder, robust high-resolution anatomical imaging of the finger was realized in cross-sectional and volumetric acquisition mode and anatomical parameters like vessel lumen were extracted. In addition, functional parameters like the blood oxygenation level were determined for individual arteries and veins via multi-spectral imaging; artery pulsing and vasoconstriction were visualized and characterized via temporal imaging. For the first time, to the best of knowledge, an indocyanine green contrast agent at clinically relevant concentrations was detected in humans by means of optoacoustic imaging. Kinetic dynamic contrast profiles were resolved on an individual vessel basis and characteristic time constants of perfusion were successfully extracted.

The proposed reconstruction acceleration techniques are expected to facilitate the long-standing challenge of video-rate, model-based live previews in 2-D imaging; in 3-D, they are expected to establish accurate model-based optoacoustic reconstructions of high computational efficiency like needed for volumetric handheld probes or dermal microscopy. The image enhancement strategies will become standard tools whenever robust auto-focusing or noise reduction is required and the signal domain analysis might be translated to similar imaging settings such as dermal microscopy or breast cancer imaging. The demonstrated high image quality in the initial finger imaging studies will be the basis for future work involving patients suffering from Raynaud's phenomenon to study significant differences compared to a control group of healthy volunteers. Most importantly, the ability to detect indocyanine green is anticipated to allow for early phase rheumatoid arthritis detection by means of optoacoustic imaging. Higher quantification accuracy is expected by using the signal domain analysis and performing pharmakinetic modeling. The diagnostic value of optoacoustic imaging for arthritis imaging will be further evaluated by cross-validation with established modalities. Furthermore, the potential for optoacoustic pulse oxymetry will have to be tested and compared with established, purely optical methods.

Zusammenfassung

Optischer Bildgebung kommt auf Grund des reichhaltigen und vielfältigen Kontrastes und der niedrigen Kosten eine zentrale Rolle in der biologischen Grundlagenforschung und in der klinischen Praxis zu. Höchste Auflösungen mit rein optischer Bildgebung sind auf Grund der starken Lichtstreuung im Gewebe auf oberflächennahe Bereiche beschränkt. Optoakustische Bildgebung hat die Grenze von geringer Auflösung in Millimeter- und Zentimetertiefe unter Ausnutzung des optoakustischen Effektes, der Umwandlung von absorbierter transienter Photonenenergie in transiente Ultraschallwellen, durchbrochen. Die immanent hybride Technologie vereint die Vorzüge beider Energieformen, d.h. den Kontrast und die spektroskopischen Möglichkeiten (bekannt als multi-spektrale optoakustische Tomographie) der optischen Bildgebung und die hohe räumlich-zeitliche Auflösung der Ultraschallbildgebung. Diese einzigartige Kombination hat neuartige präklinische Anwendungsfelder in Bereichen wie beispielsweise der Krebsforschung, der neuronalen Bildgebung oder der kardiovaskulären Diagnose ermöglicht. Optoakustische Methoden werden neuerdings auch für den klinischen Alltag getestet, wie für Brustkrebsdiagnose oder Dermatologie.

Bei der optoakustischen Tomographie wird die Verteilung der absorbierten Lichtenergie computergestützt aus den gemessenen akustischen Signalen berechnet. Moderne iterative, modellbasierierte Bildrekonstruktionsmethoden können bessere Bildqualitäten im Vergleich zu analytischen Verfahren erreichen. Jedoch sind solche Algorithmen um Größenordnungen langsamer und daher ist ihre Anwendung im klinischen Umfeld mit hohen Durchsatzraten oft nicht möglich. Die erste Zielsetzung der vorliegenden Arbeit war es, Charakteristika des optoakustischen Rekonstruktionsproblemes zu ermitteln, die eine drastische Reduzierung des Rechenaufwandes modellbasierter Rekonstruktionen erlauben und die damit einen routinemäßigen Einsatz in der klinischen Bildgebung möglich machen. Die zweite Zielsetzung der vorliegenden Arbeit war es, manuelle Benutzereingaben für modellbasierte Rekonstruktionen zu vermeiden und neue Methoden zu entwickeln, die automatisch und doch effizient eine verbesserte Bildqualität speziell für nicht-idealisierte klinische Anwendungssituationen bereitstellen können.

Im experimentellen Teil sollte ein Beitrag zur klinischen Translation optoakustischer Bildgebung geleistet werden und das Potential für akkurate Messungen im Finger evaluiert werden. Finger sind von zahlreichen prävalenten periphärvaskulären Krankheiten wie rheumatoider Arthritis oder dem Raynaud-Syndrom betroffen. Der hohe intrinsische Hämoglobinkontrast macht die multispektral optoakustische Tomographie zu einem vielversprechenden Ansatz, um die Einschränkungen der bestehenden Modalitäten zu überwinden. Die dritte Zielsetzung dieser Arbeit war daher, die Eignung der multi-spektralen optoakustischen Tomographie für klinisch relevante Anwendungen im Finger zu testen

und charakteristische anatomische und physiologische Parameter zu extrahieren.

Zur Beschleunigung des Rekonstruktionsprozesses wurden zwei auf Transformationen beruhende Methoden entwickelt. Der erste Ansatz nutzte die tomographische Symmetrie der Detektionsgeometrie aus, ermöglichte das Aufspalten und anschließende Invertieren der großen Systemmatrix und erreichte mehr als 25 hochauflösende Rekonstruktionen in einer Sekunde für das am häufigsten gebrauchte optoakustische Bildgebungssystem. Dies entspricht einer Steigerung um zwei Größenordnungen gegenüber bestehenden numerischen Methoden. Der zweite Ansatz beruhte auf einer Neuformulierung des Problems mit dünn besetzten Matrizen, die auf generellen Eigenschaften des optoakustischen Rekonstruktionsproblemes basierte. Im Vergleich zu gebräuchlichen Verfahren wurde für typische experimentelle Datensätze eine Verbesserung der Rekonstruktionszeiten um einen Faktor von 40 - 700 bei vergleichbarer Bildqualität erreicht.

Ein zum Zwecke der automatischen Verbesserung der Bildqualität entwickelter Algorithmus konnte Rauschartefakte in multi-spektralen Datensätzen deutlich reduzieren. Außerdem wurden Ursachen für unphysikalische negative Bildintensitäten identifiziert und eine effiziente Inversionsmethode mit Nebenbedingungen entwickelt. Eine auf dem Residuum basierte Autofokussierung konnte im homogenen Fall auch kleinste Bilddetails auflösen. Bei der Bildgebung von Fingern ist jedoch eine heterogene Verteilung der Schallgeschwindigkeit vorteilhaft. Durch eine neuartige Signalanalyse konnte diese Verteilung aufgefunden, verbesserte Bilder erzeugt und zusätzlich der Knochen im Finger lokalisiert werden.

Durch Erweiterung eines bestehenden optoakustischen Tomographen um einen selbstentwickelten Fingerhalter konnten hochauflösende anatomische Schnittbilder und Volumendarstellungen des Fingers erzeugt und anatomische Parameter wie das Gefäßlumen daraus extrahiert werden. Zusätzlich konnte die Sauerstoffsättigung einzelner Arterien und Venen mit Hilfe des multi-spektralen Ansatzes bestimmt werden. Das Pulsieren einer einzelnen Arterie oder die Vasokonstriktion der Fingervaskulatur konnten zeitaufgelöst dargestellt und charakterisiert werden. Zum ersten Mal überhaupt - nach unserem Wissensstand - konnte das Kontrastmittel Indocyaningrün in klinisch relevanten Dosen im Menschen mittels optoakustischer Methoden nachgewiesen werden. Die Kinetik des dynamischen Kontrastes konnte für einzelne Blutgefäße bestimmt und die Zeitkonstanten des Perfusionsprozesses konnten erfolgreich extrahiert werden.

Mit Hilfe der vorgeschlagenen Methoden zur Rekonstrutionsbeschleunigung könnte das seit langem bestehende Ziel von modellbasierten Echtzeit-Rekonstruktionen in 2-D erreicht werden; in 3-D könnten genauere, modellbasierte Rekonstruktionen mit hoher Effizienz bereitgestellt werden wie sie z. B. für portable volumetrische Geräte oder in der dermatologischen Mikroskopie benötigt werden. Die beschriebenen Bildverbesserungsstrategien könnten zu Standardwerkzeugen zur zuverlässige Fokussierung oder zur Rauschunterdrückung werden. Die Bildoptimierung per Signalanalyse könnte auf ähnliche Bereiche wie z.B. dermatologische Mikroskopie oder Brustkrebsscreening übertragen werden. Die gezeigten ersten Studien im Finger könnten die Basis für weitere Studien mit Patienten, die am Raynaud-Syndrom leiden, darstellen. Als wichtigste Folge aber könnte die Fähigkeit, Indocyaningrün nachzuweisen, eine Diagnose von rheumatoider Arthritis in der Frühphase mittels optoakustischer Tomographie ermöglichen. Eine höhere Genauigkeit der Quantifizierung wird durch Methoden wie die der Signalanalyse und die der pharmakinetischen Modellierung erwartet.

Contents

Chapter 1

Introduction

1.1 Motivation: Multi-Spectral Optoacoustic Tomography as an Emerging Clinical Imaging Modality

In our aging society, health care is playing an increasingly important role. In this context, clinical imaging has emerged as a versatile tool for gaining new insights in biomedical research as well as one of the key factors for successful diagnosis, treatment planning, and post-treatment control [1]. Biomedical imaging has probably been one of the fastest growing fields in clinical routine over the last decades, owing to the rapid progress in instrumentation and computation technologies. The use of imaging technology is highly attractive because it can visualize anatomy and molecular processes non-invasively with relatively short acquisition times and it is highly cost-efficient [2].

The family of established clinical imaging modalities has continuously grown and enabled novel visualization of a variety of anatomical structures and physiological processes [3]: Magnetic resonance imaging provides volumetric, high-resolution images with excellent soft tissue contrast and is able to determine for example brain connectivity maps [4]; X-ray computed tomography results in detailed reconstructions of bone anatomy or enables rapid angiography in interventional medicine aided by contrast agents [5]; ultrasound imaging can provide anatomical images as well as quantitative measurements of blood flow [6]; nuclear imaging methods visualize molecular and metabolic processes such as the energy uptake of certain regions in the brain or tumors [7].

Despite of the merits of each single modality in enabling specific applications, there is no single all-purpose modality. Each modality has its associated drawbacks and limitations. Consequently each application requires a specific imaging modality and often new imaging approaches have to be developed in order to facilitate certain applications. Each of the established modalities is associated with certain drawbacks such as providing low soft tissue contrast or low spatio-temporal resolution, or being carcinogenic, bulky, or costly.

Optical imaging methods are particularly attractive because of their excellent soft tissue contrast based on the distinct spectral profiles of various molecules, because of the non-carcinogenic nature of the low energy radiation,

because of the extraordinary spatio-temporal resolution possible, and because optical imaging technology is typically inexpensive and table top sized [2, 8]. For centuries the microscope has been and still is THE standard tool for biological observations and recent developments such as 2-photon, confocal, or super-resolution imaging have pushed its limits even further [9]. In gastrointestinal applications, optical endoscopes are the gold standard for examinations. Optical imaging contrast is rooted on a broad variety of molecules, ranging from intrinsic chromophores such as hemoglobin, over systemically administered contrast agents and histological staining, to genetically engineered fluorophores [10, 11]. Although light can penetrate up to several centimeters depth, optical imaging is mostly restricted to superficial imaging of a few hundreds of micrometers at most because of the strong photon scattering in biological tissues that eliminates all phase information of the photons after propagating through bulky tissue [8]. Deep tissue optical imaging is made possible aided by computational reconstruction methods. Nevertheless it comes with the disadvantage of considerably degraded resolution in the millimeter range.

Optoacoustic imaging (OA, also known as photoacoustic) has shattered the frontier of limited spatial resolution in deep tissue optical imaging and consequently facilitated a broad range of novel applications [12, 13]. As an intrinsically hybrid imaging modality it is based on the conversion of absorbed transient photon energy to ultrasonic pressure waves [14]. In this way, the spatial resolution of OA imaging is insensitive to strong photon scattering while maintaining optical contrast. Thus, the merits of optical and ultrasonic imaging, i.e. excellent contrast and high spatio-temporal resolution, are combined in this hybrid modality. OA tomography is able to provide spatially resolved maps of photon absorption by capturing the pressure waves around the imaged object and performing a subsequent reconstruction process. In addition, it is capable of providing video-rate images in two or three spatial dimensions. Especially the multi-spectral optoacoustic tomography (MSOT) technology is able to provide a spectral distinction of different absorbing molecules. OA imaging may simultaneously render anatomical, functional, and molecular images: Structures ranging from single red blood cells [15], over capillaries and blood vessels [16, 17], to whole brains and mouse torsos have been visualized [18, 19]; functional processes could be monitored from slow, longitudinal processes like tumor growth and pharmakinetics [20, 21], over hemodynamic responses and organ perfusion [22, 23], to heartbeat and even neuronal activation [24, 25]; molecular imaging studies based on OA imaging could visualize blood oxygenation saturation and glucose metabolism [26, 27], targeting of tumor and apoptotic cells [28, 29], as well as gene expression [20, 30].

The unique imaging performance of OA imaging has also led to several clinically relevant or translational animal model imaging studies [31], mostly addressing the two major causes of mortality, cardiovascular diseases and cancer. Cardiovascular imaging studies have revealed arterial plaques or hypoxic regions [32, 33], visualized stroke or perfusion [34, 35], or measured the cardiac output [24]. In OA imaging of cancer, tumor margins and angiogenesis of tumor vasculature could be visualized [36], hypoxia and metabolism studied [33], and tumor growth and treatment tracked [20, 29]. In addition, management of multiple other diseases might benefit from the advantages of OA imaging, such as gastrointestinal endoscopy [37], sentinel lymph node imaging [38], ophthalmology [39], or needle biopsy guidance [40]. Reported studies of *in vivo* OA imaging in

humans have been relatively rare to date, due to safety and legal regulations, lack of clinically approved contrast agents, or challenges in the translation from animal models to patients. But the maturity of OA instrumentation, processing, and image acquisition strategies promotes an increasing number of such studies. Among the reported studies in humans are breast cancer imaging towards OA mammography [41, 42], angiography [17], carotid imaging towards detection of arterial plaques [43], and skin imaging towards detection of melanomas [44, 45].

1.2 Objectives

OA imaging has already been demonstrated to successfully detect and characterize a broad variety of disease models in small animals, as outlined above. The next direct goal is thus to successfully translate OA imaging to a clinical setting and to evaluate its potential. Robust and versatile OA instrumentation has been designed, and dedicated reconstruction and processing frameworks have been developed as well. Consequently, a routine application of OA modalities in clinical environments has become feasible.

From a clinical point of view, there is a strong need for novel imaging technologies in many medical fields where traditional imaging modalities cannot provide sufficient image quality or flexibility. Consequently, clinical OA imaging bears promise to be successfully complement or replace existing modalities, to facilitate better diagnosis, or enable new fields of clinical imaging applications.

Many clinical OA imaging studies are on their way or are initiated, and research will address mainly three fields: first, the selection of suitable fields of application like breast cancer, dermal imaging, or endoscopy, the recruitment of volunteering patients, and the development of suitable routine imaging protocols; second, the development of dedicated hardware for the respective application such as handheld devices, or endoscopes; third, the development of robust and automated processing and reconstruction strategies that are able to provide a good image quality and that are fast enough to be used in clinical routine, i.e. they should ideally provide real-time processing.

The objectives of this work are to develop MSOT imaging towards clinical applications, which includes contributions to image reconstruction and processing as well as to investigations of application studies related to clinically relevant diseases in the human finger. The three main objectives addressed by this work are:

1. Can computationally expensive iterative OA reconstruction approaches be applied for routine clinical imaging?
   OA absorption images are not directly measured but they need to be obtained computationally from the detected acoustic signals. Analytic reconstruction methods exist and they are commonly selected due to their computational speed. However, they are also prone to significant artifacts in experimental situations. Iterative numerical reconstruction methods are more accurate, but they are orders of magnitude more computationally demanding. For the huge amount of data created during high-throughput real-time MSOT imaging, standard iterative reconstructions are often too time consuming and analytical reconstructions have to be performed at the

expense of image quality. One goal of this work is to identify particulars of the OA imaging and reconstruction problem to facilitate considerable acceleration of the image reconstruction process in typical real-time MSOT scenarios, so that high-quality numerical reconstructions can be regularly applied in clinical routine as the standard method.

2. Which reconstruction methods for good image quality in real, non-ideal clinical imaging application scenarios can be developed that work robustly, efficiently, and without user input?
   Due to the digital image formation in OA imaging, reconstruction, pre- and post-processing methods need to be applied and they directly determine the resulting image quality. In order to maintain high image quality and to avoid artifacts under non-ideal imaging conditions, multiple (pre-)processing methods need to be applied and multiple parameters need to be selected during the reconstruction process. Potential artifacts could stem from noise or arise from differences between assumed and actual acoustic properties of the imaged object and could led to e.g. negative pixel values. This work aims at developing methods which effectively suppress such artifacts and which work automatically without the need for manual expert input. In addition, high computational efficiency of such methods is also required to facilitate a routine application in clinical imaging.

3. Can MSOT be applied for imaging and diagnosing common peripheral vascular diseases in the finger?
   The intrinsic blood contrast of MSOT constitutes it an ideal candidate for imaging and diagnosing peripheral vascular diseases. Two common peripheral vascular diseases are mainly affecting the fingers: Rheumatoid arthritis results in inflamed and swollen finger joints, and is associated with large socioeconomic costs; and Raynaud's phenomenon is a class of diseases associated with hypoperfusion, ulceration, and critical ischemia in the finger. This work investigates the application and optimization of OA imaging for the diagnosis of peripheral vascular diseases in the finger. It aims at answering the question whether anatomical parameters such as vessel location, vessel diameter or vascular branching can be robustly extracted from OA measurements to potentially study anatomical changes in Raynaud's phenomenon. Further it is investigated whether functional parameters such as blood oxygenation level or the kinetic profiles of an optical contrast agent in the vasculature can be reliably extracted to potentially study perfusion or neovascularization in Rheumatoid arthritis.

1.3 Outline

The presented work is organized as follows:

The overall structure has three main parts: Chapter 2, Chapter 3, and Chapter 4 present the background and context of this work including state-of-the-art OA imaging and reconstruction approaches. Chapter 5, Chapter 6, and Chapter 7 present novel, optimized MSOT reconstruction methodologies with increased reconstruction performance and enhanced resulting image quality. Finally, Chapter 8 and Chapter 9 present initial clinical imaging studies in the

finger using the MSOT technology.

In Chapter 2, the most important optical and non-optical (pre-)clinical imaging modalities are introduced and compared with respect to their performance metrics and their specific advantages and drawbacks. Furthermore, the interaction of photons with tissue and their propagation therein is discussed and various classes of intrinsic or extrinsic targets of optical contrast are presented.

Chapter 3 introduces the fundamentals of OA imaging. The theory of the OA effect and the propagation of ultrasonic pressure waves in tissue is reviewed in detail. Then specific instrumentation considerations and approaches in existing OA systems are presented. In particular, the small animal MSOT scanner that was used for the experimental studies in this work is described.

Chapter 4 discusses the reconstruction problem associated with MSOT. Different types of reconstruction methods are presented and a particular focus is placed on the class of iterative numerical algorithms. The specific requirements for reconstruction in terms of image quality and reconstruction performance in the context of clinical imaging are discussed at the end of the chapter.

An approach for acceleration of tedious numerical inversion is presented in Chapter 5. The approach exploits the symmetries present in most OA systems. The developed method is tested with experimental data of complete and incomplete signal information and a high performance reconstruction implementation on graphics processing units is presented.

Chapter 6 presents a novel sparsity-based inversion framework. The transfor"-mation-based approach results in a sparse and thus computationally favorable formulation of the reconstruction problem. The specific transformations that render the matrix and the signals sparse are discussed in detail and the performance of the framework is tested with four different experimental data-sets.

In Chapter 7, approaches for an automated reconstruction quality enhancement in MSOT are presented. The approaches aim at noise reduction, negative image value suppression, automated image focusing, or an automated assignment of advanced acoustic propagation models. Finally, potential future directions of reconstruction approaches in the context of clinical imaging are discussed.

Chapter 8 presents initial experimental studies of *in vivo* finger imaging with MSOT. Motivated by the specific finger anatomy and associated peripheral vascular diseases therein, several anatomical and functional imaging parameters are documented. The anatomical imaging approaches include cross-sectional and volumetric angiography, while the functional imaging approaches include blood oxygenation determination, pulse detection, and vascular constriction during a thermal stress test.

In Chapter 9, clinical MSOT imaging is discussed in a context of diagnosing rheumatoid arthritis. Rheumatoid arthritis is clinically detected using multiple modalities, including pharmakinetic imaging of contrast agent profiles. For optical imaging methods, indocyanine green is one of the most common contrast agents because it is clinically approved. The first demonstration of indocyanine green detection in humans by means of MSOT is presented in a perfusion study in the finger vasculature. The results are then discussed in the context of rheumatoid arthritis detection and future steps and challenges for that purpose are outlined.

Finally, Chapter 10 summarizes and discusses the results of this work and further provides an outlook on future steps and implications of this work.

Chapter 2

Biomedical Optical Imaging

2.1 Introduction to Biological and Medical Imaging

Biomedical imaging aims at providing spatially resolved information from within the body in living organisms which exceed the information obtainable with pure vision of the eye. Its main purpose is to enable better diagnosis of diseases and planning of their treatment, or to result in a better understanding of fundamental biological processes. Early biomedical imaging modalities included mirrors, looking glasses, electrocardiography (ECG), or X-ray imaging. Since the early steps, a large armada of modalities has been developed to visualize different anatomical structures and physiological and biological processes. These developments are summarized in Sect. 2.3 and Sect. 2.4.

The exploited information might be of STRUCTURAL nature, visualizing anatomy as bone fractures in X-ray imaging; alternatively, the physiology and other FUNCTIONAL aspects might be probed as in imaging of the beating heart; in the recent years, even processes at metabolic and MOLECULAR level can be visualized based on the progress in imaging technologies.

In vivo imaging aims at providing images of organisms in their natural state without altering their structure or function for or through imaging. Noninvasive or at least minimally invasive imaging methods are thus favored as much as a non-toxic nature of the imaging process. The most important parameters of imaging are discussed in the following section.

In general, most biomedical imaging modalities are based on the following five fundamental steps (Fig. 2.1): (1) An excitation signal is generated outside of the object to be imaged. (2) The signal propagates to and within the object. (3) Harvesting a specific physical process, the signals interact with the object and thus encode information into the signal. (4) The modified signal propagates further within and from the object. (5) Outside the object, the signals are detected and further processed to result an image; either a direct physical image is formed or the images are rendered computationally in modern embodiments. The unique capabilities of OA imaging stem from the fact that signal generation/propagation ((1) and (2)) and signal propagation/detection ((4) and (5)) are rooted on entirely different physical phenomena, namely light and sound [31]. Conversion (3) between both forms of energy is achieved by means of the

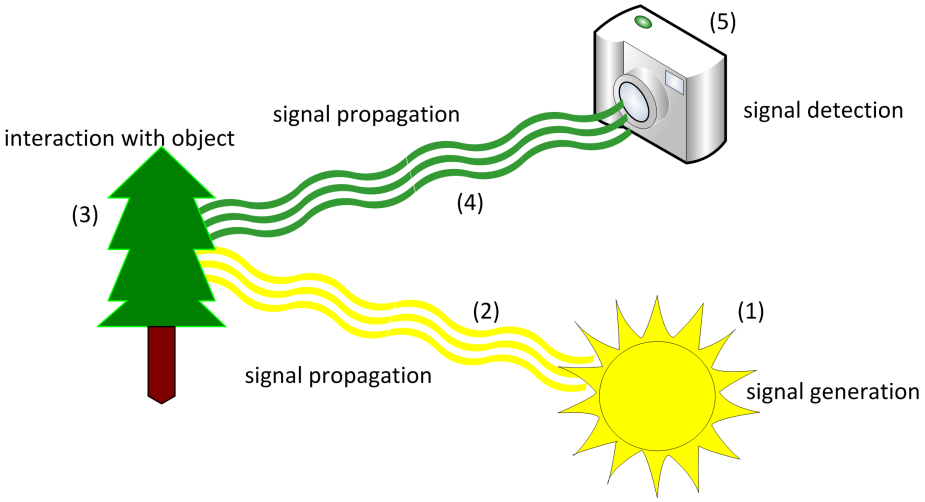

Figure 2.1: Fundamental steps involved in imaging: (1) Signal generation; (2) signal propagation towards the imaging sample; (3) interaction with the object; (4) propagation from the imaging sample; (5) detection of the signals.

OA effect, described in detail in Sect. 3.2.

2.2 Characteristic Performance Metrics of Imaging

The variety of modalities is based on multiple physical principles or technologies used; however common parameters may be used in order to characterize their imaging performance:

- Spatial resolution, which is the ability to distinguish and separate objects located closely. The achievable spatial resolution covers many orders of magnitude, ranging from tens of nm in organelles to some cm in whole-body imaging (Fig. 2.2). Spatial resolution might be anisotropic, i.e. different for different spatial dimensions, or even non-homogeneous over the field-of-view. The spatial resolution of an imaging modality can be characterized and determined by its modulation transfer function or point-spread function.

- Temporal resolution determines the acquisition time for obtaining a single image and consequently the imaging rate; for static imaging, the imaging rate is of only minor importance while in imaging of dynamic processes, the imaging rate has to be on the same order as the processes to be observed, ranging from some ns to several mins and days.

- Penetration depth is the maximum depth from which information is faithfully observed. Many of the high-resolution modalities as microscopes typically offer restricted penetration depth (Fig. 2.2). Thus their application is limited to small sized specimens or to superficial areas.

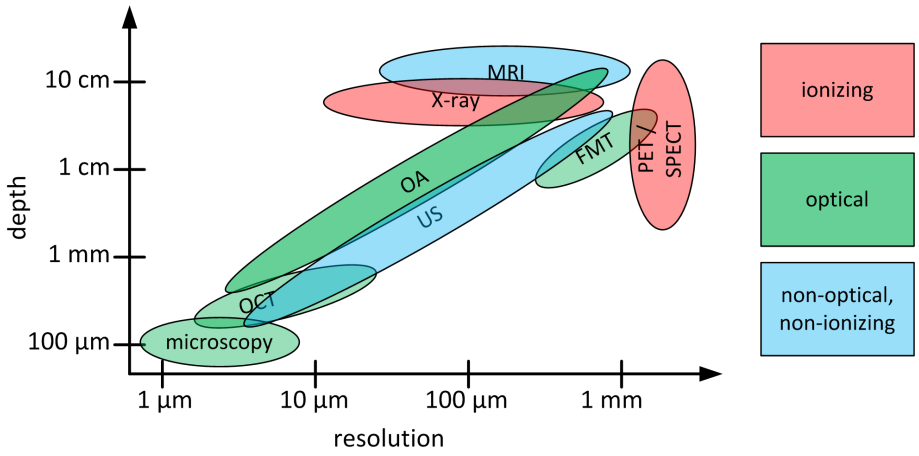

Figure 2.2: Characteristic spatial resolution and imaging depth of different biomedical imaging modalities. See Sect. 2.3 and Sect. 2.4 for details on the modalities. Diagram based on values reported in Refs. [1, 12].

- CONTRAST in the images stems from multiple physical processes, thus providing maps of heavy atoms, distribution of water, blood, or glucose.
- SENSITIVITY is the minimal number or concentration of target molecules (or of cells) that can be faithfully detected with an imaging modality at sufficient signal-to-noise ratio (SNR). Molecular imaging typically offers higher sensitivity compared to purely anatomical imaging.
- COSTS OF INSTRUMENTATION are an important economical aspect determining the practical use as much as if a modality is stationary or if it is portable and might be used at bed-side.
- Hazardous IONIZING RADIATION and involvement of toxic substances strongly limit the application of certain modalities or even prelude their usage in humans.

2.3 Overview of Preclinical and Clinical Imaging Modalities

Multiple modalities are nowadays in clinical and preclinical practice (Fig. 2.3, top). Only an exemplary selection is presented here. Optical imaging methods are presented in Sect. 2.4.

- X-RAY radiation having a photon energy of tens of keV was discovered by W.C. Roentgen in 1895 and immediately used for visualization of bones because of the strong interaction of X-ray radiation with heavy atom nuclei. In the 1970s, planar projection X-ray imaging was extended to three dimensions, acquiring images from multiple angles and reconstructing the volumetric images in the computer, known as X-ray computed tomography (X-CT) [5]. Both planar X-ray and X-CT have readily become clinical

standard modalities for structural imaging as in dental or brain visualization. The advantages of X-ray imaging are its high spatio-temporal resolution and cost efficiency. Its drawbacks are the ionizing nature of the high energy radiation involved and the low soft tissue contrast. The latter can be improved administering contrast agents (CA) in dynamic contrast enhanced (DCE) imaging in order to visualize heart function or vasculature in digital subtraction angiography [46]. Recent developments include spectral imaging using multiple photon energies or exploiting the phase-shift instead of only attenuation in phase contrast imaging in order to enhance the poor soft tissue contrast [47].

- Ultrasonic imaging (US) is based on transmitting pulsed narrow bandwidth longitudinal pressure waves in the frequency range from hundreds of kHz to several MHz and receiving their reflections from acoustically mismatching areas within the tissue [6]. Low cost of piezoelectric transducer technology and electronics and excellent spatio-temporal resolution based on the time-resolved US waves led a widespread use of US systems. In US, depth is resolved via time-of-flight (TOF) of the signals and 2-D images are created steering the acceptance angle by electronically introducing an individual time delay to the elements of the transducer arrays. Additionally, flow and motion can be easily measured based on the Doppler shift of the frequencies in the detected signals [48]. In order to increase low soft tissue contrast, highly mismatching air-filled bubbles have been used as CAs to increase elastic scattering or induce generation of higher harmonics [49]. Recently, US imaging approaches have been enriched by extension to three spatial dimensions, computational image formation, or transmission mode measurements [6].

- Nuclear imaging approaches have probably been the first to apply the molecular imaging paradigm [7, 50]. Nuclear imaging is based on the radioactive decay of an administered tracer and generation of a γ-quantum either via γ-decay in single photon emission tomography (SPECT) or 511 keV annihilation radiation of the β-decay in positron emission tomography (PET). With dextrose based radiotracers, PET is able to visualize regions of increased metabolic rate. Despite its excellent sensitivity, the use of nuclear methods is restricted to specific scenarios in practice by the limited resolution and anatomical contrast, high cost, and the ionizing nature of the radiation.

- Magnetic resonance imaging (MRI) is based on the Zeeman-energy-splitting of nuclei with a magnetic momentum exposed to a (static) magnetic field and their subsequent Bloch-oscillations of the magnetization if excited by a radio-frequency (RF) pulse [51, 52]. The penetration of RF and static magnetic fields is not limited in the human body and thus whole body imaging is possible in humans. Because of the largest gyromagnetic ratio, MRI is mainly sensitive to the distribution of hydrogen atoms, most importantly in water, and their chemical neighborhood. MRI provides excellent anatomical images of water (T1-weighted energy relaxation) and fat (T2-weighted phase correlation relaxation) at good spatial resolution owing to strong magnetic fields of up to 3 T in humans. Additionally, functional imaging is possible with MRI relying on the different

behavior of oxygenated (HbO) and and deoxygenated (Hb) hemoglobin in BOLD-MRI (blood oxygenation level dependent MRI) or anisotropic diffusion of water within directional structures in diffusion tensor imaging (DTI), providing unrivaled insights into human brain activity and anatomy [4, 53, 54]. Functional contrast might be further enhanced using Gadolinium (Gd) based CAs locally altering the relaxation rates [52, 55]. The drawbacks of MRI are typically associated with low spatio-temporal resolution at sufficient SNR and the large cost of instrumentation.

- MAGNETIC PARTICLE IMAGING exploits the non-linear behavior of paramagnetic nanoparticles in a highly inhomogeneous magnetic field [56]. It holds promise for preclinical studies owing to the high spatio-temporal resolution and unlimited penetration depth, but requires potentially toxic nanoparticles.

- ELECTRICAL IMPEDANCE TOMOGRAPHY aims at providing maps of the electrical impedance of tissue using 4-electrode measurements at multiple locations. Differences in impedance might stem from the high conductivity of blood or saline in cardiovascular applications or are associated with pathological changes [57].

- HYBRID MODALITIES have been developed in order to combine the strengths of the different modalities. Of particular interest is enriching molecular imaging with anatomical reference images, as in PET-CT or PET-MRI [58–60].

- Transmission and scanning ELECTRON MICROSCOPY have been established as validation tool for basic biological studies. Electron microscopy is restricted to *ex vivo* imaging and small sample sizes, but provides structural information on an atomic and organelle level [61].

2.4 Overview of Optical Imaging Modalities

Optical imaging roots on the vision with the bare human eye and aims at visualizing what is revealed without using additional 'tools'. Optical imaging has progressed from the first use of magnifying glasses to the live visualization of single molecules. It is appealing for mainly two reasons: First, most molecules absorb in the visible range because their excitable energy levels are comparable to the energy of those non-ionizing photons (see Sect. 2.6). Second, sophisticated technology for generation, manipulation, and detection of photons is readily available because of the ubiquitous use of optical technologies in multiple fields of science and technology; thus it is facilitating fast and cheap imaging.

- The LIGHT MICROSCOPE has enable the first visualization of single cells by A. van Leeuwenhoek in the 17th century and its resolution has since then approached its ultimate physical diffraction limit predicted by Abbe, which is approximately half the wavelength of the excitation light. The microscope has become THE standard tool of biology, but is restricted to small objects or superficial areas of some tens of μm. Multiple illumination scenarios as standard bright-field mode in transmission and reflection

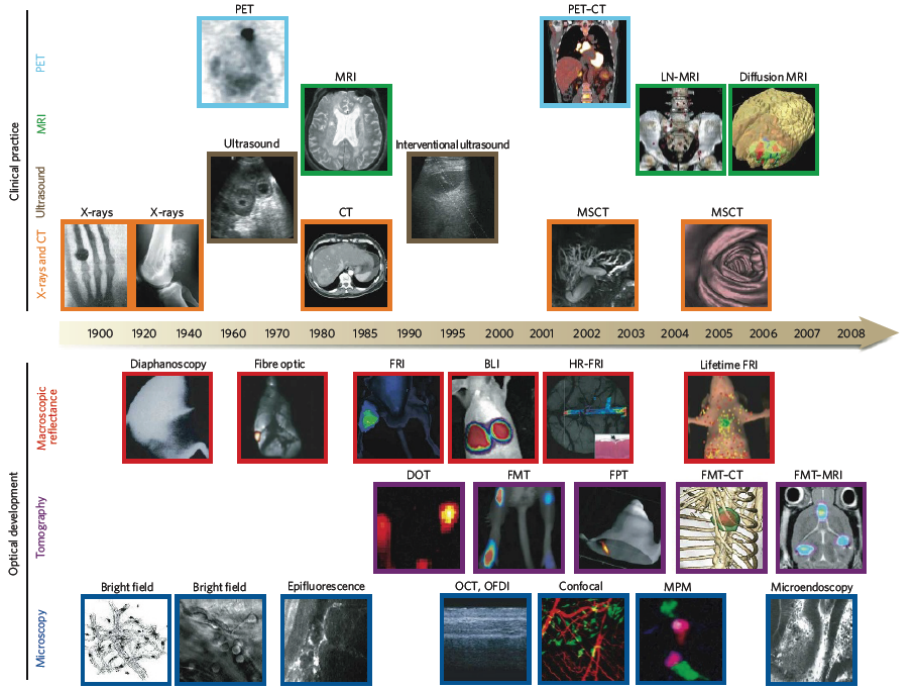

Figure 2.3: Overview of the development of clinical (top) and optical (bottom) imaging modalities.
BLI: bioluminescence imaging; FPT: fluorescence protein tomography; FRI: fluorescence reflectance imaging; HR-FRI: high-resolution FRI; LN-MRI: lymphotropic nanoparticle-enhanced MRI; MPM: multiphoton microscopy; MSCT: multislice CT; OFDI: optical frequency-domain imaging. Figure reprinted with permission from Ref. [1].

mode or dark-field mode to visualize scattering or polarization change are used [8].

- Standard microscopy has been enhanced in order to suppress undesired signals: To suppress the background absorption, FLUORESCENCE MICROSCOPY is used, where the detected photons are of slightly lower energy than the photons used for excitation (see also Sect. 2.6). Furthermore, confocal microscopy uses both focused illumination and focused detection to minimize signals not stemming from the focal plane. Similarly, 2-photon microscopy uses non-linear absorption processes which occur only for high photon densities in space and time, i.e. in the focus. 2-photon microscopy is thus basically free of background absorption [62]. Generalization to higher multi-photon processes as 3-photon microscopy is possible, resulting higher energy density thresholds and different optical selection rules, and therefore also complementary sources of contrast. However, multiphoton processes require raster-scanning and are thus relatively slow. One approach to overcome the scan-time limitation is to perform optical sectioning via restricting the illumination of the sample to an orthogonal plane, known as SPIM [63].

- SUPER-RESOLUTION TECHNIQUES have been developed to overcome the diffraction limit of optical imaging [9, 64]. One example of super-resolution is stimulated emission depletion (STED): After diffraction-limited excitation of fluorescent molecules, a second pulse depletes the excited states before emitting a photon everywhere except in the center of the annular depletion pulse [65]. Effectively, only the molecules at the arbitrarily small center are contributing to the fluorescence image; an entire image is then obtained by raster scanning of the pulses.

- Optical imaging approaches are also able to visualized the LOCAL BIOCHEMICAL ENVIRONMENT: Bioluminescence imaging is based on the expression of the reporter gene luciferase which enzymatically catalyzes the oxidation of luciferin under emission of a photon. The photon is subsequently detected by a camera system or microscope [2]. Fluorescence lifetime imaging capitalizes on the quenching of fluorescence depending on the local environment of the fluorophore which can result in a shorter fluorescence lifetime. Of particular interest is the reduction of lifetime by oxygen whose local partial pressure can be obtained from lifetime imaging [66]. Förster-resonance energy transfer imaging roots on the transfer of excitation energy from a donor molecule to a second acceptor molecule which finally emits a fluorescence photon of lower energy [67]. The efficiency of the energy transfer crucially depends on the relative distance and orientation of the two molecules and thus on the local chemical environment. Raman scattering techniques use a special type of elastic scattering, the Raman scattering from molecules with distinct vibrational energy levels, by detecting photons with an energy deviating from the excitation light energy just by the vibrational energy of the molecules [68]. In this fashion, the local molecular composition of tissues can be probed with Raman imaging methods.

- While microscopy techniques aim at enlarging tiny structures, ENDOSCOPIC IMAGING reveals regions not directly accessible to human vision as the esophagus or the colorectal tract. Early endoscopes involved systems of multiple lenses and later were based on fiber-optical imaging systems. Nowadays also CCD-based cameras directly at the tip of the endoscope are in use [69].

- Optical imaging is not restricted to superficial applications or small specimens. However, resolution considerably decreases from scattering with larger depths (see next section). OPTICAL COHERENCE TOMOGRAPHY (OCT) visualizes scattering at intermediate depths not accessible to conventional microscopy as for retinal or skin imaging. OCT is based on interferometrically detected, back-scattered photons from a light source with short coherence time [70].

- DEEP OPTICAL IMAGING is facilitated by diffuse optical tomography (DOT) which relies on a diffusive light transport model and measurements at multiple illumination-detection combinations outside the object [71]. Distribution of the absorbers is then restored numerically based on the light transport model. Fluorescence molecular tomography (FMT) aims at recovering the distribution of a fluorophore instead. FMT uses different

wavelengths for illumination and detection [72, 73]. The approach results in lower background and advantages in the reconstruction process from the additional DOT measurements at the excitation wavelength.

- Other techniques for deep tissue imaging have been developed as well: One is ULTRASONICALLY TAGGING the photons which are passing the US transducer focus. Such photons are intensity modulated with the frequency of the US. Detection of tagged transmitted or reflected photons and raster scanning of the focal point results an image then [74]. Laser speckle imaging exploits the speckle pattern generated from a highly scattering medium [75].

The above selection of the optical modalities highlights the great challenge of biomedical optical imaging: High spatial resolution is well achievable, but is restricted to very shallow depth of tens of μm. Although deep optical imaging is possible, the achievable resolution in purely optical modalities is very limited as explained in the following section. OA as intrinsically hybrid imaging modality is unique, because it is able to provide optical contrast and good resolution at up to cm scale imaging depth (see Chapter 3).

2.5 Transport of Light in Biological Tissues

The propagation of photons at arbitrary energy level (Fig. 2.4(a)) in biological tissue is in general described by the well-known Maxwell equations and their interaction with molecules by quantum field theory. However, the microscopic theory is too complex for realistic large scale problems and no useful analytical models for biomedical purposes exist besides the Mie theory to explain elastic photon scattering at spherical objects.

On large scales, photon propagation in a wave or geometrical ray model is described using effective material properties instead, mainly the refraction index, which are most often spatially varying. Effects like refraction, reflection, diffraction, and absorption are readily accounted for.

In biological tissues, scattering is the most dominant process for optical wavelengths owing to the heterogeneity on a molecular and organelle level [8, 76]. Scattering can in approximation be described by the scalar scattering coefficient μ_s, which is the inverse of the average distance a photon travels between two consecutive scattering events. The scattering coefficient exceeds the absorption coefficient μ_a considerably in most biological specimens ($\mu_s >> \mu_a$). One distinguishes between two regimes of photon transport (Fig. 2.4(b)): Ballistic photons that have not been scattered might be (re-)focused and are basis to microscopy imaging; and diffuse photons that have been scattered multiple times have (almost) no directionality and phase information and can consequently not be focused. Thus any focused beam is broadened from scattering after several mm of propagation in highly scattering tissue (Fig. 2.4(c)).

Transport of light in scattering media can be modeled in a particle model by the Boltzmann equation, an integro-differential equation. Owing to its 6-dimensional nature, the Boltzmann equation is often too complex to be solved (very few analytical solutions exist) and the lowest order approximation assuming fully diffused photons is solved instead [77], known as diffusion equation (DE):

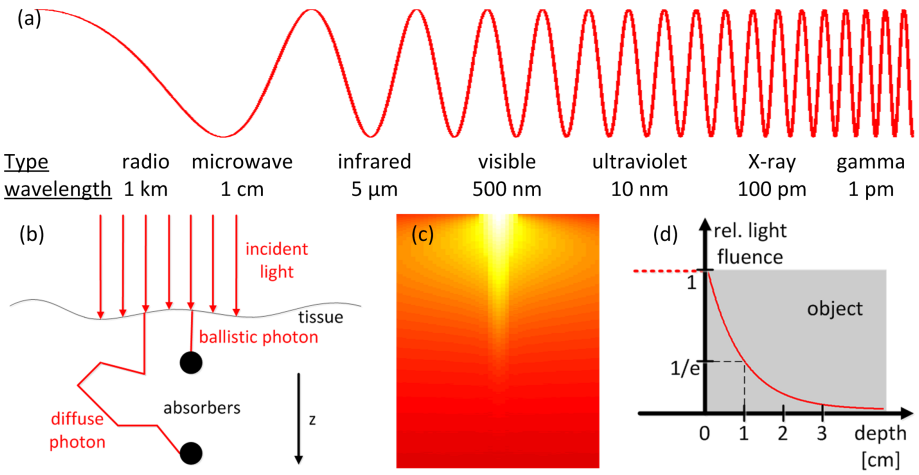

Figure 2.4: (a) Optical wavelengths as part of the electromagnetic spectrum. (b) In the ballistic transport regime, photon scattering plays only a minor role; in the diffuse regime, photons are scattered multiple times before being absorbed. (c) Illustration of light scattering when a focused beam is shed onto a scattering medium. (d) Exponential decay of light fluence for a diffusive, semi-infinite homogeneous optical medium. (b) and (d) reprinted from [14], (c) gratefully rendered by J. Turner using the MCML software package [78].

$$\mu_a(x)\Phi(x) - \nabla(D_p(x)\cdot\nabla)\Phi(x) = q_0(x), \tag{2.1}$$

where q_0 is the source term, $\Phi(x)$ is the optical fluence, i.e. the photon energy per cross section, and $D_p(x) = (3\mu_a + \mu_s')^{-1}$ is the diffusion coefficient (μ_s' is the reduced scattering coefficient deviating from μ_s for highly anisotropic scattering processes).

The 3-dimensional DE might be efficiently solved numerically using finite element methods and also analytical solutions exist, e.g. for homogeneous cylinders or semi-infinite volumes. In the latter case, the depth dependence of the fluence is given by $\Phi(z) = \Phi_0 \cdot \exp(-\mu_{eff} z)$ where $\mu_{eff} = \sqrt{\mu_a(\mu_a + \mu_s')}$ is the effective absorption coefficient, which is often in the range of a few mm^{-1} or cm^{-1} for tissue (Fig. 2.4(d)). If higher accuracy near the surface is needed, where the photons are not fully diffused yet, a higher order approximation to the DE, the δ-Eddington approximation, might be used.

In order to fully account for both ballistic and diffuse photons, Monte Carlo (MC) simulations of statistical photon random walk might be used [78, 79]. However, even when using powerful graphics processing units (GPU), MC simulations are still too time consuming for many applications.

In order to reconstruct images in DOT (or also FMT), the inverse problem of obtaining the absorption (fluorophore) distribution from the fluence values at the boundary needs to be solved. In quantitative optoacoustic tomography (often QPAT in the literature, [80]) the measured optoacoustic images consist the source term $q_0(x)$ of negative magnitude and quantitative values of the absorption coefficient are obtained by solving a light transport model for $\mu_a(x)$ as unknowns. Inverting a light propagation model is extremely challenging in

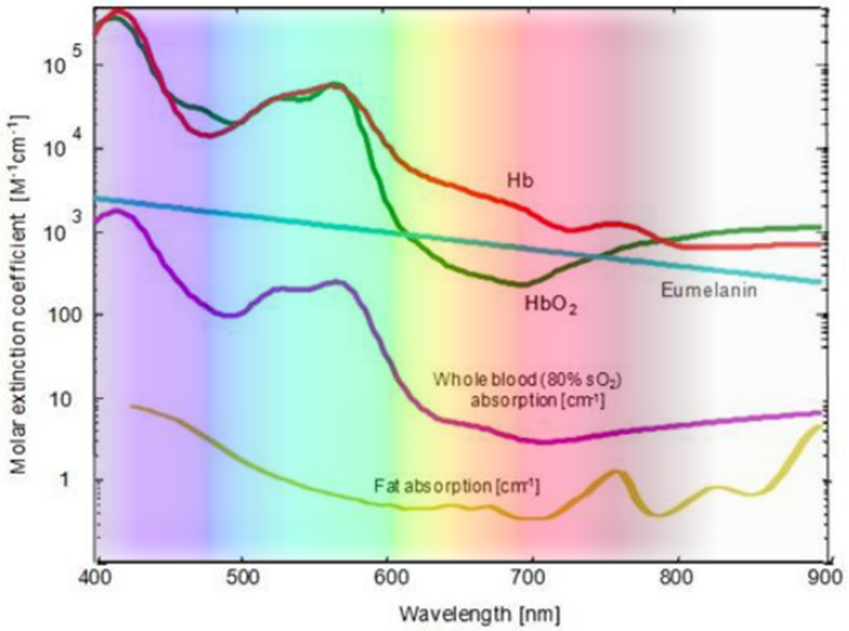

Figure 2.5: Spectra of the most dominant intrinsic tissue chromophores in the visible and near-infrared: Oxygenated and deoxygenated hemoglobin, melanin and lipids. Because of the lower absorption of hemoglobin for red and near-infrared wavelengths, light can penetrate significantly deeper into tissue, known as near-infrared window. Figure reproduced from [14].

all case for multiple reasons: (1) Although the models are linear in fluence $\Phi(x)$, they depend non-linearly on the absorption $\mu_a(x)$, requiring non-linear inversion schemes or linearized approximations. (2) The inversion consists a severely ill-posed problem (that is small relative changes in the signals might result in large relative deviations in the solution), further complicating calculations and requiring strong regularization, which limits spatial resolution. (3) The problem is non-unique as multiple solutions might exist with different scattering distributions $\mu_s(x)$; the exact scattering distributions are typically not known as well.

Despite the challenges arising from the scattering nature of tissue, deep tissue optical imaging is highly attractive to study living organisms because of the rich contrast from a variety of molecules, as described in the next section. A sufficient number of photons penetrates to considerable depths and interacts with molecules there. However, the photons are diffuse in nature and cannot be used for direct physical image generation. In the lower wavelength range of the visible spectrum, penetration depth is limited to few hundreds of μm or some mm owing to the strong background absorption of the omnipresent blood pool (Fig. 2.5). However, in the red and near-infrared (NIR) regions of the spectrum, the absorption of blood and water is comparatively low. This spectral region is known as the NIR window. Consequently, optical imaging of up to several centimeters depth is possible in the NIR.

2.6 Chromophores and Contrast Agents for Optical Imaging

The choice of targeted molecule or other source of contrast is dictated by several factors, some of which are biological and some of which are technical in nature. Yet, most importantly of course it is dictated by the biomedical questions the imaging process is supposed to answer.

The spectral properties of the target molecule play a crucial rule for the selection of excitation and/or detection wavelength(s) with implications on maximum penetration depth or the availability of laser technology or other light sources to excite signals. The spectrum of the molecule needs to be distinct to discriminate it from the background, and ideally has a pronounced peak. A large extinction coefficient ϵ allows for detection of the molecule even in small concentrations $c_i(x)$ with high sensitivity ($\mu_a(x) = \epsilon \cdot c_i(x)$). The absorption ideally depends linearly on the concentration without quenching or other nonlinear effects. For fluorescent targets, a high quantum yield is preferable (which is the fraction of absorbed energy re-emitted as fluorescent light). In OA imaging, a small quantum yield is preferred, but it is an issue of minor importance compared to the extinction coefficient.

In order to achieve a high specificity and to separate multiple target molecules (or one from the background), multi-spectral imaging is performed using several excitation wavelengths λ. In order to separate the contributions from different targets, typically a linear mixture model is assumed. All image pixels are a superposition of contributions from multiple target molecules within the pixels [14]:

$$\mu_a(x, \lambda) = \sum_i \epsilon_i(\lambda) c_i(x) \tag{2.2}$$

If the spectra of the absorption coefficients of all present molecules are known, linear un-mixing might be performed on a per-pixel basis inverting Eq. 2.2 using least-squares fitting [30]. If the spectra are not known *a priori*, blind un-mixing using principal component analysis (PCA) might be performed, concurrently resulting in the assumed target concentrations and their spectra, based on a statistical analysis of the entire images [81]. A particular challenge with un-mixing in the context of deep tissue imaging is the so-called spectral coloring: Because of the spectral dependence of the relative light fluence on strong background absorption, the optoacoustically measured spectra might considerably deviate from the physical ones. Errors in un-mixing are an unavoidable consequence, unless light propagation is accounted for or the background is modeled [82].

A large variety of possible target molecules (Fig. 2.6) is either intrinsically present, genetically encoded, or externally administered.

As most organisms are not transparent, several intrinsic chromophores present possible imaging targets. Their absorption spectra are quantum-mechanically determined in a HOMO-LUMO formulation of the molecules' states and the excitation energy decays without photon re-emission but is thermalized instead. THE most important intrinsic chromophore is hemoglobin in its oxygenated and deoxygenated forms. Their spectra are plotted in Fig. 2.5 with several isosbestic points, where both forms have the same extinction

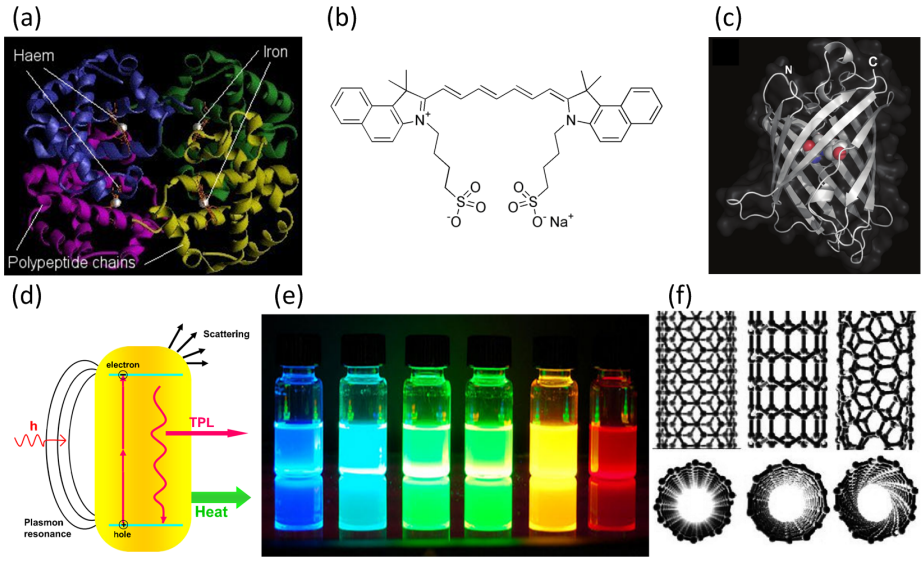

Figure 2.6: Selection of absorbers for optical imaging: (a) Hemoglobin molecule as example of an intrinsic chromophore. (b) Indocyanine green as example of a fluorescent dye. (c) Green fluorescent protein is often used as reporter gene. (d) Illustration of a gold nanoparticle. (e) Emission of different quantum dots in solution. (f) Single-walled carbon nanotubes. Figures with permission from Refs. [83] (a), [84] (c), [85] (d), [86] (e), [87] (f).

coefficient. Physiologically more relevant than the total amount of hemoglobin is its saturation level (SaO)

$$SaO = \frac{c_{HbO}}{c_{HbO} + c_{Hb}}, \tag{2.3}$$

which is the fraction of hemoglobin with oxygen bound. Spectral coloring might result in considerable error of the saturation levels of deep vasculature because of the strong blood pool background in the surrounding tissue.

Eumelanin has a strong absorption in the NIR with a relatively flat spectrum. It is highly concentrated in the skin and might limit penetration depth if absorption in the skin is too strong. Furthermore, melanin contrast facilitates optical melanoma characterization [88].

Glucose exhibits strong absorption at higher wavelengths in the NIR; thus optical imaging promises a highly attractive method for non-invasive measurements of blood glucose concentration, given a sufficient accuracy of the detection method used. At similar wavelengths, lipid absorption increases and arterial plaques in the carotids (often resulting hypoperfusion of the brain) can be detected optically [32]. Water absorption is not negligible for very high wavelengths and thus penetration depth is very limited for those long wavelengths again.

Another class of targets are small molecule organic dyes exhibiting fluorescence [89]. Fluorescence results from absorption of a photon, excitation of the molecule from the ground state, and a subsequent inter-state crossing from a

singlet state of the molecule to an excited triplet state of slightly lower energy. Based on optical selection rules the lifetime of the triplet state before spontaneously emitting a fluorescent photon is orders of magnitude larger than standard photon emission. As no fluorophore offers perfect fluorescence, i.e. quantum yield of one, a portion of the absorbed energy also thermalizes. The most important dyes are Indocyanine green (ICG) in the NIR (Fig. 2.6(b), and Sect. 9.2), methylene blue, and Prussian blue, which need to be externally administered. These dyes have been used clinically for decades and are among the limited number of FDA approved optical CAs [89].

Target absorbers might as well be genetically engineered. Genetic expression of fluorescent proteins (FP) *in vivo* was a major breakthrough in biology [84]. The first FP was the green fluorescent protein (GFP) derived from the Aequorea Victoria DNA [90], and is still one of the most important genetic tools to date (Fig. 2.6(c)). This is because it can be easily engineered in most biological model organisms. GFP has an absorption peak at around 488 nm, where light penetration depth is low, and thus GFP is most effective in microscopic imaging of small-sized organisms as bacteria. Consequently, red-shifted FPs as iRFP or mCherry [11], often GFP derivatives, with higher excitation wavelength have been developed and stable expression of FPs in the far-red or infrared is an active area of research. Genetic markers are not restricted to FPs, but might also be absorbing molecules. LacZ is a reporter gene of high biological relevance that catalyzes the reaction turning a colorless substrate X-gal into a form of bright blue color [91]. Eumelanin is not only present intrinsically, but might as well be a genetically engineered reporter system using viral vectors for its expression [88].

Nanoparticles consist another class of CAs that can be externally administered. Gold nanoparticles (Fig. 2.6(d)) exhibit strong absorption stemming from plasmonic resonances at their surface [92]. The specifics of the spectra depend on many parameters as size, aspect-ratio, or specific shape. Quantum dots (Fig. 2.6(e)) are build from semi-conducting materials and the small size in all three dimensions leads huge differences from bulky samples in the band structure and gives rise to strong photon absorption [93]. Quantum dots often exhibit fluorescence in the visible and NIR. Carbon nanotubes (Fig. 2.6(f)) are effectively 1-dimensional structures and exhibit absorption and fluorescence in the near-infrared having spectra with often multiple peaks based on van-Hove-singularities in the band structure. [87].

Given the broad variety of possible targets listed above, additional considerations have to be made for *in vivo* imaging: First, functionalization of the CA. CAs or drugs might be activatable on-site by light or US, or they might be targeting by binding to a specific receptor at the cell membrane. Functionalized CAs are often build from a compound of an antibody, a linker, and an optically active molecule. Second, the pharmakinetic behavior of CAs determines imaging performance via its half-life until its catabolized or excreted, or if it is bound to the blood pool or not. Third, toxicity of the CA and photo-bleaching limit possible applications. Last, multi-modality CAs can provide complementary image contrast in the combination of optical imaging with MRI, US, or PET.

Chapter 3

Fundamentals of Optoacoustic Imaging

3.1 Introduction

Optical imaging methods are very powerful tools for biomedical imaging, as presented in the previous chapter. The main advantages are their high spatio-temporal resolution, the excellent contrast from a large variety of intrinsic molecules which can be further extended using administered CAs, the cheap and mature technology available, and the non-ionizing nature of the radiation involved. However, photon propagation in biological tissues is dominated by scattering processes and thus high-resolution, purely optical imaging is restricted to less than 1 mm depth, mostly some hundreds of μm. However, light penetrates up to a few cm depth in the NIR and deep tissue optical imaging is possible nevertheless. Yet, spatial resolution is degraded by several orders of magnitude and is often in the range of 1 mm or worse.

OA imaging shatters the barrier of low spatial resolution in deep tissue optical imaging capitalizing on the OA effect [14, 94]: Transient pulses of incident light are absorbed and the absorbed energy is thermalized; the local temperature increase gives rise to propagating pressure waves via thermoelastic expansion. OA imaging is an intrinsically hybrid imaging modality: The excitation is performed optically while the read-out is performed ultrasonically. It combines the advantages of both kinds of modalities: It preserves the rich photon contrast of optical imaging and its spectroscopic capabilities to distinguish different molecules; however, the strong light scattering does not degrade spatial resolution because high-resolution is provided by means of US waves, which are affected two orders of magnitude less by (ultrasonic) scattering compared to photons [14]. Spatial resolution then mainly depends on the frequency content of the ultrasonic waves generated and detected. Much like the excellent spatio-temporal resolution in US imaging, OA imaging may provide video-rate volumetric images at spatial resolutions of hundreds of μm from cm depth. The two main disadvantages of the two single modalities, namely low spatial resolution in optical imaging and poor contrast in US, are thus circumvented by means of the OA effect.

Because of the beneficial combination of light and sound, OA imaging is

able to provide anatomical, functional, and molecular information from a single modality, scalable in dimension from organelles to organs [13]. For example, the complementing anatomical and functional information could visualize tumor margins and identify hypoxic areas in the tumor based on oxygenation level maps [33]. In most applications of clinical and preclinical practice, OA imaging is expected to complement existing imaging modalities rather than to fully replace them. Although many more scenarios can be thought of, the three main fields of expected application are: (1) Cancer imaging either to understand fundamental biological mechanisms of tumor growth and therapy using anatomical, functional, and molecular information; or to clinically diagnose tumors in breast or colorectal cancer [33, 95]. (2) Brain imaging to visualize structure and its pathology like detecting stroke or neuronal activity in normal and diseased state as in epilepsy [34, 96]. (3) Cardiac imaging and diseases related to the vasculature as arterial plaques, arthritis, occlusions, varicose veins, and many more [32, 52].

The hybrid nature of OA imaging opens not only the door for many new fields of imaging applications, but also requires special considerations concerning the mechanism involved in the OA effect, concerning instrumentation used for excitation and detection, and concerning the calculation of the reconstructed images from the measured signals. These aspects are discussed in the following sections of the chapter. The last section describes in detail the specific embodiment of a small animal OA scanner that was used for the experimental studies in this thesis.

3.2 The Optoacoustic Effect and the Propagation of Sound

OA imaging roots on the OA effect[1], also known as thermoacoustic or photoacoustic effect (Fig. 3.1): Transient heat $H(x)$, where $H(x)$ is the deposited energy per unit volume in the tissue, leads a local temperature increase $\Delta T(x)$. Thermoelastic expansion results in mechanical stress which subsequently propagates as pressure waves in the material [94, 97]. In most cases, the deposited heat results from the local absorption of light:

$$H(x) = \Phi(x) \cdot \mu_a(x). \tag{3.1}$$

The fractional volume expansion of the tissue is given by [94]:

$$\Delta V/V = -\kappa \Delta p_0 + \beta \Delta T, \tag{3.2}$$

where κ is the isothermal compressibility, Δp_0 is the pressure increase, and β is the thermal volume expansion coefficient. For short excitation pulses of duration τ_{ex}, the conditions of stress confinement ($\Delta X >> c \cdot \tau_{ex}$, where ΔX is pixel size and c is the speed of sound, SOS) and thermal confinement ($\Delta X >> 4 D_t \tau_{ex}$, where D_t is the thermal diffusivity) are fulfilled. Under these conditions, fractional volume expansion is negligible:

$$\Delta p_0 = \beta/\kappa \cdot \Delta T. \tag{3.3}$$

[1] Substantial parts of this and the following chapter are based on work previously published as Ref. [14].

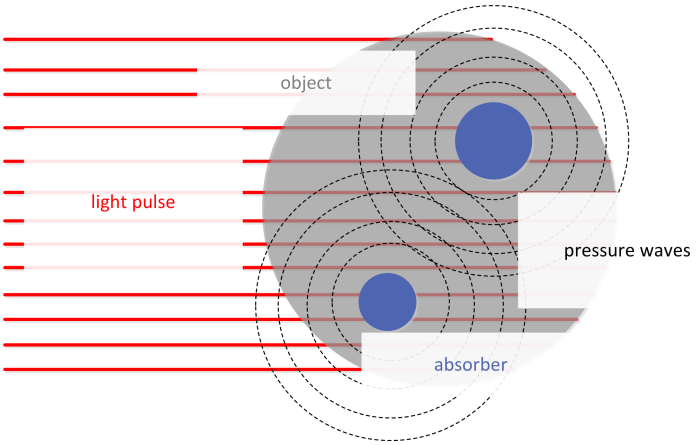

Figure 3.1: Illustration of the OA effect: Pulsed laser light is shed onto the object. When photons get absorbed within, the energy is thermalized and gives rise to an initial pressure propagating as US waves from the site of absorption. Figure reprinted from [14].

For typical tissues, temperature rises in the mK range result a pressure rise in the order of 1000 Pa. If all absorbed energy is instantly thermalized, the local initial pressure $p_0(x)$ (Δ is omitted for simplification of notation) is related to the deposited energy density via [14]

$$p_0(x) = \Gamma \cdot H(x) = \frac{\beta c^2}{C_p} H(x), \tag{3.4}$$

where Γ is the dimensionless Grüneisen parameter and C_p is the isobaric specific heat capacity. Although being different for different kinds of tissue, the spatial variations of the Grüneisen parameter are typically considered small compared to those of the deposited energy (and those of the absorption coefficient) and thus assumed to be constant and known. However, one might consider the spatial variations of Γ as complementary source of image contrast and aim at retrieving a map of $\Gamma(x)$ in QPAT [80]. Or alternatively, one might exploit the strong temperature dependance of Γ in order to either non-invasively probe the local temperature in tissue [98] or to obtain super-resolution images based on Γ-related non-linear effects [99].

The initial pressure $p_0(x)$ generated at the instant of the light pulse incidence (t=0) is spatially varying because of the spatial dependance of $H(x)$. Consequently pressure waves start to propagate. Soft tissue is typically modeled as a liquid where no shear waves propagate but only longitudinal waves. The amplitude of those waves is assumed to be in the linear propagation regime.

Pressure propagation is governed by the spatially heterogeneous wave equation [14]:

$$\nabla^2 p(x,t) - \frac{1}{c(x)^2} \frac{\partial^2}{\partial t^2} p(x,t) = -\frac{\beta}{C_p} H(x) \delta'(t), \tag{3.5}$$

where the inhomogeneous source term results from the excitation. Omitting the spatial variations of SOS ($c(x) \equiv c$), the propagation results an initial value

problem known as OA wave equation:

$$\begin{aligned} \nabla^2 p(x,t) - \frac{1}{c(x)^2}\frac{\partial^2}{\partial t^2}p(x,t) &= 0 && x \in \Re^3, t \geq 0 \\ p(x,t=0) &= p_0(x) \\ \partial_t p(x,t=0) &= 0 \\ p_m(x_d,t) &= p(x_d,t) && x_d \in S, t > 0. \end{aligned} \tag{3.6}$$

S is a (closed) surface surrounding the object on which the pressure signals $p_m(x_d,t)$ are recorded. In ideal media, the pressure $p(x,t)$ at two different spatio-temporal points (x,t) and (x',t') can be linked by a Green's function:

$$G(x,x',t,t') = \frac{\delta((t-t') - |x-x'|/c)}{|x-x'|}. \tag{3.7}$$

Substituting G into Eq. 3.6, the recorded pressure signals $p_m(x_d,t)$ on S can be explicitly stated via a Poisson-type integral:

$$p_m(x_d,t) = \frac{1}{4\pi c}\frac{\partial}{\partial t}\int dA \frac{p_0(x)}{|x-x_d|}\bigg|_{ct=|x-x_d|}. \tag{3.8}$$

The pressure signal corresponding to one specific detector location x_d is a so-called projection and the ensemble of all projections is the sinogram. SOS is most often assumed to be constant, mainly for the lack of mathematical tools to treat the heterogeneous case efficiently. In reality, the SOS varies for different types of soft tissues in a range of typically 1400 m/s - 1700 m/s [100]. Most often, c is associated with the SOS in water (1500 m/s at 25 °C) but more thorough considerations might improve the resulting OA image quality (see also Sect. 7.4 and Sect. 7.5). Similar to electromagnetic waves, acoustic waves are also absorbed. In order to account for acoustic absorption, a dissipative term is added to Eq. 3.5. The dissipative term is often modeled as fractional Laplacian, leading to power law absorption with increasing attenuation for higher ultrasonic frequencies [101, 102]. For macroscopic imaging systems, ultrasonic attenuation can be neglected in most cases. For mesoscopic and microscopic imaging applications which involve higher US frequencies, image quality might however be considerably decreased. Finally, heterogeneities might alter wave propagation, too: On a macroscopic level, mismatches of US impedance lead to an aberration and reflections at e.g. skull, bones, lungs, or organ boundaries; on a microscopic sub-wavelength level, local acoustic mismatches result in scattering of the ultrasonic waves. US scattering of phonons is nevertheless orders of magnitude lower compared to scattering of photons in tissue [13].

3.3 Optoacoustic Imaging Systems: Signal Excitation and Detection

The OA effect provides means to excite signals inside an object which can be subsequently detected outside of the object. In practice many approaches for excitation and detection have been developed that permit obtaining images with different characteristics and for different applications, depending on the instrumentation used.

Signals are mostly excited with short light pulses of high power, approximated as δ-pulses that concurrently excite all US frequencies equally. Alternatively, amplitude modulated light of frequency ω_{ex} can be used in frequency domain imaging, where the detected US frequency ω_{US} is equal to the excitation modulation frequency ($\omega_{ex} = \omega_{US}$). By measurements at multiple frequencies or chirp excitation, the full US spectrum of pressure waves can be obtained [103, 104]. Frequency domain imaging is attractive for the possibility to use cheap hardware with continuous-wave lasers and quadrature-detection signal recording.

Although light is most often used, not only the optical part of the electromagnetic spectrum can be used for excitation of OA waves. Microwaves or magnetic excitation of paramagnetic nanoparticles present alternative approaches of considerably higher penetration depth compared to light [105, 106]. Increased penetration depth of OA excitation is also possible with high energy X-ray photons [107]. They are however ionizing in nature. Generalized forms of local energy deposition might be used as well: In particular, the energy deposition of protons at the Bragg peak might be localized via the thermoacoustic effect [108].

Q-switched pulsed lasers are the main type of source to create monochromatic light for OA imaging with different active materials such as solid state Nd:YAG, dyes, or semiconductors [109]. If the provided wavelength of the laser does not match the spectrum of the targeted chromophore or multi-wavelengths imaging is performed in MSOT, an additional optical parametric oscillator (OPO) allows for tuning to arbitrary wavelength over a wide spectral range. Important laser parameters for imaging are the maximum pulse repetition rate, the minimal pulse duration τ_{ex}, and the beam shape quality if light needs to be focused onto the sample. The per-pulse energy of the lasers employed is typically in the order of tens of mJ for macroscopic imaging and of several μJ only for microscopic imaging, based on the area illuminated.

The illumination pattern on the surface varies for different imaging approaches (Fig. 3.2): In most tomographic systems, a broad illumination with maximally homogeneous light distribution is used to minimize fluence related effects. In OA microscopy, a tightly focused spot provides a beam with diffraction limited lateral resolution or in sectional imaging the origination of the signals can be restricted to approximately a plane, similar to SPIM. In both cases, imaging depth (without degradation of focusing) is limited to superficial areas owing to the strong scattering of light. Structured illumination is used to solve the ill-posed inverse problem of light propagation modeling taking measurements with different illumination patterns or in the context of compressed sensing (CS) measurements. For *in vivo* imaging, the maximum fluence is restricted by the ANSI limits to $20\,\mathrm{mJ/cm^2}$ in the NIR in order to prevent thermal damages [110]. Depending on the illumination pattern, different optical components such as diffusers, polarizers, (di-chroic) mirrors, (GRIN-)lenses, digital mirror devices, etc. are used. Free-beam optics are convenient in building prototypes in the laboratory; but in clinical systems, light is shed onto the object using light guides, single- or multi-mode fibers, or fiber bundles [111]. Handheld OA imaging is particularly based on the latter. Although light can be delivered from arbitrary directions, trans-illumination (illumination from the opposite side as the detection) and epi-illumination (from the same side as the detection) are the most prevalent configurations because of practical considerations. Many large

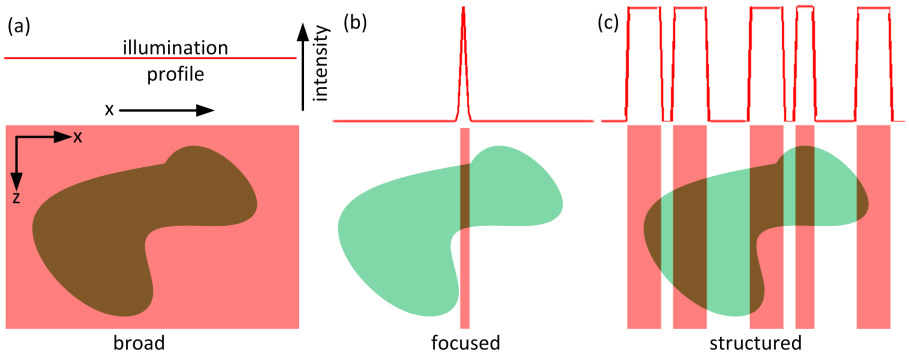

Figure 3.2: Illustration of typical illumination scenarios in optoacoustic imaging. (a) Broad illumination is used in most tomographic setups with an almost homogeneous light profile shed on the sample in order to minimize effects of light fluence. (b) Focused illumination is used in OA microscopy or selective plane imaging in order to restrict signal generation to a line or a plane. (c) Structured illumination is used with CS approaches or multi-illumination tomography in order to accelerate data acquisition or to reduce the ill-posed nature of inverting the light propagation model. In all illustrations, scattering and absorption within the object are not accounted for.

imaging objects are often only accessible from one side (e.g. in breast or skin imaging), therefore only epi-illumination is feasible. In this case, the acoustic read-out permits potentially up to double the imaging depth compared to purely optical imaging modalities with diffuse light, where the read-out is also based on photons.

After conversion of light energy to mechanical energy, the propagating US waves are measured on a detection surface surrounding the object (Fig. 3.3, top row). The temporally resolved signals are measured at multiple locations on the surface S, sampled as densely as possible. Distortions from the propagation between the object and the detection surface have to be minimized and thus water, ultrasonic coupling gel, or alternatively a solid couplant is used. The detection surface S is convex, closed and preferably of mathematically ideal shape (Fig. 3.3(a)-(c)): Either a circle in 2-D and its 3-D counterparts, a sphere or cylinder [30]; or a line in 2-D and a planar surface in 3-D [112]. More sophisticated detection patterns can improve image quality in certain experimental set-ups [113]. Similarly as for light delivery, large objects allow only restricted access for detection. Detectors can only be placed on one side of the object then and only an incomplete part of the OA waves is captured. Incomplete signal detection is known as limited view. Limited view with detection surfaces covering $<2\pi$ ($<4\pi$ in 3-D) generally results in degraded image quality compared to full view because of the lack of information on parts of the signals; artifacts appear due to the increased condition number of the reconstruction problem [114]. A large numerical aperture (NA) is thus preferred because of visibility criteria [115]: With planar sensors, their finite size results a limited NA and structures (almost) perpendicular to the detector do not emit waves that are captured by the detectors. Thus such structure cannot be reconstructed. In order to recover them, the effective NA needs to be increased, either by enlarging the surface or

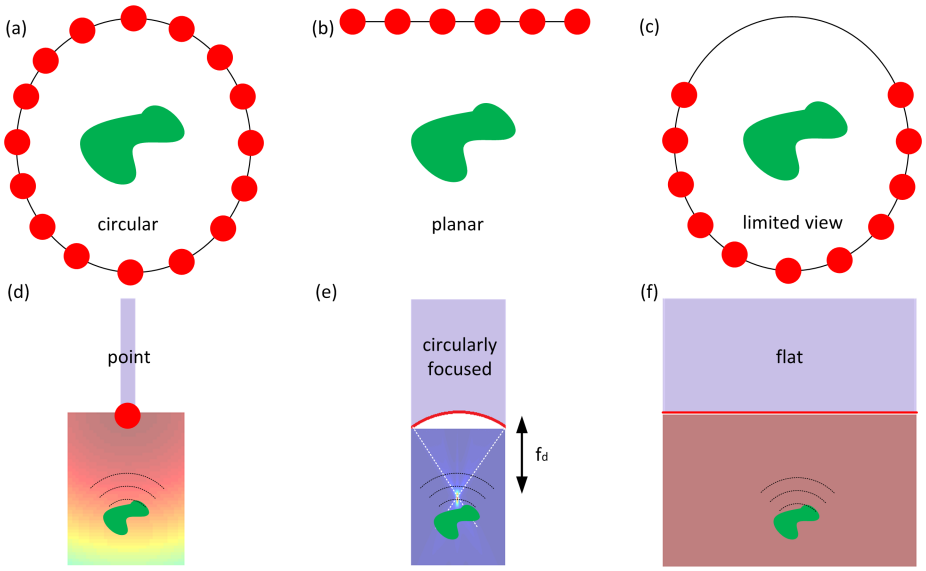

Figure 3.3: Detection surfaces and detector shapes in OA imaging. (a) Detectors (red) are evenly distributed on a circle fully surrounding the object. (b) Detection locations distributed on a line as used in OA microscopy, known as planar detection. In 3-D imaging, combinations of the two are used: Spherical, cylindrical, and planar. (c) Because of technical considerations in experimental practice, signals can often only be measured on parts of the circle, known as limit view. (d) Point-like detection element (red) with spatial dimensions smaller than characteristic US wavelengths. (e) Circularly focused detection element with focal distance f_d. (f) Flat detection element with spatial dimensions larger than characteristic US wavelengths. Illustration of the sensitivity field of the different detection elements as background. In 3-D imaging, combinations of the shapes are used: Point-like, spherically focused, cylindrically focused, line detection, and flat detection.

using acoustic reflectors [116].

The shape of the sensor element(s) is ideally point-like with dimensions much smaller than the wavelengths to be detected (Fig. 3.3(d)-(f)). Most reconstruction methods assume point-like detectors for simplicity of calculations (see also Chapter 4 and Sect. 7.3). In practice however, all detection elements are finite-sized because of the SNR decreasing with shrinkage of the detector dimensions and the technical limitations in miniaturizing sensor elements. Large spherically or cylindrically focused elements place a viable alternative [59]. Unlike point-like elements with a wide acceptance angle, the received signals originate in first order approximation only from a line or a plane and the reconstruction problem is thus reduced in its spatial dimensions. However, limited focusing capabilities for broadband signals result in a much lower lateral than axial resolution and artifacts from detected signals outside of the focal area appear, unless computationally accounted for [117]. Finally, huge line or planar sensors are attractive for their high SNR, their easy compatibility with interferometric detection, and the simple inversion based on the inverse Radon transformation [118].

The requirements of OA detection technology differ from those of standard US technology. OA transducers are not operated in transmit/receive mode. Thus OA sensors require only efficiency in detection of the low intensity signals. Most importantly, OA signals are broadband spanning many octaves from tens of kHz to several MHz in the same experiment as direct consequence of the imaged structures at multiple scales. Besides a suitable central frequency, detectors employed consequently need to have a large fractional bandwidth in order to avoid frequency bandwidth related artifacts (see also Sect. 7.2). Commercially available or customized standard piezoelectric ultrasonic technology is most often used [45]. Despite being inexpensive and robust for operation in water environments, piezoelectric technology based on PCT or polymers offers either not an optimal bandwidth or not an optimal sensitivity, respectively. Recent developments in the field of capacitive micro-machined ultrasonic transducers (CMUT) facilitate miniaturization at high bandwidth and sensitivity [119]. Particularly interesting for OA are optical sensing methods. With exception of beam deflection methods, optical sensing is based on interferometry: The effective optical path length is changed by the incident US wave either by changes in length or changes in refractive index. The intensity of the interference pattern from different paths is thus modulated by the incident US waves. Particular embodiments of interferometric detectors are Fabry-Pérot interferometers [112], Mach-Zehnder interferometers [120], micro-ring resonators [121], or fiber Bragg-gratings [122]. Interferometric detectors can provide unrivaled detection bandwidth and offer a huge potential for miniaturization.

Single element detectors have been employed in early OA systems and still are for high frequency applications [22, 123]. Time consuming mechanical raster or continuous scanning is required in order to obtain all projections. Conversely, off-shelf or custom-made parallel detection arrays have enabled real-time imaging at video-rate from single laser shots [124]. Scanning is in this case only required to capture entire volumes from slices or to improve image quality from more available projections [18, 42].

The measured electrical signals are further processed by a data acquisition system (DAQ): The DAQ pre-amplifies the weak signals in the μV range, applies anti-aliasing frequency filtering, and digitized the signals in parallel using analog-digital converters. The sampling frequency is typically much higher than the bandwidth of the transducer and a 12-bit or more vertical resolution is used to capture the dynamic range of the signals without discretization artifacts. In combination with large detection arrays, digitization can be done fully in parallel or using multi-plexing at the price of longer acquisition times (averaging of signals for a better SNR results also in increased acquisition times). The data acquisition is triggered with firing of the laser and the recorded signals are sent to a PC for storage and further on- or offline processing.

With the listed technical possibilities at hand and application scenarios in mind, several standard types of OA systems have been developed:

- Optical resolution optoacoustic microscopes compromise low power, ultrahigh repetition rate lasers that are focused through a microscope objective onto the sample to lead optically diffraction limited lateral resolution in ballistic depths. Signals are captured by a large area high frequency transducer element and images are generated by scanning the excitation beam without mechanically moving parts. Optical resolution

systems have been applied in imaging of red blood cells, superficial brain vasculature, and many more [16].

- Acoustic resolution optoacoustic microscopes provide acoustically diffraction limited spatial resolution by using a large NA spherically focused detector. Illumination is provided by means of broad or dark-field illumination with diffused photons. Through the greater imaging depth compared to ballistic photons only, several studies reported on imaging of deep brain vasculature, skin, tumors, and many more [125].

- Optoacoustic tomography (OAT) systems compromise high power, low repetition rate lasers of tunable wavelength with broad illumination for imaging at mm or cm depth. Signals are captured with parallel curved arrays on an arc or a sphere and images are formed computationally using reconstruction approaches (see next chapter). OAT has found many applications in whole-body small animal imaging, including cancer, brain and cardiovascular disease models as well as in clinical studies [22, 33, 42, 126]. Variants of OAT use standard ultrasound arrays with a planar detection surface instead of a custom-made curved array [45, 113].

- Endoscopic or intravascular systems enhance the portfolio of imaging locations to regions that are not accessible from the outside but from cavities inside of the body. Main challenges arise from the need to miniaturize light delivery and signal detection, thus favoring fiber optical systems. Furthermore, reconstructions are complicate as signals are measured form within the object instead of from the outside. Anticipated fields include colorectal cancer, arterial plaques, and many other cardiovascular diseases [32].

3.4 Small Animal Multi-Spectral Optoacoustic Tomography Scanner

In the past decade, a variety of OAT imaging systems have been developed and reported on for both small animal preclinical imaging and clinical imaging for humans [42, 95, 127, 128]. The improved OAT small animal scanner technology developed at the Institute of Biological and Medical, particularly the MSOT technology, has been commercialized. For the OA imaging performed in the context of this thesis, a commercially available small animal MSOT scanner was used (model: inVisionTF256, iThera Medical GmbH, Munich, Germany).

The scanner is designed as a compact, movable cart system (Fig. 3.4(a)), where all components are mounted in a single rack. Imaging is performed within a tank build from stainless steel (Fig. 3.4(b)). The tank is filled with warm distilled water and accessible through a tray from the top. The key component for OA image generation is the integrated illumination-detection unit (Fig. 3.4(c)).

A schematic of the system components and the workflow during imaging is shown in Fig. 3.5. Image acquisition is controlled by a computer running the proprietary ViewMSOT software. The software allows for convenient image acquisition using a graphical user interface in real-time. The acquired raw signals (approximately 10 MB/s) are dumped to a hard drive for storage including

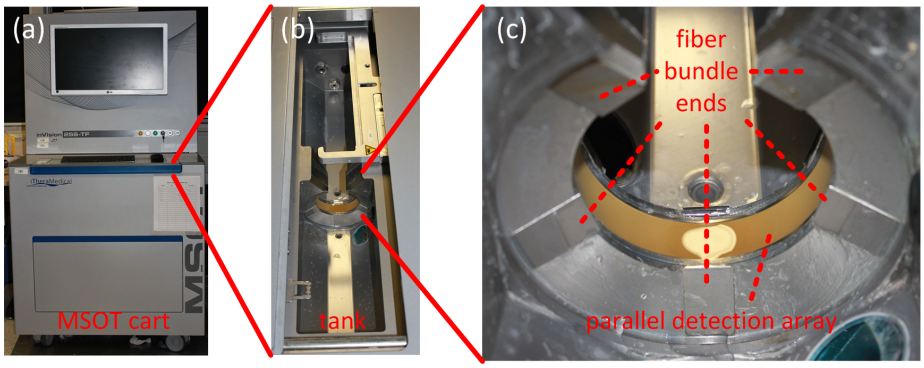

Figure 3.4: Small animal scanner mainly used in this work. (a) Photograph of the MSOT cart system (model: inVisionTF256, iThera Medical GmbH, Munich, Germany) with all components mounted in a rack. (b) Photograph of the imaging chamber with tank to be filled with water, illumination-detection unit and translation stage to move the object. (c) Photograph of the illumination-detection unit with the 10 fiber-bundle ends to provide an approximately homogeneous illumination on the object's surface and the cylindrically focused detection array to confine the imaged volume to approximately a plane.

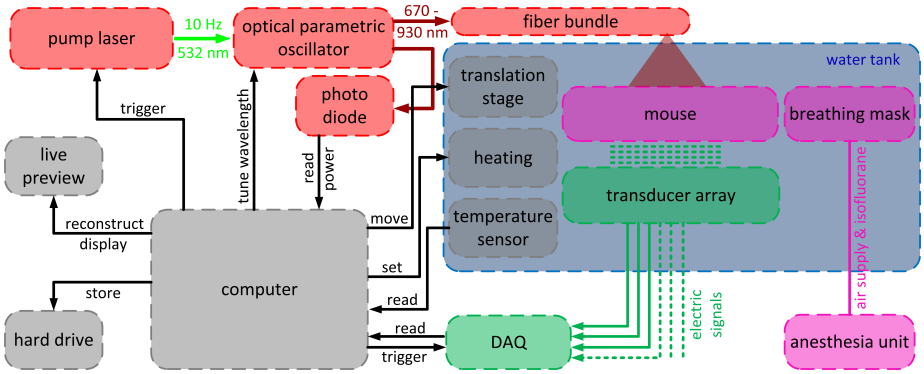

Figure 3.5: Schematics of the interplay between hardware components of the MSOT system: The computer triggers the firing of a pulsed laser and the wavelength is modified to the target wavelength via an OPO. Passing a fiber bundle, the light is shed onto the imaging object. Ultrasonic pressure waves propagate through object and water and are captured by the transducer array. The electrical signals are digitized by a DAQ and stored in the computer. Additional components include live preview, water temperature control, and positioning of the object.

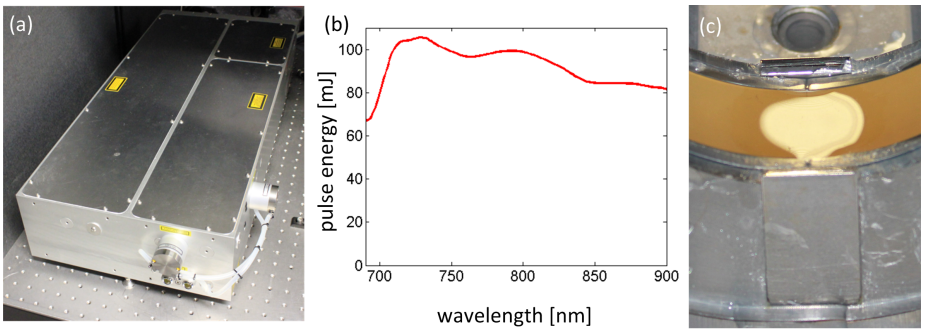

Figure 3.6: Optical excitation components. (a) Photograph of the pulsed OPO laser with fast wavelength tuning capabilities. (b) Spectrum of the average pulse energy of the OPO laser in the NIR. (c) Photograph of two of the ten fiber bundle ends to provide an approximately homogeneous illumination pattern on the surface of the object.

meta-data of the image acquisition process. The software provides a live preview from backprojection reconstructions (see Sect. 4.1) in order to chose an imaging region, identify image artifacts, and monitor the animal well-fare. The imaging region can be chosen by moving the object in all three spatial dimensions with a linear translation stage (IAI Industrieroboter GmbH, Schwalbach, Germany). Translation in z-direction allows for selection of the cross-sectional imaging plane with sub-millimeter accuracy. A volumetric image of a whole mouse or entire organs can be obtained from z-scanning and stacking of the reconstructions from multiple slices. In-plane translation in x- and y-direction permits to center the object for minimization of light fluence heterogeneity and maximal image quality in the focus of the transducers. The temperature of the water bath in the tank is controlled to ensure controlled, normal physiological conditions of the mouse at approximately 34 °C water temperature. The water temperature is increased by a heating pad attached to the tank wall and controlled by a temperature sensor and a feedback loop. Thus, the temperature can be stabilized within $<$0.2 K.

Optical signal excitation is provided by a pulsed, tunable OPO laser (model: SpitLight, InnoLas Laser GmbH, Krailling, Germany, Fig. 3.6(a)). The Nd:YAG pump laser provides short pulses ($<$10 ns) at 532 nm wavelength and 10 Hz pulse repetition rate. The pulses are shed on an OPO crystal which exhibits strong non-linear optical effects and thus allows altering the resulting wavelength. The OPO unit allows tuning of the wavelength in a range from 670 nm to 930 nm on a per-pulse basis, i.e. it allows for fast MSOT imaging without idle times from changing wavelengths. A small fraction of the pulse is shed onto a photo diode to measure the per-shot energy variations in order to eliminate image intensity fluctuations or spectral artifacts through normalization. Fig. 3.6(b) plots the resulting, wavelength dependent pulse energy with a maximum of approximately 100 mJ at the output of the laser. The light is coupled into a fiber bundle with ten output arms. The arms are equally distributed on both sides of the illumination-detection unit (Fig. 3.6(c)) to provide an approximately homogenous light fluence on the surface of the object. For an object of 2 cm diameter, the illumination stripe is ca. 8 mm wide and the light fluence does

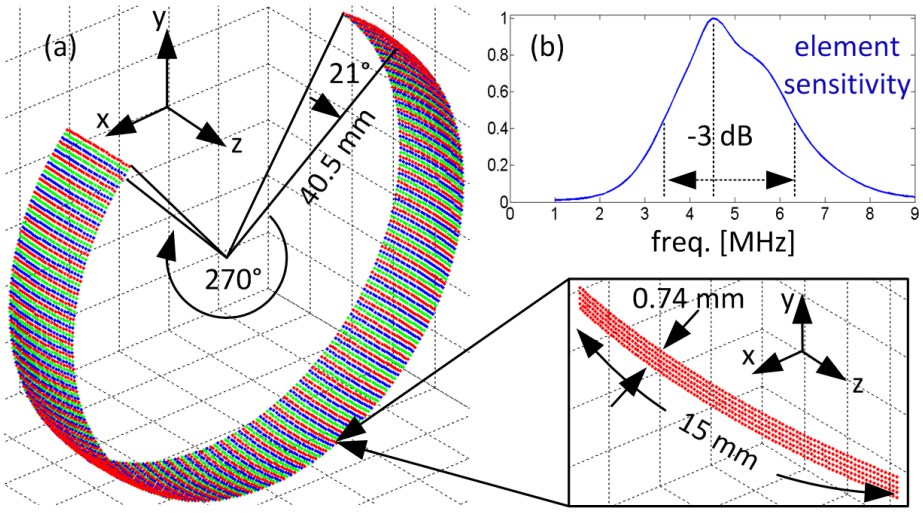

Figure 3.7: Parallel US detection array. (a) Sketch of the transducer geometry with the 256 individual elements plotted in alternating color. The array is cylindrically focused to confine received signals to approximately a plane. It coveres a solid angle of 270° and has a radius of 40.5 mm. (b) Element sensitivity in dependance of frequency with a center frequency of 4.2 MHz and lower and upper limits (-3 dB) of 3.5 MHz and 6.2 MHz, respectively.

not exceed 18 mJ/cm^2, which is below the maximum permissible exposure limit of 20 mJ/cm^2 in humans [110].

Detection of the resulting OA pressure waves is performed using a custom-made parallel detection array (Imasonic SaS, Voray, France). The circular detection array is cylindrically focused, covers a solid angle of 270°, and has a radius of 40.5 mm (Fig. 3.7(a)). It covers 21 ° in elevation direction to restrict the detected signals to originate from approximately a plane. The 256 individual elements are manufactured from a piezocomposite material. Each element has a central frequency of 4.2 MHz and a lower and upper frequency limit (-3 dB) of 3.5 MHz and 6.2 MHz, respectively (Fig. 3.7(b)). The relative bandwidth is 64 % and the transducer sensitivity result electrical signals in the μV range for OA signals of some Pa. The possible in-plane resolution is as good as 200 μm. The z-resolution is object frequency dependent, but in the order of 1 mm. A high detection sensitivity (FWHM) is achieved over a region with 3 cm diameter [129]. The detection geometry allows for a large visibility region [115] and singularities in all directions can be recovered. Consequently, a nearly isotropic resolution over a large area is achieved.

The electrical signals from the transducer are digitized by a custom-made DAQ system (Falkenstein Microsysteme GmbH, Taufkirchen, Germany). 256 parallel input channel prevent the need for multi-plexing. The low voltage electrical signals are pre-amplified, frequency filtered by a 15 MHz anti-aliasing filter, and sampled at 40 MSa/s with 12-bit vertical resolution.

With this small animal MSOT scanner, *in vivo* longitudinal mouse imaging studies are possible with high throughput. The achievable spatial in-plane resolution is up to 200 μm for centimeter-sized object. Whole image frames can

be acquired from a single laser shot at video-rate and fast spectroscopic imaging in the NIR is possible, also for fast kinetic processes. Volumetric scanning allows to acquire MSOT images of a whole mouse or entire organs. Several pre-clinical small animal imaging studies have been conducted with this scanner, ranging from gastric emptying to *in utero* imaging of the mouse embryo heartbeat [21, 130].

For parts of the experiments, a second scanner was used (model: inVisionTF128, iThera Medical GmbH, Munich, Germany). The sytem is similar, but there are two major differences: A deviating OPO laser is employed (Phocus, Opotec Inc., Carlsbad, CA) that has the same pulse repetition rate and a similar power spectrum but lacks fast wavelength tuning capabilities. Second, a different transducer array is used that has 128 detection elements only and thus a slightly decreased in-plane lateral resolution.

Chapter 4

Reconstruction Approaches for Optoacoustic Tomography

This chapter introduces to different types of reconstruction approaches with focus on model-based (MB) reconstructions in the context of (pre-)clinical imaging. Specific challenges arise from reconstruction time and image quality constraints, which are also discussed in detail.

4.1 Types of Reconstruction Algorithms

Measured OA signals might directly result spatially resolved images: In the case of OA microscopy raw images are basically stacked projections. Images might be improved using post-processing algorithms, performing reconstructions is however not mandatory. The converse is the case in OAT. The measured signals do not directly correspond to spatially localized image structures. Spatially resolved images need to be computationally rendered from the signals in a reconstruction process.

Reconstructions in OAT are one example of an inverse problem: A cause determines an effect based on a set of known physical laws. Predictions of effects based on known causes can be easily made using the governing equations. These calculations are termed forward problem or forward modeling. If only the measured signals are known, an inverse problem arises, which is to find the corresponding causes based on the same physical model. Most often, an inverse problem is much harder to solve than the associated forward problem.

The inverse problem in OAT is to find the spatially resolved initial pressure distribution $p_0(x)$ from the measured acoustic signals $p_m(x_d, t)$. The governing equations are given by Eq. 3.6 and an inverse initial value problem has to be solved. For quantification of OA reconstructions, a second inverse problem has to be solved, which is the inversion of a light transport model. The associated problem is known as quantitative OAT (or also QPAT) and is briefly discussed in Sect. 2.5 or in detail in the literature [14, 77, 80].

Multiple reconstruction approaches have been developed in order to solve the acoustic inverse problem, starting from 'classical' X-CT Radon transformation based methods [97]. The approaches can be categorized in mainly four different classes [14]. In practice, the selection criteria of which class to use in-

clude computational aspects and especially assumptions made such as detection geometry or propagation medium. In experiments, the actual forward problem can deviate from the formulation of Eq. 3.6 and the chosen type of algorithm needs to be adaptable to the real experimental conditions to retain a good image quality. Deviations from the idealized case include the spatial impulse response (SIR) of finite-sized transducers, the electrical impulse response (EIR) of limited-bandwidth transducers, limited detection coverage, a heterogeneous SOS distribution, and acoustic wave dissipation and scattering.

Series Expansion

One class of reconstruction approaches is the so-called series expansion method. The solution of the inverse problem is represented as a converging series of summed basis functions:

$$p_0(x) = \sum_k \psi_k \cdot \Psi_k(x). \tag{4.1}$$

The basis functions $\Psi_k(x)$ are the Eigenfunctions of the Laplacian in Eq. 3.5; the associated coefficients ψ_k are integral-functionals depending on the measured signals p_m. The set of Eigenfunctions is only explicitly known for a few ideal geometries (sphere, cylinder, infinite plane, cube) and an ideal dispersion relation $\omega^2 = c^2(k_x^2 + k_y^2 + k_z^2)$. Therefore, series solutions are restricted to idealized settings. Additional challenges are associated with a potentially large number of series coefficients needed for sufficient convergence or with the numerical integration of functions with singularities.

In experiments, series expansion reconstruction methods are mostly used with planar sensors: To perform reconstructions, the measured signals $p_m(x, y, t)$ are Fourier transformed in all variables $p_m(k_x, k_y, \omega) = F_{x,y,t}\{p_m(x, y, t)\}$. The transformed signals are mapped to the initial pressure $p_0(k_x, k_y, k_z)$ using the dispersion relation $\omega^2 = c^2(k_x^2 + k_y^2 + k_z^2)$. The sought-after initial pressure is then obtained by inverse Fourier transformation in all variables $p_0(x, y, z) = F_{x,y,z}^{-1}\{p_0(k_x, k_y, k_z)\}$. The algorithm is simple to implement, fast in practice, and offers an optimal computational complexity of $O(n^3 \cdot log(n))$ in 3-D and $O(n^2 \cdot log(n))$ in 2-D, where n is the number of image voxels per spatial dimension (assuming time-samples and number of projections per dimension of the detection surface being proportional to n).

Backprojection

Based on the analogy with the Radon transformation of X-CT imaging [97], backprojection (BP) type reconstruction formulas have been developed [131, 132]. First, the measured signals are frequency filtered in the temporal dimension and subsequently backprojected to all locations of possible origination (Fig. 4.1(a)). Based on TOF considerations for a homogeneous medium, the locations are distributed on an arc of radius $r = ct$ centered around the detector location of the currently processed projection. The final image is then obtained by superposition of backprojected signals for all time-instants and projections. The universal BP formula is the BP formula most often used in practice [132]:

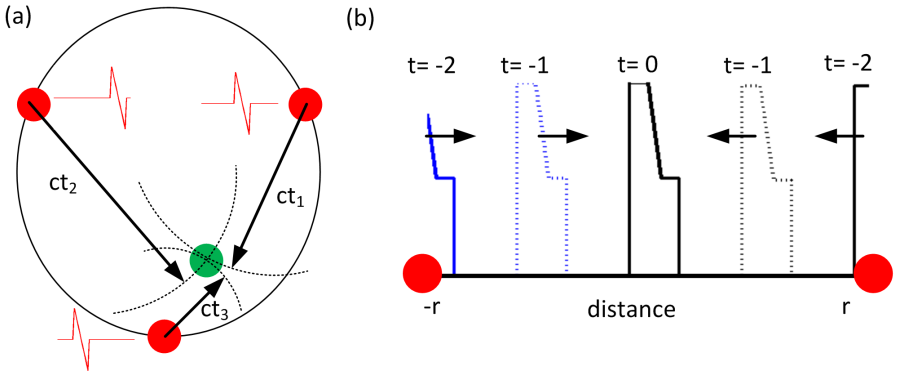

Figure 4.1: (a) Backprojection reconstruction: Signals from a specific detector and specific time-instant t_1 are frequency filtered and backprojected on an arc of radius $r = ct_1$ centered around the detector. A superposition for all detectors and time-instants results the final reconstructed image. (b) Illustration of time reversal reconstruction (1-D): The measured signals are reversed in time, re-emitted from the detector locations and propagated in time. The final reconstruction is obtained at $t = 0$.

$$p_0(x) = \frac{1}{4\pi c^2} \int dA \frac{1}{t} [\frac{p_m(x_d, t)}{t} - \frac{\partial p_m(x_d, t)}{\partial t}] \Big|_{ct=|x-x_d|} . \tag{4.2}$$

BP reconstruction formulas have been derived for ideal geometries (sphere, planar, cylinder, and 2-D counterparts), but are approximately valid in the far-field (detector distance much lager than ROI) for arbitrary geometries, too. Because of the filtering process involved, BP reconstruction might lack low spatial image frequencies. Being based on the Green's function in Eq. 3.7, exact BP formulas require ideal media and detection geometries, but several modifications have been developed like weighting for limited view or scattering, or introducing delays for heterogeneous SOS maps [133, 134]. BP reconstructions have a theoretically unfavorable complexity of $O(n^5)$ (or $O(n^3)$ in 2-D); but owing to their simplicity they are fast in practice. BP reconstructions on GPU are relatively simple to implement for the parallel, one-step nature of the algorithm and they even allow for video-rate reconstructions of volumetric data [17]. Thus BP reconstructions can generally be seen as current gold standard reconstruction method in the context of real-time live visualization and feedback.

Time-Reversal

Time-reversal (TR) reconstructions are rooted on two general observations of OA imaging [135, 136]: (1) The wave equation Eq. 3.5 is invariant under time reversal $t \to t' = -t$. (2) According to Huygens' principle, all waves have left the volume surrounded by the detectors after a finite time T_{tr}, i.e. $p(x, t) \equiv 0$, for $x \in ROI$ and $t > T_{tr}$. Although Huygens' principle is not valid in 2-D, the signal have sufficiently decayed for sufficiently large T_{tr} so that no considerable errors are induced in practical 2-D reconstructions. The reconstruction problem consists thus in a boundary value problem with the initial value at T_{tr} given as

$p(x, -T_{tr}) \equiv 0$ and boundary conditions as the time-reversed measured signals at the detection locations $p_m(x_d, -t)$. Numerical wave propagation in time then results the reconstructed initial pressure image $p_0(x) = p(x, t = 0)$ (Fig. 4.1(b)).

Unlike the other three classes of reconstruction methods, TR reconstruction is not based on a specific, known Green's function solution (Eq. 3.8). Instead, it directly operates on the level of the wave equation (Eq. 3.5). Non-ideal acoustic properties are therefore easily accounted for such as arbitrary SOS maps, heterogeneous acoustic impedance distributions, or power-law absorption terms. However, incomplete signal information in limited view geometries requires corrections to avoid reconstruction artifacts.

Most TR implementations use finite differences time domain (FDTD) methods with stencils that relate variables at neighboring points in space and time and that provide explicit or implicit solutions with only a few neighbors contributing. k-space methods obtain the spatial derivatives in the Fourier domain via the Fast Fourier Transformation (FFT). The increased accuracy of the spatial derivative permits courser time-steps and decreases the algorithm's run time in this way. Despite a low complexity of $O(n^4)$ (or $O(n^3)$ in 2-D), TR reconstructions are considered slow, especially compared to BP; but because only a few image variables need to be stored and the variable updates are local, good performance of TR implementations on GPU can be to obtained [136].

Model-Based

The fourth class of reconstruction algorithms are the so-called MB algorithms. MB algorithms directly relate the discretized images to the corresponding discretized signals and solutions are obtained by numerical inversion. MB algorithms are described in detail in the following section.

4.2 Model-Based Reconstructions for Optoacoustic Tomography

MB algorithms aim at obtaining reconstructions numerically instead of analytically. The reconstruction procedure is formulated as a discretized minimization problem and the reconstructed image is the minimizing solution or a sufficient approximate. Solving an inverse problem, like the reconstruction of OA images, typically involves five major steps [137]:

1. Discretization of the unknown, i.e. the image $p_0(x)$, and representation as a discrete vector f[1]. Image discretization can be performed on Cartesian or polar grids, with spherical voxels, or with arbitrary (potentially adaptive) meshes (see Sect. 5.2 and Refs. [138, 139]). The number of unknowns is chosen large enough for a good image quality and small enough for manageable computational cost of the reconstruction. The knowns, i.e. the pressure signals p_m, need to be discretized, too. The dimension of the measured signals is finite because of the finite number of projections and

[1] Throughout this work, discrete vectors are denoted by a plain symbol without an additional index. Their continuous counterparts are denoted with an explicit argument instead, if a differentiation between the two is required.

time-samples. If needed, they might be re-sampled to further minimize computational cost.

2. A physical MODEL M is implicitly or explicitly calculated that relates the unknowns to the knowns, i.e. the image f to the signals p:

$$p = M \cdot f. \tag{4.3}$$

For OA imaging with an ideal propagation medium and point-like detectors, the linear model M is given by Eq. 3.8 in discretized form (see also Sect. 5.2). The discretization of the model results from image and signal discretization.

3. For statistical inversion, a NOISE MODEL of the data is formulated, either based on assumptions or derived from system noise measurements. Although different noise models like an exponential distribution are possible, too, Gaussian noise is mostly assumed. Correlation in the noise can be equally modeled by the use of the covariance matrix; but most often independent white noise of uniform amplitude is assumed and the covariance matrix is the identity matrix. Further, priors on the distributions of signal and image intensities might be included, too.

4. A SCALAR COST FUNCTIONAL $F(f)$ to be minimized with respect to the unknowns f. The functional might be heuristically derived or statistically chosen based on expectation maximization. For linear models with Gaussian noise and Gaussian priors, the corresponding functional is quadratic and the resulting inverse problem is relatively easy to solve compared to deviating functionals.

   Further, regularization might be required because of the ill-posed nature of the model with a large condition number κ (which is the ratio of largest to smallest Eigenvalue) or because of incomplete input data (limited view). In CS, adding a non-quadratic penalty term like the $l1$-norm of the image in a specific basis or like the total variation of the image enables reconstructions even from under-sampled input data [140]. For regularization of ill-posed problems, a quadratic penalty term (Tikhonov regularization) might be added. The resulting quadratic functional is:

$$F(f) = \|Mf - p_m\|_2^2 + \|Pf\|_2^2, \tag{4.4}$$

   where P is a linear operator to e.g. suppress stripe artifacts or to enforce smooth images [114]. If P is a multiple of the identity matrix, the functional corresponds to statistical minimization assuming uniform Gaussian noise.

5. Finally, a MINIMIZATION ALGORITHM to numerically minimize the cost functional $F(f)$ and to obtain the reconstructed image as minimizing solution of Eq. 4.4. Depending on the functional, different minimization strategies might be employed. With convex non-quadratic functionals like $l1$-regularization, non-linear optimization algorithms like gradient descent algorithms or recently developed faster alternatives like FISTA might be

used [141]. In the quadratic case, the resulting problem is a least squares problem and many simple and fast inversion methods exist.

If the problem considered is sufficiently small in size, direct methods can be employed and the solution can be directly obtained. Without an additional regularization term, the solution is obtained by the Moore-Penrose pseudo-inverse $M^\dagger$ of the model M [142]. The inverse[2] of a full-rank matrix is $M^\dagger = (M^T M)^{-1} M^T$ and $MM^\dagger$ is an orthogonal projector. The inverse is independent of the measured data and only depends on the system geometry and on the acoustic properties. Unlike the original model, the inverse matrix is however not sparse (with most entries equal to zero) and inversion of the matrix requires $O(n^9)$ FLOPs ($O(n^6)$ in 2-D). For realistic problem sizes, the inverse is often computationally too burdensome to obtain. But if once calculated for a given experimental system, the reconstruction process reduces then to a simple and fast matrix-vector multiplication:

$$f = M^\dagger p_m. \tag{4.5}$$

Another possibility to obtain the inverse matrix is singular value decomposition (SVD), a generalization of Eigenvalue decomposition for non-quadratic matrices. SVD factorizes M as product of a diagonal matrix D with entries in decreasing absolute order and two unitary matrices: $M = U_1^T D U_2$. The inverse is then obtained as $M^\dagger = U_2^T D^{-1} U_1$ with the reciprocal of the non-zero elements on the diagonal of D^{-1}. By truncating small valued diagonal entries, inversion can be regularized (TSVD, [143]). Another alternative based on the full-rank of M is to solve the associated normal equations:

$$F(f) = \left\| M^T M f - M^T p_m \right\|_2^2. \tag{4.6}$$

Direct inversion can then be efficiently performed based on LU or QR factorization of $M^T M$. Because $M^T M$ is symmetric and positive definite, Cholesky factorization can be performed with half the asymptotic computational cost in both FLOPs and memory consumption.

For large-scale problems where direct methods are not feasible, iterative inversion is performed instead (Fig. 4.2). The solution f is obtained as converging series $f_{n_{iter}}$ with increasing iteration number n_{iter}. An initial guess f_0 can be obtained via BP or an empty image is assumed. Iterative updates are performed by first retrieving a suitable search direction and then choosing a suitable step size in the search direction. With gradient descent algorithms, the search direction is always the steepest descent direction at optimal step size. For ill-posed problems, a zig-zag behavior can be observed and convergence is thus slow. Conversely, conjugate search directions with conjugate gradient descent (CG) or its variant LSQR [144] in the Krylov sub-space spanned by

$$(M^T M)^{n_{iter}} M^T p_m, \ldots, (M^T M)^1 M^T p_m, M^T p_m \tag{4.7}$$

[2]For the remainder of this thesis, the inverse of a matrix refers to its Moore-Penrose pseudo-inverse.

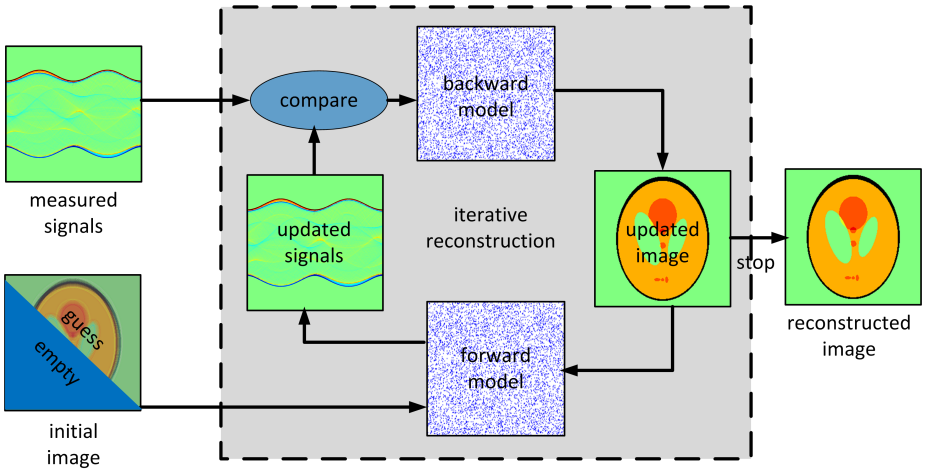

Figure 4.2: Schematic of iterative reconstruction. Starting from an empty image or an initial guess, the image is forward modeled using the model matrix M. The corresponding signal vector is compared to the measured signals and the difference vector is calculated. The difference vector is backward modeled using the transpose of the model matrix, M^T, to update the current image. After several iterations, depending on the problem considered and the algorithm used, the image obtained is a good approximation of the solution. Basically, iterative inversion works similar to a discrete vector-valued feedback loop. Figure based on Ref. [145].

are explored.

After each iteration, the solution estimate $f_{n_{iter}}$ is the minimizing vector in the current Krylov sub-space. The at least linear convergence, depending on the condition number κ_{cond}, is much faster than for gradient descent algorithms. Convergence might be further improved if a suitable preconditioner can be found that reduces κ_{cond}. In exact arithmetic, CG and LSQR converge to the exact solution when the number of iterations is equal to the number of unknowns.

Iterative algorithms are efficient for OA reconstructions because the main steps, matrix-vector multiplications Mf and M^Tp, do directly benefit from the sparsity of the model. Also memory foronly a few additional image and signal vectors needs to be allocated. The number of iterations determined by either a preset maximum or by a threshold for the relative change of $(F(f_n) - F(f_{n-1}))/F(f_0)$ is typically much smaller than the number of unknowns. Performing only a limited number of iterations also serves as intrinsic regularization.

The advantages of MB algorithms are manifold. Optimal reconstructions (in the sense of minimizing Eq. 4.4) can be obtained for arbitrary detection geometries and not only for ideal geometries and for full view [114]. The finite size of arbitrarily shaped detection elements can be taken into account [146]. Further, propagation medium properties are not restricted to ideal acoustic media: Heterogeneities, attenuation, or scattering processes can also be taken

into account [134, 147]. Regularization is possible to suppress artifacts from the ill-posed nature of the reconstruction problem, statistical noise modeling might improve image quality, or additional constraints like non-negativity of the solution estimate might be seamlessly included [148, 149]. Reconstructions from under-sampled data in a CS framework is only possible with MB algorithms, not with analytical approaches.

The drawbacks of MB reconstructions are typically associated with their high computational cost. For typical problems, inversion times are much longer than for BP and the algorithmic complexity is much higher, $O(n^6)$ in 2-D and $O(n^9)$ in 3-D. The actual run time however also depends on the sparsity of the model matrix M and the number of iterations performed. The memory requirements to store the model matrix M can also limit the possible applications of MB reconstructions, especially in 3-D ($O(n^6)$), unless matrix-free approaches are used [139, 150].

4.3 Model-Based Reconstructions for (Pre-)Clinical Optoacoustic Imaging

OA imaging has matured in the recent years from a pure proof-of-concept laboratory technology to a useful tool for biologists and pharmacologists in preclinical research and for medical doctors in clinical diagnosis. Myriads of studies have been reported like hepatic clearance [23], brain activity and damage [34, 96], cardiac function [151], tumor morphology and physiology [33], and dermal imaging [45]; and even more application fields are currently under investigation.

(Pre-)Clinical and proof-of-concept studies also impose different requirements on the reconstruction methods used. Both share the necessity to provide the maximum image quality achievable. While in experimental studies image quality can be optimized manually, this cannot be done in clinical routine where reconstruction needs to be performed automatically. More importantly, tedious offline reconstructions are not practical. Current clinical studies typically involve real-time video-rate frame acquisition at multiple wavelengths, yielding large amounts of raw data to be reconstructed[3]. For feedback to the examiner, a live visualization is needed to find the suitable image ROI or to detect dynamic processes such as heartbeat and pulsation [151, 152]. Currently, live visualization is only possible using BP methods [17]. Unfortunately, BP is not capable of providing the same image quality as MB reconstructions in typical imaging scenarios [14]. MB reconstructions are computationally demanding and video-rate MB reconstructions at sufficient resolution and image quality are an active area of research.

The amount of raw data to generate images from has been steadily increasing over the recent years, owing to the rapid progress in instrumentation technology that enabled video-rate imaging. The key advancements were:

- Parallel detector arrays replaced scanned single element transducers. The number of time-resolved electric pressure signals measured in parallel increased from 1 to up to 512 [22, 127].

[3] For example, the perfusion experiment in Sect. 9.2 took 3 min to acquire, but required more than 1 day of reconstruction time on a standard desktop computer.

- Parallel data acquisition electronics with the same number of channels prevented the need for multi-plexing. Consequently the achievable frame rate is not reduced by the multi-plexing ratio of e.g. 1:3 or 1:8 [23, 127].
- Low noise electronics in the DAQ increased the resulting SNR which had to be previously improved performing frame averaging. Assuming uncorrelated Gaussian noise, the SNR increases $\propto \sqrt{n_{avg}}$, where n_{avg} is the number of averages. Typically, tens of averages were taken for sufficient SNR whereas low noise electronics allow for single laser pulse imaging without averaging [22, 129].
- High power pulsed diode lasers are capable of delivering higher pulse repetition rates compared to Nd:YAG lasers. The former limit of 10 Hz or 20 Hz has been shifted to up to 100 Hz [25]. Furthermore fast per-pulse wavelength tuning capabilities of the OPO has eliminated idle times while changing the wavelength. Consequently, video-rate imaging with >25 frame-sets per second in multi-spectral imaging and 100 frames per second in monochromatic imaging has been realized [111].

Figure 4.3 illustrates the impact of the listed improvements on the (maximum) number of projections that can be acquired per second. The rate is compared to the increase in computational power based on Moore's law (assuming a doubling ever 18 months). In the seminal work by Wang *et al.*, a single element transducer and a 10 Hz repetition rate laser were employed and 40 averages taken. The system consequently was able to provide 0.25 proj/s [22]. The recent generation small animal scanner described in Sect. 3.4 employs a 256 element transducer array in combination with a 10 Hz repetition rate OPO laser and low noise electronics, i.e. frames can be acquired from single laser pulses. In this way, acquisition of 2560 proj/s is possible [129]. A state-of-the-art, commercially available clinical scanner compromises a 512 element volumetric probe and a 100 Hz pulse rate laser. It is thus able to acquire 51200 proj/s [25]. It is obvious that the increase of data generation capabilities of the recent systems have considerably outperformed the progress of standard computational technology. Reconstruction time however even increases super-linear with number of projections per image or the transition from 2-D to 3-D imaging increases the image reconstruction duration even further.

After the data are acquired, an inversion framework is required that reconstructs the best possible images from the measured data. For the small animal scanner used in this work, the image ROI with sufficient image quality and sensitivity is approximately 3 cm in diameter. The frequency content of the measured signals is determined by the upper sensitivity limit of the detector at f=6.2 MHz (-3 dB) which corresponds to a wavelength of $\lambda = 230\,\mu$m and ultimately an axial resolution of $0.8\lambda = 180\,\mu$m [155]. Thus, typically an image discretization with n_x x n_y=300 x 300 pixels is chosen to support the maximum resolution over the entire ROI. The number of projections n_{proj} is mostly fixed, but the number of time-samples n_t typically increases linearly with the number of image pixels per dimension. BP methods compromise three independent, nested loops (in 2-D) that are simple to implement. Because no significant additional memory is required and the algorithm is intrinsically parallel, BP reconstructions are readily implemented on GPUs and achieve real-time video-rate renderings at 50 Hz in cross-sectional and volumetric imaging.

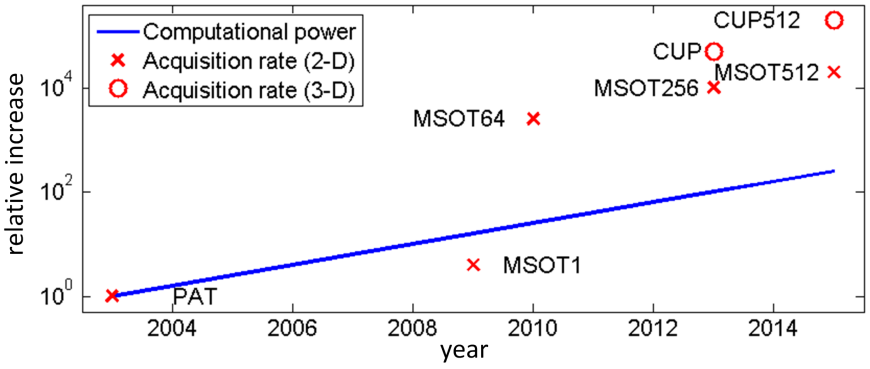

Figure 4.3: Progress in computational technology and optoacoustic detection technology. In the semi-logarithmic plot, the progress of computational technology is plotted (blue line) according to Moore's law (a doubling every 18 months). The red symbols indicate the effectively acquired projections per second for selected 2-D scanning (PAT: [22]; MSOT1: [30]), 2-D parallel detection (MSOT64: [30]; MSOT256: [129] and Sect. 3.4; MSOT512: [153]) and 3-D parallel detection systems (CUP: [154]; CUP512: [25])

Conversely, MB reconstructions impose much greater computational challenges. A typical problem consists of 300 x 300 = 90000 unknown image pixels and 1273 x 256 = 325888 measured signal samples. The overdetermined system of equations is thus governed by a model matrix M sized 325888 x 90000. In double precision format, the matrix would require 218 GBs of memory. Luckily, most of the matrix elements are zero because of the explicit time-distance relationship of US pressure propagation. The fraction of finite valued coefficients which have to be effectively stored is in the order of $1/n_x$, here 0.6 %. In a sparse algebra format (Compressed Sparse Row), storage takes only 4.7 GBs which is easily handled by standard desktop computers. Matrix vector multiplication Mf of the model M with the image vector f takes in the order of 1 s then. However two matrix-vector multiplications (Mf and $M^T p$) have to be performed in each of the n_{iter} iterations during reconstruction. Typically, at least $n_{iter} = 15$ iterations are needed to obtain a good estimate of the optimal solution. The total reconstruction time per image on CPU is thus in the order of 30 s, much slower compared to 0.1 s for a BP reconstructions. Naive implementations of MB inversion are approximately a factor of 300 x slower and thus not suited for live visualization.

Because of the increasing gap between data acquisition capabilities and computational power, the computational aspects of MB reconstructions for fast or even video-rate imaging have become increasingly important. The three most important computational characteristics of MB reconstructions are:

- Generation time of the model. If the model is stored as matrix and re-used for multiple reconstructions, model generation (in the order of several minutes for 2-D standard models) is not the crucial factor of reconstruction. The opposite is the case, if the model is barely re-used; with matrix-free on-the-fly methods, model generation alone determines the run time and thus feasibility of the method in practice.

- MEMORY REQUIREMENTS to store the model. Matrix storage becomes impossible for problem sizes with memory consumption that exceeds the size of available memory. Memory considerations are crucial especially for GPU implementations. The available memory on GPU is much smaller compared to CPU RAM. If large problem sizes demand too much storage like in 3-D imaging, matrix-free methods have to be used.

- INVERSION TIME to solve Eq. 4.4 for one image or a given set of images at a specific level of quality (with a certain number of pixels n_x and a certain number of iterations n_{iter}). In most cross-sectional MSOT imaging scenarios, inversion time is the most important factor of reconstruction performance, because the model can be stored in the memory and is reused multiple times.

Several approaches have been developed to address the above three limitations. First, matrix-free methods have been developed to shatter the memory limit [139]. The images are discretized using spherical voxels (blobs, or higher order generalizations) and the corresponding pressure signals can be calculated efficiently independent of the voxels orientation. Alternatively, signals can be obtained from FDTD methods that link signals via wave-equation based time-propagators and require only a minimal amount of storage (a few copies of the image ROI). Reconstruction is then performed on GPU by repetitive application of that operator in combination with an iterative solver [150]. Such FDTD methods offer a maximum flexibility for selection of both acoustic properties and regularization techniques, but are much slower than matrix-based methods. Generally, GPU methods can greatly accelerate reconstruction speed, if memory restrictions are met and the algorithm structure fits the GPU architecture (see also Sect. 5.5). Moreover, transformation-based methods have been shown powerful in overcoming computational limitations. First, most state-of-the-art OA systems offer geometrical symmetries (see Sect. 5.1). Whereas all analytic reconstruction formulas are rooted on these symmetries, MB algorithms mostly do not exploit them, but rather aim at solving an arbitrary inverse problem. Chapter 5 presents a transformation-based approach that renders the model matrix separable and thus enables fast reconstructions. Furthermore, sparsity and transformations to sparse representations have been shown effective for reconstructions. A sparsity-based inversion framework is presented in Chapter 6. Finally, Sect. 7.6 discusses possible future approaches for fast and yet accurate MB reconstructions.

Chapter 5

Acceleration of Model-Based Reconstructions: Exploiting Tomographic Symmetries

Reconstruction time with MB inversion does typically not meet the requirements of clinical imaging because of the large size of the inversion problems considered and the need for multiple iterations. Symmetries are among the most powerful tools to solve equations, but are most often not considered in MB inversion algorithms. In the current chapter, symmetries in multiple OA scanners are identified, and a reconstruction framework is described with huge acceleration gains from exploiting tomographic symmetries[1]. Performance is tested in full and limited view detection geometries as well as the performance on CPU and GPU architectures.

5.1 Symmetries in Optoacoustic Imaging

Standard MB reconstruction methods aim on solving the general inverse problem of minimizing Eq. 4.4. No specifics of the OA models are included and no other assumptions are made except that the model M is sparse with mostly zero elements.

Symmetries and transformations are the probably most powerful tools to solve equations in general. Transformations allow to reformulate the problem in a different, more suitable basis via a coordinate transformation; symmetries allow to reduce the dimensionality of corresponding (sub-)problems. For the physical laws of motion, Noether's theorem states every continuous symmetry implies conservation of the conjugate variable [157]. The resulting conservation of energy and momentum are the most fundamental principles of physics.

For OA imaging, multiple symmetries are present in the governing equations and the imaging systems:

- TIME-SHIFT invariance with a change of variable $t \rightarrow t' = t + \Delta t$ for arbitrary Δt results in conservation of the conjugate variable, the temporal

[1] Large parts of the results presented in this chapter have been previously published in Medical Physics as Ref. [156].

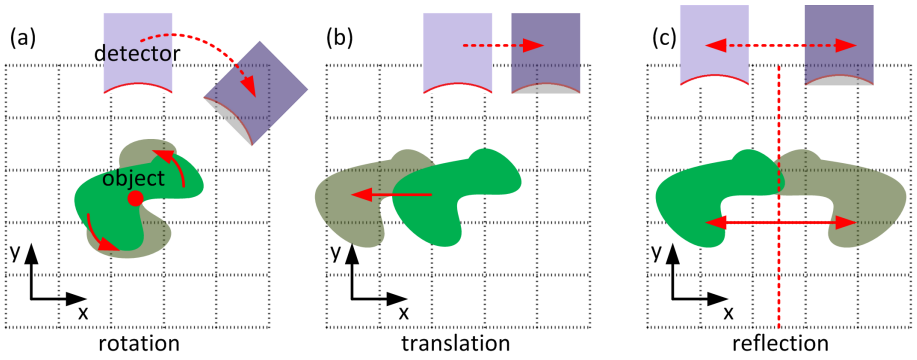

Figure 5.1: Illustration of symmetries in OA imaging. (a) For a tomographic imaging system with rotation invariance, the detected signals are the same if either the detector (blue) is rotated (dashed arrow) or the object (green) is rotated by the same angle in opposite direction (solid arrow). (b) Corresponding illustration for translation symmetry as found in e.g. raster scanning systems. (c) In a system with reflection symmetry, flipping the detector location and flipping the object with respect to the symmetry axis result the same measured signals.

frequency ω. The left-hand side of Eq. 3.5 is time-shift invariant, the source term in pulsed illumination imaging is however not ($t = 0$ is distinct). For frequency domain imaging however, time-invariance leads to conservation of frequency. Then acoustic signals are only detected at the same frequency as the light modulation frequency $\omega_{US} = \omega_{ex}$ [103, 158].

- TIME-REVERSAL invariance, $t \to t' = -t$, enabled time-reversal reconstruction methods with same wave propagation for forward propagation and reconstruction [136]. Similar to the initial value problem of forward propagation, reconstruction is transformed into a boundary value problem with the boundaries given by the measured signals at the detection surface and zero initial value. If time-reversal symmetry is broken, for instance in the presence of a dissipative term, unmodified time-reversal reconstructions result in incorrect images [102].

- ROTATIONAL invariance of space (isotropy) with $x \to x' = R_{\Delta\varphi}x$ where x represents the two dimensional space vector $x = (x, y)^T$ and $R_{\Delta\varphi} = \begin{pmatrix} +\cos\Delta\varphi & +\sin\Delta\varphi \\ -\sin\Delta\varphi & +\cos\Delta\varphi \end{pmatrix}$ with arbitrary $\Delta\varphi$ (Fig. 5.1(a)). Rotational symmetry is found in circular scanning geometries in 2-D and spherical and cylindrical geometries in 3-D. If a reconstruction problem is rotationally invariant (in continuous formulation), the following three conditions hold:

  1. The wave equation Eq. 3.6 is rotationally invariant. While this is the case for homogeneous media, it is in general not valid in heterogeneous media.
  2. The detection surface S is invariant under rotation: $x \in S \Rightarrow x' = R_{\Delta\varphi}x \in S$. The detection surface is circular (or multiple concentric circles).

3. The image ROI is invariant under rotation: $x \in ROI \Rightarrow x' = R_{\Delta\varphi}x \in ROI$. The image ROI is circular and concentric with the detection surface.

Although governing equations, image region, and detection surface must obey rotational invariance, no such assumptions are made on the imaged object and the measured signals themselves. Reconstructions of absorption distributions that are not rotationally symmetric can equally benefit as long as the reconstruction problem itself is rotationally invariant.

- TRANSLATION invariance $x \to x' = x + \Delta x$ leads to conservation of spatial frequencies and is found with planar sensors as in raster scanning microscopy, linear arrays, or z-scanning in cross-sectional imaging (Fig. 5.1(b)). The corresponding conditions as for the rotational case must be met for translational invariance, too.

- REFLECTION symmetry for a given axis $x \to x' = -x$ is often found as implication of translational and rotational symmetry in multiple systems (Fig. 5.1(c)). Reflection is a discrete, 2-fold symmetry and the number of independent symmetry axes is limited by the dimension of space.

The most important consequence of symmetries is regarding the solutions of the wave equation: If the initial pressure $p_0(x)$ corresponds to signals $p_m(x_d, t)$, symmetries imply that $p_0(R_{\Delta\varphi}x)$ corresponds to $p_m(R_{\Delta\varphi}x_d, t)$, i.e. the solution of a rotated version of the image is a rotated version of the signals. Image and opposite detector rotation have the same effect.

The above symmetry considerations can be translated from the continuous to the discrete case to be used with MB reconstructions: Images, signals, and propagation model are discrete. For a given experimental system, signals are sampled at a finite number of projections φ_i and time-instants t_τ. Geometrical symmetries are not affected by the choice of time-samples. Typically equidistant sampling is chosen, $t_\tau = \tau \cdot \Delta t$ ($\tau = 1 \dots n_t$), with angular spacing of $\Delta\varphi$ between two consecutive projections: $\varphi_i = i \cdot \Delta\varphi$ ($i = 1 \dots n_\varphi$). Thus a n_φ-fold discrete symmetry with $\Delta\varphi \cdot n_\varphi = 2\pi$ is resulting. The discrete symmetry of the detection surface requires the image discretization to have the same discrete symmetry. For the circular case, the image grid is a polar grid with the same angular spacing $\Delta\varphi$ as the detection surface (in principle, a rational multiple could also be chosen). Typically, equi-distant radial spacing is chosen, too: $r_i = i \cdots \Delta r$ ($i = 1 \dots n_r$). For the planar case, an ordinary equi-spaced Cartesian grid can be used.

For the rest of the section, the circular symmetry in 2-D is considered because of the tomographic symmetries of the system used in this work and described in Sect. 3.4. In principle, implementations for reflection and translation symmetry are alike; some of the special considerations and applications are discussed in Sect. 5.6.

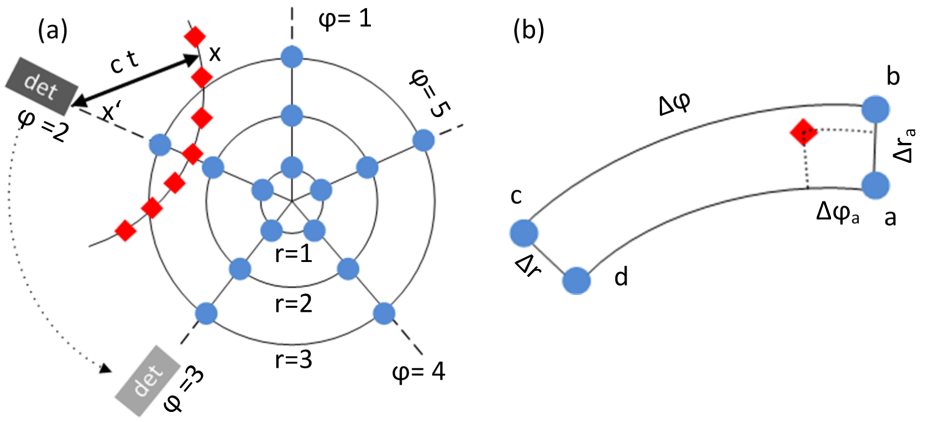

Figure 5.2: Discretization of the image grid and model calculation. (a) The image ROI is discretized to n_φ=5 and n_r=3 equi-spaced pixels (blue circles). Projection angles coincide with the grid discretization angles. For a given projection φ and time t, signals originate from an arc of radius ct which is discretized by n_{int} equi-spaced points (red diamonds). (b) Each integration point is bilinearly interpolated in angle and radius to the four neighboring pixels of the image grid.

5.2 Model Generation in a Polar Grid Formulation

In order to exploit rotational symmetry for MB reconstructions, Eq. 3.8 needs to be discretized and modeled in such a way that the resulting matrix equation $p = Mf$ (Eq. 4.3) preserves the underlying symmetry of the continuous equations, i.e. $p = Mf \Rightarrow R_{\Delta\varphi,sig}p = MR_{\Delta\varphi,img}f$. The orthogonal matrices $R_{\Delta\varphi,sig}$ and $R_{\Delta\varphi,img}$ effectively perform a rotation of signals and images by an angle of $\Delta\varphi$. In two dimensional representation of vectors p and f sized n_φ x n_t and n_φ x n_r, respectively, rotation is a circular shifting of the rows by one. If restricted to the first time-sample, $R_{\Delta\varphi,t=1}$ is the permutation matrix with entries on the lower diagonal. For the entire data vector, the rotation matrix is the Kronecker product with the identity matrix in time samples: $R_{\Delta\varphi,sig} = R_{\Delta\varphi,t=1} \otimes 1_t$ (and $R_{\Delta\varphi,img} = R_{\Delta\varphi,r=1} \otimes 1_r$; for simplicity of notation, the indices indicating data and image domain are omitted). $R_{\Delta\varphi}$ is an orthogonal matrix,i.e. $R^{-1}_{\Delta\varphi} = R^T_{\Delta\varphi}$.

With the above definitions and rotational invariance, it follows[2]:

$$R^T_{\Delta\varphi} M R_{\Delta\varphi} = M \Leftrightarrow [M, R_{\Delta\varphi}] := MR_{\Delta\varphi} - R_{\Delta\varphi}M = 0 \qquad (5.1)$$

The procedure[3] to obtain the model M is as follows [159]: For one projection at fixed position φ_i and time-sample t_τ, a fan of rays originating from the current detector location is generated that covers the entire image ROI (Fig.

[2] $R_{\Delta\varphi}$ is not well-defined, operating in both data and image domain with different identity matrices in the Kronecker product. For a rigor definition, all the arguments hold also for the normal equations in the image domain only.

[3] The standard polar forward model generation code has been initially developed by. X.L Déan-Ben.

5.2(a)). The n_{int} rays are equi-spaced in angle and the maximum distance between two rays in the ROI is smaller than the maximum grid spacing. Low n_{int} result in discretization errors and large n_{int} result in long calculation times. The integration in Eq. 3.8 along the arc with $d = ct_\tau$ is approximated as discrete sum of points on the rays (red diamonds). The points of the integration curve do in general not coincide with the points of the image grid. Thus bi-linear interpolation in angle and radius to the four neighboring grid points (a-d in Fig. 5.2) is performed. Alternative interpolation schemes, as nearest neighbor or spline interpolation are possible, but decrease accuracy or increase computational cost respectively. Finally, the temporal derivative $\partial/\partial t$ in Eq. 3.8 is obtained numerically as central finite difference with time difference Δt_{deriv}:

$$M_{\varphi=1,t=t_\tau} = \frac{M_{\varphi=1,t=t_\tau+\Delta t_{deriv}} - M_{\varphi=1,t=t_\tau-\Delta t_{deriv}}}{2\Delta t_{deriv}} \tag{5.2}$$

Although bi-linear interpolation functions are not continuously differentiable, the errors can be neglected if the time difference Δt_{deriv} is considerably below the sampling time of the signals. The complete model M is then obtained by repeating the above calculations for all time-samples (in parallel in vectorized form) and projection numbers (sequentially).

The resulting complete matrix for all projections M is highly redundant owing to the underlying rotational symmetry. The modeled signals of the second projection can be obtained by the corresponding sub-matrix $M_{\varphi=2}$ via $p_{\varphi=2} = M_{\varphi=2}f$. Alternatively, rotation of the image by $-\Delta\varphi$, forward-modeling with the sub-matrix of the first projection, $M_{\varphi=1}$, and rotation of the signals by $\Delta\varphi$ results identical signals. Consequently, all sub-matrices for higher projections are entirely redundant and can be obtained from the first:

$$M = \sum_{i=0}^{n_\varphi-1} (R_{\Delta\varphi})^i M_{\varphi=1} (R_{\Delta\varphi}^T)^i. \tag{5.3}$$

In practice, the left-hand side is not calculated explicitly, but the sequence of matrices on the right-hand side is applied to the image vectors. The rotation by $R_{\Delta\varphi}$ is implemented as row shifting of the corresponding signals and images.

Matrix compression approaches similar to those described in Refs. [160, 161] have been proposed for different symmetries and offer computational advantages for two of three previously described aspects:

- Model computation only requires $1/n_\varphi$ of the time as only one projection of the entire model is calculated. This is important in scenarios when the matrix is barely re-used as for on-the-fly methods where it needs to be recalculated in each iteration. Accounting for the finite detector size in the model considerably increases model generation time and then restriction to only one projection is particularly beneficial.

- Memory consumption is equally decreased to $1/n_\varphi$ of the original amount by the matrix compression approach. Memory saving is crucial if the number of projections is very large (particularly in 3-D imaging, [162]) and the model would else not fit the total memory; or on architectures with comparably limited memory as on GPUs.

- Inversion performance of the reconstructions is however not directly accelerated because the same number of FLOPs for matrix-vector multiplications has to be performed; benefits might however arise from the possibilities to use GPUs or caching effects since effectively multiple right-hand sides to $M_{\varphi=1}$ are used.

For a realistic example, polar and Cartesian discretization matrices were calculated using a 3 GHz dual-core processor and 8 GBs of RAM and implemented as Matlab code. Data consisted of $n_t = 1768$ time-samples and $n_\varphi = 360$ projections over 360°. The image ROI was 18 mm x 18 mm and 18 mm in diameter, respectively, and discretized using $n_x = n_y = 300$ pixels in the Cartesian case and $n_r = 250$ x $n_\varphi = 360$ pixels in the polar case. The resulting matrix dimension, memory consumption, and calculation time can be found in Tab. 5.1. The advantages of exploiting the rotational symmetry can be readily seen there.

metric	dimension	memory [GBs]	time [s]
Cartesian (M_{cart})	636480 x 90000	6.2	1517
complete polar (M)	636480 x 90000	7.1	1173
compressed polar ($M_{\varphi=1}$)	1768 x 90000	0.02	3.3

Table 5.1: Comparison of matrix dimension, memory requirements, and calculation time for three different MB approaches.

5.3 Expediting Reconstructions Using Symmetries

Most often, inversion is the crucial part in reconstruction of video-rate data-sets, not the model generation (the model can be pre-calculated and used thousands of times) or memory consumption (in 3-D it might become crucial, in 2-D reduction to 20 MB is not of crucial importance). Expediting reconstructions can be achieved exploiting conservation of variables in the presence of symmetries. Like the (temporal) frequency in the case of time, invariance under rotation by an angle $\Delta\varphi$ implies a conserved variable L. L can be associated with the angular frequency, or with the angular momentum in mechanics.

Conservation of L can be applied to the problem at hand to invert M, i.e. to find its Eigenvectors (EV). According to Eq. 5.1, M and $R_{\Delta\varphi}$ are commutating matrices. From basic linear algebra it is known that commutating operators share a common set of EVs. Finding the EVs of $R_{\Delta\varphi} = R_{\Delta\varphi,t=1} \otimes 1_t$ then simplifies finding the EVs of M. As first step, one needs to obtain the EVs of $R_{\Delta\varphi,t=1}$ while 1_t will result in a degenerate subspace for each Eigenvalue. The EVs of $R_{\Delta\varphi}$ are:

$$\Psi_l = \exp(-i\frac{l \cdot \varphi}{2\pi n_\varphi}) \otimes 1_t, \tag{5.4}$$

for $l = 0 \ldots n_\varphi - 1$. From those EVs, one can derive the corresponding operator

$$L = \sum_l \Psi_l \cdot l \cdot \Psi_l^T =: U_L^T D_L U_L, \tag{5.5}$$

where D_L is the diagonal matrix with $0...n_\varphi - 1$ elements and U_L is a unitary base transformation matrix, namely the Fourier transformation in angle φ operating independently on each time-sample. L is also commutating with M:

$$[M, L] = 0. \tag{5.6}$$

The next step is to transform the model matrix to the base of EVs of L. With

$$p = Mf = U_L^T(U_L M U_L^T)Uf =: U_L^T M_L U_L f \Leftrightarrow: p_L = M_L f_L \tag{5.7}$$

we obtain the model matrix in the new basis

$$M_L = U_L M U_L^T. \tag{5.8}$$

Knowing the EVs and Eigenvalues of M or M_L results the possibility to perform reconstructions with a direct inverse approach (Eq. 4.5). However, the model matrix M_L is too large to calculate all EVs in straight forward matter. However, the EVs of L are known with n_φ different Eigenvalues. Each EV of M_L corresponds to one Eigenvalue l of L and is a linear combination in the corresponding Eigenspace of L. For M_L, a direct consequence is that the matrix elements m_{ij} linking different Eigenvalues of L, $l_i \neq l_j$ must vanish due to the conservation of l, i.e. $l_i \neq l_j \Rightarrow m_{ij} = 0$. So non-zero (NZ) elements of M_L are restricted to m_{ij} where $l_i = l_j$. For each l, this is a block of size n_t x n_r and n_φ blocks exist in total for the block-diagonal matrix M_L. The number of NZ matrix elements is bound to n_t x n_r x n_φ elements compared to n_t x n_r x n_φ^2 for M (however, as sparse matrix, a large fraction of the elements of M is zero), as illustrated in Fig. 5.3(a) and (b). Each block of the matrix (and corresponding parts of signal and image) can be stored and processed independently. If the number of NZs in each block is sufficiently large compared to its size, dense algebra formats can be used instead of sparse formats. Dense formats can be more memory efficient by up to a factor of two and much higher computational efficiency might be achieved.

Fourier coefficients of real-valued data are redundant with the coefficient at negative frequency being the complex conjugate of the corresponding coefficient at positive frequency. For the image vector elements of f_L it implies $f_{L=l} = (f_{L=-l})^*$ where $(*)$ denotes the complex conjugate and l is taken modulo n_φ. Equal considerations can be made for signal vectors p_L. The consequence for reconstruction is that negative frequency $-l$ elements of p_L can be obtained computationally efficient from the corresponding positive frequency elements of p_L ($p_{L=-l} = p^*_{L=l}$) instead of from explicit calculations ($p_{L=-l} = M_{L=-l} \cdot f_{L=-l}$). Thus almost 50 % of model storage and calculation time corresponding to negative frequencies can be spared. Furthermore, reflection symmetry leads conservation of parity and sin- and cos-terms are not mixed. So M_L is not complex but real-valued.

The work-flow of the forward modeling with the proposed polar approach is illustrated in Fig. 5.3(c). Instead of directly calculating the pressure signals via M (1'), the right-hand side of Eq. (5.7) is used and the calculations are split in three parts (1)-(3):

1. Calculate the (angular) transformed image $f_L = U_L f$ via parallel, column-independent FFT.

2. Perform batched block-wise multiplications with the sub-blocks of M_L for non-negative frequencies l in a dense algebra format to obtain the signal vector $p_{L=l} = M_{L=l} f_{L=l}$. Coefficients at negative frequencies are obtained by complex conjugation of their positive frequency counterparts.

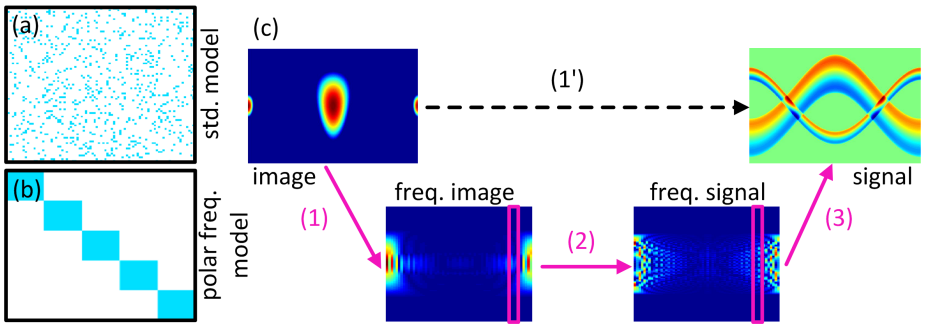

Figure 5.3: (a) Illustration of the structure of a standard forward model matrix M. In the large matrix with dimensions determined by the number of unknown image pixels and known signal samples, most of the coefficients are zero. The NZ coefficients (often only 1 %) are distributed almost throughout the entire matrix. (b) Although the dimensions of the matrix are the same after transformation to the angular frequency domain, the distribution of non-zero coefficients is different. Owing to rotational invariance, NZ coefficients are restricted to dense blocks on the diagonal. Each block corresponds to an angular frequency and the total number is given by the number of projections. (c) Workflow with the polar model: Instead of directly calculating the signals corresponding to an image with the standard forward model M (black arrow, (1')), calculation is alternatively performed in three steps: (1) Transformation of the image to angular frequency domain using FFT on a per-row basis. (2) Forward modeling step independently performed for each angular frequency using the corresponding block of the matrix. (3) Transformation of the obtained signals on a per-row basis using inverse FFT.

3. Calculate the signals $p = U_L^T p_L$ via column-independent inverse FFT.

The advantages of the proposed approach compared to direct implementation of M are less memory consumption, fewer total FLOPs for matrix-vector multiplications, and the possibility to work with dense algebra formats; the two additional transformations U_L are efficiently implemented using FFT. Reconstructions are subsequently performed using standard LSQR iterative inversion calculating M in Eq. 4.6 in the proposed manner. M_T is implemented in the same way using the Hermitian conjugate operations in reversed order.

The polar model formulation does not only allow to expedite iterative reconstructions but also to efficiently obtain the direct inverse $M^\dagger$ for realistic-sized problems. Because U_L as unitary matrix has full rank, $M^\dagger$ can be obtained as product of inverse matrices:

$$M^\dagger = U_L^T M_L^\dagger U_L. \tag{5.9}$$

M_L is block-diagonal and thus also its inverse must be block-diagonal with blocks of the same size. Each block of the inverse corresponding to a specific l is then obtained by independently inverting the block of the forward model M_L with same l. Inversion is efficiently performed using e.g. SVD and the block size size, n_t x n_r is sufficiently small compared to the size of the total matrix M_L so that the inversion is even fast for the sum of all n_φ blocks. If regularization

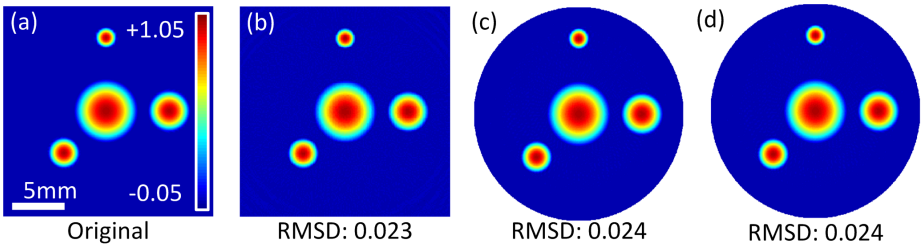

Figure 5.4: Performance of the polar reconstruction approach with simulated paraboloid absorbers. (a) Original image. (b) Iterative reconstruction using Cartesian image discretization. (c) Reconstruction with polar model using iterative inversion. (d) Corresponding reconstruction using the direct inverse.

is needed, truncated SVD can be performed, thresholding small singular values zero. Note that the resulting inverse $M_L^\dagger$ requires only the same amount of memory for storage as the forward model M_L[4]; this is orders of magnitude less than the inverse matrix of M would require (also calculation times for $M^\dagger$ are not feasible on standard computers).

Inversion is again performed in a three step manner: (1) $p_L = U_L p$, (2) $f_L = M_L^\dagger p_L$, (3) $f = U_L^T f_L$. The reconstruction process becomes extremely efficient and fast: The total reconstruction process has the same cost as one application of M_L^T; this is 50 % of a single LSQR iteration and also much faster then applying $M^\dagger$ in standard basis - if it could be obtained at all - because of the large number of NZs in the inverse in standard representation.

To test the proposed method, a numerical and an experimental[5] set of OA data was selected. The numerical phantom consisted of four paraboloid absorbers of 1 mm, 1.5 mm, 2 mm, and 3 mm diameter, respectively, with peak absorbance of unity (Fig. 5.4(a)). The shape of the corresponding signals is known analytically [138] and sampled at $n_t = 400$ time-instants at 1° spacing over 360°. Images to be reconstructed were discretized with $n_x = n_y = 250$ pixels in the Cartesian case and $n_r = 200$ and $n_\varphi = 360$ in the polar case. Using the same computational environment as described in Sect. 5.2, images were either reconstructed using Matlab's LSQR-function with $n_{iter} = 150$ iterations and no penalizing term or using the direct inverse. For quantitative image quality assessment, the root mean square difference (RMSD) was calculated:

$$RMSD = \left\| f - f^{(ref)} \right\|_2 / \left\| f^{(ref)} \right\|_2, \tag{5.10}$$

where f is the image to be compared against the reference image $f^{(ref)}$.

Fig. 5.4 shows the reconstructions obtained by LSQR with Cartesian image discretization (b), by the proposed polar approach with M_L in iterative manner using LSQR (c), and the reconstruction from the direct inverse $M_L^\dagger$. All methods correctly reconstructed the absorbers and RMSD errors versus the original in

[4]Solving the normal equations, the block size (and memory consumption) could be decreased to n_r x n_r. During reconstruction, then $M_L^T p_L$ needs to be calculated which takes as long as applying $M_L^\dagger$ directly.

[5]The mouse data were kindly provided by T. Jetzfellner and originally published in Ref. [163].

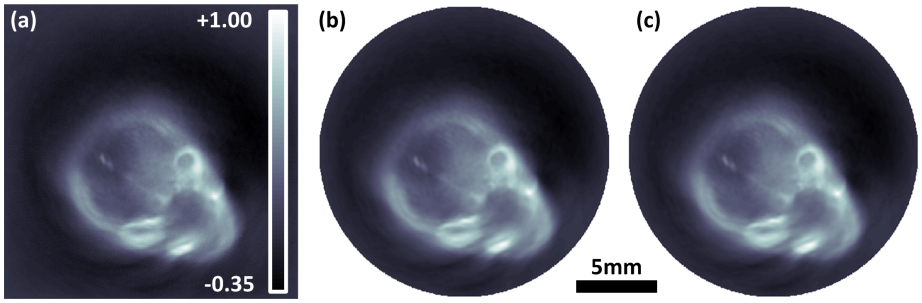

Figure 5.5: Experimental reconstructions of a baby mouse head using 360 projections in 1° steps. (a) Conventional model-based reconstruction. (b) Iterative reconstruction using the polar algorithm. (c) Reconstruction using the direct inverse. The calculation times for per slice were 104 s, 15.3 s & 0.1 s, respectively.

(a) were similar: 0.023 (b), 0.024 (c), and 0.024 (d), respectively.

Performance of the method was also tested with experimental data of 34 cross-sectional slices of an *ex vivo* baby mouse head acquired with a scanning tomographic system. The details on the system and acquisition of the data-set can be found elsewhere [163]. In brief, illumination at 650 nm was provided by an OPO laser at 10 Hz pulse repetition rate and the light beam was converted into a 2 mm thick light sheet at the surface of the head. OA pressure signals were captured by a cylindrically focused transducer element that was scanned around the object in 3° steps. Signals were then digitized by an analog-digital-converter at 100 MSa/s sampling rate and stored on a PC for further processing. $n_{img} = 34$ different planes were acquired by translation of transducer and illumination in z-direction. Prior to image reconstruction, measured projections were interpolated to 1° spacing and signals band-pass filtered from 0.1 MHz to 8 MHz to minimize noise and artifacts. Then, the whole data-set was reconstructed in parallel with effectively multiple right-hand sides for the reconstruction algorithms considered.

The data-set consisted of $n_\varphi = 360$ projections per slice, each having $n_t = 1768$ samples. To obtain high-resolution images, the ROI of 18 mm x 18 mm (or diameter) consisted of $n_x = n_y = 300$ or $n_r = 250$ pixels and $n_{iter} = 50$ iterations were performed in case of LSQR. Calculation of the conventional sparse model matrix M_{cart} took 1517 s and its storage required 6.2 GBs of RAM. Calculation of the polar model M_L was much more efficient, taking only 364 s and requiring only 0.64 GBs of RAM. Calculation of the inverse matrix $M_L^\dagger$ required additional 32 s and 0.64 GBs of RAM.

The reconstructions of one slice are shown in Fig. 5.5. Fig. 5.5(a) presents the reconstruction using conventional Cartesian image discretization with iterative LSQR inversion. Solving the inverse problem required 104 s per slice. Fig. 5.5(b) presents the same slice using the proposed polar approach in iterative manner. At no loss of image quality, reconstruction took only 15.3 s per slice. Finally, Fig. 5.5(c) presents the reconstruction from the direct inverse. Preventing the need for multiple iterations, reconstruction was fastest, taking only 0.1 s per slice.

The reconstruction speed increase by factor of 6.8 x can be attributed to two

key aspects of the iterative implementation: First, the number of operations in the forward model step is reduced from 397 MFLOPs in the Cartesian case to 159 MFLOPs in polar case, a factor of 2.5 x. Storage and thus also memory transfer was decreased by even a factor of 9.7 x. Second, using dense algebra increased the achievable performance. The actually resulting total performance depends then on problem size including number of images considered, if arithmetic operations or memory transfer bandwidth are limiting performance, on Matlab's overhead in executing multiple serial tasks instead of a single large task, and on the overhead from the two additional FFT operations. The latter is considerably increased if the number of projections can not be factorized in small prime numbers as exploited by Cooley-Tukey FFT implementations. When performing reconstructions with the direct inverse, an additional speed-up by 150 x resulted simply from reducing the workload to effectively 0.5 iterations instead of 50 iterations during LSQR.

5.4 Polar Reconstructions in Limited View Geometries

In the last section, only full view tomographic imaging geometries spanning 360° have been considered. Full view detection systems are in general preferable due to the higher image quality and the smaller condition number of the model matrix [114]. However, in many real system embodiments only an incomplete detector coverage with less than 360° is available. Accessibility of the object as in skin imaging, limitations imposed during the manufacturing process of the US probe, or simply cost of instrumentation are among the determining factors of available coverage [42, 111, 154]. In the limited view configurations considered here, only $n_{\varphi,lim} < n_\varphi$ projections are available with the same angular spacing $\Delta\varphi$ as in the corresponding full view case.

A polar forward model M_{lim} for the limited view case can be readily formulated, too. Also, memory efficient implementations using only $M_{\varphi=1}$ are directly applicable, restricting the summation in Eq. 5.3 to available projections $(0 \dots n_{\varphi,lim} - 1)$. The angular frequency case, M_L, is rooted on the conservation of angular frequency from the image to the signal domain and vice versa. However, there is a distinct direction with no detection elements. Rotational symmetry is thus broken. The signals p at projections $n_{\varphi,lim} + 1$ to n_φ are required to obtain the correct Fourier transformed signals p_L, but they are not known from the measurements. Extra-/Interpolation to not covered areas induces considerable errors and might thus only be performed for very small angles not covered.

Rotational symmetry is broken through the shape of the covered detection surface; wave propagation itself is still conserving l and the signals obtained by $M_{lim}f$ are the same as for $U_L^T M_L f_L$ for the projections available. Thus, the limited view case model can be reformulated using M_L:

$$M_{lim} = V_{lim} M = V_{lim} U_L^T M_L U_L. \tag{5.11}$$

V_{lim} is the restriction from all (full view) projections to the actually available projections. The $n_{\varphi,lim} n_t$ x $n_\varphi n_t$ matrix V_{lim} is obtained from the identity matrix by dropping rows corresponding to non-available data. Images can then

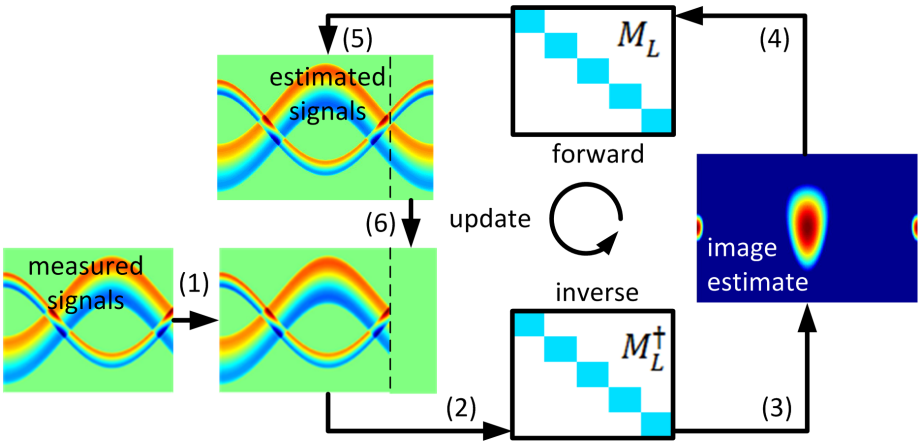

Figure 5.6: Polar reconstruction workflow in a limited view scenario. For the first iteration the measured signals are zero-filled for projections with unknown signals. The signals for all projections are then used for reconstruction of an image estimate using the inverse $M_L^\dagger$. The obtained image is then forward modeled by M_L. The resulting signals are kept for those projections not covered by the detection surface, while the signals for the other projections are replaced by the original measurement data. With the updated signals, additional iterations are performed until sufficient convergence.

be reconstructed by iteratively minimizing $\|M_{lim}f - p_{lim}\|_2^2$ via LSQR with the above implementation of M_{lim} using M_L.

The direct inverse $M_{L,lim}^\dagger$ cannot be as simply reformulated. The symmetry of the detection surface is broken and l of the data cannot be uniquely obtained. Consequently, $M_{L,lim}^\dagger$ is not expected to be block-diagonal and thus not as simple to obtain as $M_L^\dagger$. Although all factors are easily invertible, the inverse of the matrix product in Eq. 5.11 is not the product of its inverses

$$M_{lim}^\dagger \neq V_{lim}^\dagger U_L^T M_L^\dagger U_L, \tag{5.12}$$

because neither V_{lim} nor M_L have rank $n_t n_\varphi$. Images with direct inverse $M_L^\dagger$ and zero-filling p_{zero} for missing projection data or any other estimate guess will not be improved by repetitive application of the inverse and forward matrix pair because $M_L^\dagger M_L$ is by definition an orthogonal projector. However signals corresponding to $\Delta p_{zero} = U_L^T M_L M_L^\dagger U_L p_{zero}$ have larger deviations from the solution's signals $p_{lim,sol} = M_{lim} f_{sol}$ at measured projections and smaller deviations at projections only estimated or filled initially with zeros. Errors induced by initial zero-filling can then be iteratively reduced replacing the signals at unknown samples by Δp_{zero}[6]. The workflow of the reconstruction method in the limited view case based on the inverse $M_L^\dagger$ is illustrated in Fig. 5.6 and compromises the following algorithmic steps:

[6]Convergence of the algorithm to the correct solution as effective Neumann series can be proven by showing that both orthogonal projectors $V_{lim}^T V_{lim}$ and $M_L M_L^\dagger$ do not have a common set of EVs. The EVs of the first operator are in the angular space and of the second in the angular frequency space. Uncertainty principle then forbids common EVs.

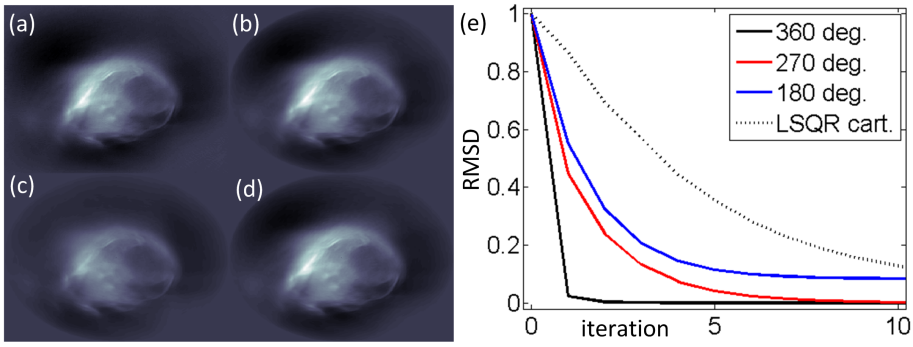

Figure 5.7: Polar reconstruction performance in a limited view scenario. (a) Iterative Cartesian reconstruction with input data from 270°using 150 iterations. (b) Corresponding iterative reconstruction using the polar approach. (c) Reconstruction from the direct inverse assuming zeros for the signals from the missing detecton angles. (d) Corresponding reconstruction using four corrective updates with forward model and direct inverse. (e) Convergence rates in limited view with the direct inverse for 360°(black), 270°(red), and 180°(blue) and convergence rate for iterative Cartesian reconstruction in (a) for comparison (black dashed).

1. Zero-padd the measured signals $p_0 = V_{lim}^T p_{lim}$.
2. Transform the signals to the angular frequency domain $p_{L,i} = U_L p_i$.
3. Reconstruct an image $f_{L,i}$ with the direct inverse from $p_{L,i}$: $f_{L,i} = M_L^\dagger p_{L,i}$.
4. Calculate the signals corresponding to the estimated solution $p_{L,i+1} = M_L f_{L,i}$.
5. Transform the signals to the original representation $p_{i+1} = U_L^T p_{L,i+1}$.
6. Update the signal vector combining original signals and improved estimates : $p_{i+1} = V_{lim}^T p_{lim} + (1 - V_{lim}^T V_{lim}) p_{i+1}$.
7. Repeat steps (2) - (6) for $i \to i+1$ until sufficient convergence.
8. Transform the solution image to the original domain: $f_{i+1} = U_L^T f_{L,i+1}$.

To test the proposed approach with limited view data, the same experimental mouse head data-set was used and either all $n_\varphi = 360$ projections (360°), the first $n_{\varphi,lim} = 270$ (270°), or the first $n_{\varphi,lim} = 180$ (180°) projections were kept, while the others were discarded prior to reconstruction. Detector coverage was chosen similar to two OAT systems in our lab (Sect. 3.4; [129]). The same image parameters were chosen like for the full view case in the previous section. Due to the missing projection information, the best MB solutions for different detector coverage angles are deviating. Limited view reconstructions are not minimizing the full view data. The respective reference images were reconstructed using LSQR inversion at $n_{iter} = 150$ iterations to ensure convergence.

Figure 5.7 presents the results of the polar approach in the limited view case. Figure 5.7(a) shows the reconstruction for a Cartesian image discretization from

$n_{\varphi,lim} = 270$ projections. Fig. 5.7 (b) presents the corresponding polar reconstruction with iterative LSQR inversion. The reduced number of projections decreased memory consumption and run time approximately proportionally for the Cartesian model (4.6 GBs and 78 s per slice compared to 6.2 GBs and 104 s for the corresponding full view case). The polar model formulation had the same requirements as for the full view case, but was still much more computationally beneficial than standard Cartesian discretization (0.64 GBs and 15.3 s per slice). Figure 5.7(c) presents the reconstruction from the direct inverse and prior zero-padding of the signals for non-measured projections. Differences in structure and intensity compared to (a) and (b) can be observed. The latter can be understood from the decreased signal energy of $p_{L,zero}$ compared to p_L. Performing $n_{iter} = 4$ iterative limited view corrections to (c) using model M_L and inverse $M_L^\dagger$, a correct reconstruction based on the inverse was obtained (d).

The convergence rates[7] of different methods and detection coverage angles are plotted in Figure 5.7(e). RMSD versus the corresponding reference LSQR reconstruction after 150 iterations is calculated as function of iterations performed. For the full view case, $M_L^\dagger$ was the correct inverse and further iterations lead no further substantial improvements (black). In the 270° limited view case, the reconstruction converged to the correct solution as the number of iterative updates is increased (red). For the smallest coverage of 180° only, RMSD decreased much slower as function of iterations and did not approach zero (blue). While reconstructions by iterating the inverse rendered correct images, iterative LSQR reconstructions exhibited visual artifacts because of the larger condition number of the problem and the absence of a penalizing regularization term. The convergence of LSQR as function of iterations in full view with Cartesian model is plotted for comparison (black dotted) and was much slower than convergence using the inverse-based updates in all cases.

The benefit of polar reconstructions[8] over Cartesian approaches in limited view geometries depends on several factors, mainly the angle covered. While the computational burdens per iteration of Cartesian reconstruction decrease proportionally to the number of projections $n_{\varphi,lim}$, polar methods have constant computational burdens irrespective of the detection angle. This leads an overhead by a factor of $n_\varphi/n_{\varphi,lim}$ per iteration compared to the full view case. For detection angles much smaller than 360° and LSQR reconstructions [111], this overhead might not be compensated by the advantages derived from computationally more efficient implementation in the full view case (less storage and run time). However, using the direct inverse convergence might be faster compared to LSQR reconstruction (again depending on the coverage); the cost per iteration are the same as for one LSQR iteration. For example, to achieve a certain image quality for the presented data-set at 270° detector coverage (RMSD < 0.15), 3 iterative updates had to be performed with the inverse matrix. To achieve the same RMSD, at least 8 LSQR iterations were needed. The overhead of $n_\varphi/n_{\varphi,lim} = 1.33$ x is compensated by the gain of 5.1 x per iteration

[7] The limited view algorithm presented here is effectively a preconditioned gradient descent minimization. Future methods might employ a version as preconditioned conjugate gradient algorithm with an enhanced convergence rate for ill-posed problems such as the limited view OA inversion.

[8] By polar an implementation of the model in the angular frequency domain is meant (M_L or $M_L^\dagger$), not a polar image discretization alone or pure matrix compression approaches ($M_{\varphi=1}$).

compared to the Cartesian model and the 2.6 x reduced number of iterations in total compared to LSQR approaches. Image reconstruction takes then 11.8 s (LSQR with Cartesian model) and 0.93 s (polar using the inverse), respectively; the polar approach with the direct inverse thus accelerated reconstructions in the most common limited view geometry by factor of 12.7 x at the same image quality on a standard PC.

5.5 Graphics Processing Units and Polar Reconstructions

In the last two decades, GPUs have developed from simple 2-D image display hardware to dedicated 3-D calculation and rendering devices with huge processing capabilities. Rendering of photo-realistic 3-D scenes at video-rate has become possible on relatively cheap standard hardware. GPUs have been optimized to perform millions of simple operations as geometrical coordinate transformations effectively in parallel. Requirements of 3-D rendering have caused GPUs to provide a much higher peak performance than CPUs in executing these operations. More than a decade ago, it was realized that GPUs might not only render graphics, but might also be used for other computations as linear algebra operations or X-CT reconstructions [164]. Since then, unprecedented growth of computational power and development of high-level, 'non-expert' GPU-APIs (application programming interface) have turned GPUs into standard computation platforms, also referred to as general purpose GPUs (GPGPU); GPGPUs are applied for numerical mathematics, cryptography, molecular dynamics simulations, bioinformatics, and many more [165–168].

GPU platforms gain their peak power through massive parallelism. Although each processing sub-unit is less powerful than a CPU, operation of hundreds or even thousands of these units in parallel results a higher peak performance. The architecture of CPUs and GPUs is fundamentally different and thus also the applicable programming paradigms. Traditional CPUs execute parallel tasks sequentially and implement the SISD paradigm (single instruction, single data), performing one operation at the time. High performance is achieved through effective hierarchical caching of memory latency; so program branching and interdependency of variables do not considerably decrease performance of SISD architectures. Conversely, GPUs only perform efficiently in highly parallel tasks according to the SIMD paradigm (single instruction, multiple data). Instead of sequentially executing the same operation multiple times, e.g. a point operation on all image pixels, the operation is concurrently performed for all image pixels by multiple processing sub-units.

The typical architecture of a GPU device is illustrated in Fig. 5.8(a). The parallel processing units (only four in the sketch) are denoted as cores. A modern GPU compromises at least several hundreds of cores. The cores are clustered into sub-groups, denoted as blocks (two in the sketch). Each core has its own set of memory registers. Read-/write-access to registers is fast, but their number is very limited compared to CPU registers. Shared memory is available to all cores within the same block and can be used for data transfer. All cores have access to the global memory of the GPU. However synchronization barriers might be needed to prevent concurrent write-access of different cores and to ensure

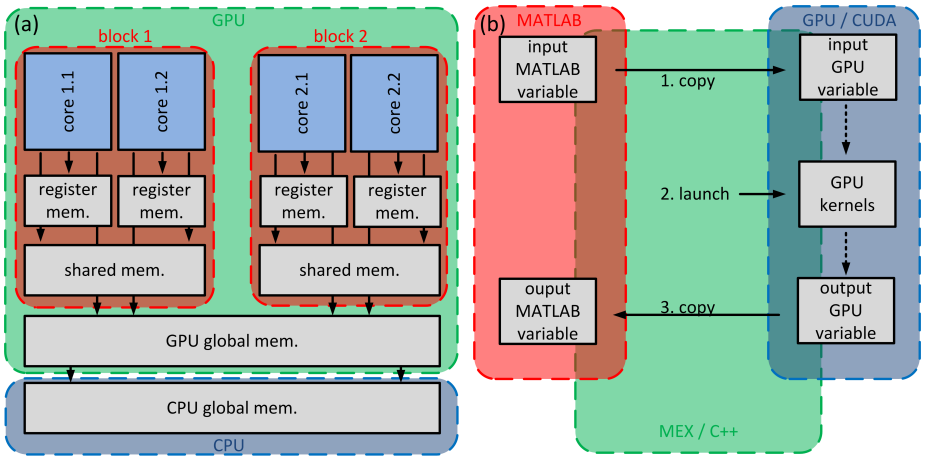

Figure 5.8: GPU architecture and workflow of reconstructions on GPU with Matlab. (a) Illustration of GPU memory hierarchy and processing unit (cores) organization. Each of the parallel processing units executes a thread and has a set of its own register memory. Several threads are grouped to blocks that can exchange data via shared memory. Access to global GPU memory is possible, but comparatively slow and potentially requires synchronization. Global GPU memory can also be accessed by the CPU. (b) Work-flow to perform reconstructions in Matlab on the GPU using the MEX interface and CUDA. The main workload is realized within the GPU kernel functions.

deterministic program behavior. Furthermore, global memory access (especially write) is orders of magnitude slower than local register memory access. Memory access latency is not hidden by a sophisticated cache hierarchy as for CPUs. High performance might be still achieved if memory access patterns are regular: If adjacent cores access adjacent memory blocks (coalesced memory access), memory transfer in the order of tens GB/s can be achieved. As consequence of that architecture, numerical problems can be solved computationally efficiently, if: (1) the problem can be split into a set (or a sequence of sets) of independent sub-problems of sufficiently small size; (2) little data exchange between sub-problems is needed; (3) if data exchange cannot be avoided, data transfer is local (using the shared memory per block); (4) multiple arithmetic operations are preformed per memory access operation to capitalize the peak computational power instead of being limited by memory bandwidth; (5) no data adaptive branching is needed; (6) total memory consumption is below total GPU memory.

CUDA is a proprietary[9], high-level API for GPGPU programming in C that hides the specifics of the hardware [166]. With CUDA C language extension, GPU operations and its memory can be accessed by the API and parallel tasks are implemented as so-called kernels. The code of a kernel is executed multiple times working on different sub-problems in parallel. Each kernel can access its own registers, associated shared memory, and the global memory. A parallel task is ended, once all kernels have finished and all results have been written to the global memory. One particular advantage of CUDA is that it is bundled

[9]CUDA runs only on NVIDIA GPUs; OpenCL is platform independent and runs both on CPUs and GPUs.

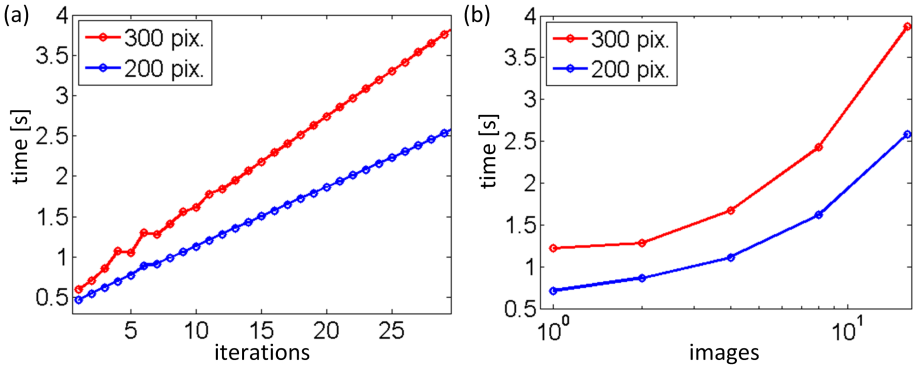

Figure 5.9: Performance of polar reconstructions on GPU. (a) Total run time for 16 concurrent images as function of iterations. Performance is linear with an offset resulting from initialization and data transfer. (b) Total run time as function of images for 30 iterations. The performance increase is super-linear because of naturally higher parallelism and reduced ratio of global memory access.

with several kernels (and APIs) for standard numerical tasks such as FFT, the BLAS dense algebra library, a sparse algebra library, and multiple more.

Multiple applications of GPU computing for the purpose of medical imaging have been reported; among them are X-CT image reconstruction and photon propagation modeling for optical imaging approaches [165, 169]. In particular, also OA reconstruction algorithms have been implemented using BP, MB in spherical voxel approximation, or k-wave FDTD approaches [17, 136, 170]. However, for a good performance both the algorithm has to be designed for GPUs and the actual implementation has to be efficient for the selected problem size, model of GPU, etc. Particularly tweaking performance with the latter might be a tedious task for custom-written kernels.

Porting of the proposed polar reconstruction algorithm to GPU was done using Matlab's build in MEX-API (Matlab executable) and CUDA libraries in C in order to provide an interface invokable from the standard Matlab environment (Fig. 5.8(b)). MEX is a C API to provide access to the Matlab variables and environment. The functionality of the MEX program code for polar reconstructions on GPU mainly consists in data handling: Memory allocation and management, data transfer to and from the GPU, library initialization and termination, and call of the CUDA kernels and CUDA library functions.

The polar reconstruction approach is particularly suited for GPU implementation because of four main reasons: (1) Reduced memory consumption compared to standard model representations; (2) coalesced memory access with dense algebra formats; (3) main operations FFT and matrix-matrix multiplication are available as highly optimized standard CUDA library functions; (4) intrinsically parallel nature of all algorithmic steps consisting of small subproblems. Besides data handling overhead in the MEX function, reconstruction work is performed by the following CUDA functions or kernels (implemention of the algorithm in Sect. 5.4):

1. Batched FFT of the signals in the angle variable for all time-points and

images (n_φ x $n_t n_{img}$) using the cuFFT library real-to-complex FFT function. Step (1) implements U_L: $p_{L,i} = U_L p_i$.

2. Permutation of signal data representation to real-valued variables with time-sample as leading matrix dimension (n_t x n_{img} x n_φ) by a custom-written kernel function. Because of the simplicity of the task as independent, one-step operation, neither memory consumption nor performance is crucial for the overall performance of the algorithm.

3. Batched, independent matrix-matrix multiplication of n_φ real-valued matrices (n_r x n_t) with (n_t x n_{img}) matrices using cuBLAS Level 3 functionality. This step implements applying the inverse $M_L^\dagger$: $f_{L,i+1} = M_L^\dagger p_{L,i}$. If no further iterations are needed, (4)-(7) are spared.

4. Corresponding batched matrix-matrix multiplications to apply the forward model matrix M_L: $p_{L,i+1} = M_L f_{L,i+1}$

5. Permutation of signal data to the same representation as prior to re-ordering in (2).

6. Batched inverse FFT as in (1).

7. Replacement of available projection signals by memory block copy of original signals, implementing $p_{i+1} = V_{lim}^T p_{lim} + (1 - V_{lim}^T V_{lim}) p_{i+1}$. If further iterations are needed, (1)-(7) are repeated.

8. Permutation of image data using a re-ordering kernel as in (5).

9. Batched inverse FFT of the image as in (6).

The above algorithm has been implemented in C using CUDA libraries (version 6) and single variable precision. Reconstruction data consisted of the mouse head data-set restricted to 270° detector coverage. Interpolated signals consisted of $n_{img} = 1 \ldots 16$ slices where each had $n_t = 1200$ (or $n_t = 1800$) time-samples and $n_{\varphi,lim} = 256$ projections. The image ROI was discretized using $n_\varphi = 346$ and $n_r = 200$ (or $n_r = 300$) pixels. $n_{iter} = 1 \ldots 30$ iterations were performed. As computational platform, an Intel i-7 multi-core CPU operating at 3.2 GHz was used equipped with 64 GBs of RAM and a Matlab 2014a environment. The GPU platform consisted of a NVIDIA GTX 780 graphics card with 2304 parallel cores and 3 GBs of total memory.

The resulting reconstruction quality is independent of performing inversion on CPU or GPU. Single variable precision did not decrease visual image quality and has a negligible influence of RMSD compared to other reconstruction parameters like number of iterations or selected SOS. Single precision formats are however improving computational performance. The resulting total GPU reconstruction run time is plotted in Fig. 5.9 as function of iterations performed (a) and number of images reconstructed in parallel (b) for the two different numbers of image pixels. Initialization and data transfer, mainly of the matrices, took 0.5 s (0.6 s for the higher resolution case). Each additional iteration then took 0.076 s (0.11 s) for all 16 images in parallel, which corresponds in average to 0.005 s/iteration/image (0.007 s/iteration/image). The performance depended non-linearly on the number of images reconstructed as can be observed from (b), showing the total run time including initialization and 30 iterations for

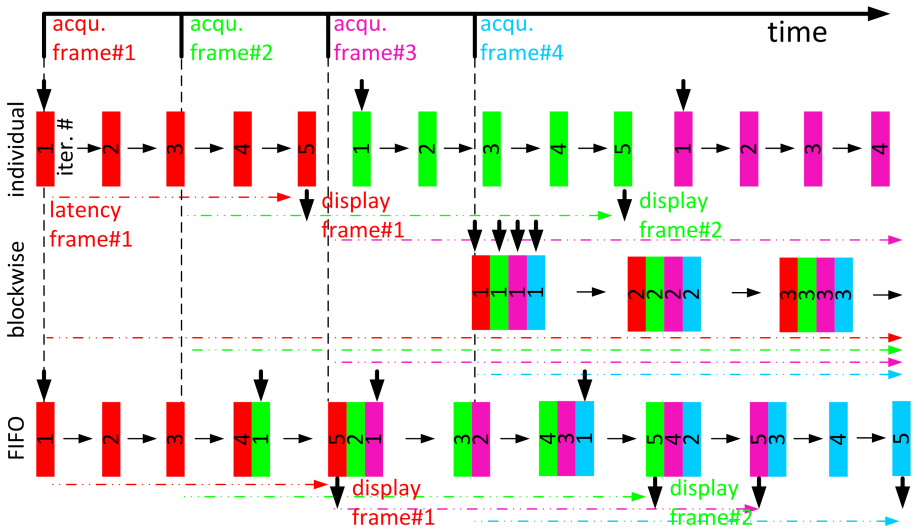

Figure 5.10: Different reconstruction schemes for multi-frame data-sets with effects on latency time and overall performance. Individual reconstruction (top row): Immediately after acquisition of the first frame (red), all 5 iterations are performed on an individual basis. After the reconstruction of frame #1 is done, reconstruction of frame #2 (green) is started. Latency time (dashed line of the same color) between acquisition and display of the first frame is smallest at a low overall performance rate. Block-wise reconstruction (middle row): All frames of the data-set are first acquired and then reconstruction is performed block-wise for all frames concurrently. Although overall performance is best, a huge delay is introduced for the first frame. FIFO reconstruction (bottom row): The reconstruction of all frames is immediately started after their acquisition. Reconstruction is then performed in parallel for all available frames that have not been fully reconstructed yet. The number of processed frames might vary and the iteration number during one step is different for different frames. With the proposed scheme, the average latency time is adaptively minimized at good overall performance rate.

different numbers of images. The achieved performance increase probably then stems from improved ratio of arithmetic operations to memory access operations. The GPU implementation presents approximately a 13 x improvement over the corresponding CPU run time for the same problem (see Sect. 5.4).

On GPU, images could thus be reconstructed with $n_{iter} = 5$ iterations in 0.024 s per image in average (0.035 s), corresponding to up to 40 FPS (28 FPS) in reconstruction rendering. The average reconstruction time was thus below the frame acquisition time of 100 ms for the scanner described in Sect. 3.4 and in the order o the 20 ms of the clinical handheld system reported in Ref. [111]. Video-rate MB reconstructions at >10 FPS and high-quality have thus been enabled using the proposed polar approach on a comparably cheap standard GPU architecture. However, the current approach did not yet provide real-time visualization and feedback. Besides minor implementation related details as a non-recurring library initialization and matrix transfer as well as an

asynchronous signal and image transfer, the fundamental obstacle is the latency time: If 16 images are reconstructed at a time, 1.5 s have passed since the first frame has been completely acquired until the other 15 frames are acquired as well. Performance on individual frame basis might lower reconstruction performance below 10 Hz. Real-time reconstruction and visualization approaches will have to find a suitable compromise between image quality, achievable overall frame rate, and latency. Two strategies might however decrease the latency time: First, in multi-spectral imaging scenarios, typically ≥ 5 wavelengths are employed. For anatomical visualization, only one wavelength is needed and the required frame rate is reduced by the corresponding factor. For spectral analysis, all wavelengths need to be reconstructed before spectral decomposition; thus all frames for different wavelengths can be reconstructed block-wise with advantages in achievable performance. Second, latency can be reduced by using an interwoven reconstruction scheme (see Fig. 5.10). Because the proposed inversion method (similar for LSQR) is non-adaptive with respect to the input data and stationary (operations do not depend on iteration number), calculations might be performed concurrently for frame #1 in its 5th iteration, for frame #2 in its 4th iteration, and so on. Using a FIFO (first in, first out) scheme with variable number of concurrent images, additional frames are added to the parallel inversion scheme as soon as the acquisition of the frame is finished and frames are removed as soon as sufficient iterations have been performed for those frames. Using such a schemes on a per MSOT-frame basis or in continuous mode, latency is minimized and overall reconstruction frame rate automatically adjusted to the system's frame rate (unless reconstruction is already faster on individual basis or slower for arbitrary large number of concurrent images).

5.6 Discussion, Outlook, and Conclusions

The proposed symmetry-based reconstruction approach has been demonstrated for a cross-sectional circular tomography geometry. The proposed methodology is also applicable to other detection geometries in 2-D and 3-D. Suitable geometries in 3-D are:

- Spherical detection surfaces with rotational invariance in two axes [19, 42, 154]. The applicability of the approach might however be limited by the often very limited coverage well below 4 π, the irregular detector distribution in at least one of the axes in the discrete case, and the computational burden of the basis transformation to spherical harmonics (although asymptotically fast algorithms exist).

- Cylindrical detection surfaces with rotational in one and translational symmetry in the other axis. Matrix compression methods with one projection only follow at hand from a suitable image discretization. For the z-axis (translation), FFT is equally applied to split the problem in independent spatial z-frequencies. After splitting based on both symmetries, sufficiently small sub-blocks are left for further processing and inversion. One fundamental difference between rotational and translational symmetries however exists: In real systems, the number of translational steps is finite and the detection aperture is thus naturally limited. The assumption of periodic boundary conditions might induce artifacts. Enlarg-

ing the reconstructed area over the actual ROI (with signal repetition or zero-padding), using focused detection elements, or applying iterative corrections via an exact forward model are potential means to reduce those artifacts. The proposed inversion approach is particularly advantageous for 3-D z-scanning with cylindrically focused detectors [18, 30]. In first approximation, all signals originate from a plane and reconstructions can be performed independently in 2-D. Corrections of out-of-plane artifacts require a 3-D model accounting for the correct detection element shape (see Sect. 3.3 and Sect. 7.3). Detected signals are elongated compared to point-like detectors and sparsity of the resulting model matrix and consequently the reconstruction speed are considerably decreased. Unlike in conventional MB representations, matrices of the proposed approach are dense and signal elongation does not influence the memory requirements of the matrix or the corresponding reconstruction speed. It is thus expected possible to perform high-resolution 3-D reconstructions of a whole mouse body at multiple wavelengths within several minutes on a standard PC. This would be even faster than conventional approaches in performing stacked 2-D reconstructions and orders of magnitude faster than conventional 3-D SIR-MB algorithms.

- PLANAR detection surfaces offer translational symmetries in two axes and the above considerations also apply there. Image quality in microscopy set-ups or linear array scanning might considerably benefit from 3-D detector shape modeling compared to first order 'in-plane' approximations [45, 171]. Currently, realistic-sized MB reconstructions are complicated by long inversion times, even for precomputed matrices in a matrix compression approach. The proposed method with its inversion efficiency might put high resolution MB reconstructions in routine application for those set-ups. In addition, MB inversion might also easily include the effects of focused illumination as found in optical and hybrid focus OA microscopy setups [59].

- NON-CONVENTIONAL scanning geometries often offer rotational and translational symmetries [113]. Because both symmetries are mixed in the same axis, the resulting blocks are too large to be efficiently inverted. Matrix compression approaches are however still applicable and image discretization different than Cartesian, as for example hexagonal closest packed lattices with hexagonal voxel shape might exploit existing symmetries best to minimize model generation burdens.

Regularization might also be applied in the context of the proposed inversion approach. While CS approaches using $l1$ regularization are possible but not favorable because of the regular detection pattern, quadratic $l2$ regularization can be seamlessly included in the inversion process. Using the direct inverse, only rotationally symmetric terms as standard Tikhonov or smoothness operators are possible. Directional terms as horizontal derivative operators [114] or general operators in polar discretization are restricted to be used with LSQR inversion only.

Furthermore, acoustic properties on which the model is based do not necessarily need to be ideal. Any propagation medium might be modeled as long as rotation symmetry is still given. This might be the case for a layered SOS

distribution of concentric rings as first order circular approximation, for statistical acoustic wave scattering [134], or for frequency dependent acoustic attenuation [101, 102]. For non-symmetric problems, approximations of the correct inverse might be obtained from the symmetry based inverse matrix and the non-symmetric forward model using linear perturbation theory for Moore-Penrose pseudo-inverse matrices or Neumann series.

Symmetry-based approaches might also be combined with the multi-scale and sparsity-based approaches presented in the following chapter: Either aiming at an acceleration of polar reconstructions by data and model sparsity or by reducing the effective workload by a multi-scale model formulation.

In summary, polar symmetry was shown to be an efficient means to expedite MB reconstructions for standard cross-sectional MSOT systems. Compared to conventional iterative MB reconstructions in the most common limited view geometry, the reconstruction process could be accelerated by a factor of 5.1 x using advantageous dense algebra model formulation and by an additional factor of 2.6 x using the direct inverse of the the model. The total acceleration on CPU by 12.7 x could be further increased by porting the reconstruction approach to GPU: Rooted on standard GPU libraries, reconstructions could be accelerated by a factor of 13 x compared to the corresponding CPU implementation. With the inversion methodology developed, high-resolution MB reconstructions of frame sequences have been enabled at video-rate of up to 40 FPS for a modern OAT system. With the achieved performance and an additional latency time minimizing inversion scheme, high fidelity MB live visualization at $>$10 Hz is ready to hand.

Chapter 6

Acceleration of Model-Based Reconstructions: Sparse Problem Formulation

For acceleration of MB reconstructions towards real-time performance in clinical imaging, a polar reconstruction approach has been developed as described in the previous chapter. It is based on a space transformation and an associated problem reformulation in the angular frequency domain that exploits the underlying rotational symmetry. The model matrix is 'compressed' losslessly, i.e. the effective matrix size (or the number of entries) size is reduced without introducing errors into the solution. The reduced reconstruction time results from a decreased matrix size and from beneficial structural properties in the new representation after transformation.

Similarly, 'lossy compression' in the reconstruction process can be aimed for, i.e. reduction of computational burdens at the price of small deviations in the reconstructed images. If the problem formulation is sufficiently sparse - similar to the good compressibility of certain digital images - computational benefits can be much greater compared to the errors introduced. 'Lossy compression' can be likewise achieved by suitable transformations of the representation space. Yet, 'lossy compression' is not restricted to systems with strict geometrical symmetries where coefficients exactly vanish to zero. Instead, the proposed 'lossy compression' approach is based on general (inexact) properties of OA data-sets and of the OA wave equation (Eq. 3.5).

In this chapter, a 'lossy' transformation-based approach with a highly sparse problem formulation is presented in order to provide fast MB reconstructions towards clinical OA image reconstruction for arbitrary detection geometries[1].

[1] The work presented in this chapter has been published in IEEE Transactions on Medical Imaging, Ref. [172]

6.1 Sparsity, Wavelet Packets Transformation, and Principal Component Analysis

The expected efficiency of compression approaches can be characterized by the resulting sparsity after transformation. A signal representation is considered sparse, if its energy is distributed to mainly a few large valued coefficients while the large majority of coefficients is low in value[2]. Sparsity can be quantified by several measures as entropy, ratio of $l1$ to $l2$ norm, or the Gini-index g used in this work [173]:

$$g = g(f) = 1 - 2\sum_{i=1}^{N} |f_i| \, (\frac{N-k+0.5}{N}) / \sum_{i=1}^{N} |f_i| \,, \tag{6.1}$$

where N is the length of vector f with entries f_i ordered in increasing absolute order. The Gini-index is the normalized area between the Lorentz-curve of the signals and the main diagonal. The Lorentz-curve is the fraction of maximum relative signal energy as function of relative number of coefficients used for its representation. The more non-linear the Lorentz-curve of a signal representation, the higher is its Gini-index: A constant signal is least sparse with $g \approx 0$, while a signal with one finite coefficient and zero everywhere else is maximally sparse with $g = 1$. To calculate Gini-indices of multi-dimensional signals and matrices, quantities are vectorized, a representative sub-set is taken (for large matrices) and the vector is restricted to NZs for reasons explained further below.

One particular transformation that might result a sparse representation is the Fourier transformation implemented as FFT in 1-D, 2-D, or 3-D. FFT does not only render sparse representations for exact spectral localization, but also when certain frequencies are dominant as exploited in image compression. However, due to exact localization of Fourier coefficients in frequency domain, corresponding signals become completely de-localized in the time domain. For signals with a certain localization in time, the Fourier signal representation might not render those signals particularly sparse.

The wavelet (WL) transformation in contrast offers localization properties in both time and frequency domain [174]. As for the Fourier transformation, coefficients are obtained from integration with a kernel which is the so called mother WL at different scales (typically a power of two) and translation positions. Many mother WLs have a finite length in time with the first few moments vanishing and a band-limited frequency range; for instance, the *daubechies1* WL (Fig. 6.1) is a bi-polar, piecewise constant WL with vanishing mean. The WL coefficients obtained by decomposition (also called analysis) can be used to recompose the original signal (synthesis) from the coefficients like in the case of the Fourier transformation.

In the discrete case, WL decomposition of a signal is implemented as filter-banks and coefficients are obtained from convolution of the filter with the signal. The filter-banks have band-pass characteristics: The so-called approximation coefficients are obtained from a low-pass filter and the so-called detail coefficients are obtained using the matching high-pass filter. Resulting coefficients are redundant and can be down-sampled (typically by a factor of two) without loss

[2]Signals are not only considered sparse, if they contain mostly zero elements resulting a small $l0$ norm; signals are also sparse, if most elements can be replaced by zero without significant error, i.e. if the $l1$ norm is small.

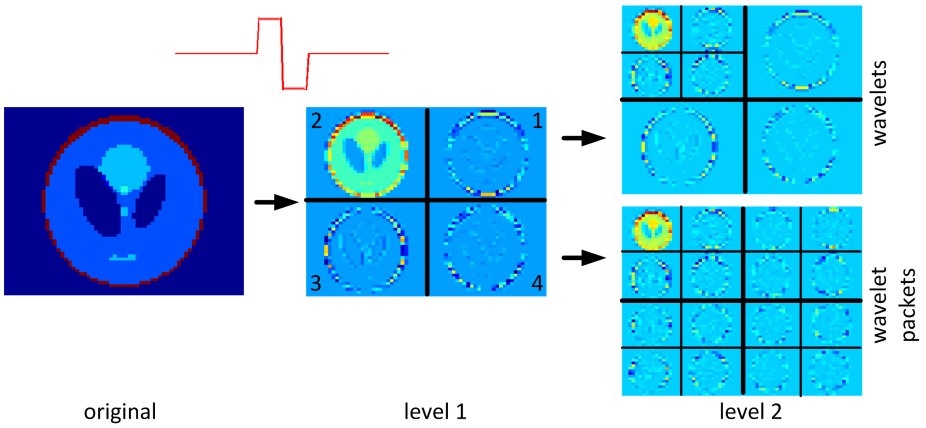

Figure 6.1: Illustration of wavelet and wavelet packets decomposition. The original image (left) is decomposed to four sub-bands applying a *daubechies1* mother wavelet (red) in both horizontal and vertical direction. The resulting horizontal (3), vertical (1), and diagonal (4) detail coefficients and the approximation (2) coefficients are visualized in the different quadrants of the central image. For a wavelet decomposition, higher level decomposition is achieved by recursive application on the approximation coefficients only (right top). In contrast, all sub-bands are further decomposed for higher level wavelet packets decomposition (right bottom).

of information. Higher level L_{dec} WL decomposition can be realized by recursive application of WL decomposition on the approximation coefficients and is resulting in an asymmetric binary tree of coefficients. Each leaf corresponds to one of the overlapping frequency bands and the number of coefficients of a node is approximately half of its parent node. A 2-D WL decomposition is obtained applying the 1-D decomposition for both axes and results in a asymmetric quad-tree of approximation, and vertical, horizontal, and diagonal detail coefficients (Fig. 6.1). The detail coefficients are not further decomposed in WL analysis and their number is, opposed to the approximation coefficients, not further reduced for higher decomposition levels.

Wavelet packets (WP) analysis extends WL decomposition to a full, symmetric binary tree (quad-tree in 2-D), where also detail coefficients are recursively decomposed to all levels [175]. Finally, all $4^{L_{dec}}$ leafs in 2-D have approximately the same size of $4^{-L_{dec}}$ of the original (Fig. 6.1). Analysis is performed using the same filter-banks but recursively applying both low-pass and high-pass filter.

Sparse representations in the WL domain (and also in the WP domain in both 1-D and 2-D) of real signals from audio, video, and also OA applications have been achieved [176]. The signals can then be effectively represented by storing only the few high valued coefficients and the large majority of low valued coefficients can be discarded. Besides a reduced storage and a more efficient processing, sparse signal representations are also often effective for noise reduction.

WP (and WL) transformation of signals (e.g. images) f to its coefficients f_{WP}, containing the coefficients of all leaves, are linear operations in 1-D and 2-D

and can be formally denoted as $U_{WP}^{(dec)}$[3]. The coefficients of only one specific leaf i can be obtained via $U_{WP,i}^{(dec)}$, which is the corresponding sub-block of the matrix $U_{WP}^{(dec)}$. After processing, the signals can be recomposed from the coefficients using $U_{WP}^{(rec)}$.

In signal processing, PCA is another important transformation for multi-dimensional data sets [177]. PCA, also known as Karhunen-Loève transformation, is the basis transformation that represents a set of observations with minimal error in energy when the number of basis vectors (components) is truncated. In multi-spectral imaging, each pixel x of an image $f = f_{\lambda,x}$ is assumed a superposition of a few different, unknown spectra measured at different wavelengths λ. PCA aims at finding the spectra and their magnitude in each pixel by de-correlation of the correlated images. PCA is a blind method and not based on assumptions on the spectra, but rather on their statistical distribution in the pixels. Formally, the PCA transformation $f_{PCA} = f_{PCA,\lambda',x}$ of a signal $f = f_{\lambda,x}$ is

$$f_{PCA} = U_{PCA} \cdot f. \tag{6.2}$$

The unitary matrix U_{PCA} is the matrix that diagonalizes the covariance[4] matrix G with EVs in decreasing order:

$$G := (f - < f >_x)(f - < f >_x)^T =: U_{PCA}^T D_G U_{PCA}, \tag{6.3}$$

where $< f >_x$ denotes the mean of f with respect to x for each λ.

Application of PCA increases sparsity of the signals through de-correlation. As optimal linear transformation, data might be sufficiently well represented by a few of the principal components only. PCA can thus be efficiently used for dimensionality reduction of multi-dimensional data to a few statistically relevant uncorrelated components.

Multi-dimensional signals can also be transformed to a combined representation, applying both WP decomposition and PCA. First, all signals are 2-D WP decomposed via U_{WP}. Then the multi-dimensional set of WP coefficients is PCA transformed with respect to the spectrum, resulting in a combined PCA-WP representation $f_{PCA-WP} = U_{PCA} \cdot (f \cdot (U_{WP}^{(dec)})^T)$. Both transformations operate independently in the spatial ($U_{WP}^{(dec)}$) and spectral (U_{PCA}) dimension of the signal. However, the resulting representation is not independent of the order because U_{PCA} depends non-linearly on the signals.

Combination of the two transformations leads potentially a sparser signal representation than achievable with a single transformation. Combined representations have been used for signal de-noising purposes (see Sect. 7.1) and are crucial for the reconstruction acceleration presented in the rest of this chapter.

[3] $U_{WP}^{(dec)}$ is however not a unitary basis transformation matrix if the WL coefficients are an over-complete representation and not a basis.

[4] Deviating definitions of PCA use the correlation matrix instead.

6.2 From a Sparse Wavelet Packets Model to a Sparse Inverse Matrix

In order to accelerate the reconstruction process, a formulation of OA reconstructions in the WP is sought that is sparse and approximately separable; then the corresponding WP inverse[5] can be obtained and reconstructions are rendered computationally efficient owing to the sparse problem formulation.

The underlying model matrix is obtained similar to the polar forward model, but with Cartesian image discretization instead [159]. The problem is then reformulated in the WP domain: The new image vector f_{WP} is obtained from a 2-D WP transformation of the original image f in x- and y-direction ($f_{WP} = U_{WP}^{(dec)} f$)[6]. Equally, the corresponding signals p are transformed to the WP domain (p_{WP}) via a 2-D WP decomposition in projection angle φ and time-sample t ($p_{WP} = U_{WP}^{(dec)} p$). The model representation in the WP domain is given by:

$$M_{WP} = U_{WP}^{(dec)} M U_{WP}^{(rec)}. \tag{6.4}$$

The sub-block of the model corresponding to one specific image domain leaf i can be obtained via $M_{WP,i} = U_{WP}^{(dec)} M U_{WP,i}^{(rec)}$. $U_{WP,i}^{(rec)}$ denotes the linear operator corresponding to WP synthesis with coefficients of leaf i only, assuming all other coefficients to be zero. Each of the sub-blocks $M_{WP,i}$ is approximately $1/4^{L_{dec}}$ of the total size of M_{WP}. If higher level WL decomposition ($L_{dec} > 1$) was performed instead, three blocks from the detail coefficients with a quarter of the size would remain, compared to 1/16, 1/64, etc. for all WP decomposed blocks. The resulting matrix still contains many zero elements because of the localization properties of WP transformation.

The sparsity in the WP formulation arises also in a non-exact manner from a general property of the governing OA wave equation. In the time domain, origination of signals and measured TOF at a certain detector are related via the distance ct. In the frequency domain, detected frequencies are related to spatial frequencies of the image via the acoustic dispersion relation $\omega^2 = c(k_x^2 + k_y^2)$. In combined space of the WP transformation, the OA wave equation also exhibits a localization property. The integration curves in Eq. 3.8 are arcs and related to the spherical Radon transformation. In the far-field approximation, when structure size is much smaller than the distance to the detector, integration arcs can be approximated as lines and the governing OA equations become a 'localized Radon transformation'. Localization properties then follow from the Fourier-slice theorem: A localized cosine image structure with modulation of a certain image frequency in x-direction creates a directional OA wave pattern [178]. Signals will be mainly detected for lateral elements at the left and at the right and only minor will be measured by the elements at top and bottom. The detected signals will be localized around the time-frequency corresponding to the image modulation frequency. If a much higher modulation frequency was chosen, resulting signals would be at much higher frequencies and thus in a different signal WP leaf. Similarly, if modulation was along y-direction, signals

[5] The framework for calculation of the WP inverse has been developed by A. Rosenthal and has been previously published as Ref. [178].

[6] With a similar notation like in the previous chapter, corresponding operators are denoted by the same symbol without explicit index for image or signal domain.

would be mainly measured by top and bottom elements and thus be localized with respect to the angle φ. In this case, such an image would mainly excite signals in another different signal WP leaf.

Consequently, different WP signal coefficients are excited by different image WP coefficients. Coefficients of the same image leaf do however excite similar signal coefficients. The cross-talk, i.e. excitation of the same signal coefficients by different image leaves, is expected to be low.

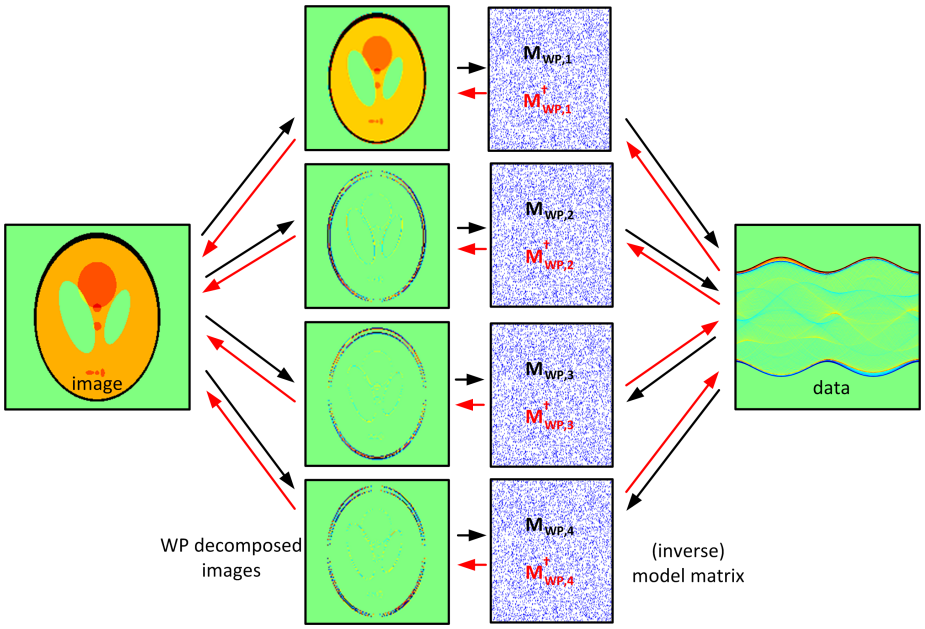

Figure 6.2: Illustration of direct reconstructions in the WP domain: The model matrix is formulated in a WP representation and split into sub-matrices $M_{WP,1...4}$ (illustrated for $L_{dec} = 1$), each representing a sub-band of the image (black arrows). Subsequently, each sub-matrix is inverted independently ($M^{\dagger}_{WP,1...4}$). Images can be reconstructed efficiently from the direct inverse matrices in a non-iterative procedure (red arrows). Adapted with permission from Ref. [172].

Sparsity of M_{WP} is a direct consequence of this localization. Although there are many NZ elements, the matrix is expected to be sparse with few large valued elements and most of the NZs small in value. quantitatively, the Gini-index of the matrix is thus expected to be high. Because image coefficients belonging to one leaf i excite only a subset of similar signal coefficients and do not significantly excite the other coefficients, specific signal coefficients can be omitted for a certain leaf i and the corresponding rows can be discarded in the matrix $M_{WP,i}$. Notably, signal coefficients kept might belong to different signal leafs and respectively the sub-sets of retained signal coefficients are different for each image leaf. The rows kept are calculated from the matrix elements that belong to the largest predefined fraction in at least one of the columns. The reduced forward model matrices in the WP domain are then obtained by

$$M_{WP,i}^{(red)} = V_i^{(red)} M_{WP,i}, \tag{6.5}$$

where $V_i^{(red)}$ is obtained from the identity matrix by discarding columns that belong to signal coefficients to be discarded by the above procedure. Finally, each of the reduced matrices $M_{WP,i}^{(red)}$ is much smaller than $4^{L_{dec}}$ in size and particularly only a tiny fraction of the size of the original matrix M_{WP} because the number of columns is reduced as well.

The structure of the sparsity from localization properties and the low cross-talk render the blocks relatively independent. Signal coefficients can be mainly attributed to one specific image leaf. Therefore it is expected that the inverses $M_{WP,i}^{\dagger}$ of the reduced model blocks $M_{WP,i}^{(red)}$ are reasonable approximations of the corresponding blocks of the inverse $(M_{WP})^{\dagger}$ of the entire model M_{WP}. The size of each of the blocks is sufficiently small to invert them with reasonable computational resources.

Reconstructions can be efficiently performed on a per image leaf basis where the respective inverse is applied to the corresponding signals coefficients and the entire image is obtained by summing the images from all leafs (Fig. 6.2).

The inverse of the full matrix in WP formulation is approximately the concatenation of the the inverse sub-matrices $M_{WP}^{\dagger} \approx \sum_i V_i^{(exp)} M_{WP,i}^{\dagger} V_i^{T,(red)}$, where $V_i^{(exp)}$ is obtained from the identity matrix by keeping all columns that belong to image leaf i.

Generally, most elements of the inverses are NZs. However, they are expected to be sparsely distributed. Most of them can thus be effectively thresholded to zero without introducing a significant error in the reconstructed image. All inverse matrix elements (n, m) that contribute less than a fraction of τ_{mat} to the absolute sum of their respective row and column are set to zero:

$$M_{WP}^{\dagger,(thresh)}(n,m) = \begin{cases} M_{WP}^{\dagger}(n,m) & \text{if } \frac{|M_{WP}^{\dagger}(n,m)|}{\sum_{n \text{ orm}} |M_{WP}^{\dagger}(n,m)|} < \tau_{mat}, \\ 0 & \text{else.} \end{cases} \tag{6.6}$$

WP transformation and related sparsity of the matrix have thus contributed in two ways to simplify and accelerate the reconstruction problem: First, by rendering the problem approximately separable with sufficiently small block size so that direct reconstruction can be performed instead of iterative reconstruction. Second, by reducing the computational requirements of applying the inverse by sparsification of its elements. The memory requirements of $M_{WP}^{\dagger,(thresh)}$ can be reduced using sparse algebra formats, if at least 50 % of the elements can be thresholded. And equally important, the number of arithmetic operations to be performed during reconstruction is equally reduced then.

6.3 Sparsity in Optoacoustic Data-Sets

Transformations can likewise be used to create sparse signal representations, not only sparse model representations. For example, the PCA transformation in the temporal domain has been shown to render the signals sparse in the context of 4-D PET reconstructions [179]. In OA imaging, sparse signal representations have been exploited for the purpose of de-noising and reconstruction acceleration, too.

WL decomposition in time-samples of all projections of a single frame leads a particularly sparse signal representation [176]. OA signals are well represented in a WL basis as broad-band, well-localized, non-stationary signals. In contrary, independent random Gaussian noise is non-sparse irrespective of the basis used. Soft or hard thresholding of coefficients below a certain, statistically optimal level thus effectively retains most of the signal energy while rejecting large portions of the noise. In multi-spectral imaging, signals for different excitation wavelengths are highly correlated while the corrupting noise is mostly uncorrelated between frames. This statistical prior improves de-noising performance compared to individual frame de-noising [180]: The multi-frame data are even more sparsely represented in a combined WL-PCA domain, where the WL coefficients are PCA transformed with respect to the spectral dimension. Hard thresholding of those coefficients results in noise rejection (see Sect. 7.1).

For the purpose of reconstruction acceleration of a single frame, signal representation in the frequency domain with subsequent small coefficient hard thresholding was proposed [181]. Reconstruction is then performed iteratively with a model in the frequency domain. Improved computational efficiency was achieved by discarding all rows of the model matrix where the corresponding signal coefficients is rejected. One disadvantage of that iterative method, however, is that the Krylov-subspace explored (see Eq. 4.7) is changed and the set of equations is effectively modified. Resulting reconstructions might thus be affected by artifacts unless additional regularization is performed. For the case of multi-frame data-sets (temporal series), a low-rank matrix estimate of the measured data has been proposed for reconstruction acceleration [182]. The low-rank matrix estimation is basically performing PCA in the frame-dimension and rejecting low energy principal components. Reconstruction is then performed with an effectively reduced number of frames and savings are given by the fraction of rejected principal components.

The multi-frame approach has the advantage of being compatible with arbitrary reconstruction approaches as only a reduced number of single frame reconstructions is performed. But as only whole frames can be discarded, performance under both aspects of de-noising and time saving might not be optimal when only a small number of frames is considered. Furthermore, sparsity on a single frame basis is not accounted for at all. Single frame accelerations rely on the sparsity of the data in the representation of the model. If a sparse signal in a certain domain is transformed to the deviating domain of the model prior to reconstruction, sparsity is typically lost and no time savings can be expected. If the signals are sparse in a specific domain, the corresponding model might not be computationally disadvantageous in this domain: For example, frequency domain signals might be much sparser than in the time domain signals, but the frequency domain model has 100 % NZ elements (with accordingly increased memory consumption), while their fraction is often only 1 % in time domain. Vice versa, if the model is particularly sparse like in the time domain, signals might however not be sufficiently sparse for substantial benefits based on signal sparsity methods.

Substantial savings can only be obtained if the inversion method employed does directly benefit from increased sparsity of the signals in terms of run time. If multiple iteration steps are performed and sparsity decreases in each step, overall savings might be marginal. One extreme case are TR reconstructions with sparse data, where the sparsity would decrease in each time-

propagation step; having to perform hundreds of those steps, no substantial benefit is achieved. So the overall possibilities of effective acceleration from sparse signal representations are restricted to domains where (1) signals are sufficiently sparse; (2) an accurate and efficient model is available at given computational resources; (3) the inversion method benefits from sparsity; and (4) overall performance exceeds that of standard iterative approaches.

Based on the above considerations, the proposed high-throughput sparsity-based inversion scheme aims at performing reconstructions in the domain of the inverse described in the previous section, the WP domain. A good performance is expected, because (1) signals are assumed to be sparse not only in the WL domain but also in the related WP domain; (2) efficient formulations of forward and inverse model are available; (3) reconstructions in sparse algebra directly benefit from sparsity of the signal vector; and (4) with the available direct inverse matrix, no iterations need to be performed but only one matrix-vector multiplication.

In order to increase signal sparsity of single frames, signals p are transformed to the 2-D WP domain, $p_{WP} = U_{WP}^{(dec)} p$. Then a hard signal thresholding operator T_{sig} is applied, setting all small valued coefficients to zero $p_{WP}^{(thresh)} = T_{sig}(U_{WP}^{(dec)} p)$. The threshold level is chosen in such a way that approximately a fraction τ_{sig} of the signal energy is retained ($\left\| p_{WP}^{(thresh)} \right\|^2 \approx \tau_{sig} \left\| p_{WP} \right\|^2$). The reconstruction approach with single frames represented in the WP domain (p_{WP}), subsequent hard thresholding ($p_{WP}^{(thresh)}$), and reconstruction with the sparsified inverse matrix ($M_{WP}^{\dagger,(thresh)}$) is referred to as WP-T. Reconstruction without signal (p_{WP}) and matrix thresholding ($M_{WP}^{\dagger}$) as originally proposed in Ref. [178] is referred to as WP-O.

In multi-spectral, multi-frame data-sets, all frames are first transformed to the WP domain[7] via $U_{WP}^{(dec)}$. Then the PCA transformation U_{PCA} is applied in the spectral dimension λ and all principal components are kept. The signal representation in this combined domain is then $p_{PCA-WP} = U_{PCA} U_{WP}^{(dec)} p = U_{PCA-WP}^{(dec)} p$. In combined representation, hard thresholding of small valued coeffcients is performed to reduce the number of NZ coefficients ($p_{PCA-WP}^{(thresh)} = T_{sig}(p_{PCA-WP}) = T_{sig}(U_{PCA-WP}^{(dec)} p)$). Like in the single-frame case, the threshold level is chosen based on the fraction of overall signal energy kept, τ_{sig}. Reconstruction with this data is referred to as PCA-WP-T strategy. For comparison, a strategy of pure PCA thresholding is also implemented, thresholding entire principal components λ' to zero, i.e. effective frames, not only single coefficients ($p_{PCA}^{(thresh)} = T_{PCA}(U_{PCA} p)$). The thresholding operator T_{PCA} retains approximately τ_{sig} of the original signal energy. This thresholding strategy is referred to as PCA-T.

Overall, inversion is boiled down from iterative reconstruction to one matrix-vector multiplication

$$f_{PCA-WP} = M_{WP}^{\dagger,(thresh)} p_{PCA-WP}^{(thresh)}, \tag{6.7}$$

[7] Like in Chapter 5, all operators involving multi-frame signal and image vectors are obtained from their single-frame counterparts as Kronecker product with the identity matrix. For simplicity of notation, the same symbols are however used without explicit indices.

where most of the coefficients of both the matrix and the signal vector are zero. By sparing calculations involving zero coefficients, a considerable gain over a non-sparse problem formulation is thus expected. The achievable gain of the method depends on the effectiveness of the thresholding approach to create zero entries without introducing significant errors.

Thus, the individual steps of the sparsity-based reconstruction approach are:

1. Calculate the forward model M (see also Ref. [159]).
2. Calculate the inverse model $M^{\dagger}_{WP}$ in the WP domain (see also Ref. [178]).
    1. Calculate the forward model $M_{WP,i}$ in WP formulation for one WP image leaf i from the original forward model M.
    2. Discard the insignificant columns of $M_{WP,i}$.
    3. Calculate the inverse model matrix $M^{\dagger}_{WP,i}$ for leaf i.
    4. Repeat (2.1)-(2.3) for all remaining leafs.
    5. Concatenate all inverses $M^{\dagger}_{WP,i}$ ($M^{\dagger}_{WP}$)
3. Sparsify the inverse model $M^{\dagger}_{WP}$ by hard thresholding of insignificant coefficients ($M^{\dagger,(thresh)}_{WP} = T_{mat}(M^{\dagger}_{WP})$)
4. Transform the multi-frame signals p to the combined representation ($p_{PCA-WP} = U^{(dec)}_{PCA-WP} p$).
5. Sparsify the signals ($p^{(thresh)}_{PCA-WP} = T_{sig}(p_{PCA-WP})$).
6. Reconstruct the images via direct, sparse reconstruction ($f_{PCA-WP} = M^{\dagger,(thresh)}_{WP} \cdot p^{(thresh)}_{PCA-WP}$)
7. Transform the images back to the standard domain ($f = U^{(rec)}_{PCA-WP} f_{PCA-WP}$).

6.4 Performance Evaluation with Multi-Spectral, Volumetric, and Temporal Data-Sets

The proposed sparsity-based inversion framework has been tested with three experimental OA data-sets: A multi-spectral data-set of a cross-sectional slice of a finger, a volumetric scanning of multiple cross-sectional slices of the same finger, and the temporal monitoring of an ICG injection in the digital vasculature[8]. Another application scenario of the sparse WP reconstruction approach in the context of SOS auto-focusing is discussed in Sect. 7.4. All code was implemented in Matlab 2014a and was executed an Intel i-7 CPU at 3.2 GHz using 64 GBs of RAM.

Multi-Spectral Data-set

The *Multi-spectral Data-set* was acquired with the setup described in Sect. 3.4

[8] The experimental details of data-set acquisition and anatomical and physiological background are discussed in Chapter 8 and Chapter 9.

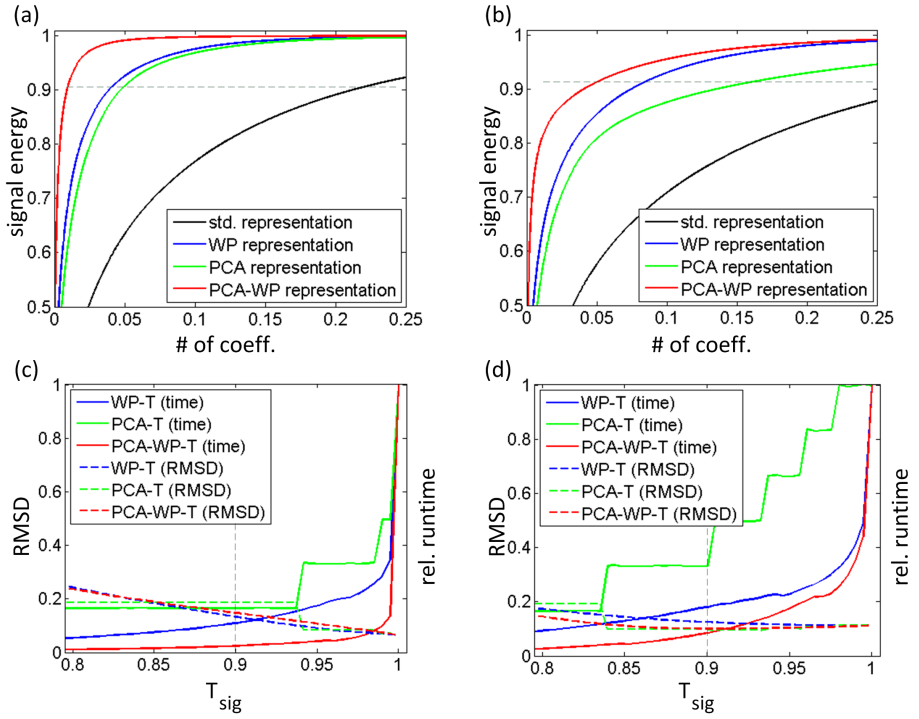

Figure 6.3: Effects of signal representation and thresholding. (a) Lorentz-curves of the *Multi-spectral Data-set* signals in standard (black, Gini-index g=0.56), WP (blue, g=0.87), PCA (green, g=0.81), and PCA-WP representation (red, g=0.92) visualize the extraordinary sparsity of the proposed signal representation. (b) Corresponding Lorentz-curves after superimposing 10 % Gaussian noise, resulting reduced sparsity (g=0.51, g=0.57, g=0.80, and g=0.79). (c), (d) Reconstruction time (solid curves) and deviation from the reference reconstruction (dashed curves) as function of signal energy thresholding level τ_{sig} for the signal representations in (a) and (b), respectively. Adapted with permission from Ref. [172].

and Sect. 8.3. The imaging location at the author's right index finger was chosen between both interphalangeal joints (see Sect. 8.1). The data-set consisted of six frames at different wavelengths of λ={715 nm, 730 nm, 760 nm, 800 nm, 850 nm, and 900 nm}. The model matrix M used for reconstruction featured 256 x 256 pixels and 256 projections with 1086 time-samples. SOS was selected as $c = 1537\,m/s$. For comparison, iterative reconstruction was performed using LSQR at $n_{iter} = 50$ iterations (LSQR-50). Alternatively, LSQR was stopped prior to convergence, performing $n_{iter} = 15$ iterations only (LSQR-15) as potential means to reduce the reconstruction time compared to LSQR-50. For WP reconstruction, a *daubechies6* mother wavelet and a decomposition level of $L_{dec} = 2$ were selected. A *Noisy Data-set* was created from the *Multi-spectral Data-set* by superimposing random Gaussian noise with 10 % peak-to-peak signal magnitude to show the intrinsic de-noising capabilities of the algorithm.

The matrix M contained only 0.6 % NZ elements and its Gini-index was

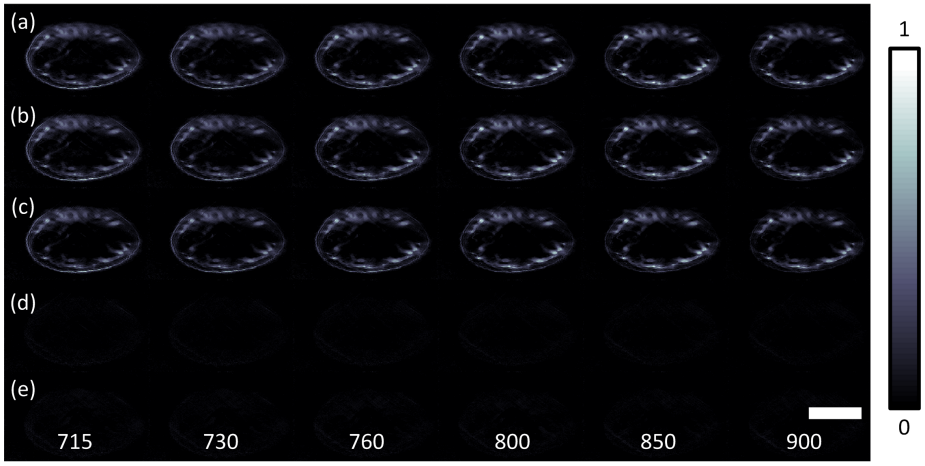

Figure 6.4: Cross-sectional finger reconstructions of the *Multi-spectral Data-set* at different wavelengths (from 715 nm at the left to 900 nm on the right) using different reconstruction methods. (a) Fully converged iterative reconstruction (LSQR-50). (b) Unthresholded direct WP inversion (WP-O). (c) Proposed sparse inversion with model and data thresholding (PCA-WP-T). (d), (e) Absolute differences of the reconstructions in (b) and (c) with respect to (a), respectively. Scale bar 1 cm. Adapted with permission from Ref. [172].

$g = 0.31$. Owing to the relatively low sparsity, attempts to reduced the number of NZ matrix elements further by discarding the lowest 50 % of entries resulted in a reconstruction error of RMSD=0.53 and are thus not a viable option for substantial acceleration. Conversely, the model in WP formulation had a much higher sparsity of $g = 0.92$, suggesting low cross-talk. The inverse matrix obtained was similarly sparse with $g = 0.84$, and could thus be efficiently thresholded. Choosing a threshold of $\tau_{mat} = 0.001$, the remaining fraction of NZs was reduced by a factor of 6.3 x. The RMSD induced through matrix thresholding was comparatively low (RMSD=0.06). Focus can either be shifted on reconstruction speed-up or on image quality upon selection of τ_{mat}.

The effects of signal representation and thresholding on sparsity, reconstruction speed, and RMSD error are visualized in Figure 6.3. Fig. 6.3(a) plots the Lorentz-curves of the *Multi-spectral Data-set* for standard (black), WP (blue), PCA (green), and combined PCA-WP representation (red), respectively, with corresponding Gini-indices of g=0.56, g=0.87, g=0.81, and g=0.92. Combining WP and PCA transformation resulted an extraordinary sparse signal representation with most of the energy (>95 %) confined in a very small fraction of the coefficients (<5 %). Fig. 6.3(b) plots the corresponding Lorentz-curves of the *Noisy Data-set*. Although the overall characteristics were similar, signals were less sparse (g=0.51, g=0.80, g=0.57, and g=0.79). Thresholding at a suitable level results then in rejection of most of the noise while keeping most of the signals.

Figure 6.3(c) plots the relative reconstruction time and the RMSD errors vs. unthresholded WP-O reconstructions of the *Multi-spectral Data-set* as function of thresholding level τ_{sig} for WP-T (blue), PCA-T (green), and PCA-WP-T

(red). The RMSD (dashed curves) increased almost linearly with rejected signal energy and the behavior was similar for all representations. Conversely, reconstruction time (solid curves) benefited highly non-linearly from thresholding. Both WP-T and PCA-T considerably led reconstruction speed-ups for increased thresholds; but both were considerably outperformed by the combined PCA-WP-T approach. Fig. 6.3(d) plots the reconstruction savings and errors for the *Noisy Data-set*. The trends in time saving were similar, but yet less non-linear. Notably, a minimum in RMSD was found (vs. noiseless WP-O reconstruction) at a level $\tau_{sig} < 1$, i.e. thresholding resulted in de-noising of the signals. For the remaining part of the chapter, a signal thresholding level of $\tau_{sig} = 0.98$ was chosen in combination with a matrix thresolding level of $\tau_{mat} = 0.001$.

Figure 6.4 presents the reconstructed images of the *Multi-spectral Data-set* with different methods. The reconstructions for different wavelengths with LSQR-50 (a), WP-O (b), and the proposed PCA-WP-T (c) approach are shown, while the respective absolute differences vs. LSQR-50 are shown in (d) and (e). RMSD compared to LSQR-50 was 0.14 and 0.16 for WP-O and PCA-WP-T, respectively. Most of the difference can be attributed to the approximate character of the inverse, not to thresholding of signals and inverse. Reconstruction times for the data-set were 108 s, 7.2 s, and 0.58 s, respectively. WP-based methods exceeded the performance of LSQR by a factor of 15 x and 186 x, respectively. Thus, exploiting sparsity of signals and inverse resulted an additional factor of 12 x compared the already fast WP-O inversion. For comparison, when reconstruction time is limited in time-critical contexts like in clinical imaging, performing LSQR with a reduced number of iterations ($n_{iter} = 15$, LSQR-15) resulted a comparable error (RMSD=0.19). The time required was however still 33 s for LSQR-15. Consequently, WP methods (WP-O and PCA-WP-T) presented a benefit by factors of 4.5 x and 56 x at comparable reconstruction quality.

Volumetric Data-set

The *Volumetric Data-set* (see Sect. 8.4 for further experimental details) was obtained from cross-sectional z-scanning of the first author's index finger and consisted of 219 consecutive slices at 0.2 mm distance at an excitation wavelength of 715 nm. The multi-frame data-set with λ=219 frames was reconstructed with the same parameters as the *Multi-spectral Data-set*. PCA was calculated in z-direction instead of in spectral dimension.

Figure 6.5 shows the reconstructions of volumetric scanning of the finger. A photograph of the finger is shown in Fig. 6.5(a) as anatomical reference. The maximum intensity projections (MIP) obtained with LSQR-15 (b), WP-O (c), and PCA-WP-T (d) are shown besides in similar orientation. Mainly skin and vascular tree can be found on the images. The Gini-indices of the signals were g=0.58 in standard and g=0.91 in combined PCA-WP representation. Reconstruction errors compared to LSQR-50 were 0.19 (b), 0.17 (c), and 0.18 (d), respectively. The inversion process took 1200 s, 245 s, and 29 s, respectively. Thus the proposed PCA-WP-T method exhibited a performance increase by factors of 8.4 x compared to WP-O and of 41 x compared to standard iterative LSQR-15 inversion at similar image quality.

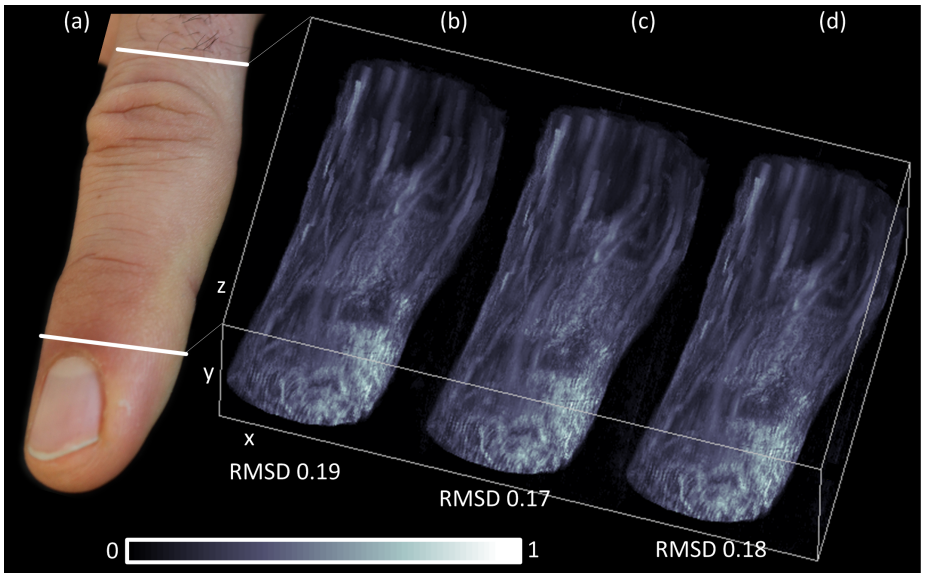

Figure 6.5: 3-D maximum intensity projections of the *Volumetric Data-set* showing a 3-D finger angiogram obtained by cross-sectional z-scanning of the finger. (a) Photograph of the finger and reconstructions obtained with (b) LSQR-15, (c) WP-O, and (d) PCA-WP-T. Reprinted with permission from Ref. [172].

Temporal Data-set

The *Temporal Data-set* was obtained from monitoring an ICG bolus in the finger vasculature of a volunteer at video rate of 10 Hz, as described in detail in Sect. 9.2. The total duration of the experiment was 180 s, and 10 consecutive frames were averaged in order to reduce the total size of the data-set to 180 effective frames. The scanner employed had 128 elements, and signals were interpolated to 256 virtual projections prior to reconstruction. SOS was selected as c=1528 m/s and all other reconstruction parameters were the same as for the *Multi-spectral Data-set*. PCA was calculated in the dimension of the λ=180 temporal frames.

Figure 6.6 presents the results with the *Temporal Data-set*. Fig. 6.6(a) shows the reconstructions of the first frame at t=0 s for LSQR-15 (top) and PCA-WP-T (bottom). Fig. 6.6(b) presents the MIPs along the y-direction as function of time. The ICG inflow into the vessels started around t $\approx$ 85 s. Sparsity of the data-set was g=0.58 in standard and g=0.95 in combined PCA-WP-T representation. Due to the extraordinary sparsity of the data-set, reconstruction time was reduced from 980 s (LSQR-15) to 1.4 s only (PCA-WP-T). RMSD versus LSQR-50 was similar in both cases (0.16 and 0.17). PCA-WP-T offered a performance increase by a factor of 700 x compared to standard iterative inversion. WP inversion in its original formulation (WP-O) required 206 s at a similar RMSD of 0.16.

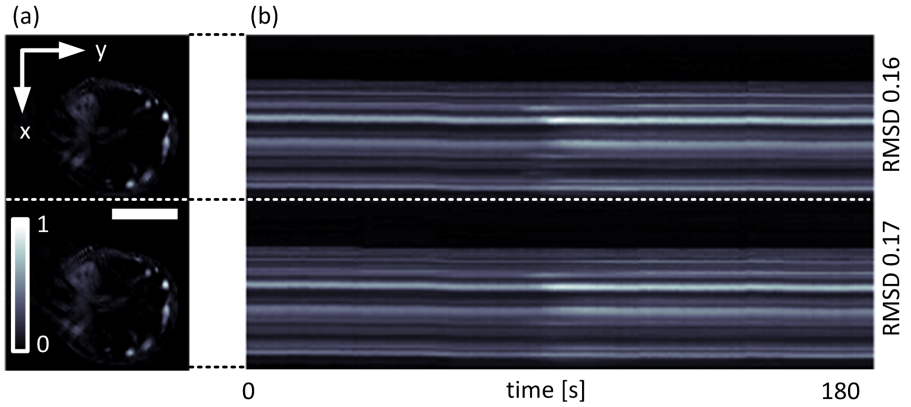

Figure 6.6: Monitoring of the ICG bolus injection in a human finger (*Temporal Data-set*). (a) Cross-sectional image at the start of the experiment reconstructed by LSQR-15 and PCA-WP-T. (b) Corresponding y-direction maximum intensity projections as function of time with ICG inflow in the different vessels at $t \approx 85$ s. Scale bar 1 cm. Adapted with permission from Ref. [172].

6.5 Discussion and Conclusions

The proposed sparsity-based reconstruction approach was rooted on transformations to and a reformulation of the reconstruction problem in a combined PCA-WP domain. The PCA-WP domain both rendered the model approximately separable and thus directly invertible, and also represented inverse model matrix and signals sparse. Thus the number of effective operations for inversion could be drastically reduced.

Four different sources of sparsity contributed to achieving fast reconstructions. First, the WP reformulation of the problem led a problem approximately represented by a set of much smaller and independent problems, based on the 'local Radon transformation' character of the OA inversion problem. Each of the blocks was sufficiently small to be directly inverted numerically. Deviations from the assumption of strict separability were the main source of reconstruction errors induced in the method. Second, the obtained inverse matrix elements were sparsely distributed and most of them could be effectively thresholded to zero. Not only was the effective number of operations in the matrix-vector multiplication reduced, but more importantly the overall memory requirements for matrix storage were reduced by the same factor. Porting to GPU might become possible for larger problem sizes, too. Third, single frames were sparse in WP representation owing to their broad-band character and their localization in time. Sparsity of the signals in the WP domain was similar to WL representations. Then signal domain thresholding in the WP domain exhibited intrinsic de-noising capabilities, too. Fourth, multi-frame data-sets were highly correlated between frames. Application of PCA de-correlated the frames. In combined PCA-WP representation, signals were extraordinarily sparse and only a tiny fraction of coefficients (throughout all principal components) needed to be retained for effective signal representation.

Experimental multi-frame data-sets are typically highly correlated. For

multi-spectral data-sets, the number of major chromophores is mostly limited (like Hb and HbO). The number of excitation wavelengths employed is typically larger than the number of those chromophores. Signals are then concentrated to the first principal components, provided spatial coloring can be neglected [77]. In volumetric imaging, adjacent slices are often correlated when imaging elongated structures such as blood vessels or from the integration over a finite effective slice thickness, if the thickness is larger than the distance between adjacent scanning positions. In temporal monitoring, a high correlation is expected because the distribution of the agent often affects only a small portion of image pixels, has a small number of different kinetic profiles, or the time constants involved are often slower than the video-rate frame acquisition rate. Conversely, correlation is decreased from strong motion like breathing or heart beat. Naturally, multiple sources of multi-frame signal sparsity can be combined, e.g. in real-time MSOT or time-lapse multi-spectral scanning of volumes [21, 111].

Limitations of the proposed method stem from the approximation character of the inverse matrix used. To improve image quality, iterative corrections using forward model and inverse can be applied. However, the residual signal then contains almost only NZs and thresholding needs to be performed again in each iteration. Also the forward modeling step needs to be similarly accelerated to achieve an overall performance benefit. Alternatively, an improved inverse can be obtained by linear perturbation theory of matrices and improved reconstructions can be obtained in one-step manner with an optimized inverse matrix. Furthermore, approaches rooted on direct inverses are restricted to quadratic regularization terms. $l1$-functionals or total variation regularization is restricted to be used with iterative inversion. In addition to the ideal models used here, the proposed approach can also be used with more sophisticated models modeling for a heterogeneous SOS distribution or the SIR of the detectors.

Future work on the sparsity-based approach will focus finding suitable strategies towards real-time imaging. While the achievable performance greatly increases with the number of frames, latency equally increases. The size of the data-set for reconstruction has to be selected to satisfy needs of both latency and peak performance. Furthermore, more sophisticated thresholding strategies might be developed. Instead of one preselected global threshold level based on a fraction of the signal energy, levels might be varying with scale and principal component and are ideally selected on the inherent signal noise to maximize denoising capabilities. To directly adapt WL de-noising strategies, signals might also be thresholded in the PCA-WL representation and prior to fully decomposing them to the PCA-WP representation. Most importantly, WP approaches might also be used to address the long standing need for efficient 3-D reconstructions. The WP methodology can also be employed in 3-D with 3-D WP decomposition and suitable 2-D parametrization of the detection surface. Once 3-D WP inverse models are available, sparsity-based methods are expected to greatly improve the performance achievable.

In summary, an inversion approach for MSOT was presented that exploits four sources of sparsity in both signal and model domain conjointly. High performance in reconstructions was enabled through a unique combination of sparsely represented direct inverse and sparse signal representations in the WP domain. The new approach accelerated computations by a factor of 41 x to 700 x as compared to the commonly applied model-based iterative reconstruction schemes with similar quality of the reconstructed images.

Chapter 7

Efficient Automatic Quality Enhancement of Model-Based Reconstructions

In the previous two chapters, two approaches based on symmetries and sparsity of the reconstruction problem have been presented in order to accelerate MB reconstructions. In a clinical setting, it is not only important to provide the clinician feedback within a given time frame. Providing an improved image quality might potentially facilitate better diagnosis and treatment. Improvements of image quality can be achieved by suppressing noise, avoiding negative pixel intensity artifacts, or using more appropriate reconstruction models. Unlike in proof-of-concept studies, reconstructions in a clinical environment should not necessarily require user input but rather need to be performed automatically. The current chapter presents solutions to these issues and further discusses potential future approaches for MB reconstructions in the context of clinical imaging.

7.1 De-Noising of Signals in Multi-Spectral Optoacoustic Tomography

Every experimental OA data-set is to a certain level corrupted by noise p_{noise}. If the SNR of the measured data is not sufficiently high, the resulting images are corrupted by visible artifacts. Then image quantification in spectral or temporal monitoring studies is complicated or even impossible. Improving the SNR by hardware means using low-noise electronics, improved transducer technology, or shielding from electromagnetic interference is always performed in practical experimental systems. Because the maximum light fluence is limited for *in vivo* studies, the noise equivalent pressure needs to be decreased to be able to detect low intensity signals. A further increase of the SNR by means of pure hardware improvements gets more and more challenging. Software methods are an additional means to increase the effective SNR.

The detected signals p_{noisy} are typically modeled as linear superposition of noise p_{noise} onto the true signals p:

$$p_{noisy} = p + p_{noise}. \tag{7.1}$$

The sources of noise can be manifold, including electronic shot noise, thermal noise, or electromagnetic interference. Consequently, the resulting noise can be statistically uncorrelated white Gaussian noise, noise with inherent frequency characteristics, or coherent noise simultaneously superimposed onto all projections.

Software approaches aim at eliminating as much of the noise term while most of the true signals. The simplest de-nosing method is averaging of frames to eliminate the statistically uncorrelated noise. SNR increases only with the square root of averages taken, and possibilities for frame averaging might be limited in clinical settings where immediate feedback is required or fast temporal processes need to be captured. Frequency filtering of the signals is only able to reduce (the part of the) noise that has no overlapping frequency bands with the signals. Spatial single frame de-noising via soft or hard coefficient thresholding in the WL domain has been previously proposed [176]. Statistical inversion with more realistic noise models than the usually assumed uniform, uncorrelated Gaussian noise might also be a future method to improve OA reconstructions [183].

To improve de-noising capabilities in MSOT data-sets for clinical settings, a spatio-spectral de-noising framework is proposed that also takes into consideration the spectral (i.e. the signal dependence on excitation wavelength) characteristics of MSOT signals and noise[1] [180]. The procedure in spatio-spectral de-noising is similar to PCA-WP-T signal thresholding: Finding a maximally sparse representation of a multi-spectral data-set and then thresholding of small valued coefficients to zero. The sparser the signal can be represented, the better the noise can be rejected, because white Gaussian noise is non-sparse independent of representation.

For de-noising, the signals are transformed to a combined PCA-WL representation: First, signals are WL transformed on a per-projection basis, resulting in the WL coefficients for different scales. Spectral frame de-correlation is then performed on a per-wavelet-scale basis via PCA, where the correlation matrix of the high valued coefficients is used to obtain the transformation matrix U_{PCA}[2]. The noise level is obtained from the standard deviations of experimentally measured signals without the presence of OA absorbers. De-noising of the signals is then performed by hard coefficient thresholding with individual threshold levels per scale, projection, and principal component (based on the experimentally determined noise level). The de-noised signals are then re-transformed to the original domain and used as input to an arbitrary reconstruction method.

The spatio-spectral de-noising was tested with artificially superimposed and purely experimental noise. In the case of artificial parasitic noise (i.e. coherent noise over all projections) superimposed onto a data-set of n_λ=21 wavelengths, RMSD after spatio-spectral PCA-WL de-noising was less than a fourth (0.024) of the corresponding RMSD after purely spatial WL de-noising not taking spectral information into account (0.10). The de-noising performance achieved ex-

[1] The results presented in brevity here are based on a project led by S. Tzoumas. A comprehensive description can be found in Ref. [180].

[2] The correlation matrix is obtained from the covariance matrix by normalizing the elements by the respective standard deviations.

Figure 7.1: De-noising of multi-spectral OA data. (a) Signals of one projection at 900 nm laser wavelength before (cyan) and after (blue) de-noising. (b) Corresponding original, noisy reconstruction. (c) Corresponding reconstruction from the de-noised signals with improved image quality.

ceeded the noise reduction capabilities of 10 frame averaging. Finally, Figure 7.1 presents the results with purely experimental noise in a mouse data-set consisting of n_λ=8 wavelengths. Fig. 7.1(a) plots the measured original signals of one projection (cyan) and the same signals after spatio-spectral de-noising (blue). Fig. 7.1(b) shows the corresponding reconstruction from the original signals at 900 nm, while Fig. 7.1(c) shows the corresponding reconstruction after spatio-spectral de-noising. An obvious artifact reduction is visible, mainly from suppression of parasitic noise in all projections.

The spatio-spectral de-noising framework can be routinely applied in clinical MSOT data-sets since the method is computationally cheap compared to the MB reconstruction process. It is most beneficial whenever the SNR is comparatively low, e.g., when using diode-pumped lasers of relatively low pulse energy and thus limited SNR. Synergistic effects arise when combined with the proposed sparsity-based reconstruction method (see Chapter 6). Efficient video-rate de-noising can be easily achieved on GPU because of the naturally parallel structure of the method and because no large matrices or iterative processes are involved.

7.2 The Causes and Effects of Negative Image Values in Optoacoustic Tomography

The presence of negative image values is another kind of image artifacts besides noise. The reconstructed image values are associated with the energy density of the deposited heat H. Because absorption followed by thermalization is a thermodynamically irreversible process, photon emission from pressure does not occur. In addition, H is a strictly non-negative quantity as product of strictly non-negative variables absorption μ_a, light fluence Φ and Grüneisen parameter Γ (Eq. 3.4).

Although no physical interpretation exists, negative values are regularly found in experimental OA reconstructions nevertheless. The obtained OA reconstructions are a result of the combination of the signals measured, of the model and its underlying acoustic propagation properties, and of the reconstruction algorithm employed.

The signals themselves are bi-polar while the initial pressure they linearly depend on is strictly uni-polar. Arbitrary bi-polar signals do not necessarily lead

uni-polar images but only a certain subset does for given inverse model. OA signals can, as discussed, be corrupted by strong noise. Furthermore, measured signals cover only an incomplete frequency band of the theoretical full bandwidth signals. The limited frequency content of the signal might either result from the limited detector frequency bandwidth or simply from frequency filtering to remove dominating parasitic signal parts such as low frequency drifts.

Deviations in the assumed model from the physical ground truth can lead negative image values. First, all model matrices are affected by discretization errors like from the calculation of the temporal derivative in Eq. 3.8. More importantly, the acoustic properties assumed in the model might be considerably deviating from experimental reality and thus errors in the reconstructed images might be present. Deviations might concern assumption on the SOS map, the SIR and EIR of the transducer, attenuation, dissipation, and acoustic scattering.

The inversion algorithm employed can also be source of negative values. BP reconstructions are only exact in the far-field approximation and errors exist for finite distances to the detectors. However, negative values might be present in MB reconstructions, too. However, one of the main advantages of the MB algorithms is that regularization and constraints can be included in the inversion process. Particularly, a non-negativity constraint ($f \geq 0$) can be added during the minimization procedure.

A non-negativity constraint can only be complied with in iterative reconstructions. A linear direct inverse can per definition not deal with an inequality constraint. With the additional constraint, inversion becomes slower because more iterations need to be performed than in unconstrained inversion. Thus, a fast inversion algorithm was developed that is almost as fast as unconstrained iterative LSQR inversion[3]. Instead of performing slowly converging gradient descend steps followed by the application of projectors or re-starting CG inversion with different active pixel constraints, the rapidly converging CG steps of LSQR are concurrently performed with the updates of the set of active pixel constraints [149]. Then, the optimum solution without negative pixel values is found (in the sense of minimal quadratic norm of the residual signal). If the residual is large, the solution can considerably deviate from the physical ground truth despite the physically motivated non-negativity constraint.

A numerical study was performed in order to test several sources of negative values and their suppression via non-negativity constraints. Fig. 7.2(a) shows the original image compromising ten paraboloid absorbers of varying diameter. The corresponding signals are analytically known and sampled at 256 full view projections at a virtual sampling rate of 40 MSa/s [138]. The resulting sinogram is shown in Fig. 7.2(b). Fig. 7.2(c)-(l) show the central horizontal and vertical cross-sections (white lines in (a)) of the reconstructions for variations of input data, model, and inversion algorithm. Images were reconstructed on a ROI of 3 cm x 3 cm with 201 x 201 pixels using LSQR or the proposed fast constrained inversion algorithm ($n_{iter} = 80$). When using signals of full bandwidth ((c) and (d)), original (black), unconstrained LSQR reconstruction (blue), and constrained reconstruction (red) were almost identical. The RMSD with respect to the original was 0.03 in both cases. When applying a low-pass filter of 1.5 MHz to the signals (8th order chebyshev), mimicking the limited sensitivity of a transducer above a certain frequency, both reconstructions in (e) and (f)

[3]The project was led by L. Ding and the detailed results can be found in Ref. [149].

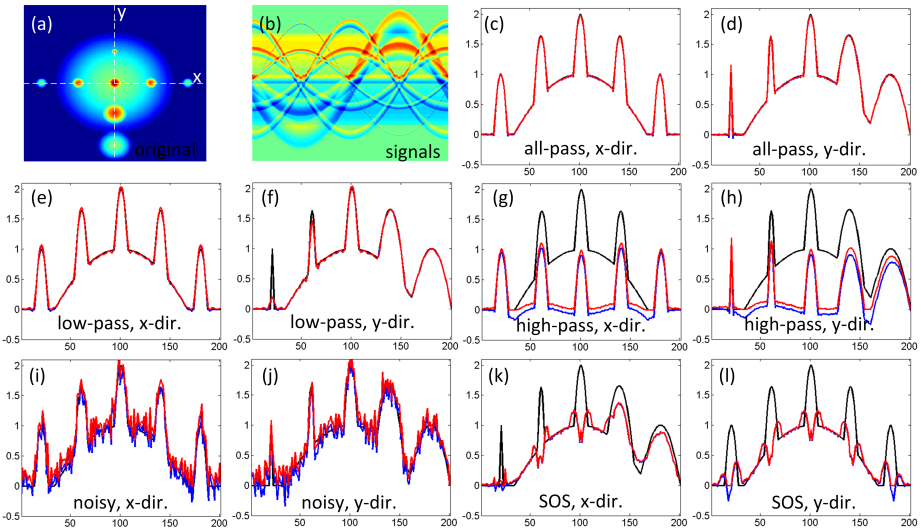

Figure 7.2: Simulation demonstrating different causes of negative image values. (a) Original numerical phantom compromising ten paraboloid absorbers of different location and size. (b) Corresponding analytically calculated pressure signals. (c),(d) x- and y-direction cross-sections of original phantom (black), standard unconstrained LSQR reconstruction (blue), and reconstruction with non-negativity constraint (red). (e),(f) Corresponding cross-sections after low-pass filtering of the pressure signals. (g),(h) Profiles after high-pass filtering of the signals. (i),(j) Profiles reconstructed from signals with 30 % Gaussian noise superimposed. (k),(l) Profiles with deviating SOS assumed during the reconstruction process.

deviated from the original in amplitude and width of the small absorbers, but did not deviate for the larger structures. No larger difference was found between the two inversion methods (RMSD 0.04). However, when a corresponding high-pass filter of 0.1 MHz was applied, mimicking a reduced sensitivity of transducers at very low frequencies, both reconstruction methods considerably deviated from the original ((g) and (h), with RMSD of 0.91 for LSQR and 0.78 for the constrained inversion). Strong image variations, i.e. high spatial image frequencies, were recovered with correct relative magnitude by both algorithms. Yet, the gradual intensity increase, i.e. low spatial image frequencies, could not be restored by either of the two methods. The additional constraint induced an overall shift of the image to the positive. The shift was not globally constant, but smoothly varying with space. Compensation for the missing input data at low frequencies could only be partially achieved. In the presence of strong Gaussian noise of 30 % of the peak-to-peak signal amplitude, large deviations could be found, too ((i) and (j)). All peaks could still be observed since the noise was not restricted to a specific frequency band. The non-negativity constraint led an overall shift of the image to positive image values and deviations from the original were thus effectively increased (RMSD of 0.36 and 0.46). Finally, a 2 % too large SOS was chosen for reconstruction. Only the large, gradually varying structures could be correctly reconstructed in this case ((k) and (l)).

None of the small- or medium-sized structures were reconstructed in correct shape. Although there were small differences between both algorithms, RMSD was equally large (0.24).

Overall, constrained inversion 'successfully' suppressed negative values in all cases; however, qualitative and quantitative difference compared to the ground truth did not improve in all cases. With undisturbed signals and correct model, (almost) no pixel constraints were active and reconstruction results were identical. In the case of band-limited signals, effectively signals needed to be extrapolated to frequencies not measured. A correct extrapolation was not possible over more than relatively small frequency ranges. For low-pass frequency characteristics, high-frequencies were not available and the resolution was consequently degraded. In non-negative inversion, the number of active constraints was very low and differences to unconstrained inversion were small. Conversely, with high-pass characteristics, the number of active constraints was relatively large, and results thus considerably deviated from unconstrained inversion. However, the restoration of low frequencies was only partial and errors were still large. In the presence of strong noise, an additional initial de-noising stage as presented in the previous section would most probably be more effective than a constraint alone and might decrease the overall RMSD rather than increasing it. Finally, deviating acoustic properties degraded image quality irrespective of a constraint applied. Instead, accurate retrieval of correct acoustic properties is more important than a non-negativity constraint to provide good image quality (see also Sect. 7.4 and Sect. 7.5).

Negative pixel values impose great challenges and lead strong restrictions to the quantitative analysis of OA images. First, in individual images considerable qualitative anatomical information might be lost, if hard thresholding of negative image values is performed. Further, ratios of different image ROI intensities might be prone to relatively large errors. Non-negativity constraints might on one hand partially solve the associated problems, but on the other hand cloak the need for other image quality improvements. In temporal monitoring of kinetics, profiles might be inverted or be a complex mixture of different profiles at multiple spatial scales. Most importantly, in multi-spectral imaging, absorption spectra might also be inverted. Blood oxygenation can only be calculated from strictly positive image values. However, even small absolute changes can lead to drastic relative errors in the oxygenation estimation when some of the pixel values at one wavelength are close to zero.

Particularly image quantification with large structures or with smaller structures in their vicinity is very challenging; this is for example the case when measuring perfusion of entire organs. Luckily, small dominating features like blood vessels are less prone to quantification error as long as the resulting resolution in the reconstructions is sufficiently high. Rigorous fitting of analytical light fluence functions or even numerical inversion of a light transport model in QPAT are likely to fail with typical experimental data because of the described low frequency errors. Improvements might be based on hardware using either multiple detection elements with different bandwidth such that also a sensitivity in the low frequency range is achieved [184, 185] or novel detection technology like interferometric detectors. Algorithmic improvements might be based on a combination of residual weighting towards the detector's bandwidth, a (non-)quadratic regularization term, and a non-negativity constraint.

7.3 Efficient Modeling of Finite-Sized Transducers

All detection elements of experimental systems are not point-like, but finite in size as depicted in Sect. 3.3. Reasons for using large detectors are either focusing to approximately a line or a plane, manufacturing restrictions, and/or increase of SNR by increasing the active detection area.

Most reconstruction algorithms do however assume infinitely small, point-like detectors. Image quality with such algorithms is deteriorated, unless the detector shape, the resulting spatially dependent SIR, and the altered signal shape are accounted for during reconstruction. This is particularly the case for image regions where the difference in time-of-flight, ΔTOF, to different regions of the detection element surface becomes comparable to or larger than the time corresponding to the maximal measured OA frequency. Image quality is not drastically affected sufficiently close to the detector's focal point or for much larger distances. But particularly shallow of the focus, the artifacts with large detector elements considerably reduce the achieved spatial resolution and overall image quality.

Multiple approaches have been proposed to reduce SIR-related artifacts like de-blurring methods, deconvolution, or synthetic aperture focusing [117, 186]. Particularly MB algorithms can seamlessly correct for the effects of SIR. Signals measured with finite-sized transducers $p_{SIR}(x_d, t)$ are ordinary OA signals $p_m(x'_d, t)$ integrated over the surface S_{trans} of the transducer, $p_{SIR}(x'_d, t) = \int S_{trans} p_m(x_d, t)$. A corresponding model M_{SIR} can be obtained by convoluting the signals of a point-like model M with the analytically known, spatially dependent SIR of a flat transducer [146]. Alternatively, the detector surface integral can be discretized and M_{SIR} obtained as sum of multiple point-like models M_i, $M_{SIR} = \sum_i M_i$. The latter approach is also possible for matrix-free models. However model generation time scales linearly with the number of detection surface integration points.

Figure 7.3(a) illustrates a typical geometry with a focused transducer of 1.5 cm focal length in 2-D, an image ROI spanning 2 cm around the focal point and a pixel size of Δx_{pix}=400μm. Fig. 7.3(b) plots the corresponding signals excited by the *Pixel1*, *Pixel2*, and *Pixel3* for a point-like (red) and the finite-sized transducer (blue). While *Pixel2* at focus created almost the same signals for both transducer models, *Pixel3* (deeper than focus) created a slightly elongated signal of different shape. In opposite, *Pixel1* shallow of focus excited an entirely altered signal shape and amplitude. Signal duration from the pixel was elongated by a factor of 3.1 x and the sub-matrix of the model M_{SIR} corresponding to pixel1 had memory requirements increased by the same factor of 3.1 x. Overall, the entire SIR model M_{SIR} required 4.6 x the memory of the corresponding point-like model M and reconstruction time was increased by a similar factor.

If a single element of finite size is circularly scanned around the object or a circular parallel array of identical elements is employed, the resulting reconstruction problem is rotationally invariant. The proposed polar reconstruction approach (see Chap. 6) can be used to create a model also accounting for the SIR. Re-using the sub-matrix corresponding to the first projection spares the tedious generation of the SIR model for all further projections. More impor-

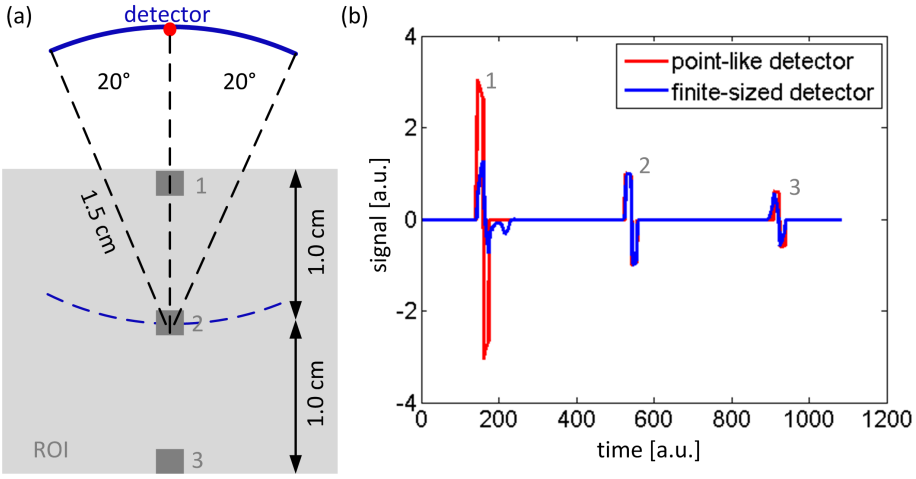

Figure 7.3: Effect of finite-sized transducers. (a) Imaging geometry with a point-like (red) and a circularly focused (blue) transducer and three pixels of interest above, in, and below the focus. (b) Corresponding sensitivity and signals shape. In focus, the signals deviate only slightly, but particularly signals originating from near the surface are considerably elongated and altered in shape.

tantly, the polar model formulation uses dense algebra instead of sparse algebra. Elongation of signal length does thus not decrease model sparsity or increase memory consumption. Images can thus be reconstructed from signals of finite-sized transducers with the same computational performance as in the point-like case (omitting a tiny increase from an altered total span of time-samples considered for the ensemble of all pixels). Moreover, an inverse can be calculated and reconstructions are much faster compared to standard iterative inversion.

To demonstrate the advantages of the polar MB approach for SIR models, a phantom consisting of solidified agar gel containing 200 μm spheres was imaged[4]. The cylindrically focused transducer element was sized 1.3 cm in active diameter with a focal length of 2.5 cm. The element was fully scanned around the object in 2.25° steps. Further experimental details are described elsewhere [187]. To model the in-plane SIR for 2-D reconstructions, the transducer surface was discretized by 120 virtual point-like detectors, and the image consisted of n_r=300 x n_φ=256 pixels in polar discretization. It took 5.1 GBs to store the entire point-like model in standard Cartesian representation and the corresponding SIR model would have required 57 GBs of storage. With the proposed angular frequency polar model, both point-like model M_L and SIR model $M_{L,SIR}$ required only 0.56 GBs of memory for storage.

Figure 7.4 shows reconstructions of one of the slices with the direct inverse for point-like (a) and SIR model (b), respectively. Both polar models required only 0.35 s per slice in total for reconstruction. A comparable iterative LSQR reconstruction (n_{iter} = 15) with standard Cartesian point-like and SIR model formulation would have required approximately 55 s and 600 s per, respectively. The proposed polar reconstruction approach becomes thus even more beneficial

[4]The data-set was kindly provided by D. Queirós.

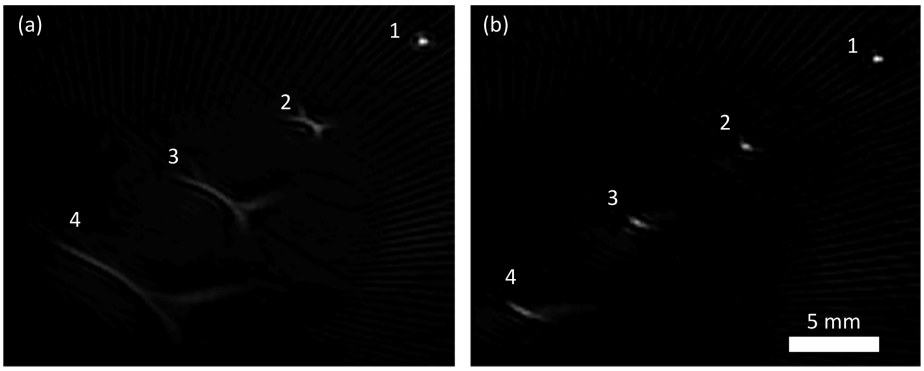

Figure 7.4: Microspheres imaged with a cylindrically focused transducer. (a) Reconstruction assuming point-like transducers. (b) Reconstruction modeling for the detector shape with drastically improved image quality. Signal elongation decreases model sparsity by a factor of 11 x, and standard reconstructions have correspondingly longer inversion times and memory consumption for storage (tens of seconds per iteration and tens of GBs). With the polar approach, reconstructions can be performed at the same computational cost as for point-like detectors (here 0.35 s in total).

when a large detector shape has to be modeled, providing a speed-up of up to two or three orders of magnitude.

The efficiency of the proposed symmetry-based inversion can also be exploited in other geometries. As discussed in Sect. 5.6, z-scanning of a cylindrically focused array (see Sect. 3.4) or raster-scanning of a spherically focused detector in dermal microscopy are commonly employed imaging approaches using large focused detectors. In the latter case, the proposed algorithm might replace commonly employed synthetic aperture focusing algorithms with similar or even improved reconstruction performance. Reconstruction is then limited by the size of the domain for which the model can be stored in memory, not by computational time for inversion. For example, an inverse model for reconstructing a 512 x 512 x 512 voxel grid can be calculated and stored on a modern PC with 128 GBs of RAM. Once the inverse model is obtained (mainly depending on the time to calculate one projection including SIR), inversion can be performed in tens of seconds per volumetric data-set instead of hours or days.

When using matrix-free approaches, the model is re-calculated twice in each iteration. In addition, modeling the SIR requires a repeated model generation for all transducer surface discretization points. Although memory consumption is not affected, performance of the inversion is decrease by at least one order of magnitude (depending on the number of transducer discretization points). In such a situation, reconstructions can be considerably accelerated based on the observation that the Green's function in Eq. 3.7 is translation invariant and depends only on the absolute difference in space. In this case, the order of integration over the transducer surface and signal origination integration can be reversed:

$$
\begin{aligned}
p_{SIR}(x_d,t) &= \int dS'_{trans} p_d(x_d + x', t) \\
&= \int dS'_{trans} \frac{1}{4\pi c} \frac{\partial}{\partial t} \int dA'' \left. \frac{p_0(x'')}{|x'' - x_d - x'|} \right|_{ct=|x''-x_d-x'|} \\
&= \frac{1}{4\pi c} \frac{\partial}{\partial t} \int dA'' \left. \frac{\int dS'_{trans} p_0(x''-x')}{|x''-x_d|} \right|_{ct=|x''-x_d|} \\
&=: \underbrace{\frac{1}{4\pi c} \frac{\partial}{\partial t} \int dA'' \left. \frac{\widetilde{p_0}(x'')}{|x''-x_d|} \right|_{ct=|x''-x_d|}}_{\text{point-like model for modified image } \widetilde{p_0}(x)}
\end{aligned}
\tag{7.2}
$$

Basically, the signals $p_{SIR}(x_d, t)$ for a finite-sized transducer can be obtained using the a point-like detector model, if the original image $p_0(x)$ is integrated along the mirrored detector shape prior to modeling (dashed blue line in Fig. 7.3 (a)), resulting $\widetilde{p_0}(x)$. The applied integration shape can be deviating for different projections because the detector orientation might vary. The additional image convolution operation per projection (inner integral) is computationally fast compared to multiple re-calculations of model projections (outer integral). Thus a considerable acceleration is expected from inverting the integration order. Drawbacks of the method result from an effectively enlarged image ROI and image discretization to square (cubic) pixels. Effects on quality and runtime are as well subject to future work as the approximating character of the method if used with heterogeneous SOS distributions, where an exact translation invariance is not given.

7.4 Residual-Based Speed-of-Sound Auto-Focusing

For highly accurate reconstructed images, the acoustic propagation properties underlying the reconstruction model must match the real acoustic properties of the object. These properties are most often not exactly known. In OA imaging of soft tissues at low frequencies below 10 MHz, the SOS distribution is the acoustic property effecting final reconstructions the most. Although in principle tables with average SOS values for different types of tissue are available [100], the SOS can be measured in transmit/receive or passive US experiments [188], or a linearized SOS map can theoretically be concurrently obtained with the absorption distribution in numerical inversion [189], only a crude SOS range is known for the large majority of experiments. Further, the SOS strongly depends on temperature, which is most often not known or controlled in experiments either. Because of the uncertainty about spatial SOS variations and because of algorithmic limitations, SOS is most often assumed spatially constant. SOS is thus a scalar parameter of the reconstruction problem and needs to be optimally chosen. Visual inspection can be a crude approach, but it is often biased and subject dependent.

The importance of and solution to choosing an appropriate SOS can be understood from the reconstruction process as the result of US wave interference. In the 1-D case of Fig. 7.5(a) with two opposing detectors (red circles), the reconstruction can be thought of as TR at a given single frequency (dashed

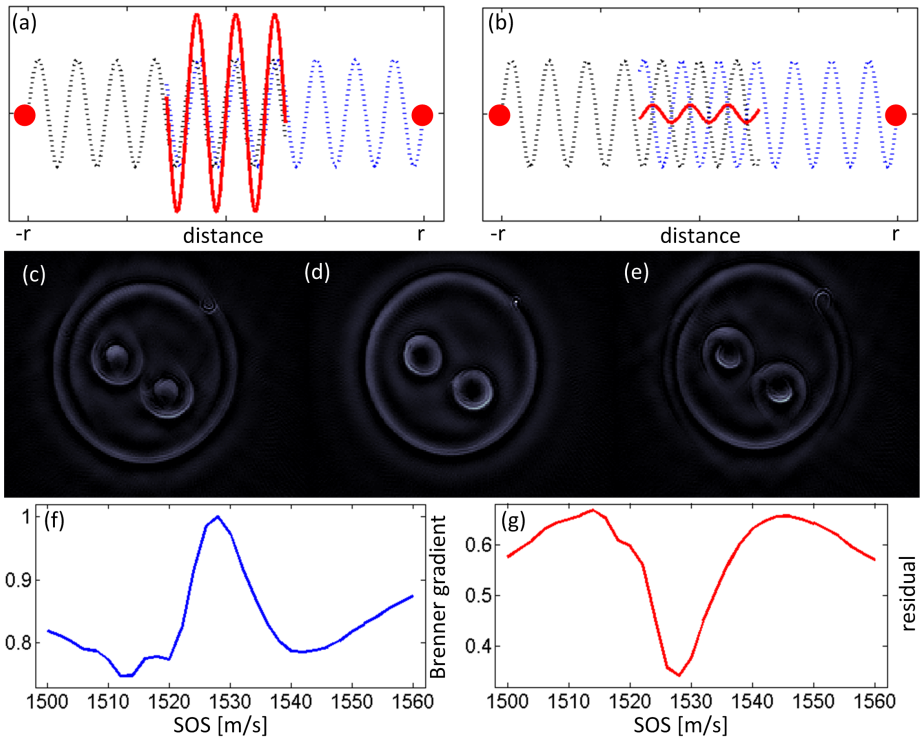

Figure 7.5: Effects and retrieval of SOS. (a) 1-D illustration of the superposition of signals from multiple detectors (red circles): US signals from left (black) and right (blue) with almost fully constructive superposition in the image region (red) for a given SOS. (b) Corresponding illustration with 5 % higher SOS resulting in mainly destructive superposition. Reconstruction of an experimental phantom at (c) 1500 m/s, (d) optimum of 1528 m/s, and (e) 1560 m/s. When using a deviating SOS, small absorbers are smeared and object boundaries do not match. (f) The corresponding Brenner gradient image domain focus functional as function of SOS increasing for images containing more high spatial frequencies. (g) Signal domain residual functional as function of SOS with minimum when signals from multiple projections overlap constructively.

lines). Using an almost optimal SOS value, signals from both detectors interfere almost entirely constructive in the ROI (red). If a 5 % higher SOS is chosen (Fig. 7.5(b)), a relative phase shift in the order of $\pi/2$ is induced and signal interference is almost entirely destructive. The intensity of the resulting image in the ROI is therefore much lower. Real OA signals are broadband and they are a superposition of signals at multiple frequencies. The respective phase shift from a SOS change is proportional to the respective frequency and the high frequency components in the ROI are much more affected by small SOS changes than the low frequency components (see also Fig. 7.2).

Automatic auto-focusing strategies originating from light microscopy have also been applied to OA imaging [190, 191]. They commonly exploit the fact that high frequency image components are much stronger affected by the choice of SOS than the low frequency components. Typically, an image functional is calculated that relates the high frequency content of an image to the low frequency content (or to a constant). One commonly applied measure, the Brenner gradient, is obtained from the norm of the spatial derivatives for all image pixels [191]. Auto-focusing then consists in reconstruction of the set of images for all SOS values within the considered range, calculation of the image focus functional for all images of the set, and selection of the image maximizing the focus functional.

Auto-focusing algorithms in the image domain are based on the reconstructed images and can consequently be sensitive to image artifacts induced (e.g limited view of when using BP reconstructions). MB algorithms offer an additional possibility for auto-focusing that does not rely on image domain focus measures. MB reconstructions are obtained from minimizing the signal residual functional $RR(f) = F(f)/F(f = 0)$ (Eq. 4.4). The optimum SOS can be obtained by concurrently minimizing the functional $RR(f, c)$ for both f and c. The dependence of the residual signals on c can be equally understood from Fig. 7.5(a) and (b). When forward-modeling the image (red) with the correct SOS, the corresponding signals at the detectors are almost identical to the original signals (black and blue) in phase and amplitude. The residual is relatively small then. Conversely, at a deviating SOS, the forward-modeled signals are considerably deviating from the original signals in phase and amplitude. The residual RR becomes relatively large then. SOS auto-focusing via residual minimization thus results the optimum SOS in the same sense that images reconstructed by MB methods are optimal, i.e. the quadratic functional of Eq. 4.4 is minimized.

Figs. 7.5(c)-(e) show the MB reconstruction of a circular agar phantom with ink background absorption and two circular ink insertions of higher absorption with a reconstruction SOS of c=1500 m/s (c), the optimum of c=1528 m/s (d), and c=1560 m/s (e). At considerably wrong SOS value, both outer and inner phantom boundaries exhibited a double structure and the microsphere at the top-right is smeared. At correct SOS, all boundaries were sharp and the microsphere appeared as point-like absorber. Fig. 7.5(f) plots the corresponding Brenner gradient as function of SOS with a distinct, but broad maximum at 1528 m/s. Fig. 7.5(g) plots the corresponding residual $RR(c)$ with distinct minimum at the same value. The exact shape of the focus functional is always dependent on the frequency content of the signals and the reconstruction method employed. For the geometry of the MSOT scanner used, the SOS needs to be determined with accuracy of $\pm$ 2 m/s in order to facilitate reconstructions with resolution of 200 μm or better.

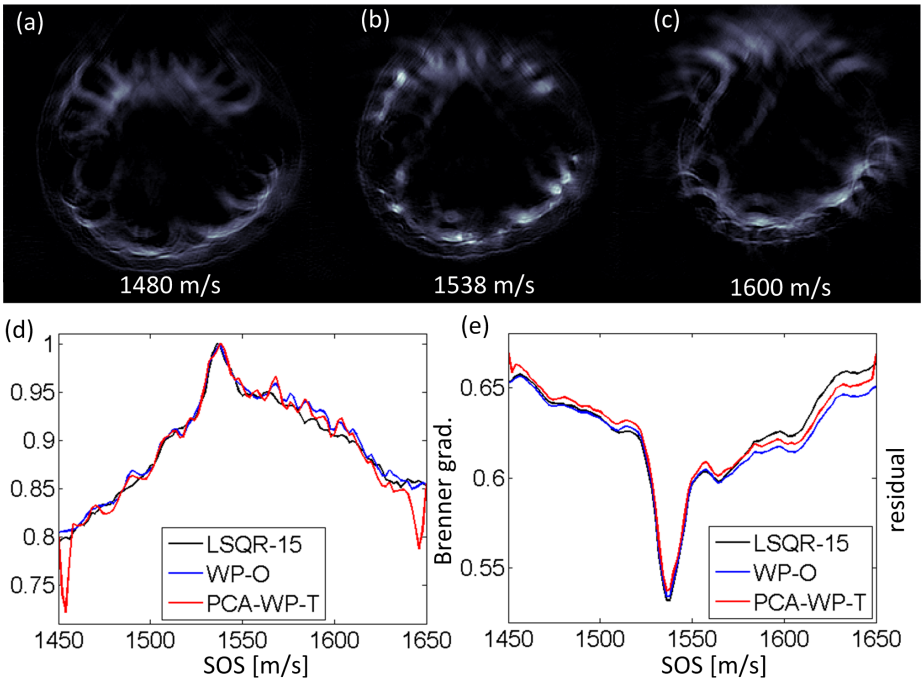

Figure 7.6: SOS auto-focusing with the WP reconstruction approach. PCA-WP-T reconstructions with c=1480 m/s (a), the manually selected visual optimum at c=1538 m/s (b), and c=1600 m/s (c). (d) Brenner gradient focus measure as function of SOS for LSQR-15 (black), WP-O (blue) and PCA-WP-T (red). (e) Corresponding plot for the residual signal energy $RR(c)$ with a minimum at 1536 m/s. Adapted with permission from Ref. [172].

The SOS can in principle be concurrently optimized with the pixel values during iterative inversion. However, the reconstruction problem becomes nonlinear in this case. Moreover, the focus functional depends not always in convex manner on SOS (e.g. for the microspheres in Fig. 7.4 or for dominating small vasculature) and a local minimum might be found instead of the global minimum. Thus the same strategy as in the image domain is employed and first all images for different SOS values are reconstructed and minimization regarding SOS is performed as second step.

However, MB reconstructions are computationally costly. Cost can be decreased by the observation that for a homogeneous medium and point-like detectors, the model (Eq. 4.3) does only depend on the product ct, not on the factors alone. Thus the same model matrix can be used for all SOS values and model re-calculation can be spared. Instead, signals are re-sampled at different time instants, corresponding to a stretching of the time axis based on the SOS change. The signal vectors for different SOS values are highly correlated then. As stated above, low frequency signals do only slightly change, while the main differences manifest in the high frequency signals. Consequently, an effective multi-frame reconstruction problem has to be solved with high inter-frame correlation in the low frequency components. The problem can be efficiently solved

using the proposed sparsity-based WP inversion framework presented Chapter 6.

Figure 7.6 presents the auto-focusing for the 900 nm frame of the *Multi-Spectral Data-set* of Sect. 6.4 with the sparse WP inversion approach. Probing a SOS range from 1450 m/s to 1650 m/s in 2 m/s steps resulted in a data-set of λ=101 effective frames. The PCA-WP-T reconstructions for c=1480 m/s (a), the manually selected visual optimum at c=1538 m/s (b), and c=1600 m/s (c) are shown in the top row. Fig. 7.6(d) plots the Brenner gradient as function of SOS for iterative LSQR (black), WP-O (blue), and PCA-WP-T (red) reconstructions. Fig. 7.6(e) plots the corresponding relative residual as function of SOS. Although there were slight quantitative differences, all reconstruction methods resulted a global extremum at c= 1536 m/s (residual) or c=1538 m/s (Brenner). Because of the complex structure of the object, the minimum of the residual was more pronounced than the maximum of the Brenner gradient. Due to the high level of signal sparsity in combined representation (Gini-index g=0.95), the PCA-WP-T reconstruction process took only 9.6 s compared to 641 s (LSQR-15) and 112 s (WP-O). This was an improvement by factors of 67 x and 12 x, respectively. However, if residual focusing is considered with WP based methods, the residual is not obtained with inversion and an additional forward-modeling step needs to be performed per image, decreasing performance benefits. But still, a sparse formulation of the auto-focusing task is much faster compared to iterative reconstructions.

7.5 Signal Domain Analysis and Segmentation for Optoacoustic Tomography

Selection of a suitable SOS has a large effect on the obtained image quality, as shown in the previous section. The range of SOS and other acoustic properties for different kinds of tissues is wide [100] and in some cases an idealized homogeneous acoustic medium is a too crude approximation. This is the case when the SOS is considerably spatially varying, in particular at the interface of imaged object and coupling medium; or if there is a strong acoustically mismatching boundary like with bones and air-filled cavities, where huge acoustic reflections are induced.

Complex, non-homogenous acoustic properties can also be accounted for during reconstruction with numerical methods. MB methods have been modified for the main effect of small SOS variations, the time delay induced, and optionally also minor effects like amplitude corrections and refraction [147, 192]. Reflections could potentially be modeled for with 'ray-tracing' model calculation [159], performing a branching of the 'rays' at acoustically mismatching, partially reflecting surfaces. FDTD methods intrinsically allow to model for arbitrary SOS and impedance maps, which are direct parameters to the wave equation used [136, 150].

The acoustic parameters used for reconstruction are often set manually in proof-of-concept studies. This approach is however not feasible in clinical practice. Concurrent recovery with the absorption map is possible in principle, but has not been demonstrated with experimental data yet [189]. Reconstructions with an approximated SOS map, segmentation of the image to different com-

partments, and subsequent auto-focusing of the SOS for all compartments is yet another approach. Not only is segmentation of OA images in an immature state compared to segmentation for traditional imaging modalities or photography, but also such an approach would be extremely computationally demanding because several rounds of heterogeneous model generation and reconstruction have to be performed during SOS optimization. Alternatively, acoustic properties might also be determined experimentally with active transmit/receive or passive OA US measurements [188]. However this approach requires specific hardware not available in all OA scanners.

The object's acoustic properties do not only affect the reconstructed images, but are also directly encoded in the signals measured. One might therefore also attempt to estimate the acoustic properties directly from OA signals instead of from the reconstructions in the image domain. In general, this task is equally complicated as in the image domain. But in some cases, like in the case of cross-sectional finger imaging, the experimental situation is sufficiently simple to extract the most relevant acoustic parameters - shape of the finger, SOS within, and outline of the bone - directly from the signals and reconstructions can be performed with the extracted parameters efficiently in one-step manner without a time consuming repeated optimization process including multiple heterogeneous reconstructions.

The proposed signal domain analysis for optimized reconstructions is based on the following assumptions on object and imaging system employed[5]:

- The model can be divided into two main compartments, each with homogeneous SOS c_0 and c_1, respectively. With a suitable choice of compartment boundary and SOS values, a better reconstruction than the best homogeneous reconstruction can be achieved.

- The boundary in the image domain is smooth and convex. In the case considered here, the boundary is characterized by $R(\varphi) = \sum_{j=n_{har}}^{-n_{har}} A_j \exp(ij\varphi)$ plus two constants for x- and y-direction offset.

- The outer compartment has minor optical absorption compared to the inner compartment and its SOS c_0 is assumed to be known (e.g. from temperature measurements). Its shape is however not known *a priori*. The outer compartment is associated with the coupling medium like gel or water.

- The inner compartment has little internal acoustic variations compared to the difference to the outer object. The inner compartment is associated with the object to be imaged. Further, strong OA absorption takes place at the boundary. Its shape and SOS c_1 are not known *a priori*.

- The optional third compartment mainly induces signal reflections and is associated with the bone.

- For a good performance of the method, a significant detection angle is covered, ideally more than 180°.

[5] A detailed description of the proposed method has been published in Photoacoustics as Ref. [193].

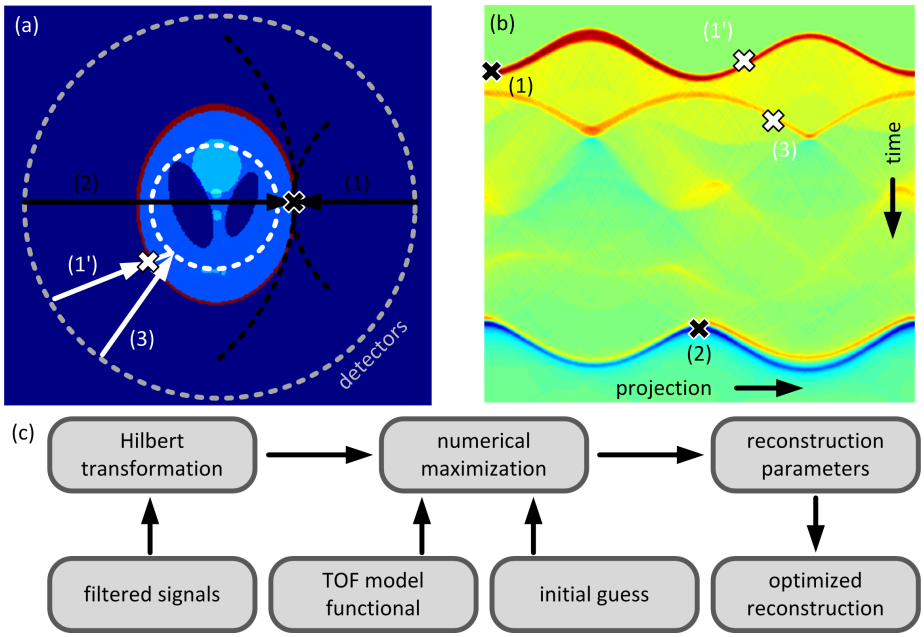

Figure 7.7: Concept of signal domain segmentation. (a) Illustration of OA signal generation with a numerical phantom. At the absorbing boundary (black cross) of the phantom huge signals will be generated for detector locations ((1) and (2)) with an integration arc (dashed black line) tangential to the boundary. Opposite detectors provide partially redundant information and consequently information on the SOS. Accordingly, boundary signals (white cross) with direct (1') and indirect (3) propagation provide information on the location of a reflecting boundary (white dashed line). (b) Corresponding sinogram with signal features corresponding to the image domain features in (a). (c) Workflow of the proposed algorithm: After band-pass frequency filtering, the signals are Hilbert-transformed with respect to time and the absolute value, $p_{hilb}(t, \varphi)$, is taken. Based on a parametrized fit function $TOF(\varphi, m)$, numerically maximizing the functional $F_{TOF}(m) = \sum_{\varphi} p_{hilb}(TOF(\varphi, m), \varphi)$ leads the optimized reconstruction parameters m. The parameters obtained in the signal domain prior to reconstruction are object shape, inner SOS c_1, and reflecting boundary. The optimized parameters are subsequently used for a single, improved reconstruction.

The concept of the signal domain analysis method is illustrated in Fig. 7.7. Distinct image features (Fig. 7.7(a)) correspond to distinct signal features (Fig. 7.7(b)). In analogy to half-time reconstructions [194], where it has been shown that the first and second half of OA signals in the homogeneous full view case are entirely redundant, signals of opposing detectors measure partially redundant information: The strongly absorbing boundary (black cross) orthogonal to the direct path to the director (1) creates large signals that are highly correlated with the signals measured at the opposite director (2). From the two corresponding arrival times and the detection surface diameter, one can calculate the average SOS. Likewise, directly propagating (1') signals from the boundary (white cross) and reflected signals (3) provide information on the location of the reflecting boundary (white dashed circle).

In order to extract the information on the acoustic properties, the original signals are Hilbert-transformed with respect to time on a per-projection basis and the absolute values of the analytical signals are taken. Thus uni-polar signal $p_{hilb}(t,\varphi)$ are generated from the bi-polar original signals $p_m(t,\varphi)$[6]. Next, model functions $TOF(\varphi,m)$ facilitate the calculation of the corresponding signal time-of-arrival of image features according to Fermat's principle at a certain detector location φ in dependance of a set of parameters m. Three different TOF model functions are used: $TOF1(\varphi,m)$ provides the shortest TOF of signals from the boundary with boundary coefficients A_j as parameters m. $TOF2(\varphi,m)$ provides the longest direct signal propagation time from the same boundary with the SOS of the inner compartment c_1 as parameter m. Finally, $TOF3(\varphi,m)$ predicts the TOF of the first signals originating from the object boundary and reflected at the mismatching boundary (characterized by $A'_j =: m$).

Optimized fit parameters m to the models TOF are then obtained numerically by maximizing a functional F_{TOF}, which is given by the sum of the Hilbert-transformed signals over all projections sampled at the time-instants predicted by TOF:

$$m = \underset{m}{\operatorname{argmax}}\left[F_{TOF}(m)\right] := \underset{m}{\operatorname{argmax}}\left[\sum_{\varphi} p_{hilb}(TOF(\varphi,m),\varphi)\right] \tag{7.3}$$

Weighting factors, initial guesses, or constraints can be seamlessly included in the maximization process. The TOF function calculations are implemented numerically and maximization is performed using Matlab's *fminunc*-function. The overall steps of the proposed method are summerized in Fig. 7.7(c).

The proposed methodology was tested with synthetic data using the k-wave toolbox [136]. The ROI consisted of 472 x 472 pixels and the central region of the original absorption map is illustrated in Fig. 7.8(a). The SOS was c_0=1500 m/s outside of the red ellipse and c_1=1700 m/s inside. An acoustic mismatch of twice the impedance on the inside is depicted by the blue circle of 80 pixel radius. Signals were detected at 720 equally spaced projections over 360°. The Hilbert-transformed signals $p_{hilb}(t,\varphi)$ are shown in Fig. 7.8(b) with the fitting curve $TOF1$ superimposed in red. $TOF2$ fitting (green) resulted in an estimate of c_1=1694 m/s, which is only 0.4 % below the actual SOS inside of the 'head'. Fitting with $TOF3$ (blue) resulted a radius of the reflector of 81

[6] Deviating from the rest of the thesis, discrete vectors in this section are denoted with their arguments in brackets, not as indices.

pixels, which was 1.8 % above the original. Fig. 7.8(c) shows the time-reversal reconstruction using the estimated SOS distribution, while Fig. 7.8(d) plots the SOS functional value $TOF2$ for homogeneous (blue) and 2-compartmental (red) model. Finally, Fig. 7.8(e) plots the central horizontal cross section of original (black), best homogeneous (blue), and 2-compartmental (red) reconstruction using the estimated parameters. The 2-compartmental reconstruction deviated less from the original in magnitude and location of the absorbers at an overall improved RMSD (0.39 compared to 0.95).

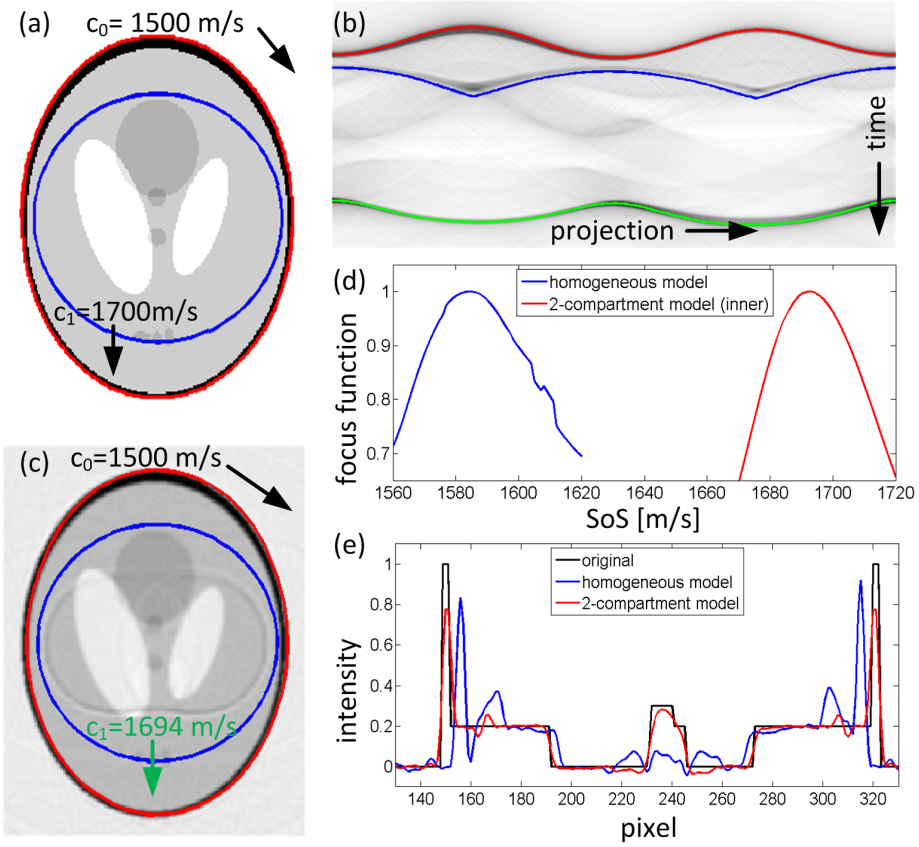

Figure 7.8: Signal domain analysis with numerical data. (a) Original absorption distribution and acoustic properties: The SOS is c_0=1500 m/s outside of the red line and c_1=1700 m/s inside. An acoustic mismatch is indicated by the blue circle (twice the impedance inside). (b) Corresponding Hilbert-transformed sinogram with fitted signal characteristics corresponding to phantom boundary (red), inner SOS (green), and reflector surface (blue). (c) 2-compartmental reconstruction using parameters from signal domain segmentation. (d) SOS focus function for homogeneous and proposed 2-compartmental model. (e) Central horizontal cross-section for original (black), optimized homogeneous (blue), and 2-compartmental model(red).

Performance with experimental, limited-view data was tested with the same ink Phantom as described in Sect. 7.4 and the results are illustrated in Figure

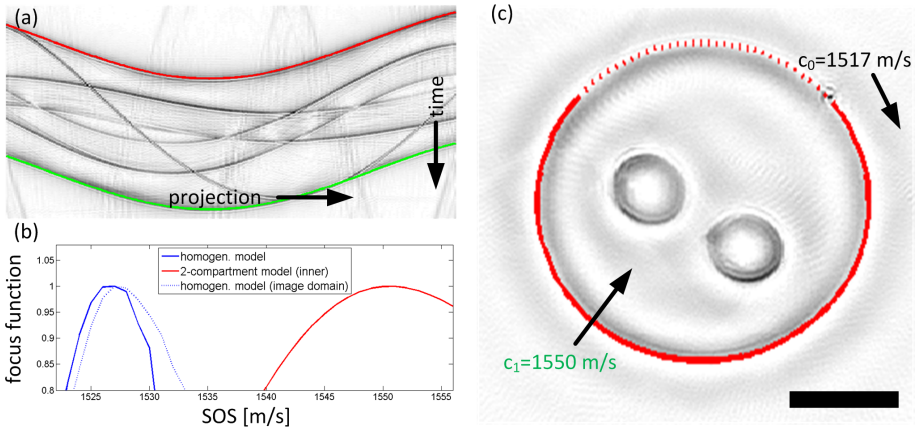

Figure 7.9: Experimental ink phantom study. (a) Hilbert-transformed sinogram of the ink phantom and fits for object outline (red) and SOS (green). (b) SOS focus function for homogeneous (solid blue) and 2-compartmental (red) model in the signal domain and homogeneous post-reconstruction auto-focus functional in the image domain (dotted blue). (c) Reconstruction of the phantom using optimized reconstruction parameters. Scale bar 5 mm.

7.9. Fig. 7.9(a) plots the Hilbert-transformed signals with fitting for object shape ($TOF1$, red) and SOS ($TOF2$, green) superimposed. Assuming an outer SOS of c_0=1517 m/s based on the water temperature of 34 °C, a diameter of 16 mm was estimated, which is 1.1 % larger than the 3/8 inch syringe used to mold the phantom. The SOS estimation curve is plotted in Fig. 7.9(b), having an optimum inner SOS of c_1=1550 m/s. For comparison, the corresponding focus function for the assumption of a homogeneous model (blue) is plotted along with its post-reconstruction residual focus function (dotted blue). Both proposed signal domain analysis and post-reconstruction auto-focusing resulted approximately the same optimum homogeneous SOS (1527 m/s and 1528 m/s). The resulting 2-compartmental MB reconstruction (the heterogeneous model was calculated with an algorithm described in Ref. [159]) is shown in Fig. 7.9(c) with the estimated object shape superimposed in red.

The proposed signal domain analysis was able to recover the most important acoustic reconstruction properties directly from the signals instead of from tedious image domain analysis that requires reconstructed images for all possible parameters tested. Thus, the algorithm was computationally very efficient. This efficiency is of particular importance because heterogeneous models are often associated with higher computational cost compared to homogeneous models. The application of the method is however limited to specific scenarios, where the assumed prerequisites hold true. Most importantly, the method can be applied to cross-sectional imaging of fingers. The corresponding results with the signal domain method are presented in Sect. 8.4. Additionally, it could also be applied in mouse torso imaging or with handheld skin, muscle, or breast imaging of tomographic or raster-scan detection geometry [111, 195]. Furthermore, the method can be easily extended to 3-D imaging based suitable TOF functionals with a 2-D parametrization.

7.6 Novel Potential Approaches of Model-Based Reconstructions for Clinical Imaging

Future reconstruction methods that meet the requirements of clinical imaging will mainly focus on numerical approaches for various reasons. BP methods cannot provide the same image quality as MB reconstructions and the application of BP methods is restricted to situations where numerical inversion is too slow even at lower resolutions or not feasible at all [185]. Series solutions are mainly important for mathematical analysis of the reconstruction problem and in practical use only with planar scanners. TR algorithms offer great flexibility, but are only exact in full view geometries and are most often much slower than matrix-based reconstructions. However, their FDTD propagation model can equally be used as core of iterative inversion at low memory requirements [150].

Novel MB approaches will most probably be diverse, because of the variety of time and quality requirements for the reconstructions, because the specific imaging scenarios considered vary a lot, and because of the high computational cost. A unified approach covering all practical situations is unlikely. Most likely, hybrid strategies based on multiple approaches will be employed. The strategies depend on the possibility to efficiently use GPU platforms instead of CPUs; whether the model can be stored in memory or matrix-free methods need to be employed; if multiple images are reconstructed using the same model (even simultaneously) or just one image per specific model; if (approximate) direct inverses are feasible or if only iterative reconstruction can be performed; if symmetries are present or if the inversion can benefit from sparsity; if and which kind of regularization is required; or if constraints need to be applied. All these considerations will determine the selection criteria of the inversion strategy employed.

Two potential future strategies are outlined in the following, one for highly efficient, high quality reconstructions in 2-D handheld imaging [111]; the other for facilitating efficient reconstructions with complex heterogeneous acoustic media and a 3-D probe with finite detection element size [35].

The first suggested approach executes multiple steps on different scales, both direct and iterative methods, and runs on both CPU and GPU (Fig. 7.10(a)). In a first step, the inverse $(M^T M)^{-1}$ for the normal equations (Eq. 4.6) is calculated on a coarser resolution scale with as much pixels as the available PC memory permits. Because the number of data samples typically exceeds the number of pixels, the size of $(M^T M)^{-1}$ is smaller $(n_x^2 \,\mathrm{x}\, n_x^2)$ than of the pseudo-inverse $M^\dagger$ $(n_x^2 \,\mathrm{x}\, n_t n_\varphi)$. Further, the matrix $M^T M$ is symmetric and positive definite. Consequently, Cholesky factorization can be applied at half the memory requirements and FLOPs involved. Further, if reflection symmetry is present, the memory can be reduced by an additional factor of 2. In single precision, the Cholesky factorized matrix for $400 \,\mathrm{x}\, 400$ pixel images can be stored with less than 30 GBs of RAM. Cholesky factorized matrices are exact inverses. Arbitrary quadratic regularization terms can be included, and reconstruction speed does not degrade in limited view settings. Initial tests revealed that peak performance might be limited by the application of M^T rather than application of the factorized matrix. When considering lower resolutions and GPU implementations, real-time reconstructions seem easily possible. In the second step, the active set of non-negativity constraints is determined. Because

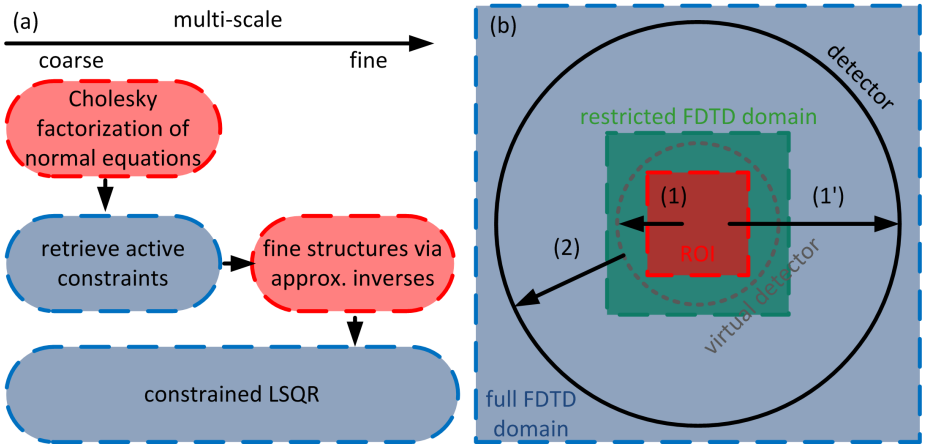

Figure 7.10: Illustration of two potential future reconstruction approaches with high efficiency. (a) Reconstruction can be performed with multiple direct (red) and iterative (blue) methods at different scales. It might include memory intensive direct reconstruction via Cholesky factorization at low scale, efficient iterative retrieval of non-negativity constraints at low scale, calculation of image details with direct WP inverses, and finalization of the image at all scales at a reduced number of constrained LSQR iterations. (b) Hybrid approach with matrix-free FDTD modeling. Instead of modeling the full domain (blue) including ROI (red) and detectors with FDTD methods, FDTD modeling can be performed on a restricted area (green) including ROI and the area of non-ideal acoustic medium. Time-resolved signals are then captured on a virtual detection surface and are in a second step directly propagated to the real detection location using a Green's function. With the hybrid approach, a higher number of ROI image pixels can be modeled at lower computational cost.

constraints are often mainly associated with low image frequencies (see Sect. 7.2), they can be effectively retrieved at low resolution. This step can possibly be performed on GPU. In the third step, details in the fine scales are added by the approximate inverses in WP domain for detail coefficient leafs (exploiting sparsity of the signals as shown in Chapter 6 results in huge performance benefits). Finally, as forth step, a reduced number of LSQR iterations is performed at full resolution to correct for the approximate character of the WP inverses and potentially update the set of constraints at fine scales. Overall, the number of computationally costly iterative steps at full resolution can be drastically reduced and thus performance considerably increased for excellent quality 2-D reconstructions.

The second method tackles reconstructions in 3-D with complex acoustic properties and finite-sized detectors. Typically, the problem is too large to store the matrix, especially because the SIR increases the memory requirements and because no symmetries are present (the model takes hundreds of GBs for storage). Recalculation of the complex model in each iteration with on-thy-fly methods would be extremely costly, even when modeling the effects of heterogeneous SOS and absorption only on much lower scales than the actual image resolution. In particular, modeling the SIR increases calculations by an additional

order of magnitude because multiple point-like models need to be calculated for transducer surface discretization (see Sect. 7.3). Conversely, FDTD methods with iterative inversion can handle complex media and detector SIR intrinsically without additional overhead [150]. They are however associated with long computational runtime. One particular reason for such long computational times with standard piezoelectric detectors (black circle in Fig. 7.10 (b)) is that the imaged object is at a considerable distance from the detection surface S and the domain to be modeled (blue) is much larger than the actual image ROI to be reconstructed (red). Furthermore, the number of time steps to be performed for the whole domain is much larger than the actual signal length. Thus time and memory consumption of full domain FDTD modeling (1') is not ideal. The space between object and detectors is typically filled with water and US propagation is ideal there. Consequently, the modeling problem can be divided in two parts: (1) Standard FDTD modeling is performed on a restricted area (green) that contains the ROI and a virtual detection surface S' surrounding the ROI. Signals are then recorded on the virtual detection surface S' for time-points until the pressure waves have entirely left the restricted domain. In the second step (2), the time-series of recorded virtual signals is propagated further to the real detection elements by means of a Green's function. For each combination of location on virtual and real detection surface, propagation is described by a time delay and an amplitude factor (a sum of several pairs for a finite-sized detector). As block-wise operations, propagation modeling is much faster than modeling it by means FDTD methods.

Considering a full view spherical detection geometry of 4 cm radius, the ROI is often only as small as 2 cm x 2 cm x 2 cm or even less, which is only 1/64th of the domain to be modeled. A restricted domain that contains a spherical virtual detection surface surrounding the ROI is only about 5 times larger than the ROI volume itself. Further, the number of time-steps can be reduced by 40 % because of the reduced propagation distance to be modeled. Most importantly, memory requirements can be reduced by factor of approximately 13 x at the same number of ROI voxels. For acceleration of FDTD methods by GPUs, memory consumption can be a limiting factor. Assuming modern GPUs can perform FDTD with 512 x 512 x 512 voxel grids, the proposed two step approach increases the number of effective ROI voxels possible from 128 x 128 x 128 to 300 x 300 x 300 at almost the same computational cost. When the same number of effective ROI voxels is kept, the proposed two-step method is expected to be up to a factor of 20 x. Thus the method could enable efficient high-resolution volumetric MB imaging with heterogeneous acoustic properties and finite-sized transducers.

7.7 Summary and Outlook

A translation of OA imaging from basic biological studies in the laboratory towards routine clinical imaging brings new opportunities for OA imaging, but also new challenges: Beside modified set-ups like handheld devices or different imaging protocols, the requirements for signal processing and reconstruction approaches become more demanding. A live preview is needed for which MB algorithms are currently too slow. Thus BP reconstruction of considerably lower image quality has to be performed. Similarly, a prompt, accurate analysis of

the data-sets acquired requires fast reconstructions of high quality without long delays. This is challenging to achieve with standard MB approaches owing to the huge size of the data-sets typically involved. Providing live MB visualization and prompt high-resolution, high quality MB images using available standard computational environments thus presents a long standing challenge for clinical OA imaging. Further, user dependent input of reconstruction parameters has to be minimized or better totally avoided in clinical settings, while in proof-of-concept studies manual tweaking might still be acceptable.

In the previous three chapters, multiple approaches have been presented that aim at providing high-quality MB images at superior reconstruction performance without user dependent inputs for clinical imaging.

Two main approaches have been shown to considerably decrease cost of MB reconstructions: First, a symmetry-based method for 2-D tomographic imaging was demonstrated to provide accurate images at a fraction of the reconstruction time of standard MB inversion. For the limited view tomographic system used in this work, a high-resolution direct inverse could be obtained and reconstructions could be accelerated by a factor of 12.7 x. The reduced memory requirements, dense algebra formats involved, and the intrinsically parallel nature of the algorithm facilitated an implementation on GPU with an additional performance increase by a factor of 13 x. High-resolution reconstructions at video-rate have thus been enabled. The basic concepts are also applicable for other kinds of OA scanners and might increase achievable image quality in e.g. dermal imaging applications. Second, a sparsity-based approach has been presented. Suitable basis transformations rendered the model matrix separable, and the obtained inverse model and multi-frame signals sparse. With the resulting direct, sparse reconstruction method, experimental data-sets could be reconstructed 40 x - 700 x faster than standard iterative inversion at comparable image quality. The sparsity-based framework is expected to find application for reconstructions in studies with huge data-sets or for a MB preview in set-ups like 3-D handheld scanners.

The reconstruction and signal processing approaches presented in this chapter have been aiming at providing improved image quality without examiner dependent input required. Some of them could also be combined with the two proposed approaches for acceleration of reconstructions. The image enhancement methods included efficient and automatic removal of noise and computationally efficient suppression of negative image values. Further, a robust residual-based auto-focusing approach was presented that provided improved image quality by retrieving an optimized homogeneous SOS for reconstruction. A generalization to the heterogeneous case with a 2-compartmental model towards superior reconstructions in finger imaging (see Sect. 8.4) was based on a signal domain analysis and automatically yielded a segmentation of the reconstruction and estimates for the reconstruction parameters used.

Several other open problems concerning reconstructions for clinical imaging still have to be tackled and solved in the future. Most importantly, high-resolution MB reconstructions in 3-D are needed, particularly for systems where no symmetries can be exploited or where computationally demanding SIR modeling is required, and thus the system matrix might be too large to be stored. Calculations are further complicated if heterogeneous acoustic properties like 2-compartmental SOS distributions have to be included in the model. More advanced, matrix-free ray-tracing methods or FDTD model are potential methods

with sufficient computational efficiency for (pre-)clinical applications. In particular, matrix-free models bundled with powerful GPU platforms of restricted available memory are expected to be computationally efficient compared to CPU implementations, where model matrices might be stored in memory. The computationally demanding nature of the OA reconstruction process hinders the use of one general approach for all systems and applications; it is rather expected that multiple dedicated methods for different kinds of systems and requirements will be developed. These will probably combine several concepts to facilitate MB reconstructions at the required image quality and with the available resources. Such methods might also be based on the concepts proposed and/or used in the context of this work.

Chapter 8

Multi-Spectral Optoacoustic Tomography of Fingers

The unique design of the human hand with its digits has provided mankind unrivaled dexterity and versatility. Conversely, loss of functionality comes with great limitations, as fingers are also affected by several prevalent diseases. Potentially, some of those diseases might also be diagnosed by MSOT. This chapter provides basic finger anatomy and common diseases in their vasculature, depicts the OA scanner used for finger imaging in humans, and presents results of initial anatomical and functional imaging studies. Finally, Chapter 9 presents the results of a perfusion study towards OA diagnosis of rheumatoid arthritis (RA).

8.1 Brief Introduction to the Anatomy of Fingers

The human hand compromises five digits with dimensions in the order of 2 cm diameter and 8 cm length. The thumb, index finger, middle finger, ring finger, and little finger are enumerated and abbreviated by D1 to D5. The basic finger anatomy is depicted in the labeled MRI slices of the index finger of the author's right hand[1] (see Fig. 8.1). Each digit compromises three bones, the phalanges, with the exception of the thumb that compromises only two bones. The bones of the finger are connected to each other and to the metacarpal bone of the hand via joints: The distal interphalangeal joint (DIP) links distal and intermediate phalanges, the proximal interphalangeal joint (PIP) links intermediate and proximal phalanges, and the metacarpal (MCP) joint links proximal phalanx and metacarpal bone. The synovia of the joint consists of the synovial membrane and the synovial fluid. Fingers can be moved with the help of tendons for flexion and extension and several nerves are facilitate motion control and haptics.

The vasculature of the fingers provides the blood supply fed through two main arteries: The radial (art. rad.) and the ulnar (art. uln.) digital artery[2] are both located laterally. Depending on which is finger considered, the arteries branch from the superficial palmar arch or the deep palmar arch. In the

[1] Dr. Reinhard Meier is acknowledged for acquisition of the MRI data at Klinikum Rechts der Isar.

[2] The radial and ulnar artery are also called proper digital arteries (art. dig. prop.).

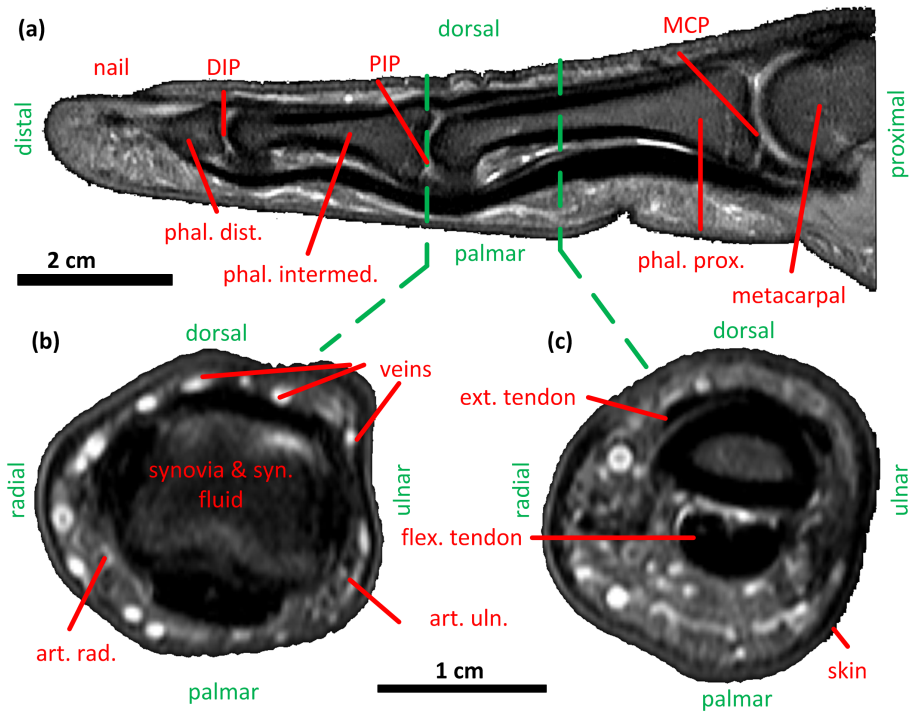

Figure 8.1: Visualization of finger anatomy using MRI slices with anatomical locations (green) and structures (red) labeled. (a) Central sagittal slice of the right index finger. (b) Corresponding transversal slice at the PIP joint. (c) Transversal slice through the phal. prox.
phal.: phalangeal bone; DIP: distal interphalangeal joint; PIP: proximal interphalangeal joint; MCP: metacarpal joint; art. rad.: radial artery; art. uln.: ulnar artery; ext.: extensor; flex.: flexor.

index finger, art. rad. branches from the deep palmar arch while the art. uln. branches from the superficial palmar arch. Close to the joints regular dorsal branches originate symmetrically from both main finger arteries. The diameter of the feeding digital arteries varies considerably between individuals and decreases for distal parts of the finger. The diameter is in the order of 1 mm [196]. Around the joint regions and in the finger tip, rich capillary structures are found. The venous structure is however less regular. Multiple major veins maintain the blood transport backward. They are located on all sides of the finger and relatively superficial compared to the main arteries. The nail is an anatomical structure with a high acoustic mismatch. However, a rich capillary structure is found in its vicinity in the nail bed and of interest in capillaroscopy [197], potentially also by means of OA imaging.

The skin is the second major source of intrinsic contrast besides the blood pool [111, 195, 198]. The skin compromises multiple layers of which the epidermis (most superficial) and the dermis are most important. At the epidermal-dermal junction small vascular structures might be identified, provided sufficient resolution. The thickness of the epidermal layer depends on site, but generally

is less than 250 μm and thus requires high frequency OA or OCT systems to be resolved [195, 199]. Melanin is an omnipresent intrinsic absorber in the skin. It has strong absorption in the entire NIR range with a flat, slightly decreasing spectrum (see Fig. 2.5). It is both a target chromophore and a major challenge for delivering light to deeper regions in the finger.

8.2 Common Peripheral Vascular Diseases in the Finger

Cardiovascular diseases are the number one cause of mortality in western societies. Although diseases directly affecting the heart are most life-threatening, widespread peripheral vascular diseases (PVD) have severe implications for patients, diagnosis, and treatment nevertheless. PVDs mainly affect the extremities and a great number is found in the finger vasculature as well. The intrinsic, label-free sensitivity to blood contrast constitutes MSOT an ideal candidate for their clinical diagnosis. An incomplete, incomprehensive overview of PVDs in the finger, mainly affecting the arteries, is given in the following (based on Ref. [200]).

The general indication ischemia is associated with a diminished blood supply and might be caused by several factors. Among them are a stenosis, a locally reduced vessel lumen, or even a complete occlusion. Ischemia might result in hypoxia of large tissue regions or ultimately even in ulceration and necrosis. Ischemic patients are typically imaged determining the vessel lumen, the blood flow via Doppler US, or by visualizing the perfusion of a CA in the vasculature.

Thromboangiitis obliterans (TAO) is an inflammatory, pathological, non-classified peripheral vascular disease often also related to Raynaud's phenomenon (RP). It is associated with strong pain, stenosis, and necrosis. To maintain the blood supply distal of a stenosis, collateral networks are formed. In TAO, the vasculature is commonly found to have a cork-screw structure near the stenosis which is an indicator for diagnosing the disease.

RP is a functional disease characterized by acute vasospasms. It mainly affects the fingers (D2-D5) and hypoperfusion results in a color change of the fingers. Up to 30 % of females are affected by RP. It might be classified primary, merely a functional disorder without anatomical changes, or secondary to another, anatomical PVD such as TAO. RP is typically diagnosed via thermal stress tests or nail bed capillaroscopy.

Atherosclerosis is characterized by the formation of fibrous or lipid-rich plaques at the vessel wall and often followed by a subsequent calcification. The reduced vessel lumen results in a stenosis and the rupture of the plaques might result in occlusions or in necrosis for very severe cases.

The autoimmune disease RA is discussed in detail in Sect. 9.1.

Besides arterial PVDs, venous PVDs are common as well. Most frequent is varicosis, where the deep veins are insufficient in transporting blood and the enlarged, crooked superficial veins need to compensate. Although varicosis imaging and monitoring its treatment is an highly interesting field for OA imaging in general,, the vast majority of varicosis is found in the lower extremities only, not in fingers.

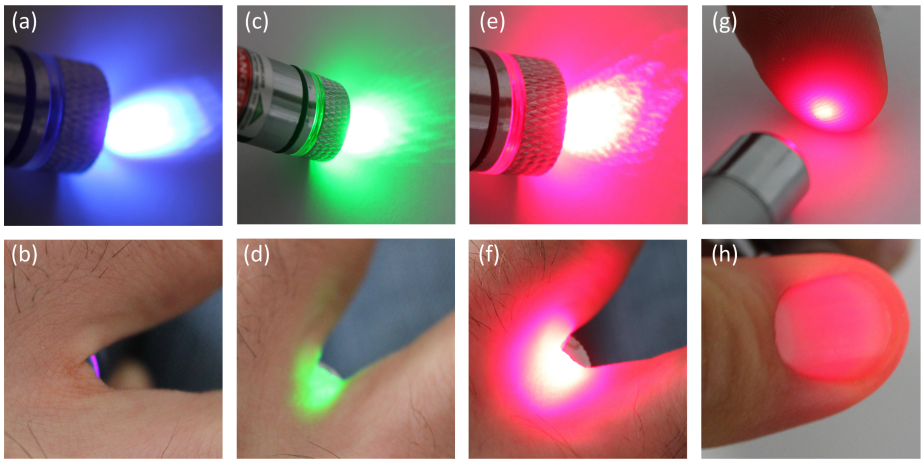

Figure 8.2: Light penetration and scattering in fingers: (a)-(f) Illustration of illumination (top row) with laser pointers and transmitted light (bottom row) through the the skin between the fingers for different wavelengths. Opposed to higher energetic photons, enhanced imaging depth is possible in the NIR window. (g),(h) Illustration of light scattering where the relatively focused beam is entirely diffused after passing through the finger tip.

8.3 Modification of the MSOT Scanner for Finger Imaging

MSOT is a highly attractive optical imaging technology for clinical finger imaging because of several reasons: The dimensions of the finger are too large for macroscopic imaging with ballistic photons and they even exceed the possible imaging depth of OCT. But diffuse light in the NIR can penetrate and thus illuminate the entire finger (Fig. 8.2). Furthermore, fingers are accessible from 360° and imaging approaches are not restricted to limited view geometries and reflection mode imaging. A variety of intrinsic absorbers of interest - Hb, HbO, and melanin - is present while fluorescence imaging technologies like FMT require administering external CAs. The major drawback of MSOT is its sensitivity to the acoustically mismatching bones in the center of the finger that induce artifacts and decrease overall image quality (see Sect. 8.4).

Cross-sectional finger imaging can be performed with the small animal MSOT scanner described in Sect. 3.4. To ensure optimal coupling, the finger is immersed in water heated to 34 °C in order to prevent cooling and to minimize SOS differences with respect to the tissue. With the OPO laser employed, the resulting light fluence on the surface of the finger is 16 $\mathrm{mJ/cm^2}$ at maximum, which is below the ANSI maximum permissible exposure limit of 20 $\mathrm{mJ/cm^2}$ in humans [110]. Orientation of the imaging planes is mainly possible in two ways. First, acquisition of sagittal or coronal slices is possible when the finger is inserted from the top through the part of the detection circle not covered by the array (Fig. 8.3(a)). Anatomical restrictions hinder imaging for large parts of the finger and thus this configuration is limited to imaging in the finger tip. Second, axial slices can be acquired when the imaged finger in stretched posi-

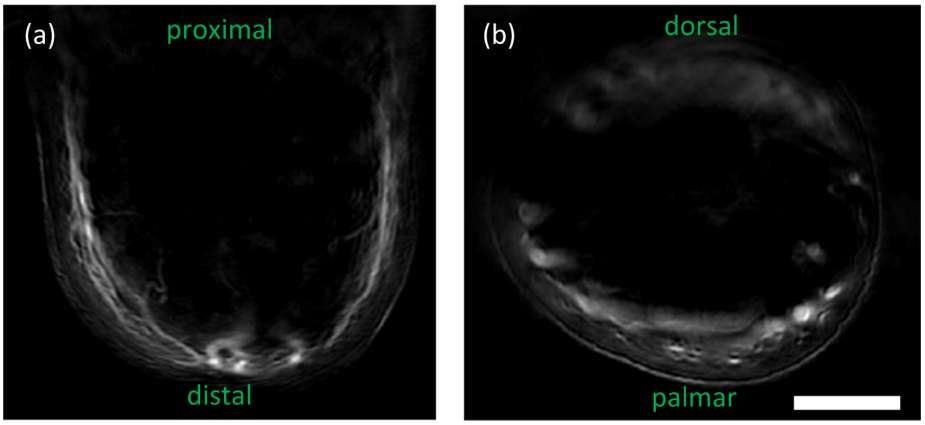

Figure 8.3: OA reconstructions of (a) a coronal slice in the finger tip and (b) an axial slice near the PIP joint. Scale bar 5 mm.

tion intersects the imaging plane, while all others are bend towards the palm (Fig. 8.3(b)). For all five digits, cross-sectional images from the finger tip to almost the area of the MCP joint can be obtained. When using a limited view scanner, the in-plane orientation of the finger can be rotated to provide best image quality at the palmar, dorsal, or lateral side of the finger.

One major challenge in finger imaging experiments is associated with the motion of the finger. Although single laser pulse imaging renders individual frames free of motion artifacts, motion hinders the correct analysis of experiments with frame averaging, temporal monitoring, spectral un-mixing, or volumetric stacking of the planes. Motion artifacts can be addressed by either using mechanical fixation of the finger during the experiment or post-reconstruction computations. With software approaches, parts of the motion related effects cannot be compensated for, including motion of the imaged plane, tilting of the finger, or non-isotropic image quality in the ROI (e.g. because of detector focusing characteristics or because of a non-homogeneous illumination pattern). Even with software processing it is thus crucial to fix the imaging plane and to reduce in-plane motion. After reconstruction, image frames have to be aligned to a reference frame regarding in-plane translation and rotation of a rigid body object. Image co-registration algorithms for this purpose can be categorized in intensity-based and feature-based approaches [201]. The first class aims at maximizing a functional like cross-correlation or mutual information of both images with respect to the transformation parameters. The second class performs co-registration in typically four steps: (1) Feature extraction from both images; (2) feature matching to find corresponding features in both images; (3) robust image transformation parameter estimation from the list of matching coordinates; (4) transformation and re-sampling of the images.

Figure 8.4 illustrates the motion related artifacts and motion correction and prevention. Fig. 8.4(a) shows the motion-free cross-sectional reconstruction of a single frame at 800 nm. Fig. 8.4(b) shows the temporal MIP over 10 s with manual fixation of the finger by the other hand. Due to strong motion, vascular features were smeared. Frames were then co-registered using the feature based ASIFT algorithm [202]. Due to the dominant vascular structure in the

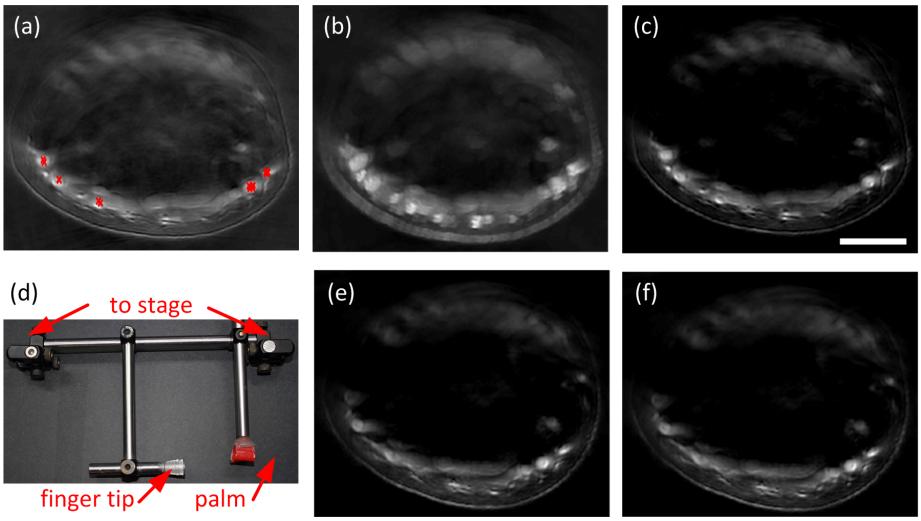

Figure 8.4: Motion correction and prevention in OA finger imaging. (a) Individual frame showing a cross-section with the detected matching key points plotted in red. (b) Uncorrected MIP over 10 s with manual fixation of the finger. (c) Corresponding MIP after motion correction via ASIFT post-processing. The in-plane motion estimated from the average distance between matching ASIFT key points was approximately 350 μm. (c) Photograph of the holder system to fix the finger at the tip, hand palm, and near the MCP joint so that a large range of axial slices is possible to acquire. (e) The MIP over 10 s with holder fixation is almost free of motion artifacts. (e) Corresponding MIP after additional motion correction with a motion estimate of 25 μm.

images, feature-based algorithms generally showed a better performance than intensity-based methods. Performance of the co-registration was mainly determined by the number of correct matches found by the algorithm (red crosses in Fig. 8.4(a)). Fig. 8.4(c) shows the corresponding MIP after ASIFT correction. The average translational motion could be estimated to be 350 μm and thus small features are much better resolved in the final MIP after correction.

Mechanical fixation was realized using a custom-designed holder system (Fig. 8.4 (d)). The fixation possibilities are limited by the minimal comfort acceptable for the volunteer during the imaging session and mainly by the restriction that no solid parts may directly intersect the imaging plane. The designed holder fixed the finger at its tip, near the MCP joint at dorsal side, and at the hand palm. Thus robust imaging for all axial slices from close to the tip to close to the MCP joint of the finger was possible. Further, fixation could be adapted to differences between individuals in length and diameter of the finger. The holder was directly attached to the linear translation stage of the scanner and thus convenient positioning in all three spatial dimensions was possible with sub-resolution accuracy. In this way, also volumetric scanning of the entire finger was possible (see Sect. 8.4). Fig. 8.4(e) shows a MIP over 10 s when fixing the finger with the holder system. Motion artifacts were considerably reduced. Results after a further ASIFT software correction are shown in Fig. 8.4(f) with an estimated average motion of 25 μm only. Motion was thus limited to a level

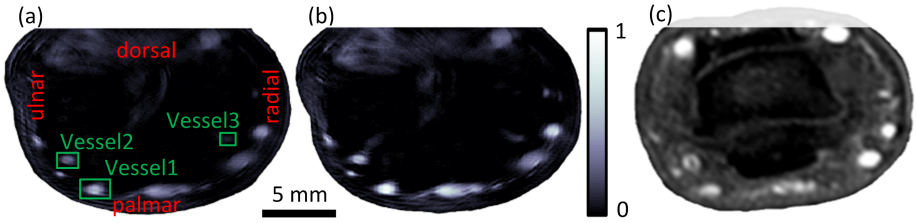

Figure 8.5: Cross-sectional OA imaging of finger anatomy. Slices in the PIP joint (a) at 730 nm and (b) 850 nm excitation wavelength show mainly the vasculature and the skin. The ROIs of three selected vessels are indicated by the green boxes. (c) Corresponding MRI slice with direct correspondence of all major vessels and good agreement in their location and size.

where computationally expensive correction could be omitted in most application scenarios and experiments, depending on motion induced and accuracy required.

8.4 Anatomical Imaging of Fingers

The OA finger renderings presented in this thesis were acquired imaging the index fingers of two volunteers: *Volunteer1* was a 28 year old healthy female and the index finger (D2) of her left hand was scanned. In this experiment, the scanner with 128 detection elements was used (see Sect. 3.4). The results of the experiment with *Volunteer1* are presented in Fig. 8.5, Fig. 8.9, and the corresponding text as well as in Chapter 9. *Volunteer2* was the author and the index finger of his right hand was scanned. All other reconstructions presented in this thesis are based on these experiments. Before OA scanning, a MRI[3] scan with in-plane resolution of 0.4 mm was first performed for the purpose of cross-validation of the images using a 3 T MRI scanner (Verio, Siemens, Erlangen,Germany) and a flexible surface coil (flex large, Siemens, Erlangen, Germany). For this, the volunteers were examined in a prone position with the hands outstretched in a praying position. All human studies were approved by the Ethical Committee of the University Hospital Klinikum Rechts der Isar of the Technical University of Munich and the experiments were conducted according to the principles of the Declaration of Helsinki, including written informed consent of the volunteers.

Anatomical cross-sectional slices

Figure 8.5 shows the anatomical finger cross-sections in the PIP joint acquired during the experiment with *Volunteer1*[4]. In the reconstruction with 730 nm excitation wavelength in Fig. 8.5(a), both skin and vasculature are equally prominent, while at 850 nm excitation wavelength the vasculature is dominant in Fig. 8.5(b). Multiple vessels of differing size could be visualized, mainly at the palmar and the dorsal side of the finger. The vessels at the dorsal

[3]The MRI scans were performed by Dr. Reinhard Meier at Klinikum Rechts der Isar.
[4]The findings of the experiment were previously published as Ref. [126].

side were considerably deteriorated due to the limited detector coverage and the presence of the bone. Fig. 8.5(c) visualizes the corresponding MRI slice. Notably, all major vessels in the OA images had a direct correspondence in the MRI slice with excellent agreement in location and size. Three exemplary vessels were selected for detailed ROI analysis (green boxes). Based on location, oxygenation (see Sect. 8.5), and perfusion (see Sect. 9.2), *Vessel1* could be identified as one of the veins, *Vessel2* as the art. uln., and *Vessel3* as the art. rad. For comparison of the vessel lumen in the OA and MRI reconstructions, a 2-D Gaussian was fitted to the respective vessels in the images and the lumen was calculated from the area at FWHM. The summarized results in Tab. 8.1 confirm the good agreement between both modalities for anatomical imaging and further a good agreement with *ex vivo* measurements reported in the literature [203].

ROI	*Vessel1*	*Vessel2*	*Vessel3*
vessel type	vein	art. uln.	art. rad.
lumen MSOT [mm^2]	4.4	2.9	1.8
lumen MRI [mm^2]	5.4	2.7	2.1

Table 8.1: Comparison of vessel lumen in MSOT and MRI slices.

Optimized reconstructions via signal domain analysis

The achievable OA image quality in the finger is degraded through the presence of the acoustically mismatching bone, as shown in Fig. 8.5 and discussed in Sect. 7.5. Also, the selection of an optimal reconstruction SOS distribution drastically influences the resulting image quality (see. Sect. 7.4 and Sect. 7.5). The previously presented reconstructions have been obtained with standard homogeneous models and manual SOS selection or residual-based auto-focusing [156, 159].

The results of the finger imaging with the developed signal domain analysis approach described in Sect. 7.5 are summarized in Fig. 8.6. In the experiment, a cross-sectional slice located approximately in the middle of the phal. intermed. was imaged at 690 nm and 900 nm excitation wavelength with the dorsal side of the finger facing the detection array. Fig. 8.6(a) shows the optimum homogeneous reconstruction at 900 nm with a SOS of c=1542 m/s which was obtained from residual-based auto-focusing. Because of the SOS difference between finger and water, deeper located vasculature was not optimally focused using a homogeneous model (green circles). The effects of the bone were mainly lack of transmitted signals from the palmar side (not covered by the detection surface) and reflection artifacts from the signals of the dorsal vasculature (blue arrows). Fig. 8.6(b) presents the Hilbert-transformed sinogram at 690 nm excitation wavelength, where melanin was prominent over hemoglobin (see Fig. 2.5). The fits for finger shape (red, $TOF1$), inner SOS (green, $TOF2$), and bone location (red, $TOF3$) are superimposed. Because of limited view detection and the alignment of skin, bone, and detector, the reflected signals could not be faithfully obtained for certain projections (dashed line segments). Fig. 8.6(c) presents the optimized reconstruction with finger shape (red) and inner SOS estimates of c_1=1610 m/s obtained from pre-reconstruction signal domain

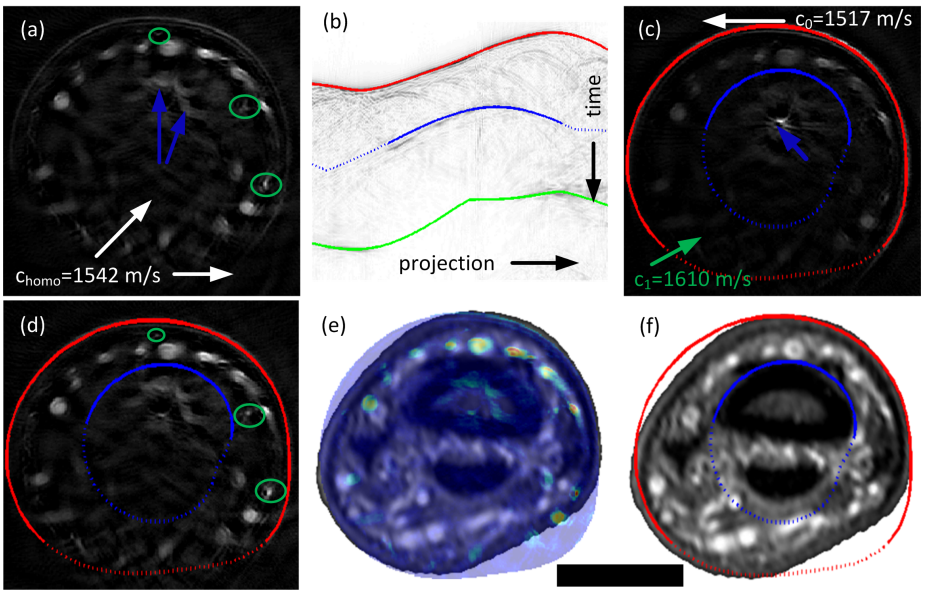

Figure 8.6: Signal domain analysis for optimized finger reconstructions. (a) Cross-sectional reconstruction of a finger at 900 nm excitation wavelength with best homogeneous SOS (1542 m/s). Blue arrows represent reflection artifacts induced by the bone. (b) Hilbert-transformed sinogram at 690 nm excitation with fits to estimate finger shape (red), inner SOS (green), and bone outline (blue). (c) Corresponding reconstruction using optimized model parameters obtained from signal domain segmentation. (d) Same reconstruction for 900 nm excitation signals with improved image quality over the homogeneous model (green circles). (e) Fusion of the corresponding MRI slice (grayscale) and the optoacoustic reconstruction (color) of (d), demonstrating the good agreement between skin and vasculature acquired with the two different modalities. (f) MRI slice with estimated skin (red) and bone (blue) location superimposed. Scale bar 1 cm.

analysis. The estimated shape of the bone with only minor optical absorption compared to the blood pool is plotted in addition (red). Fig. 8.6(d) shows the corresponding 900 nm reconstruction using the same heterogeneous model as for the 690 nm case. Opposed to the homogeneous reconstruction (a), also the small deep vasculature was properly focused (green circles), not only the more shallow vasculature. Finally, Figs. 8.6(e) and (f) present a cross-validation with the corresponding MRI slice. In Fig 8.6(e), the OA image of Fig 8.6(d) is superimposed in color onto the MRI slice (grayscale). The alignment was performed manually with dorsal vasculature as reference points. Fig. 8.6(f) plots the estimates of finger and bone shapes onto the MRI. On the dorsal part, the skin estimate agreed well with the MRI, while large deviations were found in the palmar part because of the limited detection view in MSOT and soft tissue deformations during MRI. The estimated bone shape fitted the location in the MRI equally well in the dorsal part of the finger. Deviations in the palmar part can be understood from incorrect shape extrapolation due to the limited view and from the absence of strong reflected signals for lateral parts of the detector.

The proposed signal domain analysis approach was not only computationally efficient compared to image domain methods. Moreover it was able to automatically provide improved image quality without user input required. This is an important aspect towards clinical translation of advanced reconstruction methods. With a bone estimate available, reflection artifacts might be removed using a 3-compartmental model that also accounts for reflections from an acoustically mismatching boundary [150]. The bone estimate might further ease automatic ROI analysis with RA, where inflamed sites in the synovia are located close to the bones in the finger joint (see. Sect 9.4). Furthermore, the signal domain approach is not limited to finger imaging only, but also other applications imaging through the skin like breast cancer screening or carotid imaging are possible. SOS estimates c_1 for the inner compartment cannot only be obtained in transmission mode, but also in reflection mode from dominant vessels as guide stars. Bone reflection might also occur for other imaging regions like with the ulnar bone or the ribs and the proposed methodology might be applied in this case as well.

Anatomical imaging of volumes

Using the cross-sectional OA scanner, naive imaging provided only 2-D slices as only signals approximately originating from a single imaging plane were acquired. By z-scanning of several slices of the finger, a volumetric representation of a finger ROI could be obtained by either stacking the individually reconstructed slices or by a combined 3-D reconstruction of the entire data-set. To obtain a volumetric representation, the index finger was attached to the finger holder and scanned over a range of 4.4 cm in 0.2 mm steps from the nail bed to the PIP joint. Six excitation wavelengths (715 nm, 730 nm, 760 nm, 800 nm, 850 nm, and 900 nm) were used and the total acquisition time of the data-set was approximately 6 min.

Figure 8.7 presents the result of the experiment. Figs. 8.7(a)-(d) show different slices at 715 nm excitation wavelength near the nail bed (a), in the DIP joint (b), in the middle of the phal. intermed. (c), and near the PIP joint (d). In the four slices, the different shape and dimension of the finger as well as structure and diameter of the vasculature could be observed. Because one global, homogeneous SOS was used for all slices, not all vessels in all slices were focused properly. Fig. 8.7(e) provides a photograph of the same finger with the position of the slices in (a)-(d) indicated in green. A MIP[5] of the 3-D stack of the entire 219 slices is shown in Fig. 8.7(f). The elongated structure of the whole vascular tree over the entire finger with vessels of different sizes could be visualized. Furthermore, branching of vessels at multiple sizes could be observed as well. Subsequent to performing the reconstructions for all wavelengths, the volumes were un-mixed to oxygenated and deoxygenated hemoglobin using their known absorption spectra and linear inversion with non-negativity constraint. The volumes of both components were then processed in volume rendering mode of ImageJ. The two rendered volumes are shown in Fig. 8.7(g) and (h), respectively. In the HbO channel, mainly the vasculature was visible and vessels could

[5]For volumetric renderings, the VolumeViewer of ImageJ was used. No advanced image processing besides transparency threshold level adjustment was performed. Subfigures (g) and (h) were gratefully rendered by M. Kneipp.

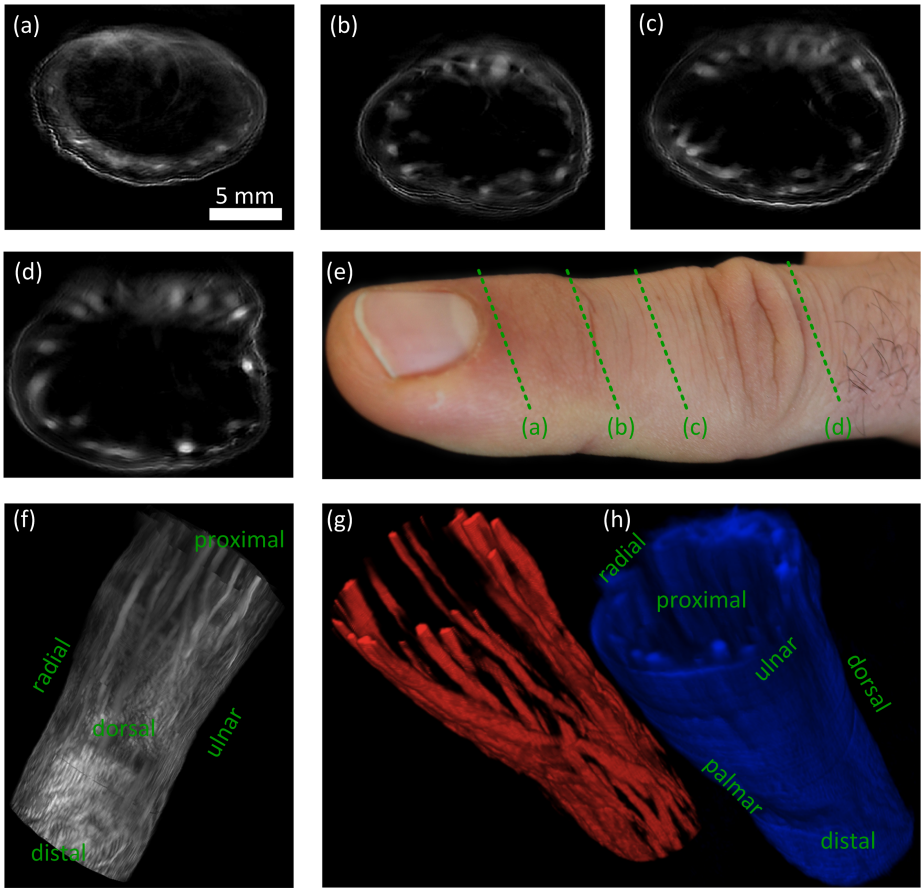

Figure 8.7: Volumetric MSOT imaging of fingers. Selected slices at 715 nm excitation wavelength (a) near the nail bed, (b) in the DIP joint, (c) approximately in the middle of the phal. intermed., and (d) near the PIP joint. (e) The photograph of the finger indicates the location of the slices in (a)-(d). (f) Dorsal view of the MIP of the volume obtained from stacking all 219 slices with both vascular and skin visible. (g) In the volumetric view of the HbO component after un-mixing, the vascular tree of the finger is visualized. (h) Corresponding Hb component with both vasculature and skin visible.

be tracked along the finger. In the Hb channel, both vasculature and skin were visible because of the similar spectra of deoxygenated hemoglobin and melanin (see Fig. 2.5). From the melanin contribution in the skin, the general shape of the whole finger could be clearly recognized.

Advanced image post-processing and segmentation algorithms might further improve the visualization quality and more importantly facilitate a detailed quantitative analysis of the data-set. For example, vessels could be automatically tracked over the entire finger, their diameter determined, and their branching studied. Concerning the multi-spectral aspect, un-mixing might be improved by explicitly accounting for further intrinsic chromophores like melanin. An explicit calculation of oxygenation level (see Sect. 8.5 and discussion therein) could allow to distinguish arteries and veins. By tracking of the recovered oxygenation saturation level over the entire vessel, effects like spectral coloring from varying vessel depth or mixing of venous blood at branches with inflow at different oxygenation saturation levels might be studied. Volumetric imaging of the vasculature might become a viable clinical alternative to DSA X-ray angiography in several applications. For example, volumetric angiography might visualize the altered vascular structure in RA (see Chapter 9) or in other diseases like TAO. In TAO, arteries are not of straight elongated shape, but they are rather winded with a corkscrew structure. At sites of highest curvature, occlusions often occur. MSOT might visualize and quantify those sites based on the altered shape of the vasculature.

8.5 Functional Imaging of Fingers

The excellent spatio-temporal resolution of MSOT in combination with its spectroscopic capabilities permits not only acquiring anatomical reference images, but also permits probing functional parameters. While many more functional parameters can be alternatively targeted for, initial imaging studies for three important functional (vital) parameters in fingers are presented in the following.

Pulse detection

The pulse rate is one of the most important vital parameters to monitor the physiological conditions and well-being of patients. The pulse can be manually palpated for example in the carotid arteries or in the radial artery of the arm. The pulse rate can also be determined during blood pressure measurements or during pulse-oxymetry. Pulse measurements provide insights in multiple cardiac diseases or the athletic fitness condition of an individual.

In order to test the capabilities of MSOT to determine the pulse rate, the right index finger was imaged at 890 nm excitation wavelength. Single wavelength imaging was performed because acquisition at multiple wavelengths was not possible with a 10 Hz repetition rate laser owing to the Nyquist theorem. Immediately before the experiment, physical exercise was performed to increase the heart rate above normal conditions and to enhance pulsating motion. Partial constriction of the finger was however not performed. The finger was then imaged over the coarse of 50 s by means of MSOT.

Figure 8.8 shows the results of a 7 s sequence of the pulse detection experiment. Fig. 8.8(a) shows a cross-sectional reconstruction of the finger at t=0 s.

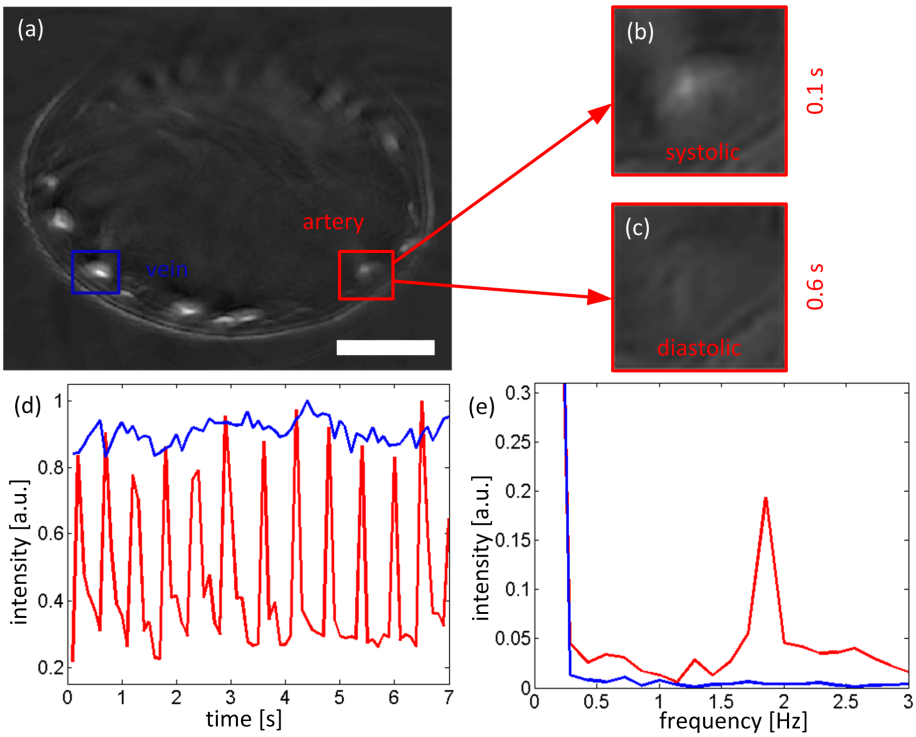

Figure 8.8: OA pulse measurement. (a) MIP of the reconstructed frames over 7 s with the ROI of the art. uln. indicated by the red box and the ROI of a vein by the blue box. The magnifications show the art. uln. (b) in systolic phase at t=0.1 s and (c) in diastolic phase at t=0.6 s. (d) The tracking of the vessel intensity from a 2-D Gaussian fitting shows the periodic motion of the artery (red), while the vein is almost constant in intensity (blue). (e) Fourier power spectrum of the vessel intensities shows an arterial pulsation rate of approximately 110 beats per minute (1.85 Hz). Scale bar 5 mm.

The red box indicates the ROI of the pulsating art. uln., while the blue box indicates the ROI of a vein for comparison. Fig. 8.8(b) shows the enlarged arterial ROI in systolic phase at t=0.1 s, while Fig. 8.8(c) shows the corresponding diastolic phase 0.5 s later. In order to quantify the pulsating motion, a 2-D Gaussian was fitted to the two ROIs in all frames and the corresponding normalized volume under the curve is plotted in Fig. 8.8(d). In the vein no pulsating motion could be detected. In the artery, a periodic motion was found with a peak frequency of approximately 1.85 Hz, which corresponds to 110 beats per minute. For later observation times, a decrease in pulse rate could be tracked and the recovery of the body towards normal physiological conditions could thus be monitored.

Pulse rate determination is most often performed in the context of pulse-oxymetry. Potential OA applications are thus discussed together with oxygenation measurements at the end of the next paragraph.

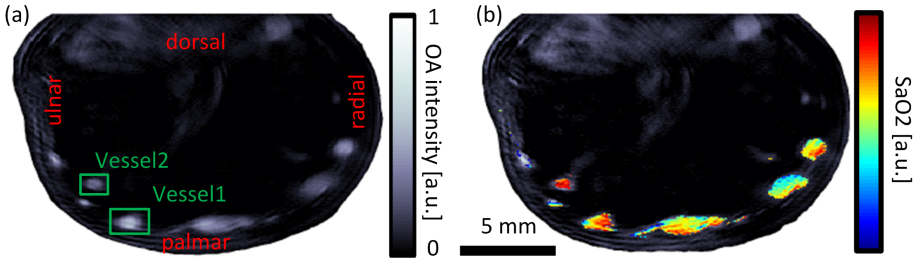

Figure 8.9: Visualization of blood oxygenation level in distinct vessels. (a) Anatomical image at 730 nm excitation wavelength with anatomical reference and ROIs of two selected vessels. (b) Blood oxygenation level obtained from un-mixing with seven different wavelengths. Art. uln. (*Vessel2*) exhibited the highest saturation level compared to the other major vessels (veins).

Blood oxygenation level determination

The blood oxygenation level SaO (Eq. 2.3) is probably the MOST important vital parameter to be monitored. Furthermore, multiple diseases such as cancer are related to hypoxia [1]. Spatial mapping of blood oxygenation enabled better tumor localization and treatment, functional mapping of brain activity in BOLD-MRI, and many more [204–206]. The oxygenation level can be determined invasively by taking blood samples or non-invasively by optical methods based on the distinct spectra of oxygenated and deoxygenated hemoglobin (see Fig. 2.5). Because of strong photon scattering in tissue, SaO can most often only be spatially mapped at very superficial depths or with extremely low spatial resolution. Pulse-oxymetry can also determine oxygenation at considerable depth like in the finger tip or in the ear lobe; it however relies on a strong pulsating motion of an artery and works typically in transmission mode, i.e. for relatively thin objects only. More importantly, pulse-oxymetry does neither provide any spatial information on the oxygenation nor the possibility to quantify the oxygenation level in venous blood.

In order to test the ability of MSOT to spatially resolve the oxygenation level, a multi-spectral data-set was acquired in the PIP joint of *Volunteer1* at seven equi-distant wavelengths ranging from 730 nm to 850 nm[6]. For each wavelength, five frames were acquired and averaged to improve SNR. Signals were normalized by the average wavelength dependent laser pulse energy prior to reconstruction. Spectral un-mixing to the concentrations of oxygenated and deoxygenated hemoglobin components was subsequently performed on a per-pixel basis by least square fitting with a non-negativity constraint, using the known spectra of the chromophores. From the un-mixed hemoglobin distributions, the oxygenation saturation level was calculated for all pixels with sufficient quality of the linear spectral fitting (R^2>0.8) and sufficiently high pixel intensity (>12 % of the maximum intensity for each wavelength).

Fig. 8.9(a) presents the reconstructed frame at 730 nm with the ROIs of two

[6]The experiment presented here is part of a study previously published as Ref. [126].

selected vessels indicated by the green boxes. Fig. 8.9(b) shows the calculated saturation level map superimposed in color onto the anatomical reference image of Fig. 8.9(a). The average saturation level in *Vessel1* was 0.88 (a.u.) and the average saturation level in *Vessel1* was 0.92 (a.u.). *Vessel2* corresponded to one of the main finger arteries (art. uln.) while *Vessel1* corresponded to one of the superficial finger veins.

Using MSOT, the oxygenation level could be individually determined for each major vessel. In principle saturation level determination is even possible on a per-pixel basis. However, major improvements need to be made before a routine clinical application of MSOT in oxymetry is anticipated: As it can be seen in Fig. 8.9(b), the saturation is considerably spatially varying within individual vessels, probably owing to reconstruction and calculation errors rather than to (unknown) physiological effect. Further, SaO could not be faithfully calculated for all image pixels. Quantification errors might have resulted from reconstruction errors in the individual images and also from considerable spectral coloring at larger depths. For accurate quantification, both issues need to be addressed.

Relative trends can however be extracted from neighboring vessels or in monitoring of temporal changes of oxygenation level. In the future, real-time MSOT could become a clinical alternative to standard pulse-oxymetry because of several advantages it offers: (1) It can provide information on the spatial distribution and resolve oxygenation level on a single vessel basis; (2) it does not necessarily require a strong pulsation of the artery; (3) thus it can also determine the saturation level of venous blood; (4) reflection mode imaging is possible and enables novel measurement sites like the carotid arteries.

Tracking the thermal response of vasculature

The peripheral vascular system is able to adopt to changes in the physiological state or also to changes in its environment [207]. Vascular constriction and dilation are a decrease and an increase in vessel lumen, respectively, and lead to a subsequent change of hemodynamics as response to an external or internal stimulus. One particular stimulus for vascular constriction and/or dilution is a thermally induced stimulus, i.e. the exposure to cold and warmth. Thermal stress tests probe the temporal response of the vasculature to e.g. exposure to cold water. An altered thermal vascular response is often associated with certain PVDs like RP [197]. Besides imaging the altered structure of capillaries in the nail bed, diagnosis and characterization of RP are often performed by thermal stress tests. OA imaging could potentially contribute to a more accurate diagnosis owing to its intrinsic blood contrast, its high scalable resolution up to centimeter depth, and its simplicity to use coupling to a water bath with adjustable temperature while imaging.

An experiment was performed to monitor the response of the finger vasculature to a thermal stress test by means of MSOT imaging. The index finger was completely immersed in a bath of iced water for 4 min. Immediately before and after exposure to the cold water, the finger was imaged at 890 nm excitation wavelength at an effective frame repetition rate of 0.4 Hz over 120 s and 260 s, respectively. Note that the tank of the scanner was filled with water heated to 34 °C. All measured frames were reconstructed using a homogeneous model and

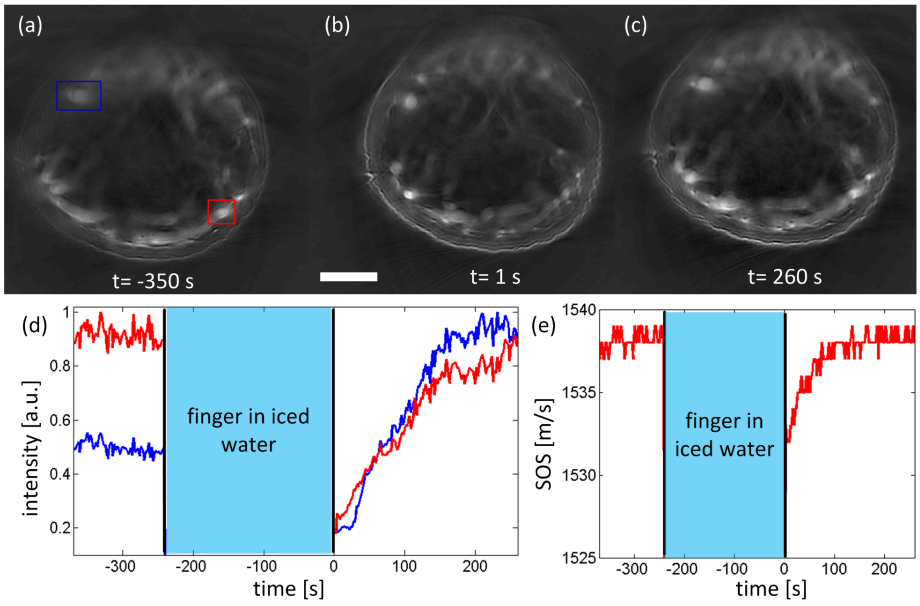

Figure 8.10: OA imaging before and a after thermal stress test with exposure of the finger to iced water. (a) Finger before the challenge with the ROIs of two selected veins indicated by the red and the blue box. (b) Corresponding reconstruction immediately after the challenge. (c) Reconstruction at the end of the experiment. (d) Intensities of the selected vessels with a rather constant baseline before the cold challenge and continuous recovery after the challenge. (e) SOS obtained from residual-based auto-focusing on individual frame basis reveals changes in tissue temperature. Scale bar 5 mm.

residual-based SOS auto-focusing in 1 m/s steps on an individual frame basis.

Figure 8.10 presents the results of the experiments. Selected frames at the beginning of the experiment (a), immediately after the cold exposure (b), and at the end of the experiment are shown (c). Differences in vessel intensity can be seen on the three images. For further analysis, two vessels were selected for detailed analysis, one palmar vein (red box) and one dorsal vein (blue box). Fig. 8.10(d) plots the intensity for both vessels obtained as amplitude from a 2-D Gaussian fitting to the ROIs in all frames. The baseline measured prior to the cold challenge was comparatively constant with variations most probably related to imaging artifacts and fitting errors rather than rapid physiological changes. From imaging after the exposure to the iced water, a large change of intensity could be observed in both vessels. The initial intensities (t=0 s) were lower than those of the respective baselines. A rapid increase in intensity with time was found for both vessels, indicating a thermal response to the 34° warm water and thus a recovery from the cold challenge. While the palmar vein reached approximately its initial intensity as before the exposure, the dorsal vein had a larger intensity than before. Fig. 8.10(e) plots the optimized SOS values used for reconstruction as function of time. The SOS was comparatively constant for the baseline and slowly increasing back to its initial value after cold challenge. Because of the exposure to the iced water, the tissue in the

finger was cooled down. The decrease in temperature resulted a corresponding decrease in SOS in the tissue of the finger from the temperature dependence of the SOS. Thus, SOS determination permitted an indirect sensing of temperature (or temperature changes) in the finger [208]. Assuming the same leading linear coefficient of 2 m/s/K like in water and the finger to be 2 cm in diameter, the global SOS change from 1538 m/s to 1532 m/s corresponded to a change of $\Delta T \approx -12$ K of average tissue temperature.

The presented experiment demonstrated the general ability of MSOT to monitor and quantify vascular constriction and dilution. Yet, several aspects need to be further addressed. It remains unclear if the detected response of the dorsal vessel is indeed an overcompensation to the stress test or an intensity drift of unknown source. Multiple subjects and all vessels in the finger need to be examined in order to answer such questions. A fully automated processing like vessel selection and a robust analysis is needed for that purpose. Optimized image quality from heterogeneous models could also ease analysis and allow to robustly track vasculature diameter directly instead of intensity. From the time series of the fittings, certain representative quantitative values (like exponential decay constant) could be extracted and in addition could be combined with thermal models for classification of the thermal response and disease progression.

8.6 Summary, Discussion, and Conclusions

In this chapter, MSOT imaging of human fingers has been experimentally demonstrated. Imaging at the site of the fingers is important because many prevalent PVDs affect the fingers and their vasculature. Examples of such PVDs with altered vascular anatomy and/or physiology are RA, TAO, sclerosis, artery occlusions, or RP. Further, vital parameter monitoring by means of purely optical pulse-oxymetry is also mostly performed in fingers.

An adaptable holder system has been developed to facilitate flexible cross-sectional, axial MSOT imaging in the finger at almost arbitrary position using a commercially available small animal scanner. Motion artifacts during acquisition could be minimized and thus spectral and temporal imaging studies were possible. Furthermore, volumetric imaging was enabled by z-scanning of the finger without the need for further post-processing. Motion could also be corrected for via an ASIFT-based registration framework that could be shown to correct with almost pixel size accuracy. For most imaging studies, the motion prevention using the mechanical holder even spared this time consuming post-processing.

High-quality anatomical images could be rendered in this fashion. The two main features in the images were the vasculature and the skin, corresponding to the two main intrinsic absorbers, hemoglobin and melanin. Blood vessels much smaller than half a millimeter in diameter could be resolved. A cross-validation of vessels with the corresponding MRI slices revealed a good agreement between both modalities in terms of size and location. Furthermore, the outline of the bone, whose absorption was not visible in the OA images, could be extracted based on its effects on US propagation (i.e. inducing signal reflections). The method was based on an automatic signal domain analysis and facilitated skin and bone outline extraction, and thus intrinsic segmentation. Further, it provided ROIs for image analysis, as well as an automated 2-compartmental SOS

distribution determination. Consequently, an improved reconstruction quality of small, deep vessels could be achieved based on an automatically generated, heterogeneous model matrix. Further work will also focus on the generation of a 3-compartmental model that also models the reflections of the acoustic signals at the bone boundary and could thus remove the reflection artifacts from the image and improve reconstruction quality.

Cross-sectional MSOT imaging was also extended to volumetric imaging of the finger based on z-scanning and stacking of the obtained images. Performing OA angiography, the overall shape of the finger could be visualized from the nail to almost the MCP joint, the vascular tree could be extracted, and the major vessels could be tracked over the entire length of the finger. Label-free OA angiography will thus be able to characterize the vascular anatomy and physiology under healthy and diseased conditions. Such diseases include TAO, where artery occlusions often occur at locations of high vessel curvature (corkscrew vessels) and might be potentially detected using MSOT in the future. In general, stenoses might be detected optoacoustically based on a reduced vessel lumen or occlusions might be visualized based on perfusion measurements (see Chapter 9). Future work will also have to address the post-processing of the reconstructions in such studies for tasks like automated vessel identification, segmentation, or quantification of parameters like vessel size or branching.

Functional OA imaging studies were as well showcased in this chapter, like monitoring of the most important vital parameters pulse rate and blood oxygenation level. The pulsation of the art. uln. could be determined to be 110 beats per minute from spatially resolved OA measurements. Conversely no pulsating motion could be observed in the veins of the finger. The blood oxygenation level could be determined based on the spectroscopic capabilities of MSOT and SaO values extracted for individual vessels. The main arteries exhibited the highest SaO level while the veins exhibited a lower oxygenation level. MSOT could thus develop into an alternative to purely optical pulse-oxymetry. For combining both pulse and oxygenation determination conjointly, faster laser technology has to be used however. OA pulse-oxymetry offers good spatial resolution, the possibility to determine the oxygenation level of both arterial and venous blood, and novel possible measurement sites such as the carotid arteries, because it does not fundamentally require pulsation and does not work in transmission mode only. Yet, MSOT technology is still expensive compared to cheap LEDs and photodiodes, and the challenge of spectral coloring of the light fluence from the background blood pool and especially from the melanin in the skin has to be managed. Error estimates for realistic *in vivo* situations will have to be obtained, and robust schemes for fluence correction based on light transport models or other advanced signal processing approaches will have to be developed. The accurate and robust correction of light fluence related effects remains one of the great open challenges for deep tissue MSOT imaging.

In a further functional imaging study, vascular constriction and dilation in response to a thermal stress test could be monitored on an individual vessel basis. Thermal stress tests are the main method to diagnose RP, usually based on superficial optical capillaroscopy or camera-based thermography. MSOT is not restricted to superficial areas and offers both high spatial resolution to monitor the changes in vessel diameter and also high temporal resolution to track temporal changes in response to environmental factors. In order to comprehensively test the capabilities of MSOT in such situations, larger cohorts of healthy and

diseased subjects will have to be examined and compared. For this purpose, an automatic processing and response quantification will have to be developed. An advanced analysis could also include modeling of a temperature map with corresponding SOS distributions in the reconstruction process or modeling of heat inflow through skin and arterial blood.

In summary, high-quality, high accuracy cross-sectional and volumetric finger imaging studies have been presented. The anatomical and functional imaging studies enabled monitoring important vital parameters like pulse rate and blood oxygenation level and enabled label-free, volumetric angiography of the entire finger. The results presented are expected to potentially influence clinical imaging and diagnosis of prevalent PVDs like RA, TAO, arterial occlusions and stenosis, and RP as well as to allow for optoacoustic pulse-oxymetry.

Chapter 9

Towards an Optoacoustic Diagnosis of Rheumatoid Arthritis

RA is an inflammatory disease affecting ca. 1 % of the total population with large associated socioeconomic cost. Molecular optical imaging approaches are, as discussed herein, promising to facilitate an early diagnosis of RA before onset of large anatomical damages. MSOT, as shown in this chapter, is highly promising for detection of the altered spatio-temporal kinetics of the ICG agent in inflamed joints, thereby overcoming the disadvantages of the current clinical imaging modalities in RA diagnosis. For this purpose, it is necessary to detect the spatially resolved ICG kinetics at the site of the finger joint at clinically relevant concentrations. Based on the findings of the associated OA ICG monitoring study presented in this chapter and the underlying molecular imaging paradigms in RA diagnosis, further steps and challenges for an early RA diagnosis based on MSOT are discussed.

9.1 Rheumatoid Arthritis and its Current Clinical Diagnosis

RA is an inflammatory auto-immune disease in multiple joints that affects approximately 1 % of the total population. Mainly females of 40 - 60 years in age at first outbreak of the disease are affected [200, 209]. Because of its expensive treatment and the considerable limitations in every day life, RA causes large economical cost in the order of billions of dollar in the United States alone [210].

Among the first joints affected by RA are typically the MCP joints in the hand, the PIP joints in the fingers, and the hips. Symptoms of RA are pain in the affected joints and considerable associated limitations performing motions like flexing. Joints are considerably swollen because of an enlargement of the synovial membrane, known as pannus. The inflammation process in the synovia results in neo-vascularization, i.e. the formation of new vasculature, and a higher density of small vasculature in that area. The newly formed vessels have an increased vessel wall permeability compared to the original vessels. This

Figure 9.1: Sequence of DCE-MR images for RA diagnosis. (a) Coronal MIP at time of first Gd agent flooding. (b) Corresponding MIP at peak intensity time. (c) After the signals in the large vasculature have decayed, strong signals are observed from the inflamed joints.

'leaky vasculature' permits smaller molecules to leave the blood stream. The leaked molecules and their kinetics can serve as basis for imaging (see Sect. 9.2). In a late state of RA, the bone around the joint might get irreversibly damaged by corrosion [211]. The molecular mechanisms of RA are still only poorly understood, and more research is on the way in order to develop effective drugs for treatment [210]. To date, RA is mostly treated with cortisone. Treatment with cortisone is effective in stopping the progression of RA and limiting the overall pain; however, treatment can only be successful when RA is diagnosed early enough when little pathological changes have taken place and the bone is not damaged yet.

Diagnosis of RA is routinely performed with multiple imaging modalities. One important indicator is still analysis of rheumatoid antibodies in the blood, regardless of additional imaging methods employed. Depending on the image modality and approach chosen, diagnosis can then be performed on an anatomical, functional, or molecular level. The clinical imaging modalities mainly used are [212, 213]:

- US imaging has a high sensitivity in diagnosing RA anatomically. Because it is a cheap and widespread imaging modality, US is still considered gold standard in RA imaging. But sonography is known to be strongly examiner dependent and diagnosis is only possible in a late state of RA, when macroscopic anatomical changes, especially bone damages, have already occurred.

- X-RAY imaging can as well provide detailed anatomical images; its drawbacks in RA diagnosis are its ionizing nature and the low soft tissue contrast. This renders X-ray imaging only effective in late phases with considerable bone damage and therapy might thus not be successful any more.

- NUCLEAR imaging methods provide functional or molecular contrast and have been used in RA diagnosis as well. The lack of anatomical reference,

the high associated cost, and the toxicity of the radio-trackers involved hinder a primary use of nuclear methods in RA imaging.

- MRI is able to provide both anatomical and functional/molecular contrast. In clinical research it has become standard, but the high associated cost prevent a general application as routine imaging modality [52]. Functional diagnosis is based on application and tracking of a Gd-based CA for DCE-MRI. Because of the different kinetics of the CA in diseased and healthy joints (see next section), the inflammation level can be quantified from the temporal profile of the CA. Figure 9.1 shows an example of RA diagnosis via DCE-MRI[1]. When most of the CA has already vanished from the main vasculature, the remaining high CA concentration indicates the PIP joints of D2 - D4 to be highly inflamed.

- Optical imaging methods are as well used in clinical research because of the functional and molecular contrast they can provide from altered optical properties of the joints in late phases of the disease or the altered CA kinetics in early phases [52, 214, 215]. Further, optical methods are cheap and non-ionizing. Besides DOT and laser speckle imaging, mostly fluorescence based imaging methods are used in combination with an administered CA whose spatio-temporal distribution is tracked and analyzed [52, 214]. Because planar fluorescent imaging is often weighted towards superficial contributions and offers no depth information, volumetric FMT imaging has recently been shown successful in diagnosis of early RA [216].

Besides the established clinical modalities, OA imaging has also been used for diagnosis of arthritis. Detection of (osteo-)arthritis was among the few reported OA studies in humans [217, 218]. Multiple studies with OA have conversely been reported in animal models [215, 219]. In animal models, imaging performance could also be enhanced be newly developed, targeting CAs, which are however not (yet) approved for the use in humans.

A translation of OA imaging technology to the clinics is foreseeable in the near future, because of the unique combination of anatomical and functional imaging contrast at high spatio-temporal resolution, because of its low cost, and because of the maturity of the technology. Diagnosis can then be performed both on anatomical basis with altered vasculature and bone structure (especially using hybrid modalities in combination with US) and on functional basis of CA tracking. The clinically CA agent ICG and its pharmakinetic model in the context of RA imaging are described in the next section.

9.2 Indocyanine Green and its Pharmakinetic Modeling

ICG is an organic dye that exhibits fluorescence in the NIR (Fig. 9.2). The cyanine group is characteristic for ICG and responsible for its fluorescence. The entire molecular structure with a weight of 751 Dalton is shown in Fig. 9.2(a). ICG absorbs and fluoresces in the NIR and the exact properties depend on its chemical environment [89, 220]. In blood serum, the peak absorption is

[1] The MRI images were acquired and gratefully provided by Dr. Reinhard Meier.

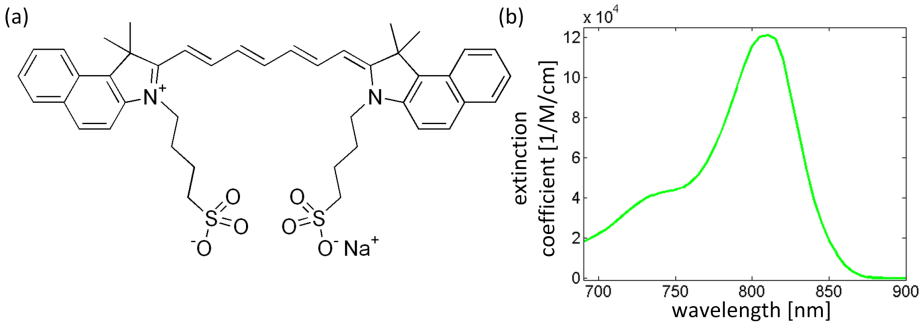

Figure 9.2: (a) Molecular structure of ICG. (b) Absorption spectrum of ICG with a peak at approximately 807 nm.

at 807 nm with an extinction coefficient of ϵ=121000 M^{-1} cm^{-1} (Fig. 9.2(b)). Peak fluorescence occurs at a higher wavelength of 822 nm with a quantum yield of 9 % [89]. However, both absorption and fluorescence depend non-linearly on concentration because of aggregation and quenching [10]. ICG is water soluble, but immediately binds to blood serum proteins like albumin with an associated change in optical properties [221]. The binding of the small ICG molecules to large proteins effectively confines them to the blood stream in healthy subjects. No metabolites of ICG are known and ICG is excreted through hepatic clearing into bile juice with a half-life of 150 - 180 s [89]. ICG has a relatively high photostability, but is degraded, if the solution is exposed to light over long periods of time [222]. Enhanced ICG-derived optical CAs have been developed recently as well, either by functionalization of ICG to bind to a specific target molecule or by using liposomal ICG with modified pharmakinetics and increased half-life in the blood stream through reduced hepatic clearing rates [223].

ICG is one of the very few clinically approved optical CAs. The main adverse reaction is due to sodium iodide and occurs as seldom as 1:10000. ICG was developed in the 1950s and is in use with multiple clinical applications since then: Among them are angiography in the context of ophthalmology, oncology, surgery, perfusion and occlusion measurements, hepatic clearing efficiency monitoring, and measurement of the cardiac output [10, 89, 220, 224, 225]. ICG has also been used for optical diagnosis of RA based on measurements of its pharmakinetics described further below [52]. ICG has as well been used in the preclinical context like for photo-thermal therapy, digestion, or stem cell differentiation [21]. Detection and imaging of ICG is mostly performed purely optically based on its fluorescence. Modalities for its detection include *in vitro* measurements of blood samples taken, camera-based planar fluorescence imaging, and FMT. ICG can as well be detected based on its absorption properties since most of the photon energy is thermalized. OA imaging technologies have successfully localized ICG in many preclinical applications [21, 23, 223, 226].

Optical diagnosis of RA with ICG is based on tracking its kinetic profile in the region of the inflammation. In healthy subjects, ICG is confined to the blood stream and cannot leak to the extravascular space. For inflamed joints however, the newly formed vasculature allows also larger molecules to penetrate the vessel walls towards the extravascular space. Finite concentrations of ICG are then found in the extravascular space with kinetics considerably deviating

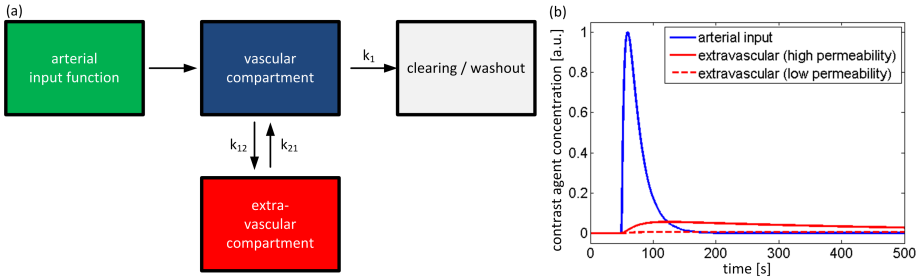

Figure 9.3: 2-compartmental modeling of CA distribution. (a) The CA concentration in the first compartment, the vascular compartment, is given by the AIF and is washed out with the blood stream. From the vascular compartment, the contrast agent can also diffuse to and from the extravascular compartment. (b) Resulting concentration profiles in the vascular and the extravascular compartment with a bolus injection: AIF (blue) and corresponding profile of the extravascular compartment (red) for high vessel wall permeability (k_{12}=0.12 min^{-1}). Although the peak concentration is much lower compared to the AIF, the decay is much slower and the concentration exceeds the concentration of the AIF for longer observation times. For low permeability (red dashed, k_{12}=0.012 min^{-1}) the peak concentration might be too small to be detected.

from those in the blood stream.

ICG with a maximum approved dose of 5 mg/kg is prepared by dissolving the powder in water (or saline) and systemically injected intravenously in one of the peripheral veins as concentrated bolus. After circulation through the heart and lungs, the ICG bolus in the arterial blood reaches the finger. The exact ICG concentration profile in the finger artery (known as arterial input function, AIF) depends on multiple factors. The shape is mainly determined by particulars of the injection and the cardiac output. Dilution of the bolus from injection site to the finger smears the profile and multiple peaks occur, when strong and fast recirculation takes place. The AIF can be modeled by a superposition of gamma functions or approximated by a dual exponential decay [227]. When the ICG has entirely diluted at later times, its concentration is approximately homogenous within the whole blood pool and slowly decays exponentially due to hepatic clearing.

The ICG kinetics within the finger are best described by linear compartmental modeling. To describe the ICG distribution in the extravascular space, a 2-compartmental model is needed (Fig. 9.3(a)): The first compartment is the vascular compartment that is fed by the AIF and ICG can be eliminated from there by washout clearing. The second compartment is the extravascular compartment where ICG can diffuse from or diffuse to the vascular compartment with a transport proportional to the concentration differences and to the rate constants k_{12} and k_{21}. The concentrations in the respective compartments are then modeled by the following linear differential equations:

$$\begin{aligned} dc_1(t)/dt &= -k_{12}c_1(t) + k_{21}c_2(t) - k_1c_1(t) + AIF(t) \\ dc_2(t)/dt &= +k_{12}c_1(t) - k_{21}c_2(t), \end{aligned} \tag{9.1}$$

where $k_{12} = k_{21}$ for passive transport. In the following, a limited exchange between compartments is assumed: $k_{12} << k_1$ and $k_{12} \cdot \tau_{AIF} << 1$, where τ_{AIF} is a time constant related to the decay of the AIF. Because the washout is fast, the concentration in the vascular compartment is equal to the AIF. For a quasi-constant AIF, the concentrations $c_1(t)$ in the vascular and $c_2(t)$ in the extravascular compartment are equal (if ICG might penetrate at all). With a rapidly changing AIF like for a bolus injection, the peak concentration $c_{2,max}$ in the extravascular compartment is much smaller than the peak concentration $c_{1,max}$ in the vascular compartment and is in the order of $c_{2,max} \approx c_{1,max} \cdot k_{12} \cdot \tau_{AIF}$. At later time points, the concentration in the extravascular compartment exceeds the concentration in the vascular compartment and decays exponentially with a rate of k_{12}.

The concentration profiles of the vascular compartment (blue) and the extravascular compartment (red) for two different sets of exchange rate constants k_{12} of vessel wall permeability are plotted in Fig. 9.3(b). For a very low constant (dashed red), almost no ICG leaks to the extravascular compartment. At higher permeability with an increased exchange constant (solid red), the ICG concentration gets significant in the extravascular compartment as well. In the example, considerable amounts of ICG are left at 200 s to 400 s, when almost all ICG has already vanished from the vascular compartment.

In RA, the 'leaky vasculature' from neo-vascularization allows penetration of ICG to the extravascular compartment with an associated exchange rate constant much higher than for health subjects: $k_{12}^{(RA)} >> k_{12}^{(healthy)} \approx 0$. Consequently, inflammation can be detected from regions with remaining CA signal at later time points, when no signal is left in the vasculature. This is also the basis of DCE-MRI with the Gd CA shown in Fig. 9.1. The slowly decaying profiles can be either detected in entire pixels, provided sufficient resolution; or the profiles can be obtained as superposition together with other, more rapid profiles in sub-resolution structures, too. For every pixel, a exchange rate constant $k_{12}(x)$ can be extracted from the corresponding kinetic profiles quantitatively or qualitatively and inflamed areas can be located in this way.

9.3 Real-Time Detection of Indocyanine Green Perfusion in a Human Volunteer's Finger

An experiment was performed to test the ability of MSOT to faithfully track the kinetic profile of an ICG bolus injection at clinically relevant doses in spatially resolved manner within the finger. The left index finger of *Volunteer1* was imaged at a slice in the PIP joint, using the 128 element MSOT scanner. Multispectral acquisition of the temporal profile was not possible because of the lack of fast wavelength tuning capabilities of the OPO employed. Consequently, single wavelength imaging at 10 Hz laser repetition rate was performed over a period of 180 s. An excitation wavelength of 800 nm was chosen, which was close to the ICG peak absorption. For kinetic profile imaging, a commercially available ICG agent (ICG-PULSION, Pulsion Medical Systems SE, Feldkirchen, Germany) was dosed to 1 mg per kg of body weight and dissolved in saline. ICG was administered into the vena mediana cubiti of the right arm of the volunteer by an experienced physician. The bolus injection started around t $\approx$ 55 s after

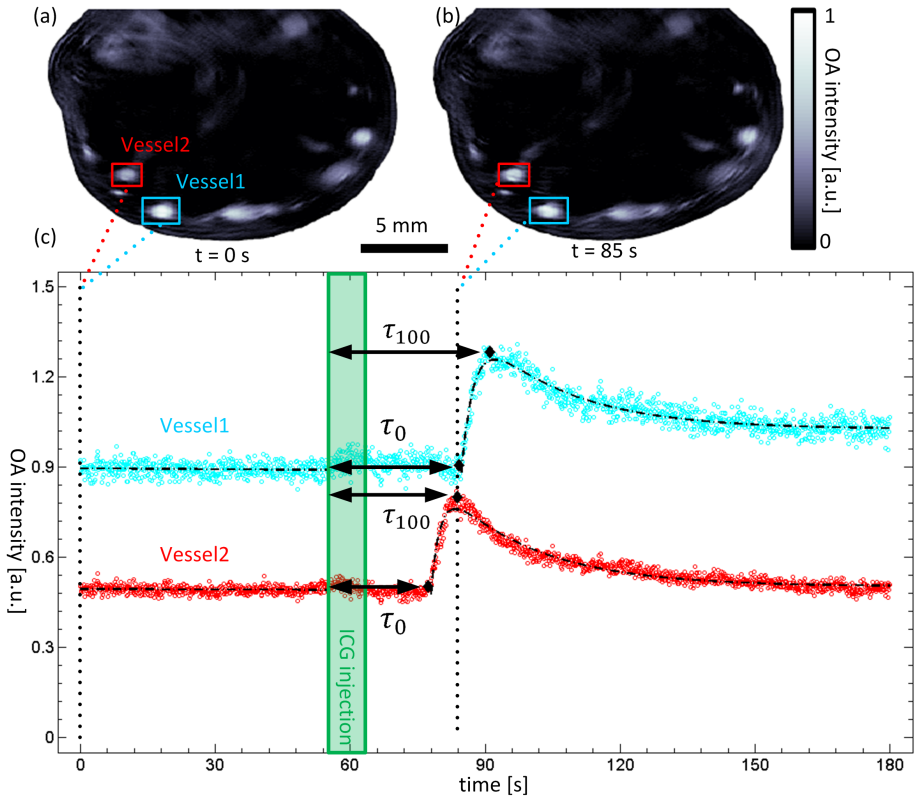

Figure 9.4: OA monitoring of ICG perfusion in the finger vasculature. (a) Cross-sectional reconstruction at the beginning of the experiment with ROIs indicated by the boxes. (b) Corresponding reconstruction at t=85 s with peaking relative dynamic contrast. (c) Profiles of ICG perfusion in the two selected vessels with fittings superimposed in black. Dynamic ICG contrast could be observed starting 22 s after the injection. In the artery (*Vessel2*), the relative dynamic contrast was higher and peak time was earlier compared to the corresponding quantities in the analyzed vein (*Vessel2*).

the beginning of imaging and the injection took around 5 s in total.

The results of the perfusion experiment are shown in Figure 9.4. Fig. 9.4(a) shows the reconstructed cross-sectional slice at the beginning of the experiment prior to ICG injection. Fig. 9.4(b) shows the reconstruction at t=85 s with peaking dynamic contrast from ICG. ASIFT corrections with respect to the first frame were performed to compensate for motion artifacts. In-plane motion could be estimated to 400 μm in average. For further analysis, two ROIs were selected: *Vessel1* was a vein (cyan box) and *Vessel2* was the art. uln. (red box). For each frame and each ROI, the 95-percentile of the OA pixel intensity was calculated. A linear baseline fitting to the values obtained was performed for the time points prior to injection in order to characterize a potential intensity drift and the fluctuation level. Based on the baseline and the peak dynamic contrast for both ROIs, the time points τ_{20} and τ_{80} were calculated, when the dynamic contrast first exceeded 20 % and 80 % of its peak value, respectively. From the

two time points τ_{20} and τ_{80}, linear extrapolation to τ_0 and τ_{100} was performed. A dual-exponential fitting was performed in order to characterize the temporal profiles in both vessels, starting from τ_0 to the end of the experiment at t=180 s:

$$f_{prof}(t) = A_1 \cdot \exp(-t/\tau_d) - A_2 \cdot \exp(-t/\tau_r) + A_3,\ t > \tau_0. \tag{9.2}$$

The extracted temporal intensity profiles of the two vessels are plotted in Fig. 9.4(c). From the baseline fittings, a potential relative intensity drift over the total duration of the experiment could be estimated to be less than 2 % of the vessel intensity. From the standard deviation of the measurements with respect to the baseline, a frame-to-frame fluctuation of less than 3 % could be estimated, potentially originating from laser power fluctuations or system noise. First inflow of ICG was observed at $\tau_0^{(2)}$=77 s in *Vessel2*, which was 22 s after the beginning of the injection. The dynamic ICG signal reached its peak at $\tau_{100}^{(1)}$=89 s and $\tau_{100}^{(2)}$=84 s, respectively, which was 34 s and 29 s after the injection. From the dual-exponential fitting using $f_{prof}(t)$, the characteristic time constants of the ICG temporal profile were calculated to be $\tau_d^{(1)}$=20 s and $\tau_r^{(1)}$=2.7 s for *Vessel1*, and $\tau_d^{(2)}$=21 s and $\tau_r^{(2)}$=2.0 s for *Vessel2*. The maximum relative signal increase with respect to the baseline was 47 % and 67 %, respectively.

The excellent spatio-temporal resolution of MSOT allowed to visualize and distinguish kinetic ICG profiles on an individual vessel basis. With the US limited spatial resolution, ICG inflow could be localized in vessels of half a millimeter diameter and less; with the video-rate image acquisition rate of 10 Hz, fast kinetics in the order of 2 s were readily captured. For individual vessels, magnitude, the time-to-peak, and the decay could be calculated from a fitting function. In the arteries, ICG inflow started and peaked earlier than in the veins[2], as expected because of the additional pass through the capillaries for venous blood. Also, the amplitude relative to the baseline was higher in the selected ulnar artery than in the selected vein[3]. Thus the ICG injection bolus was still more concentrated in the artery. After passage through the capillaries, the bolus was more diluted and consequently resulted and lower relative amplitude in the monitored vein. The shape of the observed profiles corresponded well with shapes predicted by tracer modeling after systemic injection and/or measurements with other tracers and modalities, as for example in dye-densitometry [224]. Because of the distant site of measurement, no additional peaks of recirculation were present. The delay of 22 s from systemic injection to first appearance of ICG in the finger resulted from venous flow from the injection site to the heart, passage through the lungs, and arterial flow to the finger. The delay was within the range of 10 - 31 s for a systemic circulation time in humans [228]. Further it has to be noted that the observed decay with a time constants of $\tau_d \approx 20$ s resulted from a dilution of the ICG bolus towards a homogeneous concentration on its way from the injection site to the measurement site (mainly in the heart), not from hepatic clearing of ICG with a half life in the order of 3 min. In the experimental results presented, no 2-compartmental

[2]The observation of earlier peak time was also made for the art. rad. and with respect to the other major veins.

[3]Note that the selected wavelength of 800 nm was close to the isosbestic point at 797 nm. The baseline value was almost insensitive to different oxygenation levels in arterial and venous blood. The amplitude increase was thus approximately proportional to the fraction of ICG concentration to total hemoglobin concentration.

kinetic modeling was performed. In healthy subjects, ICG is expected to be bound to the blood pool. So no extravasation occurs and the concentration of ICG in the second, extravascular compartment equals zero. Confirming the absence of extravasation, no significant dynamic ICG contrast was observed for image regions that did not have an original contrast of significant magnitude prior to injection stemming from the blood pool or melanin.

9.4 Future Steps Towards Optoacoustically Diagnosing Rheumatoid Arthritis

In Chapter 8 it was shown that high-quality anatomical and functional imaging of fingers in regions affected by RA is possible. Sect. 9.3 demonstrated - for the fist time to the best of knowledge - the ability of MSOT to track the kinetic profile of ICG in a spatially resolved per-vessel basis at clinically relevant concentrations. Yet, diagnosis of RA by means of MSOT remains challenging. RA diagnosis imposes certain requirements on the imaging technology based on the nature of the anatomy and processes to be visualized. These requirements and associated challenges, some of which still need to be solved for RA diagnosis with MSOT, are discussed in the following.

An imaging-based diagnosis of RA can be achieved based on two different paradigms, either by detecting anatomical changes or by performing functional imaging with DCE imaging (see Sect. 9.2). Anatomical imaging, also when using MSOT, is expected to visualize RA only in a late phase of the disease. Because the bones and bone damages are not directly visible in OA images, detectable changes are rather associated with the blood pool and neovascularization. However, the corresponding blood vessels are small with a size comparable to imaging resolution or even below. Vessels stemming from neovascularization might not be visible on an individual basis in the OA images, but rather the whole ensemble might result in an overall increased image intensity in the affected regions of the joint. A ROI-based analysis of the MSOT reconstructions might allow for determining a form of qualitative or quantitative measure for vascularization in that area. Bone localization (see Sect. 8.4) can provide an automatically obtained ROI for such an analysis while improved reconstruction methods can enhance the effective resolution in non-superficial areas and can thus improve and ease the analysis. Another possibility for characterization could be the average oxygenation level in certain regions, provided a hypothetical difference of SaO between healthy and arthritic joints.

Functional MSOT imaging promises to enable RA diagnosis also in early phases of the disease and is thus favorable. It is based on the distinct spatiotemporal kinetics resulting from the 'leaky vasculature' in RA. In order to qualitatively and quantitatively characterize ICG kinetics and to discriminate healthy from diseased subjects, certain imaging performance metrics associated with different domains must be met.

In the temporal domain, image acquisition must be sufficiently fast to capture the dynamics. With a frame rate of 10 Hz (and especially with 50 Hz and 100 Hz), the fast inflow of ICG in the order of few seconds can be captured at multiple excitation wavelengths. Especially with higher laser repetition rates, the amount of data to analyze from an experiment over several minutes

becomes extremely challenging, even when using fast reconstruction methods. Potentially frames need to be discarded or averaged, provided motion can be neglected. Motion in general is a huge challenge for temporal imaging studies. Although generally better avoided at all by using suitable fixation systems, it can be corrected using post-processing methods. However, motion within one multi-spectral frame-set is crucial and thus to be avoided, because erroneous artificial spectra are introduced. And most importantly, the temporal stability of the system over the entire imaging period is essential. The kinetic profiles need to be determined over the course of several minutes, dictated by the time constants for ICG dilution, ICG clearing, and compartmental exchange rates. Effects like fluctuations or a thermal drift of hardware components such as the laser, altered physiology in the finger from e.g. cooling, or simply motion from non-comfort of the volunteer would introduce artificial temporal profiles in the image pixels. The true ICG profiles and associated kinetic constants would then be (partially) masked. However, the reported perfusion experiments revealed that imaging with sufficient temporal stability over minutes is possible.

Challenges in the US frequency domain are associated with negative image values. They stem from deviating OA propagation models and particularly from the the diminished sensitivity of the detectors at low US frequencies (see Sect. 7.2). Negative image values are mainly found in deep areas with weak OA signals generated. Such signals are also expected for the ICG contrast in the inflamed regions. Negative image values might thus mask the weak, positive ICG concentration there. Besides alternative detection technologies with a higher bandwidth, alternative reconstruction approaches with a different weighting of lower US frequencies might potentially decrease negative values and increase quantification of the image values.

Multiple issues are also identified in the three spatial dimensions. The detected OA signals are a superposition of contributions from multiple locations. By performing reconstructions, the contributions are spatially localized to their origination. However, artifacts and errors cause a considerably wrong localization of contributions. Examples are limited view artifacts on the dorsal side of the finger, deviations in the assumed SOS distribution (which can however be reduced when using the proposed 2-compartmental model), and the bone reflections. For example, reflected contributions from the digital arteries can be erroneously located to within the area of the bone, induce artifacts there, and lower the reconstructed intensity of the actual site of the artery. Thus new reconstruction algorithms capable of modeling the reflections of the bones will have to be developed to increase achievable reconstruction accuracy. Additionally, the use of volumetric imaging probes is expected to avoid out-of-plane artifacts, which are decreasing image accuracy in cross-sectional 2-D systems based on focused signal detection. With volumetric probes, it is further possible to analyze larger regions of the joint and the analysis is not restricted to a single plane that has to be chosen at the beginning of the experiment.

Last, several advantages and challenges stem from the spectral domain. Most importantly, spectroscopic imaging approaches allows to selectively visualize absorption contributions from ICG contrast. With novel laser technology available, multi-spectral data-sets can be acquired at video-rate , not only single wavelength data-sets. Consequently, ICG can be detected with a much higher sensitivity compared to single wavelength imaging (as performed in Sect. 9.3), because ICG contributions can be un-mixed from the background contributions

of the blood pool. A higher sensitivity allows not only for better diagnosing capabilities, but also for a corresponding reduction of the administered ICG dose. However, multi-spectral imaging imposes also several challenges. Most importantly, spectral coloring of the light fluence in deep tissue areas alters the measured absorption spectra. If the spectral coloring is not adequately accounted for during the un-mixing process, ICG contributions are erroneously attributed to the hemoglobin concentrations and vice versa. Thus the quantity and temporal and spatial distribution of the unmixed 'ICG'-channel can be significantly corrupted. Then extraction of the kinetic parameters might be considerably complicated by the false positives or even impossible.

9.5 Summary, Discussion, and Conclusions

The auto-immune inflammatory disease RA affects approximately 1 % of the population and is associated with billion dollars socioeconomic cost. This is because of the patients' strong limitations with motion and strong pain, mainly in the joints of the fingers and of the hand, because of the high cost of treatment, and because of the irreversible anatomical damages in the late phase of the disease. Early diagnosis of RA is difficult, but then good management of the disease can be established. In clinical routine, RA is currently diagnosed with multiple imaging modalities; some of them can only detect RA in a late phases (US, X-ray), involve toxic radiotracers (PET/SPECT), or are expensive (MRI).

Conversely, optical imaging is cheap, non-ionizing, and capable of molecular and functional imaging with rich contrast. RA can be diagnosed by means of optical imaging based on the altered spatio-temporal kinetics of ICG. ICG is a FDA approved, well tolerated organic dye exhibiting fluorescence and absorption in the NIR. ICG binds to albumin and thus it cannot leave the vascular compartment. From there, it is excreted through hepatic clearing. No contributions from dynamic ICG contrast in the extravascular compartments are found in the joints of healthy subjects. In inflamed joints however, neovascularization in the joints results in vasculature with higher vessel wall permeability and ICG can therefore also leak to the extravascular space. The ICG signals from the extravascular compartment have a different kinetic profile with a long-lived, slowly decaying dynamic contrast. With the exact decay given by the corresponding exchange rate constants, significant contributions are still found in the extravascular space, when the ICG in the blood pool has already been cleared. Consequently, RA can be diagnosed by detecting the significant amounts of ICG in (the extravascular compartments of the) the inflamed joints at late time points or by directly calculating the associated larger exchange rate constant. Optical detection of RA has been realized based on camera-based planar fluorescent methods, which have a high sensitivity for ICG, but a low resolution in general and not depth resolution in particular. Recently, FMT has been shown successful in RA imaging. FMT as 3-D modality has also depth resolution, but a relatively low resolution in general and a time consuming acquisition procedure (hindering imaging of fast kinetic processes like the injection).

MSOT combines the advantages of optical RA imaging approaches with high spatial resolution in all dimensions. Its real-time capabilities allow also to capture fast temporal processes like the AIF. Here, it has been shown that MSOT is able to provide a detailed visualization of vasculature in the PIP joint.

Individual blood vessel of less than half a millimeter in diameter could be rendered (even at several millimeters depth using the signal domain segmentation 2-compartmental reconstruction algorithm) as cross-sectional slices or in a volumetric stack. In addition to the anatomical imaging, functional parameters like blood oxygenation level could be extracted on a single vessel basis. Furthermore, the location and shape of the weakly absorbing bone could be identified and might also be a potential starting point to automatically select a ROI for analysis of the potential inflammation in the synovia. Most significantly for the detection of RA, an injected ICG bolus could be detected at clinically relevant concentrations. To the best of our knowledge, this was the first demonstration of detecting ICG in humans by means of OA imaging. Its kinetics could be track and quantitatively extracted at 100 ms temporal resolution, thereby providing an AIF for compartmental kinetic modeling. Under stable imaging conditions, the ICG profile could be monitored over several minutes and different kinetics were observed on an individual vessel basis.

In order to facilitate an OA diagnosis of RA, several improvements might be implemented. Most importantly, spectroscopic imaging instead of single wavelength excitation can drastically increase the sensitivity for ICG. However, false positives from spectral coloring of the light fluence and from motion artifacts have also to be minimized then. Furthermore, the use of volumetric imaging probes will allow not only for larger imaging regions of the inflamed area, but will also eliminate out-of-plane artifacts. Algorithmic improvements might aim on reducing the negative image values present and the reflection artifacts, and thus increase overall image quality and quantification. And an automated ROI analysis and extraction of kinetic profiles and associated parameters will facilitate an application in clinical routine.

In summary, the demonstrated kinetic ICG profile monitoring paved the way towards an early diagnosis of RA by means of MSOT in the near future and to potentially establish MSOT as an alternative clinical imaging modality in this context.

Chapter 10

Summary, Conclusions, and Future Work

10.1 Summary

The aims of this work have been to advance MSOT towards its routine clinical application for finger imaging. In particular, three main goals were pursued: The first was to accelerate accurate, computationally expensive MB reconstructions towards video-rate and live visualization for clinically relevant imaging systems. The second was to automatically provide superior image quality without the need for manual user input for clinically relevant scenarios. An additional important consideration for the development of such methods was their computational efficiency so that they can be routinely applied in clinical practice. The third goal was to demonstrate the feasibility of MSOT to perform high-quality clinical imaging in the finger. For this purpose, the possibility to extract anatomical and physiological parameters related to common PVDs like RA or RP or to monitoring of vital parameters from MSOT measurements needed to be tested.

For the acceleration of MB reconstructions, two transformation-based approaches have been developed that take into account the particulars of the OA imaging and reconstruction problem.

The first approach presented in Chapter 5 was based on geometrical symmetries present in most OA systems, particularly in the case of rotational symmetries in tomographic scanners. A suitable polar image discretization revealed rotational invariance in the discrete formulation of the forward problem. Exploiting the conservation of angular frequency by a transformation of the model to this domain, the model became block-diagonal and thus separable. Consequently, a direct inverse of the model matrix could be readily obtained which is usually computationally and memory prohibitive in Cartesian discretization. An inversion scheme based on inverse and forward model in polar formulation could be developed for limited view scanners and offered a faster convergence rate and superior inversion performance. A 12.7 x higher reconstruction rate could be achieved compared to standard iterative LSQR inversion at the same image quality. The memory efficient formulation using dense algebra in combination with the intrinsically parallel nature of the algorithm enabled an ef-

ficient implementation on powerful GPU platforms using standard numerical libraries. Computations could be accelerated on GPU by an additional factor of 13 x compared to the corresponding CPU implementation and thus an overall acceleration of more than two orders of magnitude versus state-of-the-art MB reconstructions could be realized. As a consequence, peak reconstruction rates of up to 40 frames per second enabled video-rate reconstructions at high image qualities for clinical imaging. The long-standing goal of replacing BP approaches with MB live-previews is ready at hand if further approaches to decrease the frame latency time are implemented.

The second method for accelerating MSOT reconstructions is also applicable for systems where symmetries cannot be efficiently exploited. The high-throughput framework was based on a sparse and thus numerically efficient problem re-formulation. It was rooted on general properties of the OA imaging problem rather than exact symmetries. It extended a previously published, in-house developed framework that had exploited the approximate separability of the forward model in a 2-D WP representation and consequently had achieved fast reconstructions using an approximate WP inverse. In addition to model separability, three other sources of sparsity could be identified and in total four means for reconstruction acceleration were exploited conjointly: (1) model separability which enabled direct instead of iterative reconstructions; (2) the sparsity of the inverse model matrix in the WP domain allowed rejecting small valued matrix elements and thus saving memory and inversion time; (3) signals of individual frames were sparse in the WP domain and hard coefficient thresholding resulted in high computational benefits at only minor loss of image quality; (4) accounting for a high inter-frame correlation in a combined PCA-WP domain resulted in an extraordinary sparse representation of multi-frame signals and consequently an extraordinarily fast inversion performance. Under high-throughput conditions, when standard MB reconstructions are often computationally prohibitive, an acceleration of 40 - 700 x with respect to partially converged iterative reconstructions of similar RMSD error could be achieved for clinically relevant multi-spectral, volumetric, and temporal multi-frame datasets. The sparsity-based MB inversion framework enabled high-quality reconstructions at video-rate peak performance previously only achieved by BP methods of inferior image quality.

Aside to accurate model based reconstructions, a number of additional processing methods are required for achieving high image quality. Various methods have been developed in the context of this work in order to provide superior reconstruction quality at moderate computational cost and without manual user input. Negative image values are regularly found in experimental OA reconstructions although their unphysical nature hinders image interpretation and quantification. A computationally efficient iterative inversion method has been developed that suppresses the appearance of negative image values by imposing an additional inequality constraint. Moreover, multiple sources of negative value artifacts have been identified: Besides limited detection frequency bandwidth, corruption by strong superimposed noise was one of them. Effective signal de-noising of multi-spectral signals could be achieved based on an extraordinarily sparse representation in a combined PCA-WL domain and subsequent hard thresholding. A further source of negative image values was the inaccurate modeling of OA signal propagation. Particularly the assignment of the correct SOS for reconstruction was crucial, which, if not successfully performed, resulted

in a loss of resolution and an overall loss of image quality in addition to the negative pixel intensities. An optimized homogeneous SOS parameter could be accurately and robustly retrieved for MB reconstructions using a residual-based auto-focusing approach. The computationally burdensome SOS retrieval could be considerably expedited when using the developed sparsity-based, multi-frame reconstruction approach. It could also be shown that computationally favorable homogeneous models were not able to provide optimal image quality for all imaging scenarios. Therefore, suitable heterogeneous models were needed instead. The SOS distribution of a 2-compartmental model could be efficiently calculated prior to the one-step reconstruction process based on a signal domain analysis. Extraction of characteristic OA signal features allowed for conclusions on the acoustic properties; the method thus resulted in improved reconstructions and it is not limited to, but mainly applicable in OA finger imaging.

In the experimental part of this work, the feasibility of MSOT imaging for routine clinical imaging in fingers has been demonstrated in volunteers. MSOT imaging of fingers is one major field for the clinical translation of the OA technology because their vasculature is frequently affected by common PVDs such as RA, RP, or TAO. Moreover, the finger is the predominant site to monitor vital parameters. A commercially available small animal MSOT scanner system was enriched by a custom-designed adaptable finger holder system. It thus enabled real-time multi-spectral imaging of cross-sectional slices of almost the entire finger without being corrupted by motion artifacts (sub-pixel accuracy could be achieved). Aside to the stability provided by the finger holder, a robust post-processing OA motion correction framework based on ASIFT rigid body image co-registration was implemented and also allowed for the analysis of motion corrupted experimental studies.

It was shown possible to acquire anatomical cross-sectional OA images with excellent contrast from melanin in the skin and hemoglobin in the blood vessels. Vasculature of diameter much smaller than 0.5 mm could be visualized and the size and location of all major blood vessels in the finger appeared in excellent agreement with the corresponding MRI slices which were acquired for cross-validation. Cross-sectional imaging could be further extended to volumetric anatomical imaging by z-scanning of the finger, stacking of the slices, and 3-D rendering. The vascular tree of almost the entire index finger could be visualized, the course of individual vessels tracked, and their branching studied. When applying the 2-compartmental signal domain analysis, an automated and yet accurate segmentation of the finger area was obtained, superior SOS focusing and image quality compared to homogeneous models could be achieved for small, deep-seated vessels, and the outline of the optically less pronounced bone could be identified based on the induced signal reflections. The identified bone outline could be further confirmed based on the corresponding MRI slice.

In addition to the visualization of detailed anatomical features, this study also demonstrated first functional MSOT imaging of the human finger. The spectroscopic capabilities of MSOT facilitated to quantitatively distinguish the contributions of HbO and Hb and consequently to calculate the blood oxygenation level SaO on a per-vessel or even per-pixel basis for both arterial and venous blood. A higher saturation level could be demonstrated for arterial blood compared to venous blood, even without a pulsing motion of the vessels, and arteries could therefore be discriminated from veins. The video-rate temporal resolution of MSOT enabled capturing the fast, pulsating motion of the ulnar

finger artery at approximately 110 beats per minute while such a motion was absent in the accompanying veins. Another functional OA study monitored the vascular response to a thermal stress test as often performed in the context of RP using established imaging modalities. A vasoconstriction in the finger owing to the exposure to iced water could be observed and the subsequent slow recovery towards normal conditions monitored. Hypothermia and thermal recovery could as well be indirectly monitored by the temporal changes of the temperature-dependent average SOS in the finger.

For the first time, to the best of our knowledge, the detection of a clinically approved CA in humans by means of MSOT was reported. DCE OA imaging is not only of outstanding interest for general (pre-)clinical studies, but particularly in the context of diagnosing RA: Functional DCE imaging has been shown to visualize inflammation in an early phase before the onset of macroscopic pathological damages and before the possibility to detect them using anatomical imaging paradigms. For this purpose, the altered spatio-temporal kinetics of the blood-bound CA ICG need to be monitored. In this work, the profiles of an extrinsically administered ICG bolus at clinically relevant concentrations were successfully and robustly tracked in the finger vasculature over several minutes by means of MSOT. At video-rate temporal resolution, the dynamic ICG profiles could be extracted on a per-vessel basis and characteristic time constants were calculated from the temporal profiles as required for AIFs in compartmental kinetic modeling. The extracted time constants revealed a circulation time from systemic injection site to the measurement site of approximately 22 s and an appearance of the ICG in the veins delayed by 5 s compared to the arteries. Although both arterial and venous profiles were similar (2 s rise and 20 s decay), the veins exhibited a diminished peak contrast relative to the static blood pool contrast. The demonstrated results in anatomical and particularly functional ICG monitoring paved the way towards OA imaging for diagnosis of RA in early phases and future efforts based on the work performed in the context of this thesis will be devoted to achieve this goal. Some of these potential ways are outlined in the following.

10.2 Conclusions and Future Work

The methods developed in the context of this work are expected to provide novel reconstruction paradigms to provide real-time, high-quality reconstructions in clinical imaging. Moreover, the experiments performed are anticipated to stimulate further research on the fields initiated, and to pave the way for more comprehensive imaging studies in order to detect various PVDs.

The symmetry-based reconstruction approach in combination with an adaptive reconstruction scheme is expected to replace the BP live preview in clinical handheld systems by their MB counterpart. In addition, symmetry-based approaches are expected to facilitate a routine use of volumetric reconstruction approaches, thereby replacing 3-D stacking of 2-D reconstructions and eliminating the associated out-of-plane artifacts. Even when modeling the detector shape with 3-D systems, symmetry-based reconstruction approaches in the spatial frequency domain are fast because the model matrix uses dense algebra and signal elongation does not influence performance. More importantly, the feasibility of a direct inverse can accelerate the inversion process by several orders

of magnitude.

The two main future applications of symmetry-based reconstructions in 3-D are z-scanning of focused arrays and OA microscopy. When performing reconstructions for z-scanning with an entirely volumetric inversion method, out-of-plane artifacts are suppressed and a better resolution in z-direction compared to image stacking is achieved. In pre-clinical imaging, the methodology could enable fast, high-resolution reconstructions of a whole mouse body within several minutes instead of days. In clinical imaging, the method is expected to find application in z-scanning of linear arrays which are frequently used. Second, the method is expected to find application in OA microscopy, particularly in the emerging field of dermal microscopy. With the symmetry-based method, a complete modeling of the SIR of the detection element in MB algorithms becomes realistic. On modern workstations it is possible to obtain a direct inverse for image grid sizes as large as 512 x 512 x 512 voxels and the reconstruction time is thereby decreased from days to a few minutes. Consequently, it is anticipated that the methodology will be regularly applied for purposes of clinical dermal microscopy, because in this case MB algorithms are able to provide unrivaled image quality.

The second methodology developed for fast MB reconstructions, the sparse inversion framework, is expected to find application in fast reconstructions for systems with limited or no effective symmetries. It can possibly become an alternative approach for live visualization in 2-D or more importantly as fast, but yet accurate reconstruction method in 5-D imaging. The computational times could be reduced from several minutes to tens of seconds. As the methodology roots on general properties of OA signals and OA wave propagation rather than system geometries, it will mainly be applied for systems with too many voxels to calculate an inverse in a naive way but with still computationally feasible WP inverse. Using systems like the volumetric handheld probe, clinical studies comprising large multi-spectral data-sets with hundreds of temporal frames might be analyzed in a reasonable amount of time then.

In general, a broad variety of reconstruction approaches are expected to find application in OA imaging due to the complexity of the reconstruction problem involved. The method of choice is system- and application-dependent with different foci on reconstruction time constraints, resulting image quality, or complexity of the model. Because of this variety of aspects, considerable efforts will have to be undertaken in the field of MB reconstruction approaches. Novel approaches might include efficient iterative FDTD models, efficient SIR modeling by image domain convolution, or especially multi-scale and hybrid approaches: In iterative inversion methods, the first few iterations could be performed on a coarse scale only or different reconstruction methods could be employed for different scales, such as direct reconstructions from the Cholesky factorization on a coarse scale and direct WP reconstructions on a finer scale. In future methods, both concepts of symmetry and sparsity might also be powerful means to be included.

The proposed image quality enhancement strategies will also impact the standard signal processing and reconstruction routines. The unrivaled noise elimination capabilities of the developed de-noising methodology along with the comparatively low computational cost will make it a standard pre-processing routine for multi-spectral data-sets with low SNR. The proposed constrained iterative inversion can remove unphysical negative image values. However, it

is more effective, but also more challenging, to not only deal with the symptoms, but to tackle the causes such as a limited bandwidth of the detected signals, noise, or deviations of the assumed propagation model from physical reality. Potential means could be providing signals of higher bandwidth (e.g. using multiple transducers or particularly different technologies like the emerging interference detection) or algorithms that weight the missing low frequency components less. Such means could also be the application of the novel denoising method to provide sufficient SNR of the signals or propagation models that are closer to physical reality, i.e. heterogeneous models. More research will have to be conducted in order to develop fast methods to generate heterogeneous models such as efficient ray tracing or FDTD approaches. In combination with the proposed signal analysis approach, they are expected to find application in pre-clinical imaging studies like in mouse torso imaging and more importantly in clinical imaging studies in or through the skin, such as dermal microscopy, carotid imaging, or breast cancer screening.

For the experimental part of this work, the demonstration of the first clinically relevant finger imaging results is expected to motivate further studies towards clinical imaging for vital parameter monitoring and characterization of PVDs. Thereby, technical improvements in instrumentation or processing can improve the diagnosing characteristics. OA finger imaging in general requires image co-registration for motion compensation if motion cannot be totally avoided or in multi-modal imaging for image fusion or cross-validation. Alternatives to the proposed ASIFT methodology will have to be evaluated and further extended to the 3-D case to be used with volumetric probes. Novel instrumentation might also offer higher US frequencies based on broadband detection technology and thus also higher spatial resolutions, when combined with advanced, heterogeneous models like the proposed 2-compartmental model. Automatic retrieval of such models will also have to be adapted to detectors with highly limited view. Furthermore, the 2-compartmental model can be extended to a full 3-compartmental model that minimizes reflection artifacts induced by the mismatching bone. In order to quantify anatomical and functional parameters, an automated, robust feature extraction methodology is required like the signal domain analysis for the skin and the bone. More importantly, an automated processing is required in order to extract vascular parameters such as diameter or curvature and to track of vessels and their branching. Potential means could include image segmentation, feature extraction, or simply Gaussian fittings.

The proof-of-principle experiments in this thesis indicate that MSOT imaging in the finger might be used for vital parameter monitoring with pulse-oxymetry measurements. However, more efforts are still required to test its clinical potential. First, direct comparisons with the photoplethysmography or blood samples are required. Second, OA pulse detection and saturation determination need to be combined in a single experiment enabled by the available higher repetition rate laser technology and an automated processing of the temporarily varying part has to be developed to extract the heart rate. Third and most importantly, quantification of the calculated saturation levels has to be achieved, which is one of the great and long-standing challenges of MSOT in general. Although relative trends like arteries exhibiting the highest saturation are correctly determined, the obtained saturation levels do not yet match the absolute values of approximately SaO ≈ 0.98 and SaO ≈ 0.8 for arterial and

venous blood, respectively. Besides image artifacts, strong spectral coloring is the major source of deviations. Novel and robust light transport inversion schemes or advanced signal processing approaches will have to be developed in the future.

MSOT is also expected to be clinically applied for imaging of PVDs in the next years, particularly for the early diagnosis of RA. The next major step, namely a functional detection of RA based on the ICG kinetics, can be achieved qualitatively by detecting the slowly decaying ICG contrast in the joint at late phases after the injection. Alternatively, quantitative determination can be performed by extracting the ICG profile in the potentially inflamed ROI and the AIF in the feeding arteries; then both can be related via a 2-compartmental kinetic model and the exchange rate constant can be quantitatively extracted as metric to characterize the disease level. In any case, a group of volunteering patients will have to be compared with a group of healthy subjects in order to evaluate the significance of the hypothetical difference between the two groups. Furthermore, the minimal detectable dose of ICG will have to be determined, particularly when using more sensitive MSOT illumination strategies instead of single wavelength imaging. Potentially, novel instrumentation like the volumetric probe might not only be able to visualize larger regions of the joint in 3-D, but the MCP joint might also be accessible with such a limited view probe. Additionally, the use of fast 100 Hz laser technology might reduce motion-related artifacts and thus might improve quantification through a reduction of false positives in the ICG-channel.

MSOT imaging might potentially become a modality for diagnosing RP. As mainly functional disorder, RP is typically probed by functional tests similar to the thermal stress test performed in this work. For a clinical application of MSOT in this context, a standardized protocol will have to be developed and the observed thermal responses of the finger vasculature will have to be compared for patients and a control group. An advanced analysis method might also include a simplified modeling of the temperature map describing a SOS gradient within the finger. For a more direct comparison with existing superficial capillaroscopy methods, OA microscopy in the nail bed might be performed as well. Similar patient studies will also have to be performed in the case of TAO to detected the cork-screw anatomical structure of the collateral arteries by 3-D z-scanning. The required high resolution at the depth of the arteries might be achieved by means like the 2-compartmental signal domain approach.

In conclusion, clinical MSOT imaging might greatly benefit from the reconstruction and processing methods developed in the context of this work: The proposed symmetry- and sparsity-based methods are expected to develop into real-time reconstruction frameworks for clinical tomographic imaging or to facilitate OA imaging with high-throughput clinical MB reconstructions. Finger imaging with MSOT is expected to be translated towards patient imaging in prevalent PVDs such as RA, RP, or TAO within the next years. The proposed efficient processing methods such as a de-noising, auto-focusing, or advanced modeling are likely to be routinely applied to ensure sufficient the high image quality in needed for a successful routine clinical application of the MSOT technology.

Acknowledgements

Such a work would not have been possible without the support of numerous people:

I have to sincerely thank Daniel Razansky, my direct supervisor, for giving me the opportunity and the freedom to work on several exciting projects and for his constantly respectful, open-minded, and humorous character.

I am indebted to Vasilis Ntziachristos for accepting me as his PhD student and for scientific guidance throughout the entire period of my PhD at IBMI.

I have to thank Prof. Dr. Seeber, Prof. Dr. Bungartz, and Prof. Dr. Hayden for volunteering to be on my PhD committee.

Reinhard Meier is acknowledged for the fruitful collaboration on clinical OA finger imaging. Furthermore, I'm grateful for his willingness to be part of my Helena thesis committee.

I am obliged to Michal Neeman for inviting me to Rehovot and hosting me in her lab. The work of producing fascination mouse embryo images together with Reut, Marina, and Ori (and the lunches with the boys-office) was of invaluable experience to me.

I'd like to thank Luis, Amir, Stratis, and Lu for the interesting and fruitful work during the joint projects.

I appreciate the tolerance of the whole Razansky group for my low quality(?), high quantity presentations during the countless Friday mornings and the open, collegial atmosphere there. Their willingness for straight, comprehensive support in all matters is as well highly appreciated.

I have to thank the admins Susanne, Zsuzsi, Silvia, Christoph, Julia, Andi, and Karin. They are the backbone and heart of IBMI and have always been of great help while organizing all kinds of routine and non-routine tasks.

I have greatly benefited from the iThera crew - Till, Josef, Stefan, and Marcin to name just a few - and their willingness to grant me access to the most recent generation of MSOT systems. Furthermore, they always had an open ear discussing software features and software 'features'. I also like to thank Neal and Wouter for providing data-sets for testing.

Mority and Jacke are acknowledged for sharing the undomesticated biotope known as 'office' and for creating a semi-productive work atmosphere. Andrei, Amy, and Panos have always been great friends while eating Schnitzel, devouring numerous cakes, or tasting beverages.

I have to thank Christian, Stratis, and Andrei for valuable suggestions and proof-reading of this work.

And most importantly, I have to thank my family, my mother and my sister, for support throughout the years and my friends and mates for keeping me busy during the 'out-of-office' hours and days.

Bibliography

[1] R. Weissleder, M. J. Pittet, Imaging in the era of molecular oncology, *Nature* **452**, 580 (2008).

[2] R. Weissleder, V. Ntziachristos, Shedding light onto live molecular targets, *Nature medicine* **9**, 123 (2003).

[3] M. F. Kircher, J. K. Willmann, Molecular body imaging: Mr imaging, ct, and us. part i. principles, *Radiology* **263**, 633 (2012).

[4] Y. Assaf, O. Pasternak, Diffusion tensor imaging (dti)-based white matter mapping in brain research: a review, *Journal of Molecular Neuroscience* **34**, 51 (2008).

[5] W. A. Kalender, X-ray computed tomography, *Physics in medicine and biology* **51**, R29 (2006).

[6] P. Wells, Ultrasound imaging, *Physics in medicine and biology* **51**, R83 (2006).

[7] M. Pennant, Y. Takwoingi, L. Pennant, C. Davenport, A. Fry-Smith, A. Eisinga, L. Andronis, T. Arvanitis, J. Deeks, C. Hyde, A systematic review of positron emission tomography (pet) and positron emission tomography/computed tomography (pet/ct) for the diagnosis of breast cancer recurrence, *Health Technology Assessment* **14**, 1 (2010).

[8] V. Ntziachristos, Going deeper than microscopy: the optical imaging frontier in biology, *Nature Methods* **7**, 603 (2010).

[9] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, H. F. Hess, Imaging intracellular fluorescent proteins at nanometer resolution, *Science* **313**, 1642 (2006).

[10] J. T. Alander, I. Kaartinen, A. Laakso, T. Pätilä, T. Spillmann, V. V. Tuchin, M. Venermo, P. Välisuo, A review of indocyanine green fluorescent imaging in surgery, *Journal of Biomedical Imaging* **2012**, 7 (2012).

[11] K. Nienhaus, G. U. Nienhaus, Fluorescent proteins for live-cell imaging with super-resolution, *Chemical Society Reviews* **43**, 1088 (2014).

[12] V. Ntziachristos, D. Razansky, Molecular imaging by means of multispectral optoacoustic tomography (msot), *Chemical Reviews* **110**, 2783 (2010).

[13] L. H. V. Wang, S. Hu, Photoacoustic tomography: In vivo imaging from organelles to organs, *Science* **335**, 1458 (2012).

[14] C. Lutzweiler, D. Razansky, Optoacoustic imaging and tomography: reconstruction approaches and outstanding challenges in image performance and quantification, *Sensors* **13**, 7345 (2013).

[15] E. M. Strohm, E. S. Berndl, M. C. Kolios, High frequency label-free photoacoustic microscopy of single cells, *Photoacoustics* **1**, 49 (2013).

[16] K. Maslov, H. F. Zhang, S. Hu, L. V. Wang, Optical-resolution photoacoustic microscopy for in vivo imaging of single capillaries, *Optics letters* **33**, 929 (2008).

[17] X. L. Deán-Ben, A. Ozbek, D. Razansky, Volumetric real-time tracking of peripheral human vasculature with gpu-accelerated three-dimensional optoacoustic tomography, *IEEE Transactions on Medical Imaging* **32**, 2050 (2013).

[18] D. Razansky, A. Buehler, V. Ntziachristos, Volumetric real-time multispectral optoacoustic tomography of biomarkers, *Nature protocols* **6**, 1121 (2011).

[19] H. P. Brecht, R. Su, M. Fronheiser, S. A. Ermilov, A. Conjusteau, A. A. Oraevsky, Whole-body three-dimensional optoacoustic tomography system for small animals, *Journal of Biomedical Optics* **14** (2009).

[20] A. P. Jathoul, *et al.*, Deep in vivo photoacoustic imaging of mammalian tissues using a tyrosinase-based genetic reporter, *Nature Photonics* (2015).

[21] S. Morscher, W. H. Driessen, J. Claussen, N. C. Burton, Semi-quantitative multispectral optoacoustic tomography (msot) for volumetric pk imaging of gastric emptying, *Photoacoustics* **2**, 103 (2014).

[22] X. D. Wang, Y. J. Pang, G. Ku, X. Y. Xie, G. Stoica, L. H. V. Wang, Noninvasive laser-induced photoacoustic tomography for structural and functional in vivo imaging of the brain, *Nature Biotechnology* **21**, 803 (2003).

[23] A. Buehler, E. Herzog, D. Razansky, V. Ntziachristos, Video rate optoacoustic tomography of mouse kidney perfusion, *Optics letters* **35**, 2475 (2010).

[24] X. L. Deán-Ben, S. J. Ford, D. Razansky, High-frame rate four dimensional optoacoustic tomography enables visualization of cardiovascular dynamics and mouse heart perfusion, *Scientific reports* **5** (2015).

[25] G. Sela, A. Lauri, X. L. Deán-Ben, M. Kneipp, V. Ntziachristos, S. Shoham, G. G. Westmeyer, D. Razansky, Functional optoacoustic neuro-tomography (font) for whole-brain monitoring of calcium indicators, *arXiv preprint arXiv:1501.02450* (2015).

[26] J. Laufer, C. Elwell, D. Delpy, P. Beard, In vitro measurements of absolute blood oxygen saturation using pulsed near-infrared photoacoustic spectroscopy: accuracy and resolution, *Physics in Medicine and Biology* **50**, 4409 (2005).

[27] J. Yao, J. Xia, K. I. Maslov, M. Nasiriavanaki, V. Tsytsarev, A. V. Demchenko, L. V. Wang, Noninvasive photoacoustic computed tomography of mouse brain metabolism in vivo, *Neuroimage* **64**, 257 (2013).

[28] A. De La Zerda, *et al.*, Carbon nanotubes as photoacoustic molecular imaging agents in living mice, *Nature Nanotechnology* **3**, 557 (2008).

[29] A. Buehler, E. Herzog, A. Ale, B. D. Smith, V. Ntziachristos, D. Razansky, High resolution tumor targeting in living mice by means of multispectral optoacoustic tomography, *EJNMMI Research* **2**, 14 (2012).

[30] D. Razansky, M. Distel, C. Vinegoni, R. Ma, N. Perrimon, R. W. Koster, V. Ntziachristos, Multispectral opto-acoustic tomography of deep-seated fluorescent proteins in vivo, *Nature Photonics* **3**, 412 (2009).

[31] S. Zackrisson, S. van de Ven, S. Gambhir, Light in and sound out: emerging translational strategies for photoacoustic imaging, *Cancer research* **74**, 979 (2014).

[32] K. Jansen, M. Wu, A. F. van der Steen, G. van Soest, Photoacoustic imaging of human coronary atherosclerosis in two spectral bands, *Photoacoustics* **2**, 12 (2014).

[33] E. Herzog, A. Taruttis, N. Beziere, A. A. Lutich, D. Razansky, V. Ntziachristos, Optical imaging of cancer heterogeneity with multispectral optoacoustic tomography, *Radiology* **263**, 461 (2012).

[34] M. Kneipp, J. Turner, S. Hambauer, S. M. Krieg, J. Lehmberg, U. Lindauer, D. Razansky, Functional real-time optoacoustic imaging of middle cerebral artery occlusion in mice, *PloS one* **9**, e96118 (2014).

[35] X. L. Deán-Ben, D. Razansky, Adding fifth dimension to optoacoustic imaging: volumetric time-resolved spectrally enriched tomography, *Light: Science & Applications* **3**, e137 (2014).

[36] N. C. Burton, *et al.*, Multispectral opto-acoustic tomography (msot) of the brain and glioblastoma characterization, *Neuroimage* **65**, 522 (2013).

[37] J.-M. Yang, K. Maslov, H.-C. Yang, Q. Zhou, K. K. Shung, L. V. Wang, Photoacoustic endoscopy, *Optics letters* **34**, 1591 (2009).

[38] L. Song, C. Kim, K. Maslov, K. K. Shung, L. V. Wang, High-speed dynamic 3d photoacoustic imaging of sentinel lymph node in a murine model using an ultrasound array, *Medical physics* **36**, 3724 (2009).

[39] A. de La Zerda, Y. M. Paulus, R. Teed, S. Bodapati, Y. Dollberg, B. T. Khuri-Yakub, M. S. Blumenkranz, D. M. Moshfeghi, S. S. Gambhir, Photoacoustic ocular imaging, *Optics letters* **35**, 270 (2010).

[40] D. Piras, C. Grijsen, P. Schütte, W. Steenbergen, S. Manohar, Photoacoustic needle: minimally invasive guidance to biopsy, *Journal of biomedical optics* **18**, 070502 (2013).

[41] M. Heijblom, D. Piras, W. Xia, J. Van Hespen, J. Klaase, F. Van den Engh, T. Van Leeuwen, W. Steenbergen, S. Manohar, Visualizing breast cancer using the twente photoacoustic mammoscope: what do we learn from twelve new patient measurements?, *Optics express* **20**, 11582 (2012).

[42] R. A. Kruger, C. M. Kuzmiak, R. B. Lam, D. R. Reinecke, S. P. Del Rio, D. Steed, Dedicated 3d photoacoustic breast imaging, *Medical physics* **40**, 113301 (2013).

[43] A. Dima, V. Ntziachristos, Non-invasive carotid imaging using optoacoustic tomography, *Optics express* **20**, 25044 (2012).

[44] E. Z. Zhang, J. G. Laufer, R. B. Pedley, P. C. Beard, In vivo high-resolution 3d photoacoustic imaging of superficial vascular anatomy, *Physics in Medicine and Biology* **54**, 1035 (2009).

[45] L. Vionnet, J. Gâteau, M. Schwarz, A. Buehler, V. Ermolayev, V. Ntziachristos, 24-mhz scanner for optoacoustic imaging of skin and burn, *IEEE Transactions on Medical Imaging* **33**, 535 (2014).

[46] Z. Deák, M. Treitl, M. Reiser, C. Degenhart, Angiographische diagnostik akraler durchblutungsstörungen an der oberen extremität, *Der Radiologe* **50**, 879 (2010).

[47] M. Dierolf, A. Menzel, P. Thibault, P. Schneider, C. M. Kewish, R. Wepf, O. Bunk, F. Pfeiffer, Ptychographic x-ray computed tomography at the nanoscale, *Nature* **467**, 436 (2010).

[48] S. M. Daly, M. J. Leahy, 'Go with the flow': a review of methods and advancements in blood flow imaging, *Journal of biophotonics* **6**, 217 (2013).

[49] C. Kim, R. Qin, J. S. Xu, L. V. Wang, R. Xu, Multifunctional microbubbles and nanobubbles for photoacoustic and ultrasound imaging, *Journal of biomedical optics* **15**, 010510 (2010).

[50] S. R. Meikle, P. Kench, M. Kassiou, R. B. Banati, Small animal spect and its place in the matrix of molecular imaging technologies, *Physics in medicine and biology* **50**, R45 (2005).

[51] W. Zhang, J. R. Xu, Q. Lu, S. Ye, X. S. Liu, High-resolution magnetic resonance angiography of digital arteries in ssc patients on 3 tesla: preliminary study, *Rheumatology* **50**, 1712 (2011).

[52] R. Meier, K. Thuermel, P. B. Noël, P. Moog, M. Sievert, C. Ahari, R. A. Nasirudin, D. Golovko, B. Haller, C. Ganter, Synovitis in patients with early inflammatory arthritis monitored with quantitative analysis of dynamic contrast-enhanced optical imaging and mr imaging, *Radiology* **270**, 176 (2014).

[53] D. Le Bihan, J. Mangin, C. Poupon, C. A. Clark, S. Pappata, N. Molko, H. Chabriat, Diffusion tensor imaging: concepts and applications, *Journal of magnetic resonance imaging* **13**, 534 (2001).

[54] K. Schulz, E. Sydekum, R. Krueppel, C. J. Engelbrecht, F. Schlegel, A. Schröter, M. Rudin, F. Helmchen, Simultaneous bold fmri and fiber-optic calcium recording in rat neocortex, *Nature methods* **9**, 597 (2012).

[55] M. F. Reiser, G. P. Bongartz, R. Erlemann, M. Schneider, T. Pauly, H. Sittek, P. E. Peters, Gadolinium-dtpa in rheumatoid arthritis and related diseases: first results with dynamic magnetic resonance imaging, *Skeletal radiology* **18**, 591 (1989).

[56] J. Weizenecker, B. Gleich, J. Rahmer, H. Dahnke, J. Borgert, Three-dimensional real-time in vivo magnetic particle imaging, *Physics in medicine and biology* **54**, L1 (2009).

[57] E. L. Costa, R. G. Lima, M. B. Amato, *Intensive Care Medicine* (Springer, 2009), pp. 394–404.

[58] A. Ale, V. Ermolayev, E. Herzog, C. Cohrs, M. H. de Angelis, V. Ntziachristos, Fmt-xct: in vivo animal studies with hybrid fluorescence molecular tomography-x-ray computed tomography, *Nature methods* **9**, 615 (2012).

[59] H. Estrada, E. Sobol, O. Baum, D. Razansky, Hybrid optoacoustic and ultrasound biomicroscopy monitors laser-induced tissue modifications and magnetite nanoparticle impregnation, *Laser Physics Letters* **11**, 125601 (2014).

[60] T. Beyer, *et al.*, A combined pet/ct scanner for clinical oncology, *Journal of nuclear medicine* **41**, 1369 (2000).

[61] X. Shu, V. Lev-Ram, T. J. Deerinck, Y. Qi, E. B. Ramko, M. W. Davidson, Y. Jin, M. H. Ellisman, R. Y. Tsien, A genetically encoded tag for correlated light and electron microscopy of intact cells, tissues, and organisms, *PLoS Biol* **9**, e1001041 (2011).

[62] F. Helmchen, W. Denk, Deep tissue two-photon microscopy, *Nature methods* **2**, 932 (2005).

[63] J. Huisken, J. Swoger, F. Del Bene, J. Wittbrodt, E. H. Stelzer, Optical sectioning deep inside live embryos by selective plane illumination microscopy, *Science* **305**, 1007 (2004).

[64] S. W. Hell, Far-field optical nanoscopy, *Science* **316**, 1153 (2007).

[65] T. A. Klar, S. Jakobs, M. Dyba, A. Egner, S. W. Hell, Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission, *Proceedings of the National Academy of Sciences* **97**, 8206 (2000).

[66] S. Ashkenazi, Photoacoustic lifetime imaging of dissolved oxygen using methylene blue, *Journal of biomedical optics* **15**, 040501 (2010).

[67] Y. Wang, J. Xia, L. V. Wang, Deep-tissue photoacoustic tomography of förster resonance energy transfer, *Journal of biomedical optics* **18**, 101316 (2013).

[68] E. O. Potma, C. L. Evans, X. S. Xie, Heterodyne coherent anti-stokes raman scattering (cars) imaging, *Optics Letters* **31**, 241 (2006).

[69] J. Glatz, J. Varga, P. B. Garcia-Allende, M. Koch, F. R. Greten, V. Ntziachristos, Concurrent video-rate color and near-infrared fluorescence laparoscopy, *Journal of biomedical optics* **18**, 101302 (2013).

[70] A. F. Fercher, W. Drexler, C. K. Hitzenberger, T. Lasser, Optical coherence tomography-principles and applications, *Reports on progress in physics* **66**, 239 (2003).

[71] V. Ntziachristos, A. Yodh, M. Schnall, B. Chance, Concurrent mri and diffuse optical tomography of breast after indocyanine green enhancement, *Proceedings of the National Academy of Sciences* **97**, 2767 (2000).

[72] V. Ntziachristos, C.-H. Tung, C. Bremer, R. Weissleder, Fluorescence molecular tomography resolves protease activity in vivo, *Nature medicine* **8**, 757 (2002).

[73] F. Stuker, J. Ripoll, M. Rudin, Fluorescence molecular tomography: principles and potential for pharmaceutical research, *Pharmaceutics* **3**, 229 (2011).

[74] A. Hussain, K. Daoudi, E. Hondebrink, W. Steenbergen, Mapping optical fluence variations in highly scattering media by measuring ultrasonically modulated backscattered light, *Journal of Biomedical Optics* **66002**, 1 (2014).

[75] V. Kalchenko, Y. Kuznetsov, I. Meglinski, A. Harmelin, Label free in vivo laser speckle imaging of blood and lymph vessels, *Journal of biomedical optics* **17**, 0505021 (2012).

[76] W.-F. Cheong, S. A. Prahl, A. J. Welch, A review of the optical properties of biological tissues, *IEEE journal of quantum electronics* **26**, 2166 (1990).

[77] B. Cox, J. G. Laufer, S. R. Arridge, P. C. Beard, Quantitative spectroscopic photoacoustic imaging: a review, *Journal of Biomedical Optics* **17** (2012).

[78] L. Wang, S. L. Jacques, L. Zheng, Mcml–monte carlo modeling of light transport in multi-layered tissues, *Computer methods and programs in biomedicine* **47** (1995).

[79] C. Zhu, Q. Liu, Review of monte carlo modeling of light transport in tissues, *Journal of biomedical optics* **18**, 050902 (2013).

[80] G. Bal, K. Ren, On multi-spectral quantitative photoacoustic tomography in diffusive regime, *Inverse Problems* **28**, 025010 (2012).

[81] J. Glatz, N. C. Deliolanis, A. Buehler, D. Razansky, V. Ntziachristos, Blind source unmixing in multi-spectral optoacoustic tomography, *Optics Express* **19**, 3175 (2011).

[82] S. Tzoumas, N. Deliolanis, S. Morscher, V. Ntziachristos, Unmixing molecular agents from absorbing tissue in multispectral optoacoustic tomography, *IEEE Transactions on Medical Imaging* **33**, 48 (2014).

[83] N. Smith, Carbon dioxide and oxygen transport, www.diatronic.co.uk/nds/webpub/haemoglobin_structure.htm (2015).

[84] D. M. Chudakov, M. V. Matz, S. Lukyanov, K. A. Lukyanov, Fluorescent proteins and their applications in imaging living cells and tissues, *Physiological Reviews* **90**, 1103 (2010).

[85] L. Tong, Q. S. Wei, A. Wei, J. X. Cheng, Gold nanorods as contrast agents for biological imaging: Optical properties, surface conjugation and photothermal effects, *Photochemistry and Photobiology* **85**, 21 (2009).

[86] . Sigma Aldrich Co. LLC, Quantum dots, www.sigmaaldrich.com/materials-science/nanomaterials/quantum-dots.html (2015).

[87] C. L. Amiot, S. Xu, S. Liang, L. Pan, J. X. Zhao, Near-infrared fluorescent materials for sensing of biological targets, *Sensors* **8**, 3082 (2008).

[88] J. Stritzker, L. Kirscher, M. Scadeng, N. C. Deliolanis, S. Morscher, P. Symvoulidis, K. Schaefer, Q. Zhang, L. Buckel, M. Hess, Vaccinia virus-mediated melanin production allows mr and optoacoustic deep tissue imaging and laser-induced thermotherapy of cancer, *Proceedings of the National Academy of Sciences* **110**, 3316 (2013).

[89] B. E. Schaafsma, J. S. D. Mieog, M. Hutteman, J. R. Van der Vorst, P. J. Kuppen, C. W. Löwik, J. V. Frangioni, C. J. Van de Velde, A. L. Vahrmeijer, The clinical use of indocyanine green as a near-infrared fluorescent contrast agent for image-guided oncologic surgery, *Journal of surgical oncology* **104**, 323 (2011).

[90] M. Zimmer, Green fluorescent protein (gfp): Applications, structure, and related photophysical behavior, *Chemical Reviews* **102**, 759 (2002).

[91] L. Li, R. J. Zemp, G. Lungu, G. Stoica, L. V. Wang, Photoacoustic imaging of lacz gene expression in vivo, *Journal of biomedical optics* **12**, 020504 (2007).

[92] Q. Zhang, N. Iwakuma, P. Sharma, B. Moudgil, C. Wu, J. McNeill, H. Jiang, S. Grobmyer, Gold nanoparticles as a contrast agent for in vivo tumor imaging with photoacoustic tomography, *Nanotechnology* **20**, 395102 (2009).

[93] X. Michalet, F. F. Pinaud, L. A. Bentolila, J. M. Tsay, S. Doose, J. J. Li, G. Sundaresan, A. M. Wu, S. S. Gambhir, S. Weiss, Quantum dots for live cells, in vivo imaging, and diagnostics, *Science* **307**, 538 (2005).

[94] L. V. Wang, Tutorial on photoacoustic microscopy and computed tomography, *IEEE Journal of Selected Topics in Quantum Electronics* **14**, 171 (2008).

[95] S. Manohar, S. E. Vaartjes, J. C. van Hespen, J. M. Klaase, F. M. van den Engh, W. Steenbergen, T. G. Van Leeuwen, Initial results of in vivo non-invasive cancer imaging in the human breast using near-infrared photoacoustics, *Optics express* **15**, 12277 (2007).

[96] M. Nasiriavanaki, J. Xia, H. Wan, A. Q. Bauer, J. P. Culver, L. V. Wang, High-resolution photoacoustic tomography of resting-state functional connectivity in the mouse brain, *Proceedings of the National Academy of Sciences* **111**, 21 (2014).

[97] R. A. Kruger, P. Liu, Y. R. Fang, C. R. Appledorn, Photoacoustic ultrasound (paus)–reconstruction tomography, *Medical physics* **22**, 1605 (1995).

[98] I. V. Larina, K. V. Larin, R. O. Esenaliev, Real-time optoacoustic monitoring of temperature in tissues, *Journal of Physics D: Applied Physics* **38**, 2633 (2005).

[99] L. Wang, C. Zhang, L. V. Wang, Grueneisen relaxation photoacoustic microscopy, *Physical review letters* **113**, 174301 (2014).

[100] F. A. Duck, *Physical properties of tissues: a comprehensive reference book* (Academic press, 2013).

[101] X. L. Deán-Ben, D. Razansky, V. Ntziachristos, The effects of acoustic attenuation in optoacoustic signals, *Physics in Medicine and Biology* **56**, 6129 (2011).

[102] B. E. Treeby, E. Z. Zhang, B. T. Cox, Photoacoustic tomography in absorbing acoustic media using time reversal, *Inverse Problems* **26** (2010).

[103] S. A. Telenkov, A. Mandelis, Fourier-domain methodology for depth-selective photothermoacoustic imaging of tissue chromophores, *European Physical Journal-Special Topics* **153**, 443 (2008).

[104] S. Kellnberger, N. C. Deliolanis, D. Queiros, G. Sergiadis, V. Ntziachristos, In vivo frequency domain optoacoustic tomography, *Optics Letters* **37**, 3423 (2012).

[105] R. A. Kruger, K. D. Miller, H. E. Reynolds, W. L. Kiser, D. R. Reinecke, G. A. Kruger, Breast cancer in vivo: Contrast enhancement with thermoacoustic ct at 434 mhz - feasibility study, *Radiology* **216**, 279 (2000).

[106] S. Kellnberger, A. Hajiaboli, D. Razansky, V. Ntziachristos, Near-field thermoacoustic tomography of small animals, *Physics in medicine and biology* **56**, 3433 (2011).

[107] L. Xiang, B. Han, C. Carpenter, G. Pratx, Y. Kuang, L. Xing, X-ray acoustic computed tomography with pulsed x-ray beam from a medical linear accelerator, *Medical physics* **40**, 010701 (2013).

[108] W. Assmann, *et al.*, Ionoacoustic characterization of the proton bragg peak with submillimeter accuracy, *Medical physics* **42**, 567 (2015).

[109] T. J. Allen, P. C. Beard, Pulsed near-infrared laser diode excitation system for biomedical photoacoustic imaging, *Optics letters* **31**, 3462 (2006).

[110] . ANSI Standard, Z136. 1-2000: for safe use of lasers.

[111] A. Buehler, M. Kacprowicz, A. Taruttis, V. Ntziachristos, Real-time handheld multispectral optoacoustic imaging, *Optics letters* **38**, 1404 (2013).

[112] E. Zhang, J. Laufer, P. Beard, Backward-mode multiwavelength photoacoustic scanner using a planar fabry-perot polymer film ultrasound sensor for high-resolution three-dimensional imaging of biological tissues, *Applied optics* **47**, 561 (2008).

[113] J. Gateau, A. Chekkoury, V. Ntziachristos, High-resolution optoacoustic mesoscopy with a 24 mhz multidetector translate-rotate scanner, *Journal of biomedical optics* **18**, 106005 (2013).

[114] A. Buehler, A. Rosenthal, T. Jetzfellner, A. Dima, D. Razansky, V. Ntziachristos, Model-based optoacoustic inversions with incomplete projection data, *Medical Physics* **38**, 1694 (2011).

[115] Y. Xu, L. V. Wang, G. Ambartsoumian, P. Kuchment, Reconstructions in limited-view thermoacoustic tomography, *Medical Physics* **31**, 724 (2004).

[116] B. Cox, S. Arridge, P. Beard, Photoacoustic tomography with a limited-aperture planar sensor and a reverberant cavity, *Inverse Problems* **23**, S95 (2007).

[117] J. Turner, H. Estrada, M. Kneipp, D. Razansky, Improved optoacoustic microscopy through three-dimensional spatial impulse response synthetic aperture focusing technique, *Optics letters* **39**, 3390 (2014).

[118] M. Haltmeier, O. Scherzer, P. Burgholzer, G. Paltauf, Thermoacoustic computed tomography with large planar receivers, *Inverse Problems* **20**, 1663 (2004).

[119] X. Jin, O. Oralkan, F. Degertekin, B. T. Khuri-Yakub, Characterization of one-dimensional capacitive micromachined ultrasonic immersion transducer arrays, *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control* **48**, 750 (2001).

[120] G. Paltauf, R. Nuster, M. Haltmeier, P. Burgholzer, Photoacoustic tomography using a mach-zehnder interferometer as an acoustic line detector, *Applied Optics* **46**, 3352 (2007).

[121] S.-W. Huang, S.-L. Chen, T. Ling, A. Maxwell, M. ODonnell, L. J. Guo, S. Ashkenazi, Low-noise wideband ultrasound detection using polymer microring resonators, *Applied physics letters* **92** (2008).

[122] A. Rosenthal, D. Razansky, V. Ntziachristos, Wideband optical sensing using pulse interferometry, *Optics express* **20**, 19016 (2012).

[123] M. Omar, J. Gateau, V. Ntziachristos, Raster-scan optoacoustic mesoscopy in the 25-125 mhz range, *Optics letters* **38**, 2472 (2013).

[124] X. L. Deán-Ben, D. Razansky, Functional optoacoustic human angiography with handheld video rate three dimensional scanner, *Photoacoustics* **1**, 68 (2013).

[125] H. F. Zhang, K. Maslov, G. Stoica, L. H. V. Wang, Functional photoacoustic microscopy for high-resolution and noninvasive in vivo imaging, *Nature Biotechnology* **24**, 848 (2006).

[126] C. Lutzweiler, R. Meier, E. Rummeny, V. Ntziachristos, D. Razansky, Real-time optoacoustic tomography of indocyanine green perfusion and oxygenation parameters in human finger vasculature, *Optics letters* **39**, 4061 (2014).

[127] J. Gamelin, A. Maurudis, A. Aguirre, F. Huang, P. Y. Guo, L. V. Wang, Q. Zhu, A real-time photoacoustic tomography system for small animals, *Optics Express* **17**, 10489 (2009).

[128] L. Xiang, B. Wang, L. Ji, H. Jiang, 4-d photoacoustic tomography, *Scientific reports* **3** (2013).

[129] A. Dima, N. C. Burton, V. Ntziachristos, Multispectral optoacoustic tomography at 64, 128, and 256 channels, *Journal of biomedical optics* **19**, 036021 (2014).

[130] R. Avni, M. Lysenko, C. Lutzweiler, O. Six, D. Razansky, M. Neeman, unpublished.

[131] D. Finch, S. K. Patch, Rakesh, Determining a function from its mean values over a family of spheres, *SIAM Journal on Mathematical Analysis* **35**, 1213 (2004).

[132] M. H. Xu, L. H. V. Wang, Universal back-projection algorithm for photoacoustic computed tomography, *Physical Review E* **71** (2005).

[133] G. Paltauf, R. Nuster, P. Burgholzer, Weight factors for limited angle photoacoustic tomography, *Physics in Medicine and Biology* **54**, 3303 (2009).

[134] X. L. Deán-Ben, R. Ma, A. Rosenthal, V. Ntziachristos, D. Razansky, Weighted model-based optoacoustic reconstruction in acoustic scattering media, *Physics in medicine and biology* **58**, 5555 (2013).

[135] Y. Hristova, P. Kuchment, L. Nguyen, Reconstruction and time reversal in thermoacoustic tomography in acoustically homogeneous and inhomogeneous media, *Inverse Problems* **24**, 055006 (2008).

[136] B. E. Treeby, B. T. Cox, k-wave: Matlab toolbox for the simulation and reconstruction of photoacoustic wave fields, *Journal of biomedical optics* **15**, 021314 (2010).

[137] R. M. Lewitt, S. Matej, Overview of methods for image reconstruction from projections in emission computed tomography, *Proceedings of the IEEE* **91**, 1588 (2003).

[138] A. Rosenthal, D. Razansky, V. Ntziachristos, Fast semi-analytical model-based acoustic inversion for quantitative optoacoustic tomography, *IEEE Transactions on Medical Imaging* **29**, 1275 (2010).

[139] K. Wang, R. W. Schoonover, R. Su, A. Oraevsky, M. Anastasio, Discrete imaging models for three-dimensional optoacoustic tomography using radially symmetric expansion functions, *IEEE Transactions on Medical Imaging* **33**, 1180 (2014).

[140] J. Provost, F. Lesage, The application of compressed sensing for photo-acoustic tomography, *IEEE Transactions on Medical Imaging* **28**, 585 (2009).

[141] A. Beck, M. Teboulle, A fast iterative shrinkage-thresholding algorithm for linear inverse problems, *SIAM Journal on Imaging Sciences* **2**, 183 (2009).

[142] G. H. Golub, C. F. Van Loan, *Matrix computations*, vol. 3 (JHU Press, 2012).

[143] P. C. Hansen, Truncated singular value decomposition solutions to discrete ill-posed problems with ill-determined numerical rank, *SIAM Journal on Scientific and Statistical Computing* **11**, 503 (1990).

[144] C. C. Paige, M. A. Saunders., Lsqr: An algorithm for sparse linear equations and sparse least squares, *ACM Transactions on Mathematical Software* **8**, 43 (1982).

[145] M. Beister, D. Kolditz, W. A. Kalender, Iterative reconstruction methods in x-ray ct, *Physica medica* **28**, 94 (2012).

[146] A. Rosenthal, V. Ntziachristos, D. Razansky, Model-based optoacoustic inversion with arbitrary-shape detectors, *Medical Physics* **38**, 4285 (2011).

[147] X. L. Deán-Ben, V. Ntziachristos, D. Razansky, Effects of small variations of speed of sound in optoacoustic tomographic imaging, *Medical physics* **41**, 073301 (2014).

[148] K. Wang, R. Su, A. A. Oraevsky, M. A. Anastasio, Investigation of iterative image reconstruction in three-dimensional optoacoustic tomography, *Physics in Medicine and Biology* **57**, 5399 (2012).

[149] L. Ding, X. L. Deán-Ben, C. Lutzweiler, D. Razansky, V. Ntziachristos, Efficient non-negative constrained model-based inversion in optoacoustic tomography, *Physics in medicine and biology* **60**, 6733 (2015).

[150] C. Huang, K. Wang, L. Nie, L. V. Wang, M. A. Anastasio, Full-wave iterative image reconstruction in photoacoustic tomography with acoustically inhomogeneous media, *IEEE Transactions on Medical Imaging* **32**, 1097 (2013).

[151] A. Taruttis, M. Wildgruber, K. Kosanke, N. Beziere, K. Licha, R. Haag, M. Aichler, A. Walch, E. Rummeny, V. Ntziachristos, Multispectral optoacoustic tomography of myocardial infarction, *Photoacoustics* **1**, 3 (2013).

[152] M. P. Fronheiser, S. A. Ermilov, H.-P. Brecht, A. Conjusteau, R. Su, K. Mehta, A. A. Oraevsky, Real-time optoacoustic monitoring and three-dimensional mapping of a human arm vasculature, *Journal of biomedical optics* **15**, 021305 (2010).

[153] . ithera medical GmbH Munich Germany, Introducing the msot scanner family, www.ithera-medical.com/technology/preclinical-systems.html (2015).

[154] X. L. Deán-Ben, D. Razansky, Portable spherical array probe for volumetric real-time optoacoustic imaging at centimeter-scale depths, *Optics express* **21**, 28062 (2013).

[155] M. H. Xu, L. V. Wang, Analytic explanation of spatial resolution related to bandwidth and detector aperture size in thermoacoustic or photoacoustic reconstruction, *Physical Review E* **67** (2003).

[156] C. Lutzweiler, X. L. Deán-Ben, D. Razansky, Expediting model-based optoacoustic reconstructions with tomographic symmetries, *Medical physics* **41**, 013302 (2014).

[157] T. Fließbach, *Mechanik: Lehrbuch zur Theoretischen Physik I* (Springer-Verlag, 2014).

[158] P. Mohajerani, S. Kellnberger, V. Ntziachristos, Frequency domain optoacoustic tomography using amplitude and phase, *Photoacoustics* **2**, 111 (2014).

[159] X. L. Deán-Ben, V. Ntziachristos, D. Razansky, Acceleration of optoacoustic model-based reconstruction using angular image discretization, *IEEE Transactions on Medical Imaging* **31**, 1154 (2012).

[160] S. H. Bu, Z. B. Liu, T. Shiina, K. Kondo, M. Yamakawa, K. Fukutani, Y. Someda, Y. Asao, Model-based reconstruction integrated with fluence compensation for photoacoustic tomography, *IEEE Transactions on Biomedical Engineering* **59**, 1354 (2012).

[161] J. Aguirre, A. Giannoula, T. Minagawa, L. Funk, P. Turon, T. Durduran, A low memory cost model based reconstruction algorithm exploiting translational symmetry for photoacustic microscopy, *Biomedical optics express* **4**, 2813 (2013).

[162] M. Araque-Caballero, J. Gateau, X.-L. Deán-Ben, V. Ntziachristos, Model-based optoacoustic image reconstruction of large three-dimensional tomographic datasets acquired with an array of directional detectors, *IEEE Transactions on Medical Imaging* **33**, 433 (2014).

[163] T. Jetzfellner, A. Rosenthal, K.-H. Englmeier, A. Dima, M. n. A. Caballero, D. Razansky, V. Ntziachristos, Interpolated model-matrix optoacoustic tomography of the mouse brain, *Applied Physics Letters* **98**, 163701 (2011).

[164] F. Xu, K. Mueller, Accelerating popular tomographic reconstruction algorithms on commodity pc graphics hardware, *IEEE Transactions on Nuclear Science* **52**, 654 (2005).

[165] G. Pratx, L. Xing, Gpu computing in medical physics: A review, *Medical physics* **38**, 2685 (2011).

[166] E. Lindholm, J. Nickolls, S. Oberman, J. Montrym, Nvidia tesla: A unified graphics and computing architecture, *IEEE micro* pp. 39–55 (2008).

[167] S. A. Manavski, *IEEE International Conference onSignal Processing and Communications* (IEEE, 2007), pp. 65–68.

[168] J. A. Anderson, C. D. Lorenz, A. Travesset, General purpose molecular dynamics simulations fully implemented on graphics processing units, *Journal of Computational Physics* **227**, 5342 (2008).

[169] N. Ren, J. Liang, X. Qu, J. Li, B. Lu, J. Tian, Gpu-based monte carlo simulation for light propagation in complex heterogeneous tissues, *Optics express* **18**, 6811 (2010).

[170] K. Wang, C. Huang, Y.-J. Kao, C.-Y. Chou, A. A. Oraevsky, M. A. Anastasio, Accelerating image reconstruction in three-dimensional optoacoustic tomography on graphics processing units, *Medical physics* **40**, 023301 (2013).

[171] K. Maslov, G. Stoica, L. H. V. Wang, In vivo dark-field reflection-mode photoacoustic microscopy, *Optics Letters* **30**, 625 (2005).

[172] C. Lutzweiler, S. Tzoumas, A. Rosenthal, V. Ntziachristos, D. Razansky, High-throughput sparsity-based inversion scheme for optoacoustic tomography, *IEEE Transactions on Medical Imaging* **35**, 674 (2016).

[173] N. Hurley, S. Rickard, Comparing measures of sparsity, *IEEE Transactions on Information Theory* **55**, 4723 (2009).

[174] S. G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**, 674 (1989).

[175] S. Mallat, *A wavelet tour of signal processing* (Academic press, 1999).

[176] J. A. Viator, B. Choi, M. Ambrose, J. Spanier, J. S. Nelson, In vivo port-wine stain depth determination with a photoacoustic probe, *Applied Optics* **42**, 3215 (2003).

[177] I. Jolliffe, *Principal component analysis* (Wiley Online Library, 2005).

[178] A. Rosenthal, T. Jetzfellner, D. Razansky, V. Ntziachristos, Efficient framework for model-based tomographic image reconstruction using wavelet packets, *IEEE Transactions on Medical Imaging* **31**, 1346 (2012).

[179] M. N. Wernick, E. J. Infusino, M. Milosevic, Fast spatio-temporal image reconstruction for dynamic pet, *IEEE Transactions on Medical Imaging* **18**, 185 (1999).

[180] S. Tzoumas, A. Rosenthal, C. Lutzweiler, D. Razansky, V. Ntziachristos, Spatiospectral denoising framework for multispectral optoacoustic imaging based on sparse signal representation, *Medical physics* **41**, 113301 (2014).

[181] Y. Zhang, Y. Wang, C. Zhang, Efficient discrete cosine transform model-based algorithm for photoacoustic image reconstruction, *Journal of biomedical optics* **18**, 066008 (2013).

[182] K. Wang, J. Xia, C. Li, L. Wang, M. Anastasio, Fast spatiotemporal image reconstruction based on low-rank matrix estimation for dynamic photoacoustic computed tomography, *Journal of biomedical optics* **19**, 56007 (2014).

[183] A. Tarantola, *Inverse problem theory and methods for model parameter estimation* (SIAM, 2005).

[184] G. Ku, X. Wang, G. Stoica, L. V. Wang, Multiple-bandwidth photoacoustic tomography, *Physics in medicine and biology* **49**, 1329 (2004).

[185] J. Gateau, A. Chekkoury, V. Ntziachristos, Ultra-wideband three-dimensional optoacoustic tomography, *Optics letters* **38**, 4671 (2013).

[186] H. Roitner, M. Haltmeier, R. Nuster, D. P. O'Leary, T. Berer, G. Paltauf, H. Grün, P. Burgholzer, Deblurring algorithms accounting for the finite detector size in photoacoustic tomography, *Journal of biomedical optics* **19**, 056011 (2014).

[187] D. Queirós, X. L. Deán-Ben, A. Buehler, D. Razansky, A. Rosenthal, V. Ntziachristos, Modeling the shape of cylindrically focused transducers in three-dimensional optoacoustic tomography, *Journal of biomedical optics* **18**, 076014 (2013).

[188] J. Jose, R. G. H. Willemink, S. Resink, D. Piras, J. C. G. van Hespen, C. H. Slump, W. Steenbergen, T. G. van Leeuwen, S. Manohar, Passive element enriched photoacoustic computed tomography (per pact) for simultaneous imaging of acoustic propagation properties and light absorption, *Optics Express* **19**, 2093 (2011).

[189] A. Kirsch, O. Scherzer, Simultaneous reconstructions of absorption density and wave speed with photoacoustic measurements, *SIAM Journal on Applied Mathematics* **72**, 1508 (2012).

[190] A. Santos, C. Ortiz de Solorzano, J. J. Vaquero, J. Pena, N. Malpica, F. Del Pozo, Evaluation of autofocus functions in molecular cytogenetic analysis, *Journal of microscopy* **188**, 264 (1997).

[191] B. E. Treeby, T. K. Varslot, E. Z. Zhang, J. G. Laufer, P. C. Beard, Automatic sound speed selection in photoacoustic image reconstruction using an autofocus approach, *Journal of Biomedical Optics* **16** (2011).

[192] D. Modgil, M. A. Anastasio, P. J. La Riviere, Image reconstruction in photoacoustic tomography with variable speed of sound using a higher-order geometrical acoustics approximation, *Journal of Biomedical Optics* **15** (2010).

[193] C. Lutzweiler, R. Meier, D. Razansky, Optoacoustic image segmentation based on signal domain analysis, *Photoacoustics* **3**, 151 (2015).

[194] M. A. Anastasio, J. Zhang, X. C. Pan, Y. Zou, G. Ku, L. H. V. Wang, Half-time image reconstruction in thermoacoustic tomography, *IEEE Transactions on Medical Imaging* **24**, 199 (2005).

[195] J. Aguirre, M. Schwarz, D. Soliman, A. Buehler, M. Omar, V. Ntziachristos, Broadband mesoscopic optoacoustic tomography reveals skin layers, *Optics letters* **39**, 6297 (2014).

[196] J. Braga-Silva, C. Kuyven, F. Fallopa, W. Albertoni, An anatomical study of the dorsal cutaneous branches of the digital arteries, *Journal of Hand Surgery (British and European Volume)* **27**, 577 (2002).

[197] A. L. Herrick, Management of raynaud's phenomenon and digital ischemia, *Current rheumatology reports* **15**, 1 (2013).

[198] P. van Es, S. Biswas, H. Moens, W. Steenbergen, S. Manohar, Initial results of finger imaging using photoacoustic computed tomography, *Journal of biomedical optics* **19**, 60501 (2014).

[199] R. K. Wang, S. L. Jacques, Z. Ma, S. Hurst, S. R. Hanson, A. Gruber, Three dimensional optical angiography, *Optics express* **15**, 4083 (2007).

[200] H. Ackermann, *AllEx-Alles fürs Examen: Das Kompendium für die 2. ÄP* (Georg Thieme Verlag, 2012).

[201] B. Zitova, J. Flusser, Image registration methods: a survey, *Image and vision computing* **21**, 977 (2003).

[202] J.-M. Morel, G. Yu, Asift: A new framework for fully affine invariant image comparison, *SIAM Journal on Imaging Sciences* **2**, 438 (2009).

[203] B. Strauch, W. de Moura, Arterial system of the fingers, *The Journal of hand surgery* **15**, 148 (1990).

[204] P. Vaupel, D. K. Kelleher, M. Höckel, *Seminars in oncology* (Elsevier, 2001), vol. 28, pp. 29–35.

[205] S. Ogawa, T.-M. Lee, A. R. Kay, D. W. Tank, Brain magnetic resonance imaging with contrast dependent on blood oxygenation, *Proceedings of the National Academy of Sciences* **87**, 9868 (1990).

[206] Y. Murata, K. Sakatani, T. Hoshino, N. Fujiwara, T. Kano, S. Nakamura, Y. Katayama, Effects of cerebral ischemia on evoked cerebral blood oxygenation responses and bold contrast functional mri in stroke patients, *Stroke* **37**, 2514 (2006).

[207] L. B. Rowell, Cardiovascular adjustments to thermal stress, *Comprehensive Physiology* (1983).

[208] K. Daoudi, P. van Es, S. Manohar, W. Steenbergen, Two-dimensional spatiotemporal monitoring of temperature in photothermal therapy using hybrid photoacoustic ultrasound transmission tomography, *Journal of biomedical optics* **18**, 116009 (2013).

[209] D. Golovko, R. Meier, E. Rummeny, H. Daldrup-Link, Optical imaging of rheumatoid arthritis, *International journal of clinical rheumatology* **6**, 67 (2011).

[210] I. B. McInnes, G. Schett, The pathogenesis of rheumatoid arthritis, *New England Journal of Medicine* **365**, 2205 (2011).

[211] F. McQueen, Magnetic resonance imaging in early inflammatory arthritis: what is its role?, *Rheumatology* **39**, 700 (2000).

[212] A. Baillet, C. Gaujoux-Viala, G. Mouterde, T. Pham, J. Tebib, A. Saraux, B. Fautrel, A. Cantagrel, X. Le Loët, P. Gaudin, Comparison of the efficacy of sonography, magnetic resonance imaging and conventional radiography for the detection of bone erosions in rheumatoid arthritis patients: a systematic review and meta-analysis, *Rheumatology* **50**, 1137 (2011).

[213] A. N. Colebatch, *et al.*, Eular recommendations for the use of imaging of the joints in the clinical management of rheumatoid arthritis, *Annals of the rheumatic diseases* **72**, 804 (2013).

[214] D. Chamberland, Y. Jiang, X. Wang, Optical imaging: new tools for arthritis, *Integrative Biology* **2**, 496 (2010).

[215] J. R. Rajian, G. Girish, X. Wang, Photoacoustic tomography to identify inflammatory arthritis, *Journal of biomedical optics* **17**, 0960131 (2012).

[216] P. Mohajerani, M. Koch, K. Thürmel, B. Haller, E. J. Rummeny, V. Ntziachristos, R. Meier, Fluorescence-aided tomographic imaging of synovitis in the human finger, *Radiology* **272**, 865 (2014).

[217] Y. Sun, E. S. Sobel, H. Jiang, First assessment of three-dimensional quantitative photoacoustic tomography for in vivo detection of osteoarthritis in the finger joints, *Medical physics* **38**, 4009 (2011).

[218] G. Xu, J. R. Rajian, G. Girish, M. J. Kaplan, J. B. Fowlkes, P. L. Carson, X. Wang, Photoacoustic and ultrasound dual-modality imaging of human peripheral joints, *Journal of biomedical optics* **18**, 010502 (2013).

[219] N. Beziere, *et al.*, Optoacoustic imaging and staging of inflammation in a murine model of arthritis, *Arthritis & Rheumatology* **66**, 2071 (2014).

[220] G. R. Cherrick, S. W. Stein, C. M. Leevy, C. S. Davidson, Indocyanine green: observations on its physical properties, plasma decay, and hepatic extraction, *Journal of Clinical Investigation* **39**, 592 (1960).

[221] B. Jung, V. I. Vullev, B. Anvari, Revisiting indocyanine green: Effects of serum and physiological temperature on absorption and fluorescence characteristics, *IEEE Journal of Selected Topics in Quantum Electronics* **20**, 149 (2014).

[222] W. Holzer, M. Mauerer, A. Penzkofer, R.-M. Szeimies, C. Abels, M. Landthaler, W. Bäumler, Photostability and thermal stability of indocyanine green, *Journal of Photochemistry and Photobiology B Biology* **47**, 155 (1998).

[223] N. Beziere, N. Lozano, A. Nunes, J. Salichs, D. Queiros, K. Kostarelos, V. Ntziachristos, Dynamic imaging of pegylated indocyanine green (icg) liposomes within the tumor microenvironment using multi-spectral optoacoustic tomography (msot), *Biomaterials* **37**, 415 (2014).

[224] T. Iijima, T. Aoyagi, Y. Iwao, J. Masuda, M. Fuse, N. Kobayashi, H. Sankawa, Cardiac output and circulating blood volume analysis by pulse dye-densitometry, *Journal of clinical monitoring* **13**, 81 (1997).

[225] D. A. Reuter, C. Huang, T. Edrich, S. K. Shernan, H. K. Eltzschig, Cardiac output monitoring using indicator-dilution techniques: basics, limits, and perspectives, *Anesthesia & Analgesia* **110**, 799 (2010).

[226] G. Ku, L. V. Wang, Deeply penetrating photoacoustic tomography in biological tissues enhanced with an optical contrast agent, *Optics letters* **30**, 507 (2005).

[227] M. A. Horsfield, J. S. Thornton, A. Gill, H. R. Jager, A. N. Priest, B. Morgan, A functional form for injected mri gd-chelate contrast agent concentration incorporating recirculation, extravasation and excretion, *Physics in medicine and biology* **54**, 2933 (2009).

[228] G. C. Sutton, J. Karnell, G. Nylin, Studies on the rapidity of complete blood circulation in man, *American heart journal* **39**, 741 (1950).

List of Peer-Reviewed Publications

Lutzweiler, C.; Razansky, D., Optoacoustic imaging and tomography: reconstruction approaches and outstanding challenges in image performance and quantification. *Sensors* **2013**, *13*, 7345-7384.

Lutzweiler, C.; Deán-Ben, X.L.; Razansky, D., Expediting model-based optoacoustic reconstructions with tomographic symmetries. *Medical physics* **2014**, *41*, 013302.

Lutzweiler, C.; Meier, R.; Rummeny, E.; Ntziachristos, V.; Razansky, D., Real-time optoacoustic tomography of indocyanine green perfusion and oxygenation parameters in human finger vasculature. *Optics letters* **2014**, *39*, 4061-4064.

Tzoumas, S.; Rosenthal, A.; **Lutzweiler, C.**; Razansky, D.; Ntziachristos, V., Spatiospectral denoising framework for multispectral optoacoustic imaging based on sparse signal representation. *Medical physics* **2014**, *41*, 113301.

Ding, L.; Deán-Ben, X.L.; **Lutzweiler, C.**; Razansky, D.; Ntziachristos, V., Efficient non-negative constrained model-based inversion in optoacoustic tomography. *Physics in medicine and biology* **2015**, *60.17*, 6733.

Lutzweiler, C.; Meier, R.; Razansky, D., Optoacoustic image segmentation based on signal domain analysis. *Photoacoustics* **2015**, *3*, 151-158.

Lutzweiler, C.; Tzoumas, S.; Rosenthal, A.; Ntziachristos, V.; Razansky, D., High-throughput sparsity-based inversion scheme for optoacoustic tomography. *IEEE Transactions on Medical Imaging* **2016**, *35*, 674-684.

List of Abbreviations

API Application programming interface.

art. uln. Ulnar artery (of the finger).

art. rad. Radial artery (of the finger).

BOLD Blood oxygenation level dependent.

BP Backprojection.

CA Contrast agent.

CG Conjugate gradient.

CMUT Capacitive micro-machined ultrasonic transducer.

CS Compressed sensing.

DAQ Data acquisition system.

DE Diffusion equation.

DIP Distal interphalangeal joint.

DOT Diffuse optical tomography.

DTI Diffusion tensor imaging.

ECG Electrocardiography.

EIR Electrical impulse response.

EV Eigenvector.

FDTD Finite differences time domain.

FFT Fast Fourier transformation.

FIFO First in, first out.

FLOP Floating point operation.

FMT Fluorescence molecular tomography.

FP Fluorescent protein.

FWHM Full width at half maximum.

GFP Green fluorescent protein.

GPGPU General purpose graphics processing unit.

GPU Graphics processing unit.

Hb Deoxygenated hemoglobin.

HbO Oxygenated hemoglobin.

ICG Indocyanine green.

MB Model-based.

MC Monte Carlo.

MCP Metacarpal joint.

MIP Maximum intensity projection.

MRI Magnetic resonance imaging.

MSOT Multi-spectral optoacoustic tomography.

NA Numerical aperture.

NIR Near-infrared.

NZ Non-zero element.

OA Optoacoustic (also known as photoacoustic, PA).

OAT Optoacoustic tomography (also known as photoacoustic (computed) tomography, PAT, PACT).

OCT Optical coherence tomography.

OPO Optical parametric oscillator.

PCA Principal component analysis.

PET Positron emission tomography.

PIP Peripheral interphalangeal joint.

PVD Peripheral vascular disease.

QPAT Quantitative photoacoustic tomography.

RA Rheumatoid arthritis.

RF Radio-frequency.

RMSD Root mean square deviation.

ROI Region of interest.

RP Raynaud's phenomenon.

SIMD Single instruction, multiple data.

SIR Spatial impulse response.

SISD Single instruction, single data.

SNR Signal-to-noise ratio.

SOS Speed-of-sound.

SPECT Single photon emission computed tomography.

SPIM Selective plane illumination microscopy.

STED Stimulated emission depletion.

SVD Singular value decomposition.

TAO Thromboangiitis obliterans.

TOF Time-of-flight.

TR Time-reversal.

US Ultrasound.

WL Wavelet.

WP Wavelet packets.

X-CT X-ray computed tomography.

List of Symbols

$\otimes$ Kronecker product.

$\mathbf{1}$ Identity matrix.

AIF Arterial input function.

β Thermal volume expansion coefficient.

$\mathbf{c}$ Speed-of-sound.

$\mathbf{c_i}$ Chromophore i concentration.

$\mathbf{C_p}$ Isobaric specific heat capacity.

$\mathbf{D}$ Diagonal matrix.

$\delta(\cdot)$ Delta distribution.

$\mathbf{D_p}$ Photon diffusion coefficient.

$\mathbf{D_t}$ Thermal diffusivity.

$\mathbf{\Delta X}$ Pixel size.

$\mathbf{f}$ Image vector.

$\mathbf{F}$ Vector valued functional.

$\mathbf{f_d}$ Focal distance.

$\mathbf{f_{US}}$ Ultrasonic frequency.

$\mathbf{F_{x,y,z,t}}$ Fourier transformation in x-, y-, z-, and/or t-dimension.

$\mathbf{G}$ Green's function.

$\mathbf{g}$ Gini-index.

$\mathbf{\Gamma}$ Grüneisen parameter.

$\mathbf{H}$ Deposited energy density.

$\mathbf{k}$ Exchange rate constant.

κ Isothermal compressibility.

$\kappa_{\mathbf{cond}}$ Condition number.

$\mathbf{k_{x,y,z}}$ Spatial frequency in x-, y-, or z-dimension.

$\mathbf{L}$ Angular frequency.

λ Wavelength.

$\mathbf{L_{dec}}$ Wavelet (packets) decomposition level.

$\mathbf{M}$ Optoacoustic propagation model (matrix).

$\mathbf{m}$ Model parameter.

$\mu_\mathbf{a}$ Optical absorption coefficient.

$\mu_\mathbf{eff}$ Effective absorption coefficient.

$\mu_\mathbf{s}$ Optical scattering coefficient.

$\mu'_\mathbf{s}$ Reduced scattering coefficient.

$\mathbf{n_{avg}}$ Number of frame averages.

$\mathbf{n_{int}}$ Number of integral approximation points.

$\mathbf{n_{iter}}$ Number of iterations.

$\mathbf{n}_\varphi$ Number of projections.

$\mathbf{n_{x,y,z,r,t,\lambda}}$ Number of samples / pixels per dimension.

ω Temporal frequency.

$\mathbf{p}$ Pressure.

$\mathbf{p_0}$ Initial optoacoustic pressure.

$\mathbf{\Phi}$ Optical fluence.

φ Angle.

$\mathbf{p_{hilb}}$ Hilbert-transformed signals.

$\mathbf{\Psi}$ Function base.

$\mathbf{P}$ Tikhonov regularization operator.

$\mathbf{q_0}$ Optical source term.

$\mathbf{r}$ Radius.

$\mathbf{R}$ Rotation matrix.

$\mathbf{RR}$ Relative residual.

$\mathbf{S}$ Detection surface.

$\mathbf{S_{trans}}$ Transducer surface.

$\mathbf{t}$ Time.

$\mathbf{T}$ Temperature.

τ Time constant.

$\tau_{\mathbf{ex}}$ Excitation pulse duration.

$\tau_{\mathbf{sig,mat}}$ Signal (matrix) thresholding level.

$\mathbf{TOF}$ Time-of-flight functional.

$\mathbf{U}$ (Unitary) base transformation matrix.

$\mathbf{V}$ Volume.

$\mathbf{V_{lim}}$ Projection reduction matrix.

$\mathbf{V^{(red)}}$ Signal coefficient reduction matrix.

$\mathbf{x_d}$ Detector location.

$\mathbf{x}, \mathbf{y}, \mathbf{z}$ Spatial coordinate(s).

List of Figures

List of Tables