

Optimierung von Bildgebungsverfahren für die 3D-Ultraschall-Computertomographie

Zur Erlangung des akademischen Grades eines
DOKTOR-INGENIEURS
von der Fakultät für
Elektrotechnik und Informationstechnik
des Karlsruher Instituts für Technologie (KIT)
genehmigte Dissertation
von

Dipl.-Ing. Ernst Kretzek

geb. in Hermannstadt

Hauptreferent: Prof. Dr. rer. nat. Marc Weber
Korreferent: Prof. Dr.-Ing. Dr. h.c. Jürgen Becker

Tag der mündlichen Prüfung: 04.08.2015

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die zum Erfolg dieser Arbeit beigetragen haben.

Dem Institutsleiter Prof. Marc Weber und meiner Betreuerin Dr. Nicole Ruiter möchte ich für die Gelegenheit danken, mit meiner Dissertation einen Teil zum bisher weltweit einmaligen 3D-Ultraschall-Computertomographen beitragen zu können. Ebenso bedanke ich mich bei Prof. Jürgen Becker für die Übernahme des Korreferats dieser Arbeit.

Mein besonderer Dank gilt besonders Dr. Nicole Ruiter für ihre langjährige Unterstützung, Begleitung und Betreuung während meiner gesamten Promotionszeit. Mein weiterer Dank richtet sich an alle Kollegen und Studenten am Institut für Prozessdatenverarbeitung und Elektronik, die ich während meiner Promotion kennen und schätzen lernen durfte. Es war eine schöne, wertvolle Zeit mit euch.

Insbesondere möchte ich mich bei Dr. Nicole Ruiter, Dr. Matthias Birk und Peter Figuli für die gute Zusammenarbeit im DFG-Projekt bedanken. Ebenso gilt mein Dank Prof. Hartmut Gemmeke, Michael Zapf, Dr. Thorsten Hopp, Dr. Robin Dapp, Dr. Benedikt Kohout und Wei Yap Tan für den guten Austausch und die vielen Diskussionen. Bei den Studenten Aniko Szajko, Milan Tenk, Patrick Hucker, Srinivas Rao Kerekare, Christian Jülg und Christoph Elsner möchte ich mich für ihren Einsatz und ihre Unterstützung rund um das Thema dieser Dissertation bedanken.

Ein besonderer Dank von ganzem Herzen gilt meiner lieben Ehefrau, die ich während meiner Promotionszeit heiraten durfte. Durch ihre Liebe, unermüdliche Unterstützung und Motivation, gepaart mit dem nötigen Freiraum hat sie erheblich zum erfolgreichen Abschluss dieser Arbeit beigetragen.

Nicht zuletzt danke ich meinen Eltern für den Weg, den sie mir eröffnet haben sowie für ihre fortwährende Unterstützung und ihr Vertrauen in mich. Mein Dank gilt allen, die in irgendeiner Art und Weise zum Gelingen dieser Arbeit beigetragen haben.

Erklärung

Ich versichere wahrheitsgemäß, die Dissertation bis auf die dort angegebene Hilfe selbständig angefertigt, alle benutzten Hilfsmittel vollständig und genau angegeben und alles kenntlich gemacht zu haben, was aus Arbeiten anderer und eigenen Veröffentlichungen unverändert oder mit Änderungen entnommen wurde.

Karlsruhe, 01.07.2015

.....
(Ernst Kretzek)

Kurzfassung

Am Karlsruher Institut für Technologie wird mit der 3D-Ultraschall-Computer-Tomographie (3D-USCT) ein neuartiges bildgebendes Verfahren erforscht. Es hat eine verbesserte Diagnose von Brustkrebs zum Ziel und verspricht eine hohe Bildqualität. Sie ist die weltweit erste vollständig dreidimensionale Untersuchungsmethode basierend auf Ultraschall. Die Bildrekonstruktion in 3D mittels der genutzten „Synthetic Aperture Focusing Technique“ (SAFT) ist sehr rechenintensiv. Daher wurden insbesondere Grafikprozessoren, „Graphics Processing Units“ (GPUs), als parallele Zielarchitekturen eingesetzt, um eine klinische Anwendbarkeit für die Rekonstruktion mit SAFT zu erreichen. Bei dieser Standard-SAFT wurden jedoch vereinfachende Annahmen getroffen, die eine Rekonstruktion von hochauflösenden Bildern bei komplexen und inhomogenen Objekten wie der Brust nicht zulassen.

In dieser Arbeit werden Erweiterungen der SAFT mit unterschiedlicher Komplexität erforscht, die auf einer realistischeren Annahme von heterogenem Brustgewebe basieren. Dadurch erreicht die 3D-USCT erstmals die nötige Bildqualität für hochauflösende Bilder. Es werden Lösungen vorgestellt, die trotz der erhöhten Rechenkomplexität die klinische Anwendbarkeit, im Sinne einer Bildberechnung innerhalb von 15 Minuten, mit einem Hochleistungs-GPU-Server ermöglichen.

Um die einfache Weiterentwicklung durch Wissenschaftler zu gewährleisten, wird als aktuelle Entwicklungsumgebung MATLAB genutzt, und der mit paralleler Hardware beschleunigte Standard-Algorithmus effizient eingebunden. Mittels einer entwickelten Simulationssoftware wurde der SAFT-Algorithmus analysiert und die wesentlichen Einflüsse der bisher vereinfachenden Annahmen auf die Bildqualität untersucht. Es werden Methoden entwickelt, die die limitierenden Annahmen überwinden, ohne dabei die Rechenzeit wesentlich zu erhöhen und so den Kompromiss zwischen Bildqualität und Rechenzeit optimieren. Somit ist die Annahme einer konstanten Schallgeschwindigkeit und Dämpfung mit Hilfe einer Schallgeschwindigkeits- und Dämpfungs-Korrektur nicht mehr notwendig. Zudem wird mit einer geeigneten Approximation für die SAFT-Rekonstruktion eine signifikante Beschleunigung erreicht, obwohl sie für heterogene Medien geeignet ist. Eine Erweiterung der SAFT ermöglicht es, die bei der Rekonstruktion auftretenden Artefakte effizient zu unterdrücken. Mit einer neuartigen Methode in der 3D-USCT kann darüber hinaus für jeden Bildpunkt eine eigene richtungsabhängige Reflexionscharakteristik berechnet werden. Diese Richtungsinformationen stehen für eine Vielzahl von weiteren Methoden zur Verfügung und können sowohl zur Verbesserung des Kontrasts, als auch zur Klassifizierung von Gewebe genutzt werden. Neben Simulationen und statischen Phantomen wurden die Methoden anhand klinischer Daten evaluiert, die in einer Pilotstudie am Universitätsklinikum Jena aufgenommen wurden.

Die in dieser Arbeit entwickelten Algorithmen bilden die Grundlage für hochaufgelöste Rekonstruktionen, die erstmals in einer klinisch relevanten Zeit durchgeführt werden konnten. Der neu entwickelte Rekonstruktionsalgorithmus bildet über die 3D-USCT hinaus eine Grundlage zur deutlichen Beschleunigung aller Algorithmen, die auf SAFT basieren. Die Erzeugung der Richtungsinformation erweitert die 3D-USCT um eine vollständig neue vierte Modalität und eröffnet damit weitere Forschungsgebiete.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation und Zielsetzung	1
1.2	Beitrag der Arbeit	2
2	Grundlagen	5
2.1	Ultraschall	5
2.2	Häufigkeit von Brustkrebs	5
2.3	Methoden der Brustkrebsdiagnose	6
2.4	Analyse der Bildqualität	8
2.5	Grafikprozessoren	14
2.6	Stand von Technik und Forschung	19
3	Standard-SAFT in der Reflexionsbildung des 3D-USCTs	25
3.1	3D-USCT mit dem Prototypen des IPE	25
3.2	Reflexionen bei Ultraschall	32
3.3	SAFT in der 3D-USCT	36
4	GPU-basierte Beschleunigung der Standard-SAFT	41
4.1	MATLAB als Entwicklungsumgebung	41
4.2	Integration in MATLAB	42
4.2.1	Design einer effizienten Schnittstelle	42
4.2.2	Implementierung der MEX-Funktion	43
4.2.3	Implementierung des SAFT CUDA Kernels für die GPU	44
4.2.4	Performanzanalyse	46
5	Analyse der Bildqualität	49
5.1	Die Bildqualität der 3D-USCT	49
5.2	Metriken zur Analyse der Bildqualität	50
5.3	Aufbau der Simulationsumgebung	55
6	Reflexionsbildung mit SAFT für heterogene Medien	61
6.1	Korrektur der Annahme einer konstanten Schallgeschwindigkeit	61
6.1.1	Methodik und Umsetzung	63
6.1.2	Evaluation	64
6.2	Korrektur der Annahme einer konstanten Dämpfung	72
6.2.1	Methodik und Umsetzung	73
6.2.2	Evaluation	74
7	Beschleunigung der SAFT für heterogene Medien	81
7.1	Beschleunigung der korrigierten SAFT-Algorithmen	81
7.1.1	Berechnungsstrategien	82
7.1.2	Evaluation	87
7.1.3	Implementierungen der korrigierten SAFT-Algorithmen	92
7.1.4	Performanz	94
7.2	Weitere Beschleunigung der SAFT für heterogene Medien	96

7.2.1	Methodik und Umsetzung	97
7.2.2	Fehleranalyse der Approximation	99
7.2.3	Evaluation	103
7.2.4	Performanz	110
8	Erweiterung der SAFT zur Reduktion der Gratinglobe-Artefakte	113
8.1	Methodik und Umsetzung	113
8.2	Evaluation	116
9	Erweiterung der SAFT zur Gewinnung von Reflexionscharakteristiken	123
9.1	Das Reflexionsmodell nach Phong	123
9.2	Berechnung der Reflexionscharakteristik	126
9.3	Evaluation der Reflexionscharakteristik	133
9.3.1	Extraktion neuer Informationen	134
9.3.2	Klassifikation	144
9.3.3	Performanz	150
10	Zusammenfassung, Diskussion und Ausblick	153
	Literaturverzeichnis	159

Abkürzungsverzeichnis

Die folgenden Abkürzungen beziehen sich teilweise auf deutsche und teilweise auf englische Akronyme.

AP	<i>Aperturpositionen</i>
A-Scan	<i>Amplituden-Scan</i>
ATT	<i>Dämpfung, engl. Attenuation</i>
DAQ	<i>Datenerfassungssystem</i>
FWHM	<i>Full Width at Half Maximum</i>
GPU	<i>graphics processing unit</i>
MRT	<i>Magnetresonanz-Tomographie</i>
MSE	<i>mean squared error</i>
NDT	<i>nondestructive testing</i>
PSF	<i>Punktspreizfunktion</i>
PSNR	<i>peak-signal-to-noise ratio</i>
PSPNR	<i>peak-signal-to-peak-noise ratio</i>
Radar	<i>radio detection and ranging</i>
RHOA	<i>räumliche homogene Objektausleuchtung</i>
ROI	<i>region of interest</i>
SAFT	<i>synthetic aperture focusing technique</i>
SAS	<i>synthetic aperture sonar</i>
SAR	<i>synthetic aperture radar</i>
SOS	<i>Schallgeschwindigkeit, engl. Speed of sound</i>
SEP	<i>Sender-Empfänger-Paare</i>
SNR	<i>signal-to-noise ratio</i>
Sonar	<i>sound navigation and ranging</i>
TAS	<i>transducer array system</i>
TGC	<i>time gain compensation</i>
TOF	<i>time of flight</i>
US	<i>Ultraschall</i>
Voxel	<i>Volumen Pixel</i>

1. Einleitung

Brustkrebs ist die weltweit häufigste Krebserkrankung bei Frauen [1]. Eine entscheidende Rolle für die Heilungschancen spielt dabei die Größe des Tumors bei der Erstdiagnose. Von ihr hängt die Wahrscheinlichkeit ab, ob sich bereits Metastasen gebildet haben. Im Mittel beträgt die Größe 1 cm [2] und einer Metastasenwahrscheinlichkeit von 30%. Mit einer früheren Diagnose bei einer Tumorgöße 0,5 cm könnte sie auf 5% reduziert werden. Eine vielversprechende bildgebende Methode zur Brustkrebsfrüherkennung ist die 3D-Ultraschall-Computertomographie (3D-USCT) [3]. Neben der Möglichkeit reproduzierbare 3D-Abbildungen der weiblichen Brust zu erstellen, wird im Vergleich zur Standard-Untersuchungsmethode, der Röntgenmammographie, keine ionisierende Strahlung genutzt. Der Hauptvorteil der Methode besteht darin, dass sich mit nur einer Messung drei verschiedene physikalische Parameter der Brust rekonstruieren lassen: Schallgeschwindigkeit, Dämpfung und Reflektivität. Die Kombination dieser drei rekonstruierten Volumina ermöglicht eine Klassifizierung der Gewebearten und Gewebestrukturen [4, 5, 6]. Am Institut für Prozessdatenverarbeitung und Elektronik (IPE) wurde ein Prototyp eines 3D-USCT entwickelt und bereits in einer ersten klinischen Studie erfolgreich getestet [7] (siehe Abbildung 1.1).

1.1 Motivation und Zielsetzung

Die 3D-Volumenrekonstruktionen der physikalischen Parameter sind sehr rechenintensiv und können auf einem aktuellen Prozessor bei einer sequentiellen Berechnung mehrere Tage in Anspruch nehmen. Die höchste Rechenintensität besitzt dabei die Rekonstruktion der Reflexionsbilder, bei der die *Synthetic Aperture Focusing Technique* (SAFT) [8] zum Einsatz kommt. In [9] wurden daher aktuelle parallele Hardwarearchitekturen zur Beschleunigung in der 3D-USCT untersucht. Für die SAFT

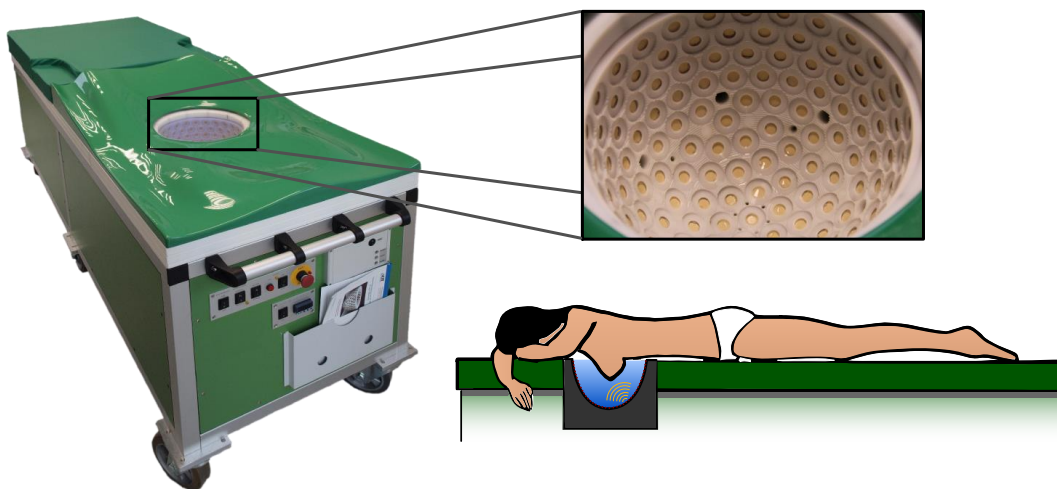


Abbildung 1.1: Photos des 3D-USCT-Prototypen am IPE (links) mit der Apertur (oben rechts) und einer schematischen Zeichnung der Messanordnung (unten rechts).

zeigte sich die höchste zu erreichende Performanz mit Grafikprozessoren, *Graphics Processing Units* (GPUs). Für einen Einsatz der 3D-USCT in einer Klinik oder Arztpraxis ist weiterhin eine Begrenzung der Hardware bezüglich der Anschaffungs- und Betriebskosten sowie des Raumbedarfs gegeben, die eine Ausführung auf einem Rechencluster nicht ermöglichen. Die zur Verfügung stehende Hardware beschränkt sich daher auf einen Rechner mit mehreren GPUs. Weiterhin ist mit einem klinischen Einsatz eine Limitierung der Rechendauer auf ca. 15 Minuten gegeben. Diese Dauer ergibt sich in etwa aus der Zeit, die eine Patientin benötigt, sich nach der Messung beim Arzt für eine Besprechung einzufinden.

Eine weitere Herausforderung ergibt sich dadurch, dass der Rekonstruktion mit SAFT vereinfachende Annahmen zugrunde liegen, die die Bildqualität begrenzen. So geht der Standard-SAFT von einem homogenen Medium mit konstanter Schallgeschwindigkeit und einer vernachlässigbaren Dämpfung aus. Eine Brust ist jedoch ein komplexes und inhomogenes Objekt. Deshalb wird für hochauflösende Bilder eine Erweiterung des SAFT-Algorithmus für inhomogene Medien benötigt.

1.2 Beitrag der Arbeit

Der Fokus dieser Arbeit liegt auf der Verbesserung der Rekonstruktion von Reflexionsbilder mittels SAFT, wobei sowohl die Bildqualität als auch die Performanz für einen klinischen Einsatz erhöht werden soll. Die Bildqualität lässt sich gewöhnlich mit komplexeren und genaueren Algorithmen verbessern, wobei jedoch ebenfalls die Rechenintensität steigt. Umgekehrt kann die Rechengeschwindigkeit oft durch ungenauere Approximationen erhöht werden, bei der wiederum die Bildqualität sinkt. Damit ergeben sich zwei gegensätzliche Ziele für die ein Kompromiss gefunden werden muss.

Für den praktischen Einsatz der GPU-beschleunigten SAFT-Rekonstruktion wird erstmals eine effiziente Schnittstelle zwischen der GPU und der aktuell eingesetzten Entwicklungsumgebung in MATLAB konzipiert und implementiert.

Für die quantitative Analyse und Bewertung der Bildqualität wird eine eigens für diese Arbeit entwickelte Simulationsumgebung genutzt. In Kombination mit vielfältigen Bildmetriken kann so die Leistungsfähigkeit der SAFT-Rekonstruktionen bewertet und die Einhaltung der wesentlichen Randbedingungen der Bildqualität überprüft werden.

Um die Limitierungen der Bildqualität durch die Annahmen der Standard-SAFT zu überwinden, wird eine Schallgeschwindigkeits- und Dämpfungs-Korrektur entwickelt. Dabei wird erstmals anhand der Bildqualitätsmetriken und Berechnungsstrategien analysiert, wie die Information der Schallgeschwindigkeits- und Dämpfungsvolumen über das heterogene Medium genutzt werden kann, um einen optimierten Kompromiss zwischen Bildqualität und Rechengeschwindigkeit zu erreichen.

Es wird zudem eine neuartige SAFT-Rekonstruktionsvariante für heterogene Medien vorgestellt, die durch eine geeignete Approximation die Komplexität der Rekonstruktion reduziert und somit einen signifikanten Performanzgewinn erreicht. Dadurch wird es erstmals möglich, die Rekonstruktion für heterogene Medien in der klinisch relevanten Zeit von 15 Minuten durchzuführen. Darüber hinaus wird die Rekonstruktion mit SAFT mit einer schnellen Reduktion der SAFT-typischen Rekonstruktionsartefakte erweitert, um die Bildqualität weiter zu verbessern.

Abschließend wird eine neuartige Idee für die SAFT getestet und evaluiert: Bei der SAFT-Rekonstruktion wird standardmäßig nur ein Reflektivitätswert pro Voxel zurückgegeben. Es lässt sich jedoch pro Voxel eine richtungsabhängige Reflexionscharakteristik berechnen, die Aufschluss darüber gibt, wie stark eine Struktur innerhalb des Voxels in einzelne Richtungen streut. Diese gewonnenen Informationen über das Medium eröffnen neue Forschungsgebiete, da es gilt diese Daten auszuwerten. In dieser Arbeit wird bereits gezeigt, dass sich die Richtungen der Oberflächennormalen einer streuenden Struktur berechnen lassen. Weiterhin zeigen die Daten das Potential, eine Klassifizierung einzelner Voxel anhand ihrer Reflexionscharakteristiken durchzuführen. Aus den Reflexionscharakteristiken können zudem reflektierende Strukturen herausgefiltert und mit einem höheren Kontrast dargestellt werden.

Die in dieser Arbeit entwickelten GPU-beschleunigten Rekonstruktionsvarianten wurden bereits im Verlauf einer ersten klinischen Studie erfolgreich getestet und ermöglichen weitergehende Forschungen zur Optimierung des 3D-USCT Prototyps.

Wesentliche Teile der Arbeit wurden auf wissenschaftlichen Konferenzen veröffentlicht [10, 11, 12, 13, 14].

2. Grundlagen

2.1 Ultraschall

Mit Ultraschall wird Schall mit einer Frequenz oberhalb der Hörgrenze von 20 kHz bis etwa 1 GHz bezeichnet [15]. Kommt der Ultraschall für ein bildgebendes Verfahren in der Medizin zum Einsatz, wird dies als Sonographie bezeichnet. Dabei wird ausgenutzt, dass eine in das Gewebe eingestrahlte Ultraschallwelle an Gewebegrenzen gestreut bzw. reflektiert wird. Der dabei eingesetzte Ultraschall ist für das Gewebe unschädlich.

2.2 Häufigkeit von Brustkrebs

Der Brustkrebs ist laut International Agency for Research on Cancer mit 1,67 Millionen neuen Krebsdiagnosen im Jahr 2012 die weltweit häufigste Krebserkrankung bei Frauen [1] (siehe Abbildung 2.1). Dabei liegt die Sterberate mit 60% der Erkrankungsrate deutlich geringer als bei anderen Krebsarten. Im frühen Stadium verläuft die Krebserkrankung in der Regel beschwerdefrei. Daher sind regelmäßige Kontroll- bzw. Vorsorgeuntersuchungen unabdingbar. Diese werden nach der Leitlinie zur Brustkrebs-Früherkennung [16] in Absprache mit dem Facharzt abhängig vom Alter der Patientin mit unterschiedlichen Methoden durchgeführt. Am häufigsten kommt das Abtasten der Brust vom Facharzt (Palpation) sowie die Röntgen-Mammographie als Vorsorgeuntersuchung zum Einsatz. Für die Heilungschance ist eine frühzeitige Diagnose bei einer noch geringen Größe der krebsartigen Strukturen ausschlaggebend. Hat der Tumor jedoch bereits Metastasen in weitere Organe oder Lymphknoten gestreut, steigt die prozentuale Sterberate erheblich an. Bei einer rechtzeitigen Diagnose kann der Tumor brusterhaltend operativ entfernt werden. Ist der Tumor bereits in einem fortgeschrittenen Stadium, kann jedoch auch die

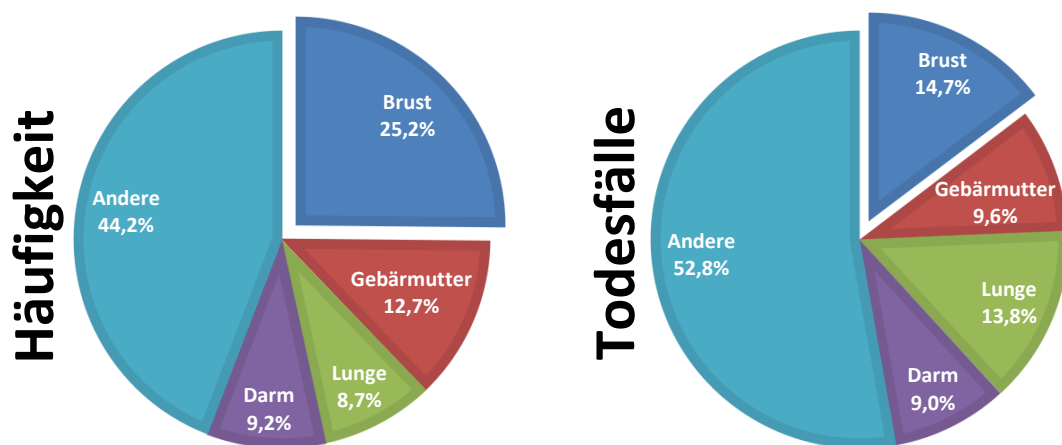


Abbildung 2.1: Prozentuale Häufigkeit der Krebsarten bei Frauen weltweit (links) und die entsprechenden krebsbedingten Todesfälle (rechts). Die Daten wurden dem GLOBOCAN Projekt 2012 entnommen [1].

Brust komplett abgenommen werden, da sie kein lebenswichtiges Organ ist. Diese Vorgehensweise wird meist in Kombination mit einer Chemo- oder Strahlentherapie durchgeführt, um das Rückfallrisiko zu vermindern [17]. Fortschritte bei den Therapie- und Diagnosemöglichkeiten sind bereits zu erkennen, worauf die sinkende Sterberate der erkrankten Frauen in den letzten Jahren deutlich hinweist [18]. Jedoch steigt die Zahl der Brustkrebsneuerkrankungen weiterhin stetig an. Nach [19] ist bei Frauen innerhalb der Europäischen Union zu erwarten, dass die Sterberate durch Lungenkrebs im Jahr 2015 erstmals höher liegt als bei Brustkrebs.

2.3 Methoden der Brustkrebsdiagnose

Palpation

Palpation (lateinisch *palpare* für „streicheln“) ist das manuelle Abtasten eines Körperbereichs. Im Bereich der Brustkrebsuntersuchung bedeutet dies das Abtasten beider Brüste durch einen Facharzt oder die Patientin selbst. Die Palpation ist die älteste und weitverbreitetste Brustkrebsuntersuchung, bei der durch Abtasten versucht wird, Auffälligkeiten wie z.B. Knoten oder Verhärtungen zu entdecken. Jedoch besteht ein entscheidende Nachteil darin, dass Tumore erst ab einer Größe von ca. zwei Zentimetern entdeckt werden können. Bei dieser Größe beträgt die Wahrscheinlichkeit, dass sich bereits Metastasen gebildet haben 60% [20]. Ein weiterer Nachteil besteht in der Nichtreproduzierbarkeit dieser Untersuchung. Dennoch ist die Palpation das Verfahren, durch welches der Verdacht auf Brustkrebs häufig erstmals aufkommt und zu weiteren Untersuchungen führt.

Röntgen-Mammographie

Die Röntgen-Mammographie ist die etablierte Standardvorsorgeuntersuchung für Brustkrebs und gilt zurzeit allgemein als einzig wirksame anerkannte Methode für die Erkennung von Brustkrebsvorstufen bzw. frühen Tumorstadien [16]. Bei dieser Methode wird die Brust mithilfe von Röntgenstrahlung untersucht. Um einen ausreichenden Bildkontrast innerhalb einer vertretbarer Strahlungsstärke zu erreichen, wird die Brust zwischen zwei Kompressionsplatten komprimiert. Hieraus ergeben sich bereits entscheidende Nachteile dieser Methode. Zum einen kann die Kompression bei der Patientin Schmerzen verursachen. Zum anderen führt die Kompression zu einer schlechten Reproduzierbarkeit der Untersuchung und einer schwierigen Lokalisierung des Tumors im Anschluss. Des Weiteren wird die Untersuchung erst bei Frauen zwischen 50 bis 69 Jahren regelmäßig durchgeführt, da bei Frauen mit dichtem Drüsengewebe, insbesondere bei jüngeren Frauen, kein ausreichender Bildkontrast erreicht werden kann [16]. Auch die relativ hohe Strahlenbelastung von 2-4 Milligray (mGy), die einer durchschnittlichen Jahresbelastung eines Mitteleuropäers entspricht [21], kommt hinzu. Bei diesem Verfahren liegt die Wahrscheinlichkeit der Metastasenbildung bei 30%. Die durchschnittliche Tumorgöße der ersten Diagnose beträgt 10 mm [22].

Konventioneller Ultraschall

Der konventionelle Ultraschall, auch Sonografie genannt, ist bei der Brustkrebsuntersuchung als nicht-invasive Untersuchungsmethode ebenfalls von großer Bedeutung. Sie stellt eine Zusatzuntersuchung dar, wenn die Palpation oder die Mammographie

keine eindeutigen Befunde aufzeigen [16]. Bei symptomatischen Befunden gilt sie bei Frauen, die jünger als 40 Jahre sind, als bildgebende Methode der ersten Wahl. Dennoch gilt sie nicht als geeignete alleinige Methode zur Vorsorgeuntersuchung [16]. Hierbei kommt ein konventioneller Ultraschall-Handscanner zum Einsatz, der vom Facharzt manuell über die Brust bewegt wird. Die Auflösung bei diesem Verfahren liegt je nach Frequenz ($\geq 7,5 \text{ MHz}$) im Submillimeterbereich, dennoch ist die Reproduzierbarkeit durch das manuelle Abfahren der Brust mit der Hand des Arztes sehr eingeschränkt. Ein weiterer Nachteil besteht darin, dass die Eindringtiefe des Ultraschalls aufgrund der hohen Frequenz in die Brust begrenzt ist, da die Dämpfung im Gewebe mit steigender Frequenz zunimmt.

Kontrastmittel-Magnetresonanztomographie

Die *Magnetresonanztomographie* (MRT), auch Kernspintomographie genannt, ist zum jetzigen Zeitpunkt in der klinischen Praxis die einzige dreidimensionale Methode zur Brustkrebsdiagnose. Jedoch kann auch für die MRT keine Empfehlung als alleinige Methode zur Vorsorgeuntersuchung ausgesprochen werden [16]. Bei der MRT wird der Patientin ein Kontrastmittel injiziert, das sich schneller in Tumorgewebe anreichert und ausgewaschen wird als bei gesundem Gewebe. In speziellen Brust-MRTs mit Magnetfeldstärken von 1,5 bis 3 Tesla wird somit der Tumor erkannt. Die erzeugten Schichtbilder haben eine Auflösung im Millimeterbereich, mit der bereits Tumore einer Größe von 5-10 mm diagnostiziert werden können [23], wobei die Metastasenwahrscheinlichkeit bei geringen 5% liegt [22]. Die Nachteile sind zum einen der hohe Kostenaufwand, insbesondere bei der Anschaffung. Deshalb wird diese Methode erst angewandt, wenn die gängigeren Untersuchungsmethoden keine klaren Ergebnisse erzielen. Weitere Nachteile ergeben sich durch die Untersuchung in einer engen Röhre, die bei der Patientin Unbehagen auslösen kann, sowie mögliche Allergien verursacht durch das Kontrastmittel. Der Vorteil liegt in der guten Reproduzierbarkeit und der hohen Bildqualität der Aufnahmen. Die MRT als einziges echtes 3D-Verfahren stellt eine gute Vergleichsmöglichkeit für 3D-USCT-Bilder dar.

Röntgen-Computertomographie

Die Röntgen-Computertomographie ist ein neues Verfahren zur Brustkrebsdiagnose, welches sich noch in der Forschungsphase befindet und somit aktuell noch nicht in der klinischen Praxis Verwendung findet. Hierbei wird die Brust mithilfe von Röntgenstrahlen dreidimensional untersucht. Der entscheidende Nachteil, der bisher dem Einsatz dieser Methode im Weg stand, liegt in der hohen Strahlendosis (7-18 mGy [24]). Diese stellt eine unverhältnismäßig hohe Belastung des Organismus dar. Die Vorteile der Methode liegen in der hohen Auflösung der Bilder im Mikrometerbereich sowie in der guten Reproduzierbarkeit der Aufnahmen.

Biopsie

Die Biopsie (Gewebeentnahme) ist eine weitere Untersuchungsmethode, bei der es sich nicht um eine Vorsorgeuntersuchung, sondern um einen minimal-invasiven Eingriff handelt. Nach der Gewebeentnahme wird das Gewebe näher untersucht, um es als gut- oder bösartig diagnostizieren zu können. Bei gutartigem Gewebe wird eine regelmäßige Kontrolle durchgeführt, bösartiges Gewebe muss dagegen zeitnah entfernt werden.

2.4 Analyse der Bildqualität

Für die Analyse eines bildgebenden Systems sind quantitative Kriterien der Bildqualität wichtig, anhand derer die Leistungsfähigkeit des Systems beschrieben und mit anderen verglichen werden kann. Das gilt auch für die zur Rekonstruktion eingesetzten Algorithmen. Will man die Bildqualität unterschiedlicher Algorithmen miteinander vergleichen, kann das am besten mit geeigneten Bildqualitätsmaßen geschehen. Dazu soll im Folgenden zuerst generell darauf eingegangen werden, wie Abbildungssysteme beschrieben werden können. Anschließend wird der Fokus auf die Bewertung der Bildqualität gelegt und Metriken beschrieben, die geeignete Maße für das Beispiel des 3D-USCT-Prototypen am IPE darstellen.

Ein Abbildungssystem wird maßgeblich von seiner Apertur bzw. Sensorgeometrie bestimmt. Die **Apertur** wird in der Optik meist als Öffnungsweite einer Blende bzw. als deren Durchmesser bezeichnet, mit der Licht empfangen werden kann. Allgemeiner wird die Gesamtheit der räumlichen Ausdehnung der aktiven Abbildungselemente als Apertur bezeichnet. Jede Apertur besitzt dabei eine Abstrahlcharakteristik, die sich aus der Abstrahlcharakteristik aller einzelnen Abbildungselemente und ihrer räumlichen Anordnung zusammensetzt. Im Fall des 3D-USCT-Prototypen am IPE ist damit die Anordnung aller *Ultraschall* (US)-Transducer gemeint, die auf der Oberfläche in Form eines Halbellipsoids angeordnet sind und als Abbildungselemente dienen.

Für die **Leistungsfähigkeit** eines Abbildungssystems sind insbesondere die Auflösung und Abbildungsfunktion, der Kontrast, das Rauschen und die Bildartefakte von Bedeutung, auf die im Folgenden näher eingegangen wird. **Das Auflösungsvermögen** lässt sich allgemein als die Fähigkeit eines bildgebenden Systems definieren, zwei Strukturen unterscheiden zu können. Die Auflösung eines Systems stellt dabei den kleinsten noch wahrnehmbaren Abstand der beiden Strukturen dar. Um das Auflösungsvermögen eines Systems zu beschreiben, wird oft die Abbildung eines unendlich kleinen punktförmigen Objekts genutzt, das unterhalb der Auflösungsgrenze liegt. Aus der Systemtheorie ist bekannt, dass eine Erregung mit einem Dirac-Impuls eine Systemantwort hervorruft, die nur vom System selbst abhängig ist [25]. Ein punktförmiges Objekt, mit einer Ausdehnung unterhalb der Auflösungsgrenze, stellt für ein bildgebendes System solch einen Dirac-Impuls dar und ist daher als Testsignal für das System geeignet. Diese Abbildung wird auch als **Abbildungsfunktion**, *Punktspreizfunktion* (PSF) oder engl. *point spread function* bezeichnet. Abbildung 2.2 zeigt solche Abbildungsfunktionen beispielhaft nach [26] in einer dreidimensionalen (oben) und einer zweidimensionalen Darstellung (unten). Wird ein (angenähert) unendlich kleiner und heller Punkt auf einem schwarzen Hintergrund (a) abgebildet, kann er in einem realen System nicht mehr genauso scharf dargestellt werden. Er wird eine systemspezifische Glättung erfahren und mit einer gewissen Ausdehnung dargestellt werden (b). Werden zwei oder mehr Punktstreuer abgebildet, ist die Abbildungsfunktion für das Auflösungsvermögen entscheidend. In Abbildung 2.2 (c) sind vier Punktstreuer gegeben: Zwei nah beieinander liegende und zwei weiter entfernte Punktstreuer. Die nah beieinander liegenden Punktstreuer in (c) können systembedingt nicht mehr unterschieden werden (d) und erscheinen als ein etwas hellerer und langgezogener Punktstreuer. Die beiden weiter entfernten Punkte können dagegen klar voneinander unterschieden werden. Um den kleinsten noch wahrnehmbaren Abstand zweier Punkte zu bestimmen, wird häufig das

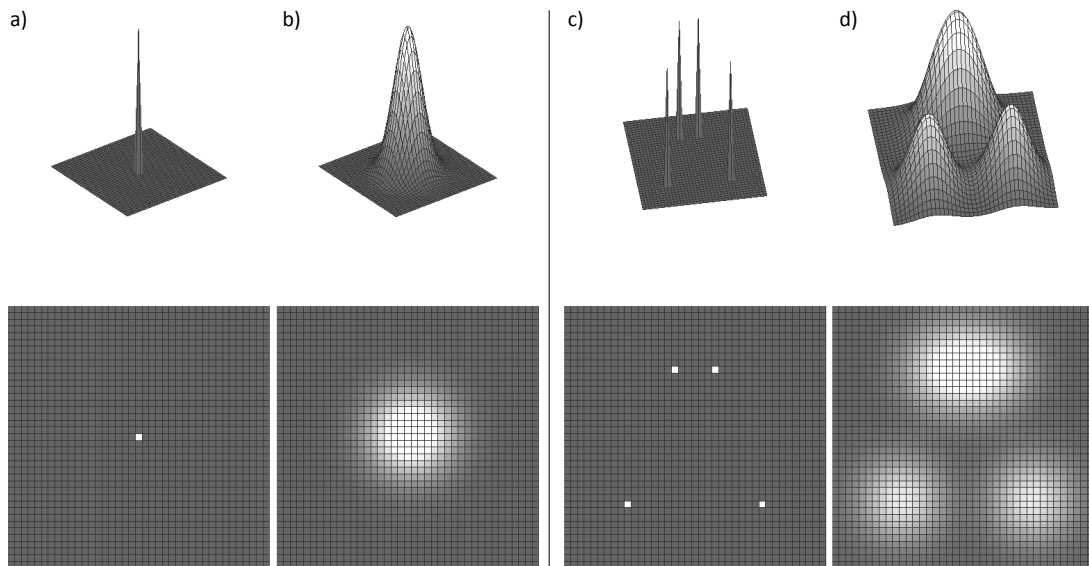


Abbildung 2.2: Beispiel eines einzelnen Punktstreuers (a) und seiner systembedingten Abbildungsfunktion (b). Beispiel zweier unterschiedlich weit entfernter Punktstreuerpaare (c) und ihre entsprechende Abbildungsfunktion (d). Oben in einer dreidimensionalen und unten in einer zweidimensionalen Darstellung.

Rayleigh-Kriterium aus der Optik angewandt. Das Kriterium von Rayleigh [29] besagt, dass zwei Bilder gerade noch auflösbar sind, wenn das Maximum des einen Bildes mit dem Minimum des anderen zusammenfällt (siehe Abbildung 2.3, links). Daran angelehnt wird allgemeiner die Halbwertsbreite

$$FWHM = |x_1 - x_2| , \quad (2.1)$$

oder auch engl. *Full Width at Half Maximum* (FWHM) als Maß für die Auflösung verwendet (siehe Abbildung 2.3, rechts). Hierbei gilt

$$f(x_1) = f(x_2) = \frac{1}{2} f(x_{max}) . \quad (2.2)$$

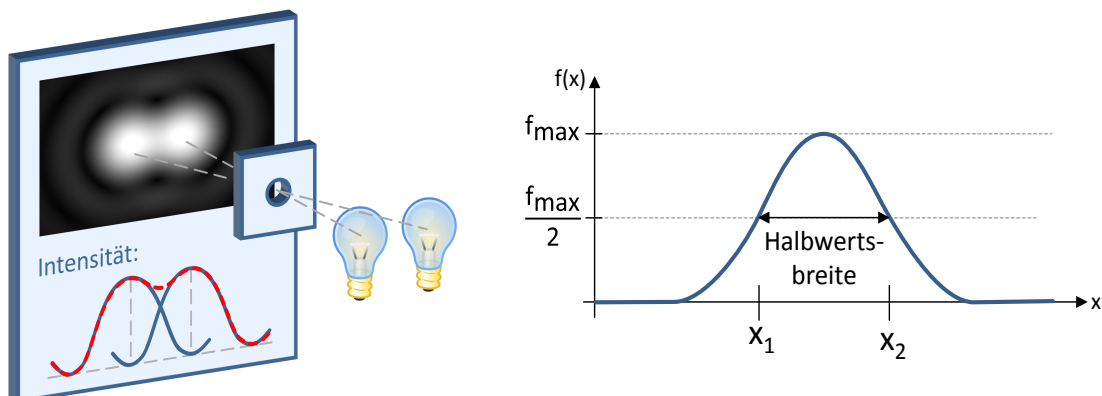


Abbildung 2.3: Das Rayleigh-Kriterium adaptiert nach [27] (links) und die Halbwertsbreite nach [28] (rechts).

Der Kontrast bezeichnet den Unterschied zwischen hellen und dunklen benachbarten Bereichen eines Bildes. Nach [26] wird der Kontrast in der bildgebenden Medizintechnik bestimmt durch

1. den bildgebenden Prozess, d.h. der Intensität der Quelle, der Abschwächung der Quellenergie und der Sensitivität des Aufnahmeegerätes.
2. die Charakteristik der Umgebung bei der Aufnahme, d.h. physikalische Eigenschaften, Größe und Form des Objekts und den Einsatz von Kontrastmitteln.
3. den Bedingungen beim Darstellen des Bildes, d.h. der Raumausleuchtung und der Ausrüstung für die Darstellung wie z.B. den eingesetzten Displays.

Dabei ist im allgemeinen ein hoher Kontrast wünschenswert. Insbesondere um gesuchte Strukturen vom Hintergrund oder anderen Strukturen deutlicher abzuheben. In der bildgebenden Medizintechnik z.B. werden häufig Kontrastmittel eingesetzt. Dadurch wird der Kontrast von Tumorgewebe zu gesundem Gewebe angehoben und kann besser unterschieden werden.

Rauschen ist als gewisser Anteil bei der Abbildung der Realität immer vorhanden. Je stärker das Rauschen wird, desto mehr besteht die Gefahr, dass die Bildinformation vom Rauschen überdeckt wird und so nicht mehr erkennbar ist. Als ein wichtiges Maß für das Verhältnis von nutzbarer Information zum Rauschen gilt das Signal-zu-Rausch-Verhältnis aus der Signaltheorie, das auch Störabstand oder *signal-to-noise ratio* (SNR) genannt wird.

Bildartefakte oder auch Bildfehler sind charakteristisch für ein bestimmtes bildgebendes Verfahren. Bildartefakte verändern das Bild und können eine richtige Diagnose erschweren oder im schlimmsten Fall verhindern. Daher ist es von entscheidendem Interesse ihr Auftreten und ihre Herkunft zu verstehen. Steht die Ursache der Bildartefakte fest, können eventuell geeignete Maßnahmen getroffen werden, um diese Artefakte zu vermeiden. Oder diese Artefakte können wie bei der konventionellen US-Bildgebung sogar genutzt werden, um Informationen über das Gewebe zu erlangen.

1. Speckle-Rauschen

Ein typisches Bildartefakt bei der US-Bildgebung ist das Speckle-Rauschen, das sich im Bild als körniges Muster darstellt. Es entsteht durch die Interferenz von Schallwellen, die an kleinen Inhomogenitäten im Gewebe gestreut werden [30]. Dabei ist die Stärke der Streuung abhängig von dem Verhältnis der Größe der streuenden Inhomogenitäten und der Wellenlänge der US-Welle (siehe Abbildung 2.4). Bei der 3D-USCT wird durch die Abbildungen aus vielen verschiedenen Sichtwinkeln das Speckle-Rauschen deutlich unterdrückt (*spatial compounding* [32]) und wird daher in dieser Arbeit nicht weiter untersucht.

2. Dämpfungsartefakte

Dämpfungsartefakte treten insbesondere auf, wenn es starke Dämpfungsunterschiede in dem zu untersuchenden Gewebe gibt. Standardmäßig wird ein durchschnittlicher Dämpfungswert des Gewebes angenommen, um eine zeitdynamische Verstärkung als Korrektur anzuwenden (*time gain compensation* (TGC)). Während es hinter stark reflektierenden oder dämpfenden Objekten zu lokalen Abschattungen kommt, führt ein schwach dämpfendes Objekt zu einer lokalen Verstärkung. In Abbildung 2.5 ist links eine Leber mit Gallenblase dargestellt, die

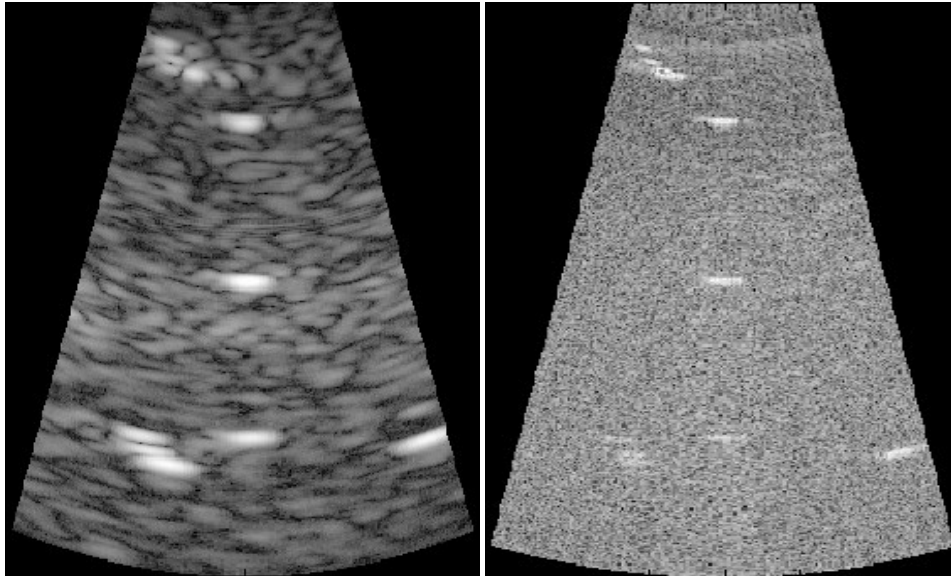


Abbildung 2.4: Typisches Speckle-Rauschen bei konventionellem US bei niedriger (links) und hoher (rechts) US-Frequenz [31].

Gallensteine enthält. Die Gallensteine mit hoher Dichte verursachen eine Abschattung. Der Rest der Gallenblase ist mit schwach dämpfender Flüssigkeit gefüllt und führt daher zu einer Verstärkung des dahinter liegenden Gewebes. Diese Artefakte lassen daher einen Rückschluss auf die Dämpfung der dargestellten Objekte zu. Wird eine Ringapertur für die Datenaufnahme genutzt, führt eine hohe Dämpfung im Gewebe zu einer Abschattung zur Mitte hin (siehe Abbildung 2.5, rechts).

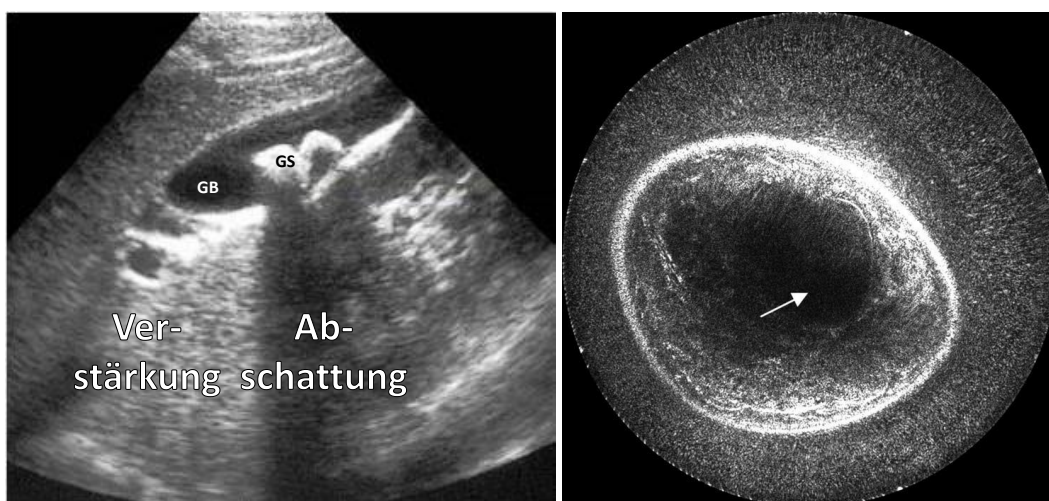


Abbildung 2.5: Eine Leber mit Gallenblase (GB) im Vordergrund, die Gallensteine (GS) enthält (links) [33]. Dämpfungs-Artefakte bei einer 2D-USCT mit Ringapertur (rechts) [34].

3. Brechungs- und Reflexionsartefakte

Trifft der ausgesendete Ultraschall auf eine Grenze zweier Gewebearten, wird er abhängig vom Winkel und den Schallwiderständen total reflektiert oder teilweise reflektiert und teilweise transmittiert. Die Art der Reflexion kann spiegelartig, diffus oder eine Mischart sein. Zudem kommt es dabei zu Brechungseffekten, die zu einer Verdoppelung der Objekte, Abschattungen oder Spiegelartefakten (siehe Abbildung 2.6) führen können. Der konventionellen US-Bildgebung liegt standardmäßig die Annahme einer konstanten Schallgeschwindigkeit zugrunde, bei der eine mittlere Schallgeschwindigkeit für das zu untersuchende Gewebe genutzt wird. Tatsächlich unterscheiden sich jedoch die Schallgeschwindigkeiten im Gewebe. Die Folge von größeren Abweichungen der Schallgeschwindigkeit ist eine Defokussierung und zeigt sich in verschiedenen Arten von Bildartefakten. Es können Artefakte in Form einer Verschiebung und fehlerhafter Größenabbildung von Objekten auftreten, wobei Oberflächen auch verformt dargestellt werden können.

4. Rekonstruktionsartefakte

In der 3D-USCT eingesetzten SAFT-Rekonstruktion für die Reflektivität, die in Kapitel 3.3 näher beschrieben wird, kommt es zu Rekonstruktionsartefakten. Diese Artefakte werden auch Grating-Lobe-Artefakte genannt [36, 37]. Wenn die Zahl der genutzten A-Scans verhältnismäßig groß ist, können diese Artefakte vernachlässigt werden. In Abbildung 2.7 sind beispielhaft drei Bilder eines Punktstreuers und seiner Umgebung von $2 \times 2 \text{ mm}^2$ für eine ringförmige Sensorgeometrie dargestellt, die mit 10 (links), 100 (mittig) und 1000 (rechts) *Sender-Empfänger-Paaren* (SEP) rekonstruiert wurden. Die Positionen der Ultraschallwandler (blau) und des Punktstreuers (rot) sind jeweils im oberen Teil dargestellt. Bei einer geringen Anzahl an SEP sind neben dem zentralen Punkt noch deutliche Artefakte zu erkennen (unten links). Diese stellen sich bei 100 SEP bereits weit geringer dar (unten mittig) und sind bei 1000 SEP nicht mehr auszumachen (unten rechts).

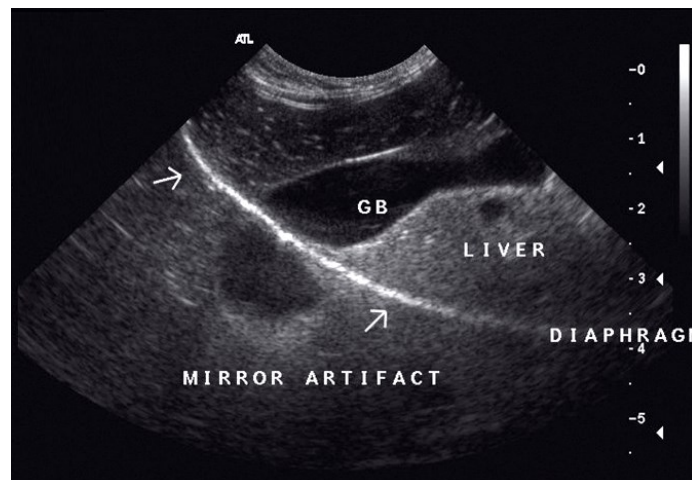


Abbildung 2.6: Reflexions-Artefakt durch das Zwerchfell (engl. diaphragm), das zu einer Spiegelung der Gallenblase führt [35].

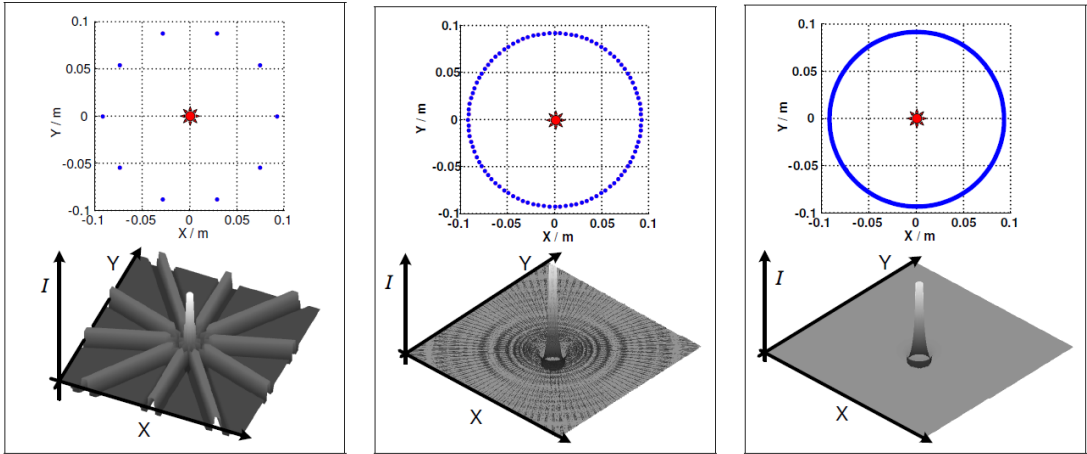


Abbildung 2.7: Grating-Lobe-Artefakte, die bei der SAFT-Rekonstruktion entstehen. Rekonstruktion bei 10 (links), 100 (mittig) und 1000 (rechts) genutzten SEP [38].

2.5 Grafikprozessoren

In dieser Arbeit werden zwei NVIDIA GPUs aus zwei GPU-Generationen eingesetzt. Die GTX590, die auf der Fermi-Architektur basiert und die GTX TITAN, basierend auf der Kepler-Architektur. Im Folgenden wird auf den Aufbau und die Funktion der Kepler-GPUs eingegangen und die Unterschiede zur Fermi-Architektur erläutert. Weiterhin wird auf die Programmierung mit CUDA-C eingegangen.

Aufbau und Funktion der Kepler-Architektur

Die parallele Verarbeitung der GPUs basiert auf der Compute Unified Device Architecture (CUDA) von NVIDIA. CUDA definiert dabei sowohl die skalierbare Hardwarearchitektur der Grafikprozessoren (GPUs) für parallele Berechnungen, als auch das Programmiermodell mit der Programmiersprache CUDA-C [39].

Die GPU als CUDA-Device besteht aus einem GPU-Chip und einem externen Speicher im Bereich von mehreren Gigabytes, dem *Global Memory*. Die Kommunikation mit der GPU verläuft über die PCIe-Schnittstelle, wobei sämtlicher Datentransfer über das *Global Memory* stattfindet. Der innere Aufbau ist schematisch in Abbildung 2.8 für die Kepler-Architektur dargestellt, die aus mehreren SMX-Einheiten (*eXtended streaming multiprocessor*) besteht. Eine SMX-Einheit stellt dabei die Weiterentwicklung der, in der Fermi-Generation noch genannten, SM-Einheit (*streaming multiprocessor*) dar. Weiterhin ist ein 1536 kB großer Level-2-Cache und mehrere Speichercontroller für die Kommunikation mit dem *Global Memory* auf dem GPU-Chip verbaut.

Der innere Aufbau einer SMX-Einheit ist in Abbildung 2.9 dargestellt. In einem SMX befinden sich 192 Recheneinheiten, sogenannte *CUDA Cores* (Core), die jeweils aus

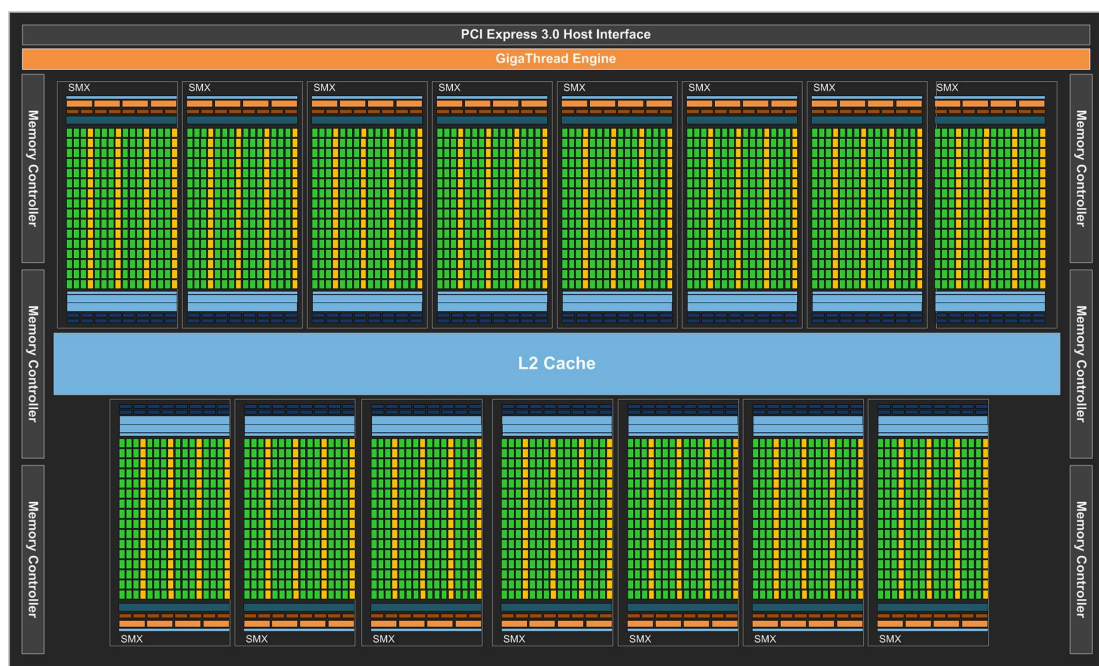


Abbildung 2.8: Schematische Ansicht des inneren GPU-Aufbaus des Kepler GK110-Chips [40].

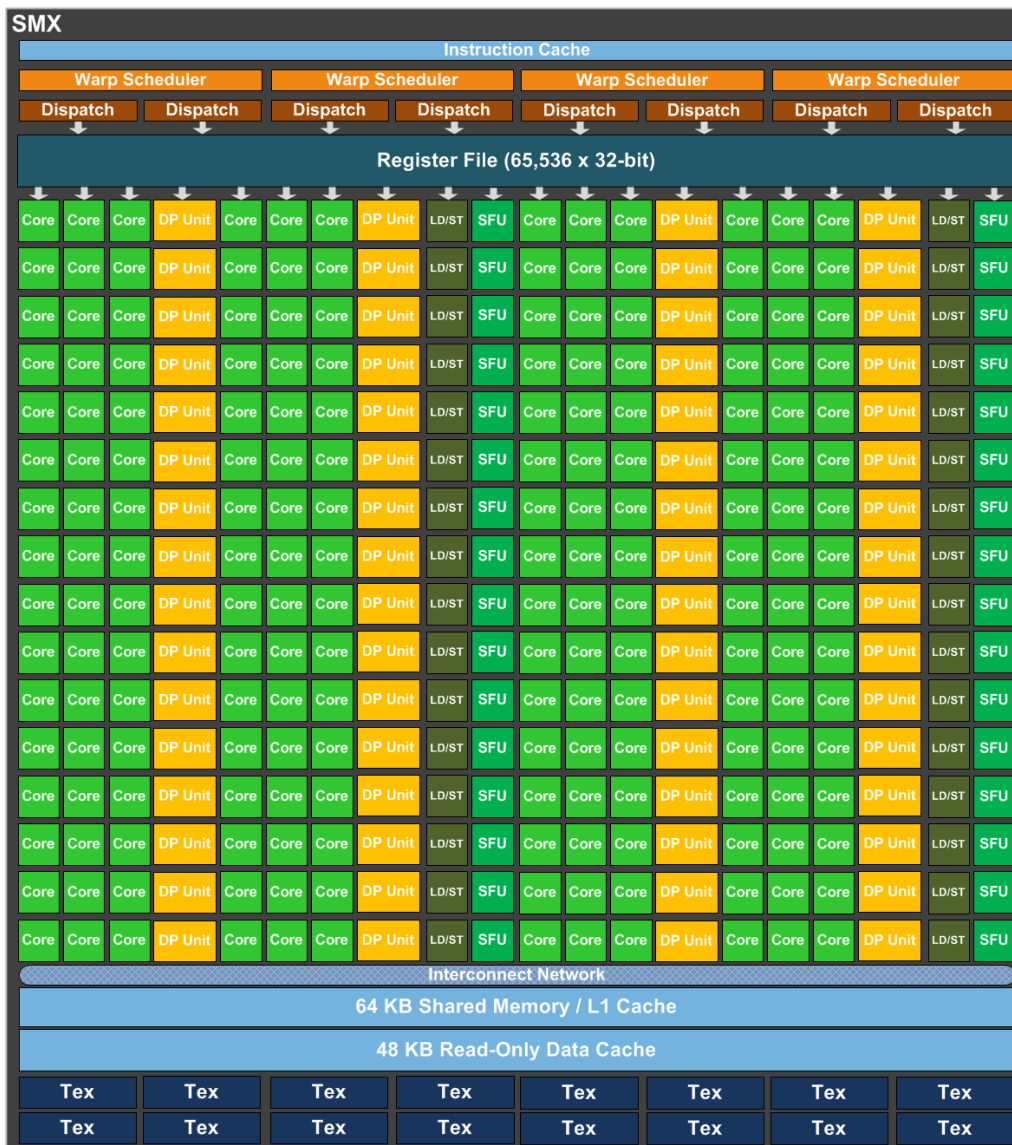


Abbildung 2.9: Schematische Ansicht des inneren Aufbaus einer SMX-Einheit [40].

einer Fest- und einer Gleitkommaeinheit bestehen und den IEEE-754-2008 Standard unterstützen. Für die Berechnungen von Gleitkommazahlen doppelter Genauigkeit sind in einer SMX-Einheit 64 eigene DP-Einheiten (*double-precision*, DP Units) verbaut. Weiterführende Funktionen wie Winkel-, Wurzel- oder Exponentialfunktionen werden über 32 Spezialfunktionseinheiten (*special function units*, SFU) abgearbeitet. Für die Speicherzugriffe stehen 32 Lese- und Schreibeinheiten (*load and store*, LD/ST) zur Verfügung. Über 16 Textureinheiten (*texture units*, TEX) sind speziell optimierte Lesezugriffe möglich, die aus der Verwendung von Texturen bei der Grafikberechnung stammen. Texturen wurden entwickelt, um eine schnelle Anzeige von Oberflächen in der 3D-Grafik zu ermöglichen. So ermöglicht eine Textur einen optimierten Zugriff auf Datenelemente, die in einem zwei- oder dreidimensionalen Format vorliegen. Es werden benachbarte Speicherstellen vorgeladen und liegen somit bereits in dem Cache vor. Zudem ermöglichen die Textureinheiten eine lineare Interpolation in 1D, 2D oder 3D, die direkt in der Hardware durchgeführt wird.

Weitere Speicher, die sich auf der GPU befinden, sind ein 256 kB großes Register (*register File*), ein 48 kB großer nur lesbarer Speicherbereich (*read-only data cache*) und ein 64 kB großer konfigurierbarer Speicher. Der *read-only data cache* fungiert dabei als ein nur lesbarer Level-1-Cache, über den auch die Texturzugriffe laufen. Weiterhin steht er für Daten zur Verfügung, die als konstant definiert wurden. Der 64 kB große konfigurierbare Speicherbereich beinhaltet einen gemeinsamen Speicher (*shared memory*) und einen Level-1-Cache. Für die Größe dieser beiden Speicher kann dabei eine von drei Einteilungen gewählt werden: 16/48 kB, 32/32 kB oder 48/16 kB.

Für die parallele Datenverarbeitung auf einer GPU kommt NVIDIAS Datenverarbeitungsstrategie namens *single-instruction-multiple-threads* (SIMT) zum Einsatz [39]. Dieser Name ist angelehnt an die *single-instruction-multiple-data* (SIMD)-Klasse der Flynnschen Klassifikation [41]. Im Vergleich zur CPU, bei der mehrere Datenelemente in einem *Thread* abgearbeitet werden, wird bei der GPU pro Datenelement ein eigener Thread erzeugt und bearbeitet. Dabei werden vom *Warp Scheduler* jeweils 32 *Threads* zu einer Gruppe zusammengefasst, die *Warp* genannt wird. Bei der Abarbeitung der *Threads* starten die einzelnen *Threads* alle an der selben Programmadresse. Alle *Threads* eines *Warps* führen strikt nach Instruktionsabfolge jeweils eine gemeinsame Instruktion aus. Jedoch hat jeder *Thread* eine eigene Instruktionsadresse und ein eigenes Statusregister, wodurch es möglich ist, dass einzelne *Threads* verzweigen und einem anderen Programmfluss folgen. Solche Kontrollflussdivergenzen führen jedoch zu einer sequentiellen Abarbeitung der einzelnen Kontrollflusspfade, wobei die jeweiligen inaktiven *Threads* ausgeschaltet werden, die einem anderen Kontrollfluss folgen. Erst wenn alle nötigen Kontrollflusspfade sequentiell abgearbeitet wurden, werden die folgenden Instruktionen wieder gemeinsam abgearbeitet.

Um Wartezeiten während der Verarbeitung, die z.B. bei Speicherzugriffen auf den *global memory* entstehen können, dennoch auszunutzen, besteht die Möglichkeit eines Kontextwechsels. Bei diesem, von NVIDIA genannten *Hardware Multithreading* [39], wird die freie Zeit genutzt, um alternative *Warps* der insgesamt vier *Warp Scheduler* weiter zu verarbeiten. Die Anzahl der *Warps*, die parallel abgearbeitet werden können, hängt dabei von dem Ressourcenbedarf der *Threads* ab, da die Anzahl an Registern in einem SMX beschränkt ist. Ist der Ressourcenbedarf jedoch zu hoch, kann nur ein *Warp* abgearbeitet werden. Für eine hohe Auslastung der GPU ist es wichtig, dass immer eine genügend hohe Anzahl an parallelen *Warps* für Berechnungen bereitstehen. Die prozentuale Auslastung der GPU durch die *Warps* wird *Occupancy* genannt.

Eine weitere Möglichkeit bietet sich dadurch, dass pro *Warp Scheduler* zwei *Dispatch Units* verbaut sind. Durch das sogenannte *instruction level parallelism* (ILP) ermöglichen sie es, zwei aufeinanderfolgende unabhängige Instruktionen der *Warps* jeweils in verschiedenen Funktionseinheiten auszuführen [42].

Unterschiede der Fermi-Architektur

Die Fermi-Architektur [43] ist der Vorgänger der Kepler-Architektur und basiert auf einem sehr ähnlichen Hardwareaufbau, wie in Abbildung 2.8 und Abbildung 2.9 dargestellt. Eine Übersicht der Hardwareparameter ist in Tabelle 2.1 aufgelistet. Der GF110-Chip der Fermi-Generation besitzt dabei 16 SM-Einheiten im Vergleich zu 14 SMX auf dem GK110-Chip. Eine SM-Einheit der Fermi-Generation besteht im

Generation GPU-Bezeichnung	Fermi GTX 590	Kepler GTX TITAN
Prozesstechnologie	40 nm	28 nm
Jahr der Markteinführung	2010	2013
Mrd. Transistoren	3,0	7,1
Grafikprozessor	GF110	GK110
Taktrate in GHz	1,54	0,837
max. Leistungsaufnahme in W	244	250
max. Rechenleistung in TFlop/s	1,6	4,5
max. Speicherbandbreite in GB/s	192	288
Anzahl von SM(X)s	16	14
Level-2-Cache in kB	768	1536
Globaler Speicher in GB	1,5	6
CUDA Cores	32	192
Spezialfunktionseinheiten (SFU)	4	32
Textureinheiten (TEX)	4	16
Lese- und Schreibeinheiten (LD/ST)	16	32
<i>Warp Scheduler/Dispatcher</i>	2/2	4/8
<i>Register File</i> in kB	128	256
<i>Shared Memory</i> in kB	64	64

Tabelle 2.1: Übersicht über die Spezifikationen der genutzten GPU-Hardware nach [9]. Der obere Teil bezieht sich auf den gesamten Chip und der untere Teil jeweils auf einen SM(X) [40, 43].

Vergleich zu einer SMX-Einheit aus deutlich weniger Funktionseinheiten. Dies begründet sich vor allem durch das Ziel, die Leistungsaufnahme der neueren Kepler-Generation deutlich zu reduzieren [40]. Dies wurde insbesondere durch eine nahezu halbierte Taktrate erreicht. Um dennoch eine höhere Rechenleistung mit der Kepler-Generation zu erreichen, ist eine höhere Anzahl an Funktionseinheiten der Kepler-Generation nötig. So besitzt die Fermi-GPU pro SM insbesondere nur ein Sechstel der CUDA Cores, ein halb so großes *register file*, ein Achtel der SFU-Einheiten und ein Viertel der Textureinheiten. Die Größe des *shared memory* ist dagegen gleich. Da gleich viele *Warp Scheduler* wie Dispatch-Einheiten verbaut sind, ist bei der Fermi-Generation noch keine ILP möglich.

Programmierung

Die Programmierung der GPU erfolgt passend zur CUDA Hardwarearchitektur durch CUDA-C, wobei auch hier ein hierarchisches Programmier- und Speichermodell genutzt wird. In CUDA-C wird ein sequentielles Programm geschrieben, das *Kernel* genannt wird. Auf der GPU wird die parallele Ausführung des *Kernels* durch den Aufruf für mehrere *Threads* gestartet. Dabei werden die *Threads*, wie in Abbildung 2.10 dargestellt, in mehrere *Blocks* eingeteilt. Die maximale Anzahl an *Threads* in einem *Block* beträgt für die Fermi- und Kepler-Generation 1024 *Threads*. Ein *Block* wird für die Zeit seiner Ausführung einem SM(X) fest zugeordnet. Mehrere *Blocks* werden wiederum zu einem *Grid* zusammengefügt. Den *Threads* und *Blocks* werden bei ihrem Aufruf eindeutige Indizes zugewiesen, die jeweils mit x-, y- und z-Koordinaten angegeben werden. Diese dreidimensionale Indizierung erleichtert die

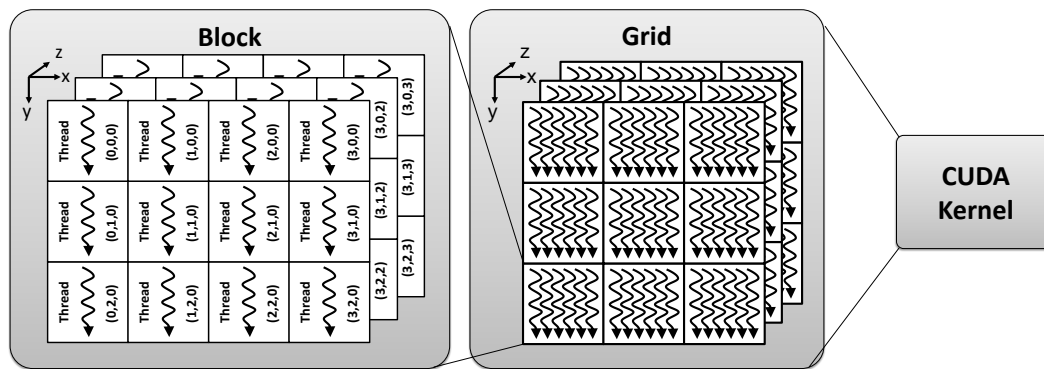


Abbildung 2.10: Hierarchische *Thread*-Gruppierung in CUDA (nach [39]).

Berechnung bei dreidimensionalen Problemstellungen, wie z.B. der Berechnung eines Volumens.

Die hierarchische Einteilung in *Threads*, *Blocks* und *Grid* führt auch zu einem hierarchischen Zugriff auf die oben erwähnten Speicherbereiche, wie in Abbildung 2.11 dargestellt. Jeder Thread kann auf die lokalen privaten Register zugreifen, wobei der Zugriff innerhalb eines Taktes stattfindet. Alle Threads innerhalb eines *Blocks* haben die Möglichkeit den *shared memory* zu nutzen, um darüber Daten zwischen Threads auszutauschen. Hierbei kann der Zugriff ebenfalls in einem Takt durchgeführt werden. Jedoch ist das *shared memory* in 32 Speicherbänke aufgeteilt und es kommt zu Bankkonflikten, wenn auf dieselbe Speicherbank zugegriffen wird. Ein Bankkonflikt führt wiederum zu einer sequentiellen Abarbeitung der Speicherzugriffe. Eine Speicherkommunikation über mehrere Blöcke hinweg ist nur über das *global memory* möglich. Der Zugriff über das *global memory* benötigt 200 - 400 Taktzyklen bei der Kepler-Generation [39].

Ein weiterer besonderer Speicherbereich ist der *constant memory*, der eine Größe von 64 kB besitzt. Da sich dieser Speicher im *global memory* befindet, ist zudem jeder SMX mit einem 8 kB großen *constant cache* ausgestattet. Jeder *Thread* kann auf die Daten des *constant memory* lesend zugreifen. Greifen mehrere *Threads* gleichzeitig auf eine Speicheradresse zu, werden die Daten allen *Threads* gleichzeitig bereitgestellt (*Broadcast*). Wird auf verschiedene Speicheradressen zugegriffen, werden die Lesezugriffe seriell ausgeführt. Dieser Speicher eignet sich daher besonders für Daten, auf die alle *Threads* gleichzeitig zugreifen.

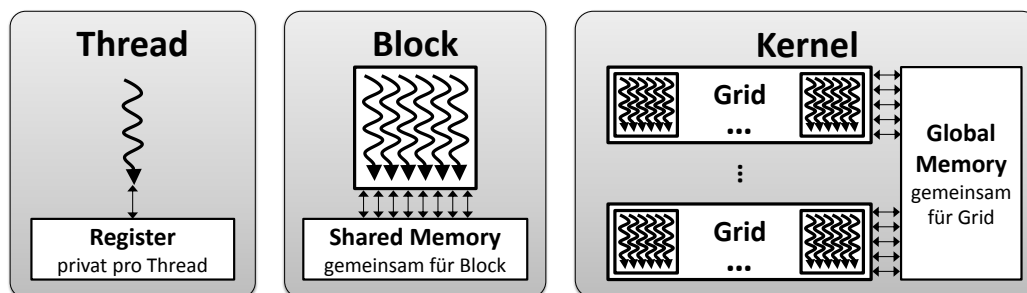


Abbildung 2.11: Speicherzugriffsmöglichkeiten innerhalb der hierarchischen *Thread*-Gruppierung in CUDA (nach [40]).

2.6 Stand von Technik und Forschung

Die am IPE entwickelte 3D-USCT ist die weltweit erste vollständig dreidimensionale Untersuchungsmethode basierend auf Ultraschall. Die Brust wird dabei nicht nur aus allen Dimensionsrichtungen beschallt. Es werden ebenso Reflexionen aufgenommen, die in alle Dimensionsrichtungen abstrahlen. Bisherige dreidimensionale bildgebende Systeme beruhen dagegen auf einer Zusammensetzung von Schichtbildern. Im Folgenden soll ein kurzer Überblick über die aktuellen Standardverfahren der medizinischen Bildgebung mit Ultraschall gegeben werden.

Konventioneller Ultraschall

Konventionelle Ultraschallsonden, wie in Abbildung 2.12 links dargestellt, bestehen heute meist aus mehreren kleineren Ultraschallwandlerelementen, die zu einem Array zusammengefasst werden (siehe Abbildung 2.12 mittig). Ein solches Ultraschallarray kann in unterschiedlichen Formen angeordnet sein, wie z.B. linear in einer Reihe (*linear array*) oder konvex (*curved array*). Die Ultraschallsonde sendet einen gerichteten Ultraschallstrahl aus. An den Gewebestrukturen kommt es dabei zu Reflexionen des Ultraschalls, die wieder empfangen werden können. Über die Reflexionen und die Laufzeit kann somit die Struktur innerhalb des Gewebes in der Strahlrichtung vermessen werden. Anschließend wird der Strahl in lateraler Richtung weitergeschwenkt, um ein 2D-Bild darstellen zu können (siehe Abbildung 2.12 rechts). Das Schwenken des Ultraschallstrahls kann dabei mechanisch oder über eine zeitversetzte Ansteuerung der Einzelelemente (*phased array*) durchgeführt werden. Bei dieser bildgebenden Methode wird das Medium nur einseitig bestrahlt. Dadurch kommt es bei stark dämpfenden oder stark reflektierenden Strukturen zu einer Abschattung von tiefergelegenen Bereichen. Weiterhin treten die Ultraschall-typischen Speckle-Artefakte auf, die zu einer Reduzierung des Bildkontrasts führen. Die Auflösung ist stark anisotrop. Die axiale Auflösung ist von der genutzten Mittenfrequenz und Bandbreite abhängig. Die laterale Auflösung ist dagegen zusätzlich von der Ausdehnung der aktiven Sensorelemente abhängig und verglichen mit der axialen Auflösung geringer.

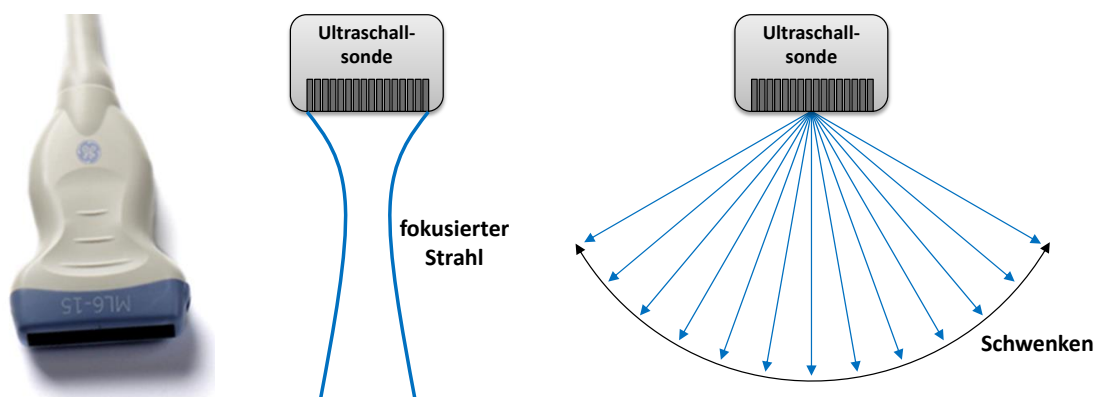


Abbildung 2.12: Beispiel einer Ultraschallsonde mit linearem Array (links). Aussenden eines Ultraschallstrahls (mittig). Schwenken des Ultraschallstrahls in laterale Richtung für eine 2D-Bildgebung (rechts).

Ultraschall-Computertomographie

Um den Nachteil des konventionellen Ultraschalls zu vermeiden, wird die Brust bei der Ultraschall-Computertomographie (USCT) anstatt aus einer aus mehreren Richtungen vermessen. Dazu können mehrere Ultraschallsensoren auf einem Ring angeordnet werden [44]. Oder konventionelle Ultraschallsonden werden mechanisch um die Brust geführt [45]. Durch die Vermessung des Mediums aus mehreren Richtungen kann die Abschattung der Bereiche hinter stark dämpfenden Strukturen verbessert und das Speckle-Rauschen unterdrückt werden (*spatial compounding*). Zudem wird das Auflösungsvermögen nahezu isotrop, da die axiale und die laterale Auflösung miteinander kombiniert werden. Ein großer Vorteil der USCT ist zudem, dass nicht mehr nur die Reflexionen gemessen werden können. Sind die Ultraschallwandler auf beiden Seiten des Objekts auf einem Sensorring vorhanden, können zudem die Transmissionsdaten aufgezeichnet werden. Dadurch ist es möglich, Informationen über die Dämpfung und Schallgeschwindigkeit innerhalb des Mediums zu gewinnen. Da die Datenaufnahme mit fokussierten Ultraschallstrahlen deutlich mehr Zeit benötigen würde, können unfokussierte Signale ausgesendet und somit direkt die gesamte Brust vermessen werden. Für die Bildgebung muss jedoch anschließend die Fokussierung in der Software durchgeführt werden (Post-Beamforming), wodurch der Rechenaufwand deutlich ansteigt. Ein weiterer Vorteil der USCT stellt vor allem die Reproduzierbarkeit der USCT-Aufnahmen dar. Da es hier, im Vergleich zur Mammographie oder dem konventionellen Ultraschall, zu keiner Deformation der Brust kommt, können die Bilder von verschiedenen Zeitpunkten verglichen und auf Veränderungen im Gewebe hin überprüft werden.

3D-Ultraschall-Computertomographie

Der konventionelle Ultraschall lässt sich durch Drehen, Schwenken oder Verfahren der Ultraschallsonde um eine weitere Dimension erweitern. Wird die Position und Ausrichtung der Ultraschallsonde verfolgt (Tracking), kann aus den 2D-Bildern ein 3D-Bild berechnet werden. 3D-Ultraschallsysteme sind vor allem durch die Pränataldiagnostik bekannt geworden, bei der eine gerenderte Darstellung [48] der berechneten Volumen eingesetzt wird (siehe Abbildung 2.13). Diese Systeme besitzen



Abbildung 2.13: 3D-Ultraschallbilder von Zwillingsembryos [46, 47].

jedoch ebenfalls nur eine Blickrichtung, weshalb keine Transmissionsbilder berechnet werden können.

Auch aus mehreren 2D-USCT-Schichtbildern, die über eine Positionsänderung des Sensorrings aufgenommen wurden, kann ein 3D-Bild berechnet werden. Die beiden aktuell wichtigsten Forschungsgruppen neben der 3D-USCT-Gruppe am IPE sind die Delphinus Medical Technologies aus Plymouth (USA) und Techniscan Medical Systems aus Salt Lake City (USA). Beide haben bereits eine Vielzahl klinischer Fälle untersucht und werden im Folgenden kurz vorgestellt.

Die Firma Delphinus Medical Technologies Delphinus entstand 2010 als eine Ausgründung des Karmanos Cancer Institute in Detroit (USA). Ihr zweiter Prototyp wird SoftVue genannt und besitzt bereits eine FDA-Zulassung für medizinische Bildgebung zu diagnostischen Zwecken [49]. Beim SoftVue System [50] sind 2048 Ultraschallsensoren mit einer Mittenfrequenz von 3 MHz auf einem Ring angeordnet, der während der Vermessung in vertikaler Richtung verschoben wird und somit den gesamten Brustbereich abdecken kann (siehe Abbildung 2.14 rechts). Damit ergeben sich 4,19 Millionen *Sender-Empfänger-Paare* (SEP), die mit 512 parallelen Empfangskanälen ausgelesen werden. Es können Transmissions- und Reflexionsbilder erstellt werden. Die Auflösung des Systems für Reflexionsbilder wird mit 2,5 mm x 0,3 mm x 0,3 mm bei einem maximalen Brustdurchmesser von 22 cm angegeben. Für die Rekonstruktion der Reflexionsbilder wird eine Brechungs- und Dämpfungskorrektur in 2D durchgeführt. Diese Korrekturen werden in [34] für mehrere Brüste experimentell untersucht. Eine Schicht kann dabei in 10 Sekunden berechnet werden. Dazu kommt eine Blade-Server-Architektur mit fünf Blades zum Einsatz [51], wobei ein Master-Blade vier Compute-Blades steuert, die jeweils mit zwei Quad-Core Intel Xeon E5620 CPUs, zwei Nvidia Tesla M2070 GPUs und 96 GB RAM ausgestattet sind. Das Master-Blade ist ebenfalls mit zwei Quad-Core Intel Xeon E5620 CPUs ausgestattet und besitzt 192 GB Hauptspeicher.

Der USCT-Scanner der Firma TechniScan Medical Systems aus Salt Lake City (USA) kann ebenfalls Transmissions- und Reflexionsbilder erstellen. Dabei wird ein getrennter Aufbau genutzt, der gehoben und gedreht werden kann (siehe Abbil-

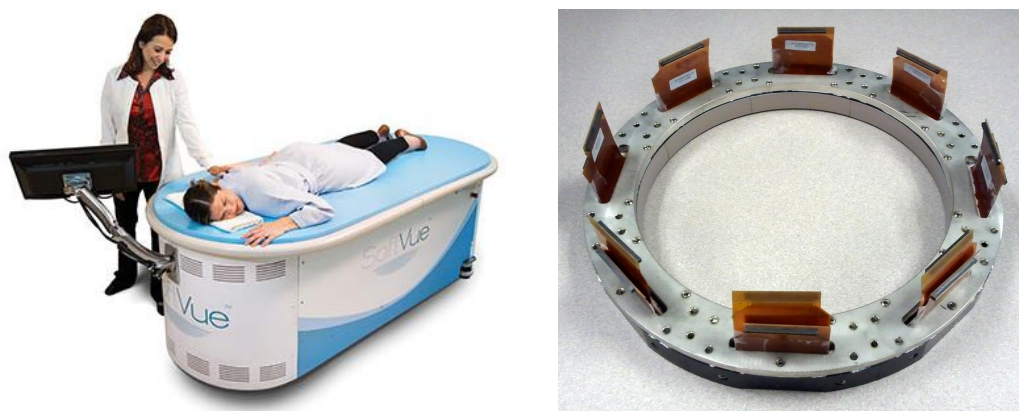


Abbildung 2.14: Das SoftVue-USCT-System von Delphinus [51]. Der Prototyp (links) und der Sensorring mit 2048 Ultraschallwandlern (rechts).

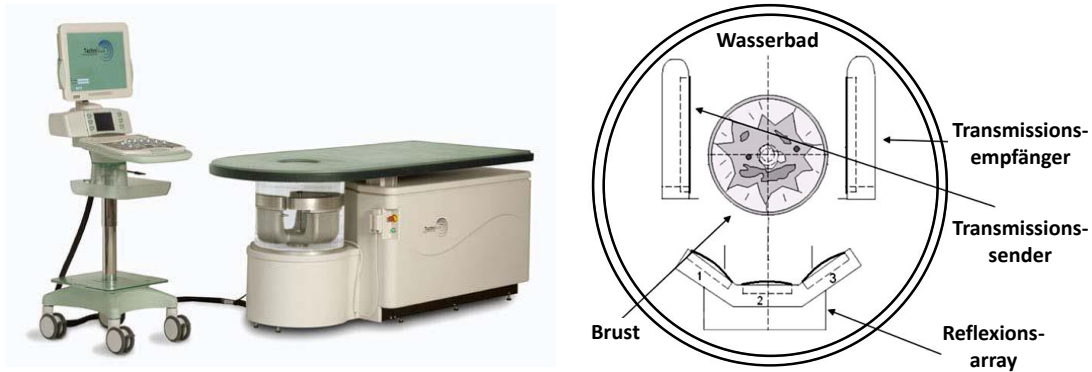


Abbildung 2.15: Der USCT-Scanner von Techniscan Medical Systems (links) und der getrennter Aufbau der Ultraschallwandler für Transmissions- und Reflexionsbilder (rechts) [52].

Abbildung 2.15). Für die Transmissionsbilder stehen zwei gegenüberliegende Sensorfelder mit 64 Sendern und 960 Empfängern zur Verfügung [52]. Die Reflexionsdaten werden mit drei Ultraschallsonden mit insgesamt 192 Ultraschallsensoren aufgezeichnet. Sie arbeiten bei einer Mittenfrequenz von 5 MHz, 80% Bandbreite und sind auf unterschiedliche Tiefen fokussiert. Für die Rekonstruktion der einzelnen 2D-Schichtbilder werden die Bilder aus unterschiedlichen Richtungen kombiniert und dargestellt (*360° spatial compounding* der B-Scans). Es wird in [53] erwähnt, dass die Rekonstruktion auf Basis der Schallgeschwindigkeitsbilder eine Brechungskorrektur durchführt und die Amplituden entlang des Strahls mithilfe der Dämpfungsbilder angepasst werden. Jedoch werden keine Informationen angegeben, wie diese Korrekturen angewandt werden. Die Auflösung der Bilder wird mit 0,8 mm x 0,8 mm x 1 mm angegeben [52]. Die Gesamtrekonstruktionsdauer für Transmissions- und Reflexionsbilder beträgt dabei ca. 40 Minuten. Ein 2D-Schichtbild wird innerhalb von Sekunden berechnet, das 3D-Bild benötigt ca. 12 Minuten bei einem Einsatz von zwei GPUs [53]. Auf Grund finanzieller Probleme wurde die Arbeit an dem System eingestellt, wodurch die Zukunft dieses Systems ungewiss ist.

Beiden Systemen ist gemeinsam, dass die Reflexionsdaten nur in einer Ebene aufgenommen werden. Dabei nutzt das SoftVue-System von Delphinus bereits unfokussierten Ultraschall in der xy-Ebene. In Elevationsrichtung ist er jedoch fokussiert. Dies führt zu einer 8,3-fach höheren Schichtdicke in z-Richtung, verglichen mit der Auflösung innerhalb der xy-Ebene. Der USCT-Scanner von Techniscan beschallt die Brust jeweils nur einseitig. Durch dieses Vorgehen können Reflexionen, die aus der Ebene herausstrahlen, nicht mehr erfasst werden. Bei beiden Systemen wird bereits eine Brechungs- und Dämpfungskorrektur mithilfe der vorher gewonnenen Transmissionsbilder durchgeführt. Diese werden jedoch ebenfalls nur in 2D durchgeführt.

Damit besitzen beide Systeme keine vollständig dreidimensionale Aufnahme und Rekonstruktion der Ultraschallsignale. Dies macht den 3D-USCT am IPE weltweit einmalig.

SAFT-Rekonstruktionen im Radar- und Sonar-Bereich

Die *synthetic aperture focusing technique* (SAFT) wird neben der medizinischen Bildgebung mit Ultraschall auch in Radar- und Sonaranwendungen genutzt und wird dort als *synthetic aperture sonar* (SAS) bzw. *synthetic aperture radar* (SAR) bezeichnet. Beim Ortungsverfahren mit *radio detection and ranging* (Radar) kommen elektromagnetische Wellen im Bereich von 3 MHz bis zu etwa 300 GHz zum Einsatz [54]. Beim *sound navigation and ranging* (Sonar) mit Ultraschall wird im Bereich von 20 kHz bis ca. 1,6 MHz gearbeitet [55, 56]. Dabei wird von einem Flugzeugradar oder Sonargerät ein Signal seitlich ausgesendet und über die Laufzeit der Reflexionen am (Meeres-)Boden ein Bild berechnet. Daraus ergeben sich 2D-Bilder als Projektion der 3D-Szene.

In [57] wird ein Überblick gegeben, wie sich SAR aus SAS entwickelt hat. Des Weiteren werden die Unterschiede zwischen SAR und SAS sowie die genutzten Rekonstruktionsalgorithmen aufgezeigt. So besteht neben der Rekonstruktion im Zeitbereich, die auch in dieser Arbeit genutzt wird, die Möglichkeit, die Rekonstruktion im Fourrierraum beschleunigt durchzuführen. Jedoch muss dafür insbesondere die Annahme einer konstanten Schallgeschwindigkeit gemacht werden. Die Rekonstruktion im Zeitbereich ist zwar zeitaufwändiger, kann jedoch alle geometrischen Veränderungen und Schwankungen des Sonargerätes berücksichtigen. Ebenso kann eine nichtlineare Schallgeschwindigkeitsverteilung kompensiert werden [57]. Daher wird, trotz des hohen Rechenbedarfs, die Rekonstruktion im Zeitbereich durchgeführt, wenn eine hohe Bildqualität erreicht werden soll. Eine hohe Rechenleistung wird dabei ebenfalls mittels GPUs erreicht. Die NATO Science and Technology Organisation (NATO STO) beschreibt in [58] ein echtzeitfähiges autonomes Sonarsystem, das Rekonstruktionen mit dem SAFT-Algorithmus bei einer Auflösung von 1,5 cm x 2,5 cm durchführt. Mithilfe einer NVIDIA Tesla C1060 GPU beträgt die dabei erreichte Bildrate 4 Hz.

Die Bildrekonstruktion bei Radar und Sonar berechnet eine einzige Schicht, in der die Reflektivität der Bodenoberfläche dargestellt wird. In der 3D-USCT werden dagegen viele Schichten rekonstruiert, wodurch sich ein erhöhter Rechenaufwand ergibt.

Beschleunigung der SAFT für die 3D-USCT

Obwohl die SAFT bereits seit über 30 Jahren als Rekonstruktionsmethode von Reflexionsbildern bekannt ist [59], konnte sie aufgrund ihrer hohen Rechenkomplexität lange Zeit nicht in der medizinischen dreidimensionalen Bildgebung praktisch eingesetzt werden. Durch die stetige Weiterentwicklung von immer schnelleren und parallelen Hardwarearchitekturen rückt dieses Ziel näher und führt dazu, dass die Beschleunigung der medizinischen bildgebenden Algorithmen ein aktuelles Forschungsgebiet darstellt. Von Matthias Birk wurde in [9] untersucht, in wieweit sich die grundlegenden bildgebenden Algorithmen der 3D-USCT durch die Berechnung auf parallelen Hardwarearchitekturen beschleunigen lassen. Dabei wurden, wie in Abbildung 2.16 dargestellt, aktuelle parallele Architekturen aus vier Klassen auf ihre Fähigkeit hin untersucht, die SAFT-Rekonstruktion beschleunigen zu können. Für die einzelnen Hardwareplattformen wurde jeweils von Hand optimierter Code geschrieben. Dies war nötig um die jeweiligen Hardwarefeatures nutzen zu können und zudem unterschiedliche Programmiermodelle benötigt wurden. Für den Vergleich der Rechenleistung wurde als Metrik das Produkt der rekonstruierten Voxel und der dafür verwendeten Druckprofile (A-Scans) pro Zeiteinheit genutzt und in GVA/s angegeben. Aus den Untersuchungen ergab sich die höchste Rechenleistung von 46,4 GVA/s bei einer SAFT-Rekonstruktion mittels einer GTX TITAN GPU, bzw. 21,9 GVA/s auf einer GTX 580. Ein Virtex-7-FPGA (XC7-VX1140T-2) erreichte 31,8 GVA/s, und ein Intel Core i7-2700K Mehrkernprozessor mit vier Kernen 2,2 GVA/s. Für die Vielkern-Prozessoren wurde das Tlera-TILE-Gx8036 untersucht, das nur für eine alternative SAFT-Rekonstruktion zum Einsatz kam. Das Tlera-TILE-Gx8036 erreichte jedoch nur eine Rechengeschwindigkeit unterhalb der CPU und wurde nicht weiter untersucht. Damit erreichte die GPU die höchste Rechenleistung mit SAFT und besitzt zudem eine geringere Entwicklungszeit für neue Algorithmen als ein FPGA, der mit VHDL programmiert wurde. In dieser Arbeit werden daher, von diesen Ergebnissen ausgehend, weiterführende SAFT-Rekonstruktionsalgorithmen für GPUs untersucht.

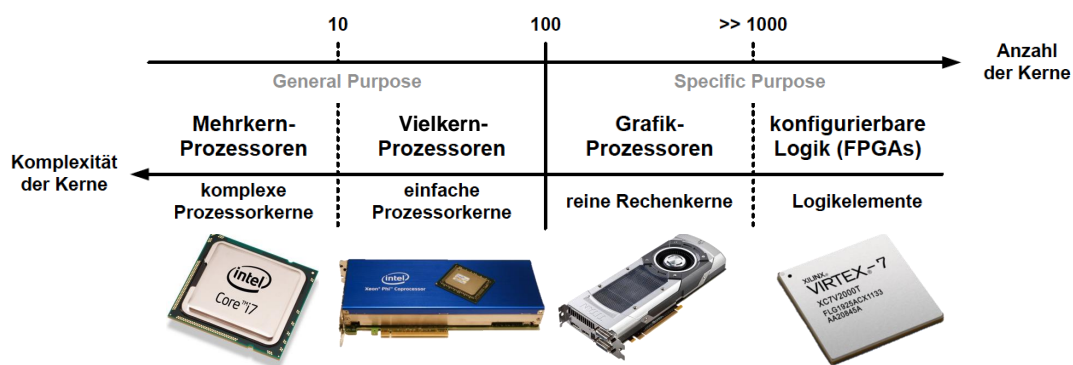


Abbildung 2.16: Klassifikation paralleler Architekturen gemäß ihrer Granularität, d.h. der Anzahl an Rechenkernen gegenüber ihrer Komplexität (nach [9]).

3. Standard-SAFT in der Reflexionsbildgebung des 3D-USCTs

3.1 3D-USCT mit dem Prototypen des IPE

Im Folgenden soll der Aufbau des 3D-USCT-Prototypen am IPE näher erläutert werden. Der 3D-USCT-Prototyp ist in Abbildung 3.1 links dargestellt. Er beinhaltet die Messapertur sowie das Datenerfassungssystem, das 480 parallele Empfangskanäle bereitstellt (siehe Abbildung 3.1 rechts). Abbildung 3.2 links zeigt die optimierte, aus einem Block Polyoxymethylen gefertigte, semi-ellipsoidale Aperturform. Der horizontale Innendurchmesser der Apertur beträgt 26 cm und die vertikale Tiefe 16 cm. Auf der semi-ellipsoidalen Oberfläche der Apertur befinden sich 157 Aussparungen. In diesen werden die sog. *Transducer-Array-Systeme* (TAS) befestigt, die in Abbildung 3.2 rechts dargestellt sind. Des Weiteren stehen Aussparungen für zwei PT100-Präzisionswiderstandsthermometer der Firma Jumo GmbH zur Bestimmung der Wassertemperatur, sowie Ein- und Auslässe für Wasser zum Anschluss der Aufbereitungsanlage bereit. Mit Hilfe von Schrittmotoren und einer Getriebeübersetzung kann die gesamte Apertur stufenlos um bis zu 3 cm angehoben und bis zu 30° rotiert werden. Damit ist es möglich bei mehreren *Aperturpositionen* (AP) Messungen durchzuführen, wodurch sich eine höhere Anzahl an *Sender-Empfänger-Paaren* (SEP) ergibt. Ein kombiniertes Heben und Verdrehen kann innerhalb von 15 Sekunden durchgeführt werden.

In Abbildung 3.3 ist links die Piezogrundfläche der Ultraschallwandler dargestellt. Diese ist in der mittleren Aussparung der inneren TAS-Elektronik angeordnet (Abbildung 3.3 rechts). Die quadratische Piezofläche hat eine Größe von $(5,5 \text{ mm})^2$ und ist in Subelemente unterteilt. Jeweils 2x2 dieser Subelemente werden zusammenge-

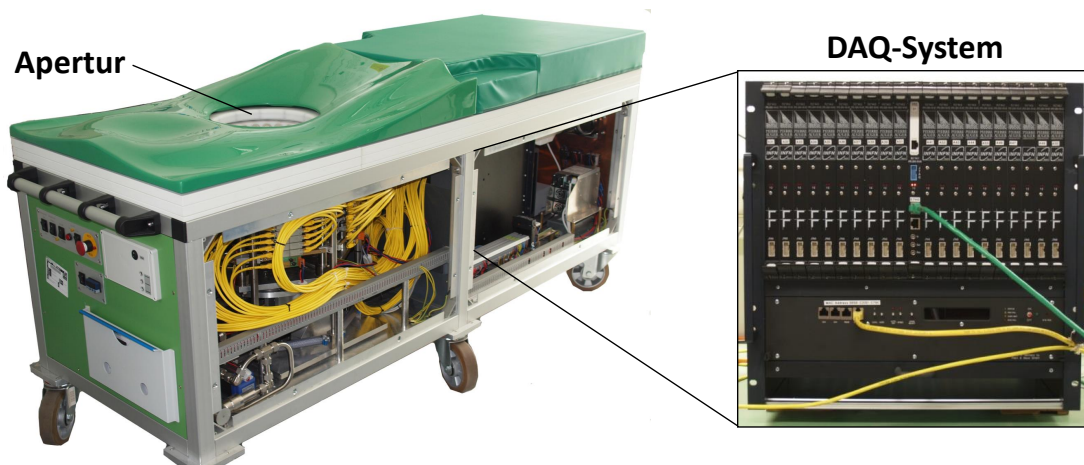


Abbildung 3.1: Der aktuelle 3D-USCT-Prototyp des IPE mit geöffneten Seitenabdeckungen (links) und integriertem Datenerfassungssystem (DAQ-System) (rechts).

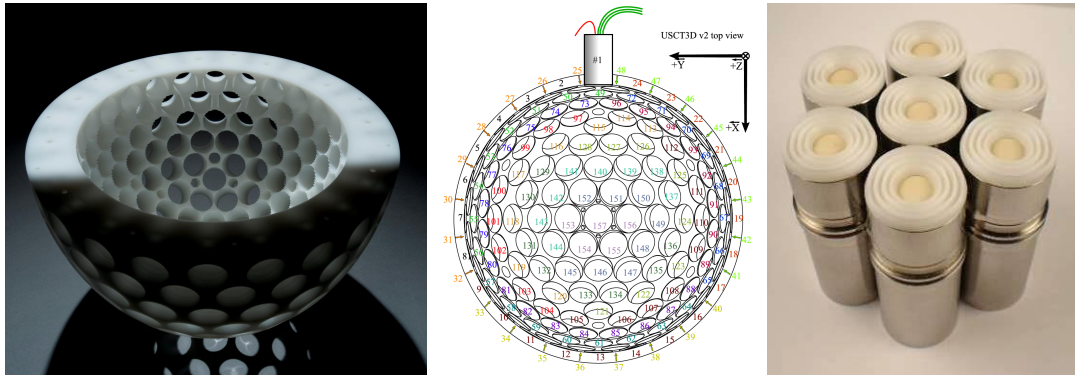


Abbildung 3.2: Apertur des 3D-USCT mit Aussparungen für die TAS (links). Anordnung und Nummerierungsschema der TAS bei Aufsicht (mittig). Transducer Array Systeme (TAS) mit Diffusorkappen (rechts).

schaltet und als vier Sender (rot) bzw. neun Empfänger (blau) betrieben. Damit ergibt sich eine Gesamtanzahl von 628 Sendern, 1413 Empfängern und 887.364 SEP pro AP. Aktuell ist es möglich bis zu 48 AP aufzunehmen, wodurch Signale für maximal 42 Millionen SEP aufgenommen werden können. Die Ultraschallwandler arbeiten bei einer Mittenfrequenz von 2,5 MHz. Die Mittenfrequenz ist dabei ein Kompromiss zwischen Eindringtiefe und Auflösung. Während sich die Auflösung mit kleinerer Wellenlänge bzw. größerer Frequenz verbessert, sinkt dagegen die Eindringtiefe, da die Dämpfung im Gewebe mit der Frequenz ansteigt. Die Bandbreite liegt bei 1,5 MHz und der Öffnungswinkel bei $38,2^\circ$ [60]. Auf den beiden Elektronikplatinen (Abbildung 3.3 rechts) befinden sich die analoge Beschaltung der Piezoelemente und ein Mikrocontroller. Der Mikrocontroller übernimmt dabei die Adressierung der TAS über einen gemeinsamen Bus. Weitere Aufgaben des Mikrocontrollers sind das Auslesen eines Temperatursensors, der sich ebenfalls auf der Platine befindet, sowie die Generierung der Ultraschallpulsform, die von den Sendern ausgesendet wird. Die seitlich angeordneten Platinen der TAS-Elektronik werden bei der Fertigung nach unten geklappt und in zylindrische Edelstahlhülsen eingeführt (siehe Abbildung 3.2 rechts). Abschließend werden die zylindrischen Edelstahlhülsen mittels Polyurethan wasserdicht vergossen und mit einer Diffusorkappe beklebt, um Rückwandreflexionen zu vermeiden.

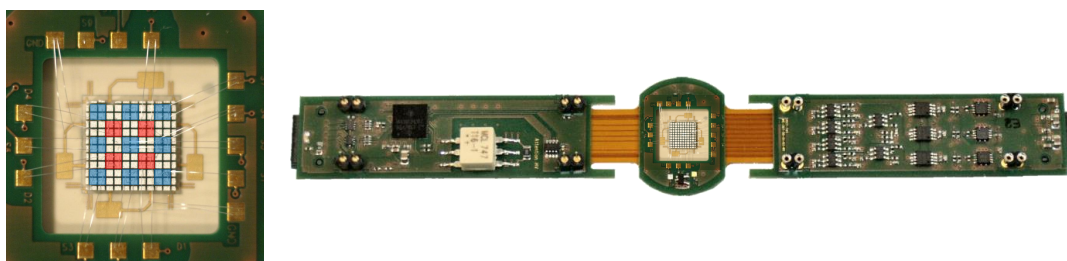


Abbildung 3.3: Abbildung der Piezogrundfläche (links), sowie des inneren Elektronikaufbaus eines TAS (rechts) [61].

Die Datenaufnahme wird mithilfe des *Datenerfassungssystem* (DAQ-System) durchgeführt (siehe Abbildung 3.1 rechts). Darin sind 80 Cyclone II FPGAs verbaut, die 480 parallele Kanäle für die Analog-Digital-Wandlung bereitstellen. Pro Messung, bei der ein Sender ein Ultraschallsignal aussendet, werden simultan für $300 \mu\text{s}$ die Signale von allen Empfängern aufgezeichnet. Ein dabei aufgenommenes Signal stellt das Druckprofil über der Zeit für ein SEP dar und wird *Amplituden-Scan* (A-Scan) genannt. Es werden 12-Bit-ADCs eingesetzt, die mit einer Abtastrate von 20 MHz arbeiten und eine programmierbare Verstärkung besitzen. Mit den 480 Empfangskanälen kann nur ein Drittel der 1413 Empfänger-Signale parallel aufgenommen werden. Daher ist ein 3-fach Multiplex für die Aufnahme nötig.

Sequenz von Aufnahme- und Rekonstruktion

In Abbildung 3.4 sind die grundlegenden Schritte von der Messvorbereitung bis zur Rekonstruktion dargestellt, auf die in diesem Abschnitt näher eingegangen werden soll. Als **Vorbereitung** für die Untersuchung legt sich die Patientin bäuchlings auf eine Liege, die sich auf dem 3D-USCT-Gerät befindet. Über eine Aussparung in der Liege kann eine Brust in die mit Wasser gefüllte Apertur des Geräts eingeführt werden.

Bei der **Signalaufnahme** senden die Ultraschallsender nun nacheinander eine nahezu kugelförmige Wellenfront aus, die durch den Behälter läuft. Beim Durchqueren der Apertur interagiert die Wellenfront mit der Brust. Simultan zum Aussenden der Ultraschallwellenfront wird die Aufnahme der Empfänger gestartet. Die an den Empfängern ankommenden Ultraschallsignale werden als Druckverläufe über der Zeit (A-Scans) vom DAQ-System parallel aufgezeichnet. Sie enthalten somit die Summe aller Interaktionen der Wellenfront mit der Brust. So wird sukzessiv verfahren, bis für alle Sender und alle AP eine Ultraschallwelle emittiert und die Signale vom DAQ-System aufgezeichnet wurden. Die Messung wird jeweils $300 \mu\text{s}$ lang aufgezeichnet. Diese Messdauer ergibt sich aus den geometrischen Abmessungen der

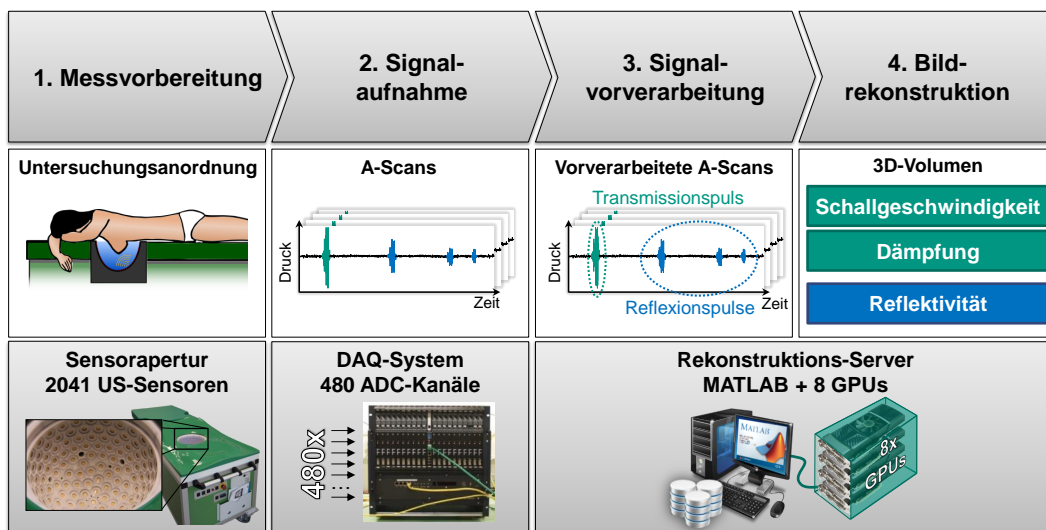


Abbildung 3.4: Darstellung der Sequenz von Aufnahme- und Rekonstruktion.

Apertur und dem möglichen Schallgeschwindigkeitsbereich sowie eines zusätzlichen Sicherheitsintervalls.

Die Signale an den Empfängern werden bei einer Auflösung von 12 Bit mit einer Abtastrate von 20 MHz digitalisiert. Um den Rauscheinfluss zu reduzieren, können die selben SEP mehrfach gemessen werden. Über eine Mittelwertbildung ergibt sich der endgültige A-Scan bei einer n -fachen Messung mit einem um \sqrt{n} -fach verbesserten Signal-zu-Rausch-Verhältnis. Nach der Mittelwertbildung liegen die Datenwerte, auch *Samples* genannt, als 16 Bit-Werte vor. Bevor sie in den insgesamt 86 GB großen Speicher im DAQ-System abgelegt werden, wird noch eine nahezu verlustfreie Datenkompression mithilfe einer dreifach digitalen Unterabtastung durchgeführt [62]. Dieses Vorgehen ist möglich, da durch die Mittenfrequenz von 2,5 MHz und der Bandbreite von 1,5 MHz das Basisband im Bereich von 0 Hz bis 1,75 MHz leer ist. Das bei der Unterabtastung auftretende Aliasing spiegelt das Nutzfrequenzband in das leere Basisband, wobei es zu keiner Überlagerung kommt. Die Datenmenge eines A-Scans beträgt somit $300 \mu\text{s} \cdot 10 \text{ MHz} \cdot 2 \text{ Byte} / 3 = 2000 \text{ Byte}$ und für eine vollständige AP 1,8 GB. Damit ergibt sich die maximale Anzahl der AP zu 47. Die aufgenommenen Daten werden anschließend über eine 100 MBit Netzwerkverbindung an einen Rekonstruktions-Server mit einer Auslesezeit von 47 Minuten übertragen. Alternativ können die Daten bei einer größeren Anzahl an Messungen auch auf einer externen Festplatte zwischengespeichert und abschließend auf den Rekonstruktions-Server übertragen werden. In Tabelle 3.1 werden die Systemparameter der 3D-USCT-Hardware zusammengefasst dargestellt.

Der dritte Schritt ist die **Signalvorverarbeitung**, die auf dem Rekonstruktions-Server nach dem Datentransfer durchgeführt wird. Sie wird direkt auf den A-Scans ausgeführt und hat das Ziel, die Messdaten so zu verarbeiten, dass idealerweise ausschließlich die für die Bildrekonstruktion relevanten Daten enthalten bleiben. Das im Signal enthaltene Rauschen soll dabei so weit wie möglich reduziert werden. Die für die Rekonstruktion der Reflexionsbilder relevanten Daten sind vor allem die genauen Zeitpunkte der Reflexionen, die einen starken Einfluss auf die Rekonstruktion der Reflektivitätsbilder besitzen. In der 3D-USCT werden daher die A-Scan-Daten

Eigenschaft	Wert
Mittenfrequenz f_0	2,5 MHz
Bandbreite f_B	1,5 MHz
Abtastrate f_s	20 MHz
Digitalisierung	12 Bit
Aufnahmedauer pro A-Scan	300 μs
Mittelwertbildung	1-32
Multiplex-Faktor	3
Anzahl TAS	157
Anzahl Sender pro TAS	4
Anzahl Empfänger pro TAS	9
Anzahl Aperturpositionen	1-47
Größe der Rohdaten	$\leq 86 \text{ GB}$
DAQ-Aufnahmezeit 1AP	10 s

Tabelle 3.1: Überblick über genutzte Systemparameter [60].

für die Rekonstruktion mit Hilfe des *Adapted Matched Filters* [63] aufbereitet. Dabei wird ein Matched Filter eingesetzt [64], der das Signal-zu-Rausch-Verhältnis der A-Scans erhöht und die für die Bildrekonstruktion relevanten Informationen der Echosignale hervorhebt. Über eine Faltung mit dem vom Sender ausgesendeten und daher bekannten Anregungspulsform kommt es zu einer Maximumsbildung am Pulsbeginn, wodurch die Ankunftszeit eines Pulses detektiert werden kann. Dafür wird die Einhüllende des Signals gebildet und die lokalen Maxima detektiert. Das resultierende Signal enthält somit nur noch an den Ankunftszeiten der Reflexionspulse einen Dirac-Impuls mit jeweiliger Größe. Da die Pulsform der zurückprojizierten Daten einen starken Einfluss auf die resultierende Punktabbildungsfunktion besitzt, und damit auch auf die Auflösung, wird der von Norton und Linzer in [65, 66] für sphärische Aperturen theoretisch hergeleitete Optimale Puls genutzt (siehe Abbildung 3.5). Der Optimale Puls $p(t)$ erzielt für sphärische Aperturen eine minimale Punktabbildungsfunktion und berechnet sich nach

$$p(t) = [1 - 2(\pi \cdot f_c \cdot t)^2] \cdot e^{-(\pi \cdot f_c \cdot t)^2} . \quad (3.1)$$

Er besitzt vom Pulsmaximum einen scharfen Abfall sowie negative Anteile, die zur Auslöschung der Rekonstruktionsartefakte beitragen. Ein weiteres Ziel der Signalvorverarbeitung ist die Behandlung von Phasenverschiebungen, die aufgrund von Ungenauigkeiten im System auftreten können, wie z.B. in der gemessenen Temperatur, der angenommenen Schallgeschwindigkeit, der Position der Ultraschallwandler, Fehlerquellen oder Rauscheinflüssen. Auch aufgrund von Pulsdeformationen durch materialabhängige Frequenzverschiebungen werden Phasenverschiebungen nie ganz vermeidbar. Ist eine zeitliche Unschärfe σ_t bei der Pulsdetektion oder durch Ungenauigkeiten im System bekannt, kann diese genutzt werden, um die Nulldurchgänge des Optimalen Pulses festzulegen. Diese Skalierung des Optimalen Pulses ermöglicht trotz Unsicherheiten bei der Rekonstruktion eine konstruktive Überlagerung. Die ursprüngliche Mittenfrequenz f_c im Puls kann an die zeitliche Unschärfe des Systems mit $f_c = 1/\sigma_t$ angepasst werden. Die Zeit t lässt sich über die Abtastfrequenz f_s mit $k = t \cdot f_s$ in Samples umrechnen. Damit ergibt sich aus Gleichung (3.1)

$$p_\sigma(t) = \left[1 - 2\left(\pi \cdot \frac{k}{\sigma_t \cdot f_s}\right)^2 \right] \cdot e^{-\left(\pi \cdot \frac{k}{\sigma_t \cdot f_s}\right)^2} . \quad (3.2)$$

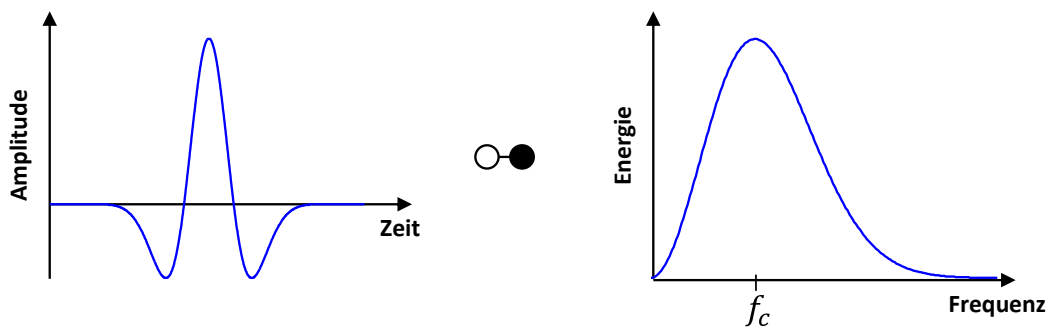


Abbildung 3.5: Optimaler Puls für SAFT in normierter Zeitdarstellung (links) mit entsprechendem Frequenzspektrum (rechts).

Weiterhin kann der Optimale Puls passend zu einer gewählten Auflösung skaliert werden. In Abbildung 3.6 links sind zwei Punktstreuer (\vec{x}_1 und \vec{x}_2) dargestellt, die in einem Abstand Δd zueinander liegen. Um diese beiden Streuer unterscheiden zu können, wird nach dem Abtasttheorem eine Auflösung von mindestens $res = \Delta d/2$ benötigt. Durch Umformulieren des Abstands Δd über die Laufwege des Ultraschalls vom Sender über die Punktstreuer zum Empfänger bei einer konstanten Schallgeschwindigkeit c ergibt sich

$$\begin{aligned} res &= \frac{\Delta d}{2} = \frac{d_2 - d_1}{2} \\ &= \frac{t_2 \cdot c - t_1 \cdot c}{2} \end{aligned} \quad (3.3)$$

$$= \frac{c}{4}(t_2 - t_1) = \frac{c}{4} \cdot \tau \quad (3.4)$$

Wird τ durch die im Signal vorkommende maximale Frequenz f_{max} ausgedrückt, ergibt sich $res = \frac{c}{4 \cdot f_{max}}$. Nimmt man aufgrund des nahezu symmetrischen Spektrums in Abbildung 3.5 links an, dass $f_{max} \approx 2 \cdot f_c$ gilt, erhält man für den Optimalen Puls in Gleichung (3.1) folgenden Zusammenhang

$$p(t) = \left[1 - 2\left(\pi \cdot \frac{c}{8 \cdot res} \cdot t\right)^2 \right] \cdot e^{-\left(\pi \cdot \frac{c}{8 \cdot res} \cdot t\right)^2} . \quad (3.5)$$

In Abbildung 3.6 rechts sind beispielhaft zwei Optimale Pulse bei verschiedenen Auflösungen res dargestellt. Da der Schwerpunkt dieser Arbeit auf der Rekonstruktion der Reflexionsbilder mit SAFT liegt, wird auf die Signalvorverarbeitung nicht weiter eingegangen. Eine detailliertere Beschreibung der Signalvorverarbeitung sowie deren Implementierung ist in [9, 63] gegeben.

Sind die A-Scans vorverarbeitet, können sie zur **Bild-Rekonstruktion** im vierten Schritt genutzt werden. Dieser Schritt wird ebenfalls auf dem Rekonstruktions-Server durchgeführt. In Abbildung 3.4 sind in den A-Scans zwei verschiedene Signaltypen markiert. Der Transmissionspuls (grün) ist der zuerst eintreffende Puls und hat folglich den schnellsten Weg vom Sender zum Empfänger genommen. Alle weiteren später eintreffenden Pulse (blau) werden Reflexionspulse genannt. Sie haben einen

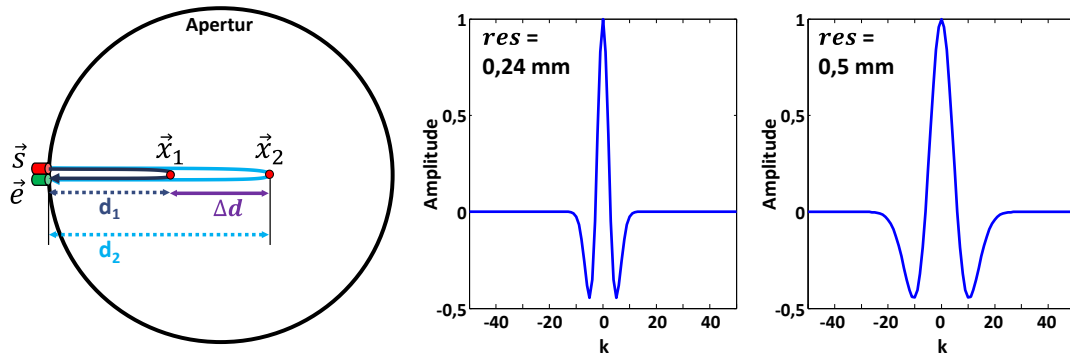


Abbildung 3.6: Berechnung der zur Auflösung passenden Breite des Optimalen Pulses (links). Optimaler Puls mit unterschiedlichen Breiten zur entsprechenden Auflösung res (rechts).

längeren Laufweg und wurden durch Reflexionen an Streuern verursacht. Diese Unterscheidung führt auch zu den beiden wesentlichen Bildgebungsverfahren, die in der 3D-USCT Verwendung finden: Die Transmissionstomographie und die Reflexionstomographie.

Mittels der Transmissionstomographie ist es möglich, quantitative Informationen über das enthaltene Gewebe zu gewinnen. Anhand des bekannten direkten Laufwegs zwischen Sender und Empfänger sowie der Signallaufzeit kann auf die Schallgeschwindigkeit des durchlaufenen Mediums geschlossen werden. Über die Amplitudenverminderung von Sendepuls zum empfangenem Puls kann zudem die Dämpfung des Mediums ermittelt werden. Diese Informationen über das Medium können genutzt werden, um zusätzlich die Qualität der Reflektivitätsbilder zu verbessern (siehe Kapitel 6). Insbesondere die Schallgeschwindigkeitsverteilung ist zur Berechnung genauer Zeitpunkte der Reflexionen ein wichtiger Faktor bei der Reflexionstomographie. Während die Transmissionstomographie quantitative Informationen über das enthaltene Gewebes liefern kann, handelt es sich bei der Reflexionstomographie um qualitative Informationen. Mit ihr kann auf die Intensität der Reflektivität der inneren Strukturen der Brust geschlossen werden, indem die Reflexion des Ultraschalls beim Übergang zwischen verschiedenen Gewebearten abgebildet wird. Auf den verwendeten SAFT-Algorithmen der Reflexionstomographie wird in Kapitel 3.3 näher eingegangen.

3.2 Reflexionen bei Ultraschall

Eine Reflexion bezeichnet in der Physik das Zurückwerfen von Wellen an einer Grenzfläche. Trifft eine ausgesandte Ultraschallwelle auf eine Grenzfläche, an der sich der akustische Wellenwiderstand (Schallkennimpedanz Z) sprunghaft ändert, wird sie teilweise oder auch total reflektiert. Dabei gelten ähnliche Gesetze wie in der Optik. Beim konventionellen Ultraschall, bei dem das Impuls-Echo-Verfahren genutzt wird, sowie bei der Reflexionstomographie werden die Reflexionen an Gewebegrenzen aufgezeichnet und daher die Unterschiede der Schallkennimpedanz Z abgebildet.

Die Schallkennimpedanz Z lässt sich nach

$$Z = c \cdot \rho \quad [kg/m^2s] \quad (3.6)$$

aus den Materialparametern der Schallgeschwindigkeit im Gewebe c in m/s und der Dichte ρ in kg/m^3 berechnen [26]. Trifft der Ultraschall unter einem Einfallswinkel θ_e auf eine glatte Oberfläche, wird er unter demselben Winkel $\theta_r = \theta_e$ reflektiert (siehe Abbildung 3.7). Der Reflexionskoeffizient R ist dabei als Verhältnis der Amplitude der reflektierten Welle A_r zur Amplitude der einfallenden Welle A_e definiert. Er berechnet sich für glatte Oberflächen nach [26] abhängig vom Einfallswinkel θ_e und Transmissionswinkel θ_t nach

$$R = \frac{A_r}{A_e} = \frac{Z_2 \cos \theta_e - Z_1 \cos \theta_t}{Z_2 \cos \theta_e + Z_1 \cos \theta_t} \quad (3.7)$$

Der Transmissionswinkel θ_t kann in strahlengeometrischer Näherung nach dem Brechungsgesetz von Snellius

$$\frac{\sin \theta_t}{c_2} = \frac{\sin \theta_e}{c_1} \quad (3.8)$$

berechnet werden. Analog zum Reflexionskoeffizienten R ist ein Transmissionskoeffizient T definiert, wobei gilt $R + T = 1$.

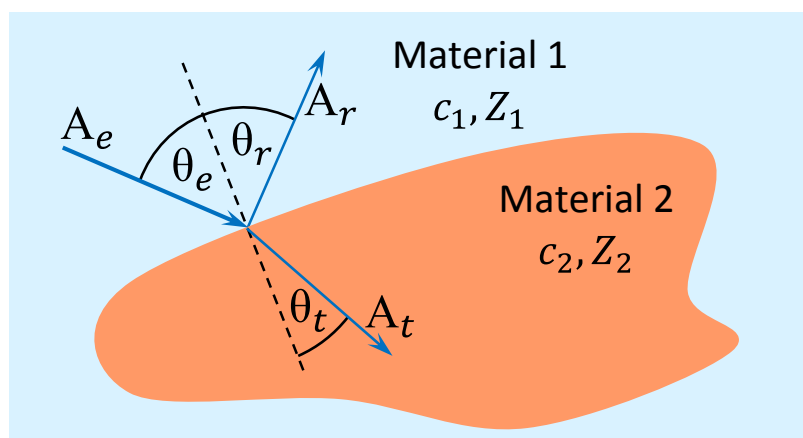


Abbildung 3.7: Schematische Darstellung von Reflexion und Transmission von Schallwellen an einer Grenzfläche nach [67].

Reflexionsarten

Die Art der Reflexion hängt von der Beschaffenheit der Oberfläche bzw. der Grenzfläche ab. Im Allgemeinen lassen sich Reflexionen, wie aus der Computergrafik bekannt, aus drei Komponenten zusammensetzen [68]: Der ambienten, diffusen und spiegelnden Komponente.

Die **ambiente** Komponente ergibt sich aus der Umgebungsbeleuchtung. Es wird angenommen, dass die Reflektivität unabhängig von Lichtquelle und Betrachtungswinkel ist und gleichmäßig auftritt. Sie wird daher allein über eine Materialkonstante modelliert. In der 3D-USCT wird der Ultraschall ausschließlich von Sendern erzeugt, die eine direkte Schallquelle darstellen. Es tritt daher keine ambiente Komponente auf und wird in dieser Arbeit nicht betrachtet.

Die **diffuse** Komponente beschreibt eine in alle Richtungen gleichmäßige Streuung und stellt ein Maß für die Mattheit bzw. Rauheit einer Oberfläche dar. Der an der Oberfläche reflektierte Ultraschall besitzt eine konstante Amplitude, die nur noch von dem Einfallswinkel θ_e abhängig ist. Umso kleiner der Einfallswinkel zur Oberflächennormale \vec{N} ist, desto höher ist die empfangene Amplitude von der Oberfläche (siehe Abbildung 3.8 mitte und rechts). Dieser Zusammenhang wird gewöhnlich durch das Lambertsche Kosinusetz [69]

$$A_{diffus} = A_e \cdot \cos \theta = A_e \cdot (\vec{S} \cdot \vec{N}) \quad (3.9)$$

beschrieben, wobei A_e die Amplitude des einfallenden Ultraschalls ist und θ der Winkel zwischen Sendervektor \vec{S} und der Oberflächennormalen \vec{N} ist.

Bei der **spiegelnden** bzw. Glanzlichtkomponente handelt es sich um eine Reflexion an einer glatten Grenzfläche. Die bisher in Gleichung (3.7) betrachtete Art der Reflexion stellt eine ideal gerichtete Spiegelstreuung dar (siehe Abbildung 3.9 links). Der ausfallenden Winkel der Reflexion entspricht dem einfallenden Winkel $\theta_r = \theta_e = \theta$. Bei einer idealen Spiegelreflexion gilt, dass die Amplitude des ausgehenden Ultraschalls $A_{spiegelnd}$ gleich der Amplitude des einfallenden Ultraschalls A_e ist. Zu einer ideal spiegelnden Reflexion kommt es jedoch selten. In der Realität weist die gespiegelte Reflexion normalerweise einen gewissen Streuanteil auf, der von der Beschaffenheit der Oberfläche abhängt (siehe Abbildung 3.9 mittig und rechts).

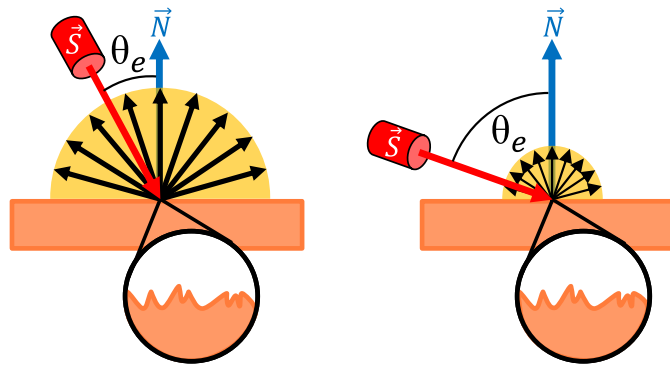


Abbildung 3.8: Schematische Darstellung einer diffusen Reflexion bei kleinem Einfallswinkel (links) und großem Einfallswinkel (rechts). Die Oberflächenbeschaffenheit ist vergrößert dargestellt.

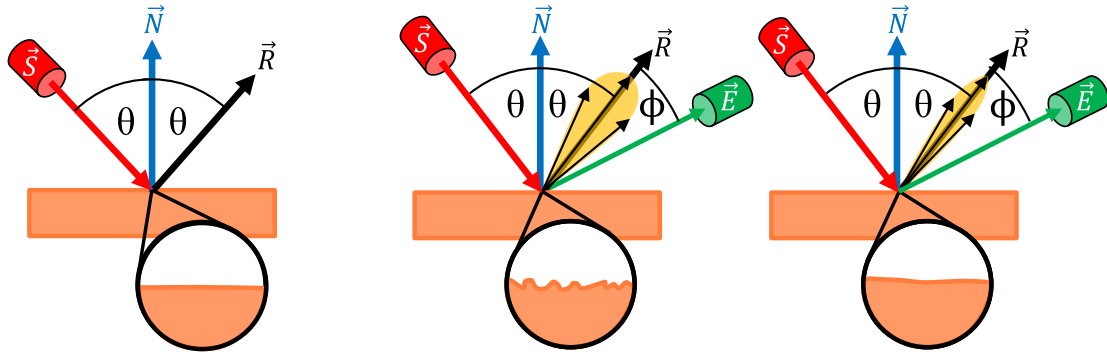


Abbildung 3.9: Schematische Darstellung von spiegelnden Reflexionen. Ideal (links), für einen kleinen Parameter n (mittig) und einem großen Parameter n (rechts). Die Oberflächenbeschaffenheit ist vergrößert dargestellt.

Ein häufig gebrauchtes Modell zur Beschreibung der spiegelnden Reflexion ist das Phong-Modell. Dabei wird die Amplitude der spiegelnden Reflexion folgendermaßen berechnet

$$A_{\text{spiegelnd}} = A_e \cdot \cos^n \phi = A_e \cdot \left(\vec{R} \cdot \vec{E} \right)^n, \quad (3.10)$$

wobei ϕ der Winkel zwischen dem Vektor \vec{E} des Beobachters und der Hauptreflexionsrichtung \vec{R} ist. Über den materialabhängigen Parameter n kann die Streubreite der Reflexion definiert werden.

Die Reflexionsart ist zudem abhängig von der Größe des Objekts, das den Ultraschall reflektiert. Ist die Größe des Streuers dabei sehr viel kleiner als die Wellenlänge, wird auch von **Punktstreuern** gesprochen (siehe Abbildung 3.10). Der Rekonstruktion mit SAFT liegt die Annahme zugrunde, dass alle Streuer als ideale infinitesimal kleine Punktstreuer wirken. Sie strahlen, unabhängig von der Richtung aus der eine Ultraschallwelle eintrifft, eine neue Kugelwelle ab. Ein Punktstreuer wird nach dem Huygens-Prinzip somit als Quelle einer neuen Elementarwelle betrachtet [70]. Zum Vergleich ist die Amplitudenverteilung der drei Reflexionsarten über den Winkel θ bei $\vec{E} = \vec{N}$ und $n = 10$ in Abbildung 3.11 dargestellt. Dabei ist die Amplitude jeweils auf die maximale Amplitude normiert. Es ist erkennbar, dass die spiegelnde Streuung einen deutlich schmäleren Winkelbereich besitzt, in dem die Reflexion detektierbar ist. Ein Punktstreuer reflektiert dagegen in alle Richtungen mit derselben Amplitude.

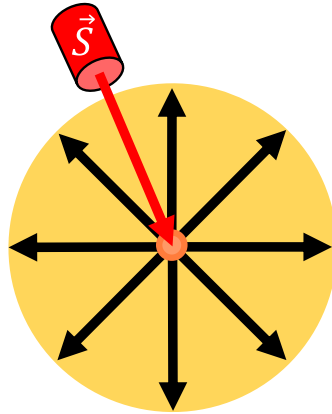


Abbildung 3.10: Schematische Darstellung der Streuung an einem Punktstreuer.

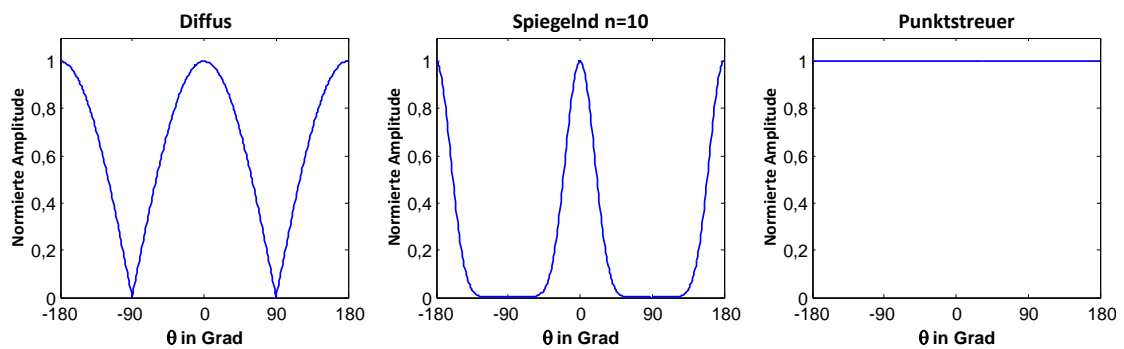


Abbildung 3.11: Amplitudenverteilung der drei Reflexionsarten über den Winkel θ bei $\vec{E} = \vec{N}$. Eine diffuse Reflexion (links), eine spiegelnde Reflexion (mittig) und eine Reflexion an einem idealen Punktstreuer (rechts).

3.3 SAFT in der 3D-USCT

In den folgenden Abschnitten wird zunächst die Funktionsweise der SAFT beschrieben und die Implementierung des Standard-Algorithmus vorgestellt.

Die Aufgabe einer Rekonstruktion ist es, aus den Messdaten wieder ein möglichst wirklichkeitsgetreues Abbild der abgebildeten Eigenschaften des Objekts zu erstellen. Dabei gehört die Rekonstruktion eines Objektes aus den (verteilt) gemessenen Echosignalen zu der Kategorie der inversen Problemen, die nicht im Allgemeinen lösbar sind. Für den praktischen Einsatz werden Verfahren benötigt, die eine Lösung in endlicher Zeit liefern können. Für die exakte Rekonstruktion eines Reflexionsvolumens bzw. der lokalen Impedanzunterschiede ist die Lösung der vollen Wellengleichung nötig. Diese Lösung benötigt jedoch einen enormen Rechenaufwand, der aktuell selbst mit großen Rechenclustern nicht zu bewältigen ist [71]. Stattdessen kommt für die 3D-USCT mit SAFT ein vereinfachter Bildrekonstruktionsalgorithmus zur Anwendung. Er hat seinen Ursprung im *nondestructive testing* (NDT), dem zerstörungsfreien Testen mittels Ultraschall [8]. Von Norton und Linzer [72, 59] wurde die SAFT in den Bereich der medizinischen Bildgebung übertragen. Die grundlegende Idee der SAFT beruht auf der Summation vieler einzelner Aufnahmen, die aus verschiedenen geometrischen Positionen erzeugt wurden, um daraus ein Bild mit höherer Qualität zu berechnen. Da der 3D-USCT-Prototyp das zu messende Objekt von nahezu allen Seiten umgibt, kann daraus ein 3D-Volumen rekonstruiert werden. Nach Norton und Linzer [72, 59] werden für dieses Verfahren folgende vier Annahmen getroffen:

1. Das Objekt ist schwach reflektierend. Die Energie in der sich im Medium ausbreitenden Ultraschallwelle ist um ein Vielfaches größer als die gesamte Energie, die auf dem Weg zum Empfänger zurück gestreut wird. Reflexionen zweiter Ordnung können vernachlässigt werden (Bornsche Näherung).
2. Die Absorption des Mediums ist im Wesentlichen gleichmäßig. Die auftretende Dämpfung kann systematisch kompensiert werden, indem das Empfängersignal entsprechend verstärkt wird (TGC). Durch Streuungen im Medium auftretende Dämpfungseffekte werden vernachlässigt.
3. Die Schallgeschwindigkeit ist im Fokusbereich, der engl. auch *region of interest* (ROI) genannt wird, im Wesentlichen konstant. Es wird somit angenommen, dass keine Fehler aufgrund von Schallgeschwindigkeitsänderungen im Medium auftreten.
4. Objekte können sich dabei aus einer Ansammlung von isotropen Punktstreuern zusammensetzen, die als eine Quelle von Huygens-Elementarwellen agieren. Die Reflektivität des Mediums, die die relative Amplitude des von einem Punktstreuer zurückgestreuten Signals angibt, kann somit als unabhängig vom Einfallswinkel betrachtet werden.

Für die **Signalaufnahme** sendet nacheinander jeder Sender ein unfokussiertes Ultraschallsignal aus, während gleichzeitig die Signale von allen Empfängern vom DAQ-System für $300 \mu\text{s}$ aufgezeichnet werden. Die analogen Signale werden dabei mit einer Abtastfrequenz von 10 MHz digitalisiert. Abbildung 3.12 zeigt das Vorgehen der Signalaufnahme im Fall der 3D-USCT für ein *Sender-Empfänger-Paar* (SEP) (\vec{s}, \vec{e}) und zwei Punktstreuern an beliebigen Positionen \vec{x}_1, \vec{x}_2 . Das dabei aufgenommene Signal $p(t)$ am Empfänger stellt das Druckprofil über der Zeit für ein

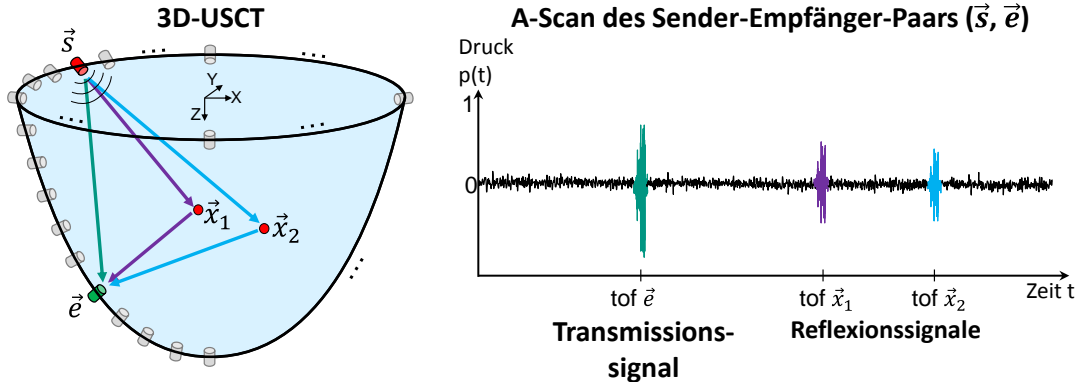


Abbildung 3.12: Schematische Zeichnung des Messprinzips für einen A-Scan mit einem SEP.

SEP dar und wird auch *Amplituden-Scan* (A-Scan) genannt. Da für viele SEP jeweils ein A-Scan aufgenommen wird, wird die Zuordnung der Signalamplituden in dieser Arbeit über

$$AScan(s, e, t) \quad (3.11)$$

beschrieben. Dabei stellen s und e die jeweiligen Indizes der genutzten Sender und Empfänger dar, die sich an den Positionen (\vec{s} und \vec{e}) befinden. Mit t wird der Zeitpunkt innerhalb eines A-Scans angegeben.

Das an einem Sender an der Position \vec{s} unfokussiert ausgesandte Ultraschallsignal durchläuft das Medium Wasser und wird an den idealen Punktstreuern bei \vec{x}_1 und \vec{x}_2 in alle Richtungen reflektiert. An der Position \vec{e} wird das Signal vom Empfänger registriert. Es setzt sich aus folgenden drei Signalen zusammen: Das Signal, das den schnellsten Weg von \vec{s} nach \vec{e} nimmt, wird *Transmissions-Signal* genannt. Weiterhin werden zwei Echos registriert, die durch die Reflexion der Ultraschallwelle an \vec{x}_1 und \vec{x}_2 entstehen. Der genaue Zeitpunkt, an dem ein ausgesendetes Signal am Empfänger ankommt, wird *time of flight* (TOF) genannt. Dieser Ankunftszeitpunkt wird durch die Länge des Laufwegs auf dem entsprechenden Pfad $P(\vec{s}, \vec{x}, \vec{e})$ vom Sender \vec{s} über den Streuer bei \vec{x} zum Empfänger \vec{e} und der mittleren Schallgeschwindigkeit \bar{c}_P bestimmt. Gleichung (3.12) beschreibt den Zusammenhang zwischen der Position eines Punktstreuers \vec{x} und dem Zeitpunkt tof im A-Scan, bei dem das Echo registriert wird.

$$tof_P(\vec{s}, \vec{x}, \vec{e}) = \frac{\|\vec{s} - \vec{x}\| + \|\vec{x} - \vec{e}\|}{\bar{c}_P} \quad (3.12)$$

Der Index P zeigt die Abhängigkeit von dem jeweils genutzten Pfad $P(\vec{s}, \vec{x}, \vec{e})$ an. Die beiden Abstandsterme im Zähler berechnen sich mittels der euklidischen Norm:

$$\|\vec{a} - \vec{b}\| = \sqrt{(a_x - b_x)^2 + (a_y - b_y)^2 + (a_z - b_z)^2}. \quad (3.13)$$

Die Amplitude der empfangenen Echosignale ist von der Dämpfung D_P abhängig, die der Ultraschallpuls auf dem durchlaufenen Pfad erfahren hat. Durchläuft der Ultraschallpuls, der mit einer Amplitude $A_0 = 1$ ausgesendet wurde, einen Weg der Länge l , sinkt die Amplitude des Echos unter der Annahme einer Dämpfung bei einem konstanten Dämpfungskoeffizienten α (2. Annahme der SAFT):

$$AScan(s, e, tof_P) = A_0 \cdot D_P = A_0 \cdot e^{-\alpha \cdot l} \quad (3.14)$$

Das Prinzip der **Bildrekonstruktion** mit SAFT ist für den vereinfachten Fall von zwei A-Scans und zwei Punktstreuern in Abbildung 3.13 dargestellt. Hier für den Fall in dem Sender, Empfänger und Punktstreuer in einer Ebene liegen. Die zentrale Idee der Rekonstruktion mit SAFT [8] ist es, die Echosignale, die aus einer Vielzahl von geometrisch verschiedenen SEP-Positionen aufgenommen wurden, nach Gleichung (3.12) in das zu rekonstruierende Volumen zurück zu projizieren und aufzusummieren. Für die Rekonstruktion wird das Volumen der Größe (X,Y,Z) zunächst in ein diskretes Voxelfeld aufgeteilt. Für jeden Voxel \vec{x} wird die Reflektivität

$$I(\vec{x}) = \sum_{\forall(s,e)} AScan(s, e, tof_P) \quad (3.15)$$

berechnet, indem für jedes SEP der Zeitpunkt $tof_P(\vec{s}, \vec{x}, \vec{e})$ im A-Scan berechnet, die entsprechende Amplitude $AScan(s, e, tof_P)$ geladen und aufaddiert wird. Dieses Vorgehen wird für alle Voxel \vec{x} im zu rekonstruierenden Volumen der Größe (X,Y,Z) durchgeführt. Die Rekonstruktion des Standard-SAFT wird als Pseudo-Code in Algorithmus 3.1 beschrieben.

Bei der Rückprojektion ist zu beachten, dass ein Echosignal innerhalb eines A-Scans nicht mehr eindeutig nur einer Position zugeordnet werden kann. Es wird durch ei-

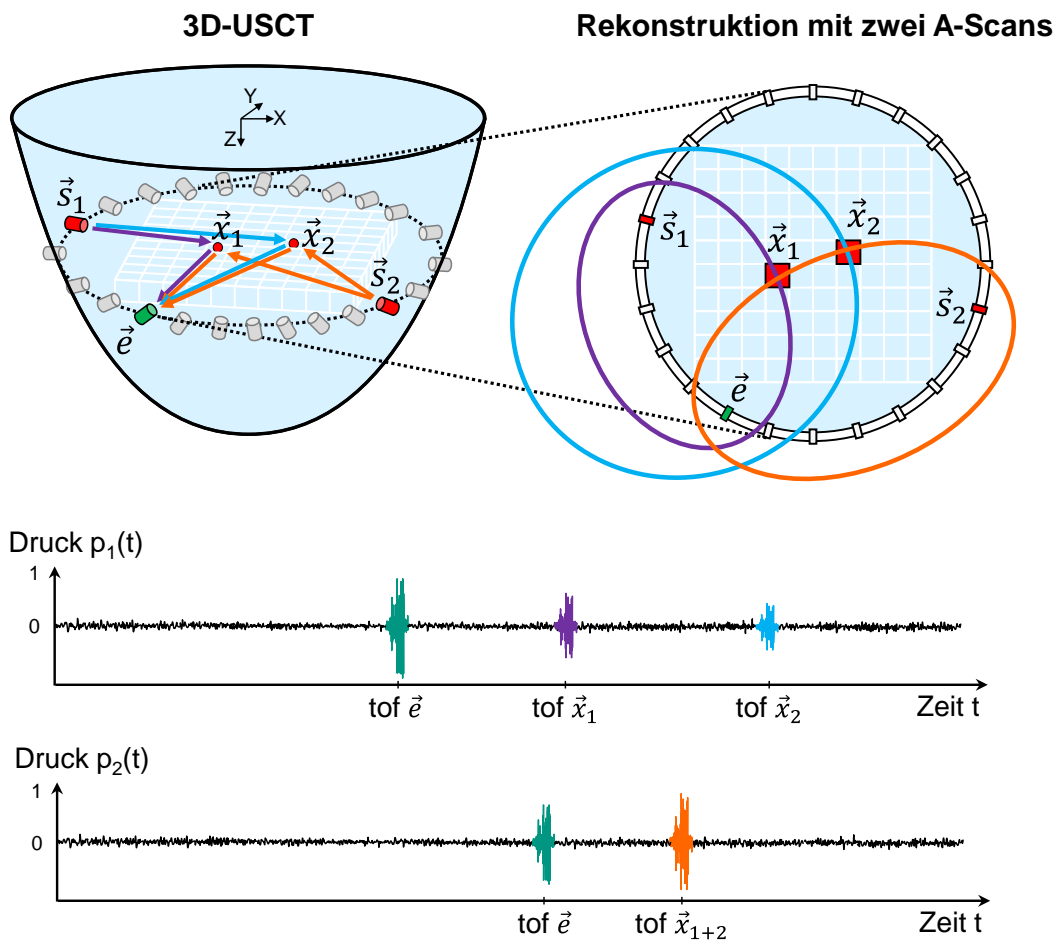


Abbildung 3.13: Schematische Zeichnung der Rekonstruktion mit SAFT für eine xy-Schicht (oben rechts) mit zwei A-Scans (unten).

Algorithmus 3.1 : Rekonstruktion mit SAFT

```

1: forall the Voxel( $X, Y, Z$ ) do
2:   Bestimme Voxel-Position  $\vec{x}$  in [m]
3:   forall the A-Scans( $e, r$ ) do
4:     Lade Sender-Position  $\vec{s}$  und Empfänger-Position  $\vec{e}$  in [m]
5:     Bestimme Abstand  $d = \|\vec{s} - \vec{x}\| + \|\vec{x} - \vec{e}\|$ 
6:     Bestimme Laufzeit  $tof = d/\bar{c}$  in [s]
7:     Lese Amplitude  $A = AScan(s, e, tof_P)$ 
8:     Akkumuliere  $Volumen(x, y, z) += A$ 
9:   end
10: end

```

ne Ellipsoidenoberfläche im Volumen abgebildet, mit den beiden Fokalfpunkten an den Positionen der dazugehörigen Ultraschallwandler \vec{s} und \vec{e} (siehe Abbildung 3.13 rechts oben). Alle Voxel, die sich auf der Oberfläche des Ellipsoids befinden, haben dabei dieselbe Laufzeit (tof_P). Für Voxel wie \vec{x}_1 und \vec{x}_2 an Positionen eines tatsächlichen Punktstreuers kommt es zu einer konstruktiven Überlagerung der Ellipsen. $I(\vec{x})$ wird mit steigender Anzahl genutzter A-Scans erhöht. Dagegen kommt es für Voxel, die keine Punktstreuer enthalten, ausschließlich zu einer Erhöhung, wenn sie für bestimmte SEP dieselbe Laufzeit (tof_P) wie ein wahrer Punktstreuer besitzen. Durch diese fehlerhafte Erhöhung entsteht der SAFT-typische Rekonstruktions-Artefakt (siehe Abbildung 2.7). Je mehr A-Scans für die Rekonstruktion genutzt werden, umso mehr verliert der fehlerhafte Beitrag in der Summe an Bedeutung und kann vernachlässigt werden. Entscheidend für die Bildqualität ist daher eine genügend große Anzahl an A-Scans bei einer entsprechend kleiner Dynamik des Objekts, damit diese Artefakte unterdrückt werden [73].

Es kann jedoch für die Reflexionstomographie auch eine Beschränkung auf eine Teilmenge der A-Scans sinnvoll durchgeführt werden. In [38] wird gezeigt, dass die Rückprojektion der Ellipsen bei SAFT nichtlinear bezüglich der Ausdehnung der einzelnen Ellipsen stattfindet. Die Ausdehnung hängt dabei stark von dem Abstand zwischen Sender und Empfänger ab und vergrößert sich, je weiter Sender und Empfänger voneinander entfernt sind. Die größte Entfernung ist für den Transmissionsfall gegeben, bei der sich Sender und Empfänger gegenüberliegen. Für die Auflösung in der Reflexionstomographie sind daher die SEP am geeignetsten, bei denen Sender und Empfänger nah beieinander positioniert sind. Für eine einfache Bestimmung der Teilmenge werden nur die A-Scans von SEP genutzt, bei denen Sender und Empfänger in einem gewissen Winkelbereich θ zueinander stehen. Der Winkelbereich wird dabei über den Winkel zwischen Sendernormalen \vec{N} und dem Vektor \vec{SE} definiert. In Abbildung 3.14 ist beispielhaft der Winkelbereich von 45° - 90° dargestellt. Diese Beschränkung auf eine Teilmenge der A-Scans wirkt sich zudem positiv auf die Rekonstruktionszeit aus, da weniger A-Scans bei der Rekonstruktion genutzt werden.

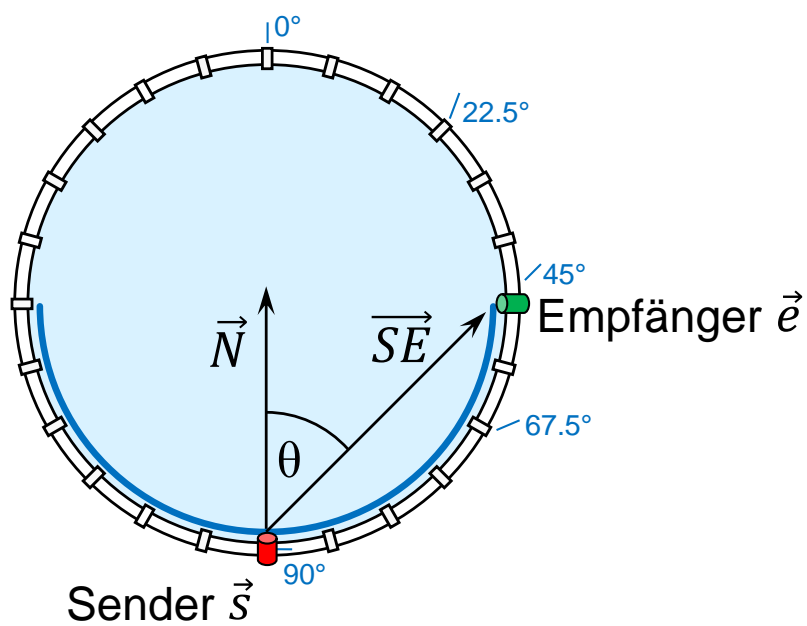


Abbildung 3.14: Winkelbereich zwischen Sender und Empfänger vereinfacht in 2D für die Teilmenge der genutzten SEP dargestellt.

4. GPU-basierte Beschleunigung der Standard-SAFT

4.1 MATLAB als Entwicklungsumgebung

MATLAB ist eine in der Wissenschaft weit verbreitete Entwicklungs-Software. In ihrer benutzerfreundlichen Umgebung stellt MATLAB eine höhere Programmiersprache zur Verfügung, die eine schnelle Entwicklung von Algorithmen ermöglicht. Weitere Stärken sind die vielfältigen effektiven Funktionsbibliotheken und ausgereiften graphischen Darstellungsmöglichkeiten. Durch die Repräsentation der Daten in Matrix-Vektorform eignet sich MATLAB zudem hervorragend für ein beschleunigtes Rechnen auf paralleler Hardware, wie z.B. auf GPUs. Aus diesen Gründen wird die MATLAB-Programmierungsumgebung auch für die Signalverarbeitung und die Bildrekonstruktion in der 3D-USCT eingesetzt.

Da MATLAB besonders in der Wissenschaft gerne genutzt wird, wurden viele Untersuchungen durchgeführt, um herauszufinden wie rechenintensive Aufgaben effizient berechnet werden können. In [74] geben Ubaidullah et al. einen guten Überblick über die Beschleunigung von rechenintensiven Anwendungen in MATLAB. Sie zeigen auf, dass die Nachteile von MATLAB bei dem langsamen Interpreter liegen, der den MATLAB-Code ausführt. Dieser zeigt insbesondere Schwächen beim Ausführen von Schleifen und besitzt zudem hohe Latenz bei Speicherzugriffen. Ein gemischtes Programmierkonzept wird vorgeschlagen, um die Vorteile von MATLAB mit der schnellen Ausführung anderer Sprachen wie C oder FORTRAN oder CUDA C zu kombinieren. MATLAB kann somit als flexible Entwicklungsplattform weiter genutzt werden. Die Ausführung der zeitaufwendigsten Teile des Programms wird dagegen direkt in C ausgeführt, um die höchste Beschleunigung zu erreichen. MATLAB selbst bietet eigene Möglichkeiten an, um GPUs direkt im m-Code zu programmieren oder den schon kompilierten CUDA C Kernel auszuführen. Die Programmierung mit m-Code ist zwar sehr einfach zu nutzen, erreicht jedoch die geringste Beschleunigung. Dieses gemischte Programmierkonzept für MATLAB wurde in [75] und [76] evaluiert und zeigte eine starke Beschleunigung. Die Arbeit in [77] untersucht die Beschleunigung einer 2D Wavelet-basierten Komprimierung für die medizinische Bildgebung. Die Beschleunigung erreichte dabei nur einen sehr geringen Faktor von 1,7. Zum einen wurde der gesamte Code in CUDA C geschrieben, zum anderen wurde kein geeignetes Konzept für den Datentransfer zwischen CPU und GPU entwickelt. Dies führte zu einer Vielzahl an Datentransfers und einer verlangsamten Berechnung.

Um die Vorteile von MATLAB sowie die Beschleunigung mittels paralleler GPU Hardware zu kombinieren, soll daher das gemischte Programmierkonzept verwendet werden. Dabei wird die rechenintensive Rekonstruktion mittels SAFT auf die Hardware ausgelagert und dort parallel beschleunigt berechnet. Bei diesem Konzept muss allerdings dem Datentransfer zwischen CPU und GPU besondere Beachtung geschenkt, sowie eine effiziente Schnittstelle entwickelt werden.

Integrationsmöglichkeiten in MATLAB

Um ein CUDA-C-Programm in den MATLAB-Ablauf einzubinden, sind drei Elemente notwendig:

1. Eine MEX-Funktion (MATLAB executable [78]), die die Schnittstelle zwischen MATLAB und C darstellt und es erlaubt, beliebigen C-Code auszuführen.
2. Eine CUDA-C-Implementierung des SAFT-Algorithmus für die GPU, auch CUDA-Kernel genannt. Diese wird aus der MEX-Funktion heraus aufgerufen.
3. Ein geeignetes und effizientes Konzept, um die Daten zwischen MATLAB und der GPU schnell zu transferieren.

Diese sollen im Folgenden am Beispiel des SAFT-Algorithmus näher erläutert werden. Dabei wurde die Schnittstelle zu MATLAB soweit optimiert, dass der Overhead des SAFT-Aufrufs vernachlässigt werden kann.

4.2 Integration in MATLAB

4.2.1 Design einer effizienten Schnittstelle

Das bisherige Bildrekonstruktionsprogramm für die 3D-USCT ist in Abbildung 4.1 (links) dargestellt. Jeder aufgenommene A-Scan wird sequentiell geladen, durchläuft die Vorverarbeitung und wird am Ende für die Bildrekonstruktion mit SAFT

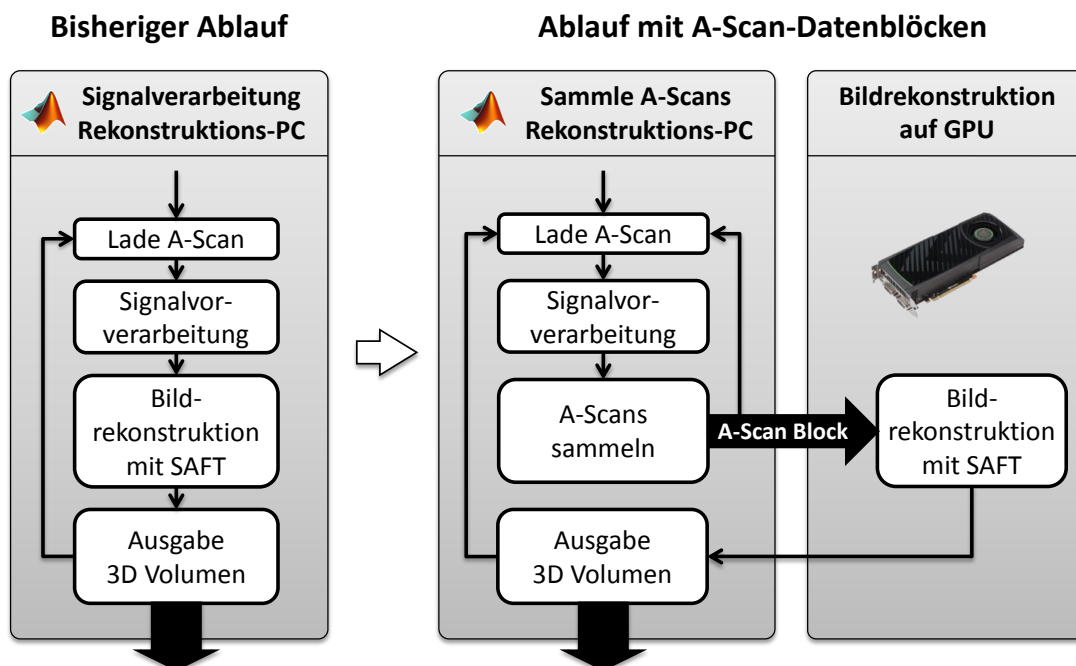


Abbildung 4.1: Übersicht der Datenverarbeitung in Matlab: Ursprüngliche Verarbeitung (links) sowie GPU gestützte Verarbeitung nach Sammeln der A-Scans in Datenblöcken (rechts).

genutzt. Dabei kann die Vorverarbeitung mit einer Vielzahl an Parametern konfiguriert und für verschiedene Durchläufe angepasst werden. Dadurch ist eine flexible und klar strukturierte Entwicklung für die Wissenschaftler möglich. Dieses Vorgehen mit jedem einzelnen A-Scan und einer nur geringen Datenmenge von 8 kB steht jedoch im Gegensatz zu einer GPU-basierten Beschleunigung. Die GPU wurde entwickelt, um einen CUDA-Kernel in vielen Threads parallel auf vielen CUDA-Kernen und Daten auszuführen. Diese Instruktionsausführung beruht auf einer Erweiterung des SIMD-Modells, welches Nvidia single-instruction-multiple-threads (SIMT) nennt. Eine GPU ist daher am besten für rechenintensive Anwendungen geeignet, die einen Algorithmus auf einer großen Datenmenge von unabhängigen Daten ausführen.

Es wurde daher der Ansatz gewählt, zuerst die einzelnen A-Scans nach der Vorverarbeitung in einem A-Scan-Datenblock zu sammeln und anschließend in einem zweiten Schritt an die GPU zu übertragen (Abbildung 4.1, rechts). Die GPU kann so das Reflexionsvolumen mit einer großen Anzahl an A-Scans berechnen, die vorrätig im globalen Speicher der GPU bereitliegen. Dieses Vorgehen minimiert sowohl den Overhead, der mit dem Aufruf der MEX-Funktion von MATLAB heraus verbunden ist, als auch den Overhead aufgrund des Datentransfers. Der Vorteil dieses Vorgehens ist die Transparenz für den restlichen Anwendungscode, da die Vorverarbeitung von dieser Anpassung unberührt bleibt. Der gesammelte A-Scan-Datenblock enthält für jeweils einen A-Scan den aufgenommenen A-Scan selbst, die Positionsdaten des dazugehörigen Senders und Empfängers sowie einen mittleren Schallgeschwindigkeitswert pro Pfad. Wurde der erste A-Scan-Datenblock vorverarbeitet, kann der nächste A-Scan geladen, vorverarbeitet, ebenfalls gesammelt und anschließend zur GPU transferiert werden.

4.2.2 Implementierung der MEX-Funktion

Um den GPU-SAFT-Kernel in die existierende MATLAB-Rekonstruktionssoftware zu integrieren, wird C-Code benötigt, der die GPU ansteuert, den Datentransfer durchführt und den CUDA-C-Kernel aufruft. Die Schritte im Programmablauf einer Volumenrekonstruktion sind in Abbildung 4.2 dargestellt und werden im Folgenden näher beschrieben:

1. Mithilfe einer MEX-Funktion kann C-Code ausgeführt werden, welche die Schnittstelle zwischen MATLAB und C-Code darstellt. Alle notwendigen Parameter und Daten für den SAFT-Algorithmus und Speicheradressen für den A-Scan-Datenblock werden hier dem C-Code übergeben.
2. Um die Vorteile der Beschleunigung mehrerer GPUs zu nutzen, wird die Arbeit auf alle im System vorhandenen GPUs aufgeteilt. Dafür wird für jede GPU ein Thread auf dem Host-PC initialisiert, der den weiteren Ablauf in den Schritten 3-6 für jeweils eine GPU parallel steuert.
3. Reicht der Speicher einer GPU nicht aus, um die Berechnung in einem Durchlauf auszuführen, muss das zu berechnende Volumen nochmals in Teilvolumen aufgeteilt und nacheinander berechnet werden. Dafür wird abhängig von der A-Scan-Datenblockgröße eine maximale Anzahl an Z-Schichten ermittelt, die in den Speicher passen und somit in einem Durchlauf berechnet werden können.

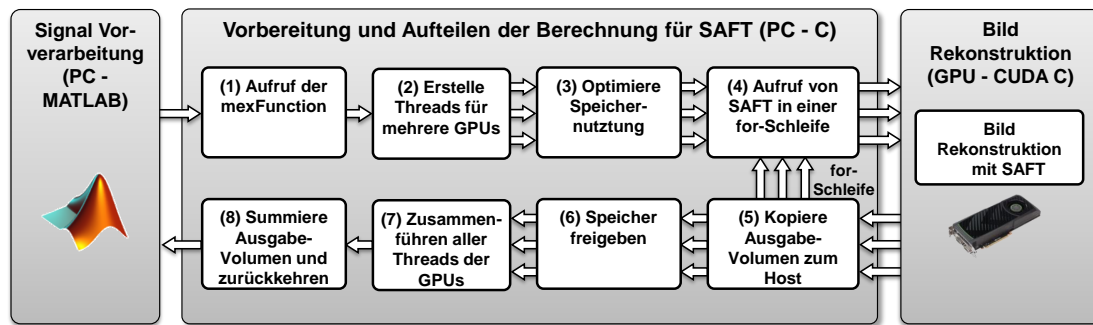


Abbildung 4.2: Überblick über die Schritte der Vorbereitung und Aufteilung der Volumenrekonstruktion im C-Code.

4. Innerhalb der for-Schleife werden alle Z-Schichten des Ausgabeteilvolumens für eine GPU nacheinander berechnet. Dafür wird auf der GPU, jeweils für die maximale Anzahl an Z-Schichten, der SAFT-Algorithmus ausgeführt.
5. Ist die Berechnung fertig, werden die Daten der Z-Schichten in den Host-Speicher zurück kopiert und SAFT wird für die nächsten Z-Schichten aufgerufen (Schritt 4). Dies geschieht solange, bis alle Z-Schichten des Ausgabeteilvolumens einer GPU fertig berechnet sind.
6. Ist die Berechnung für das gesamte Ausgabeteilvolumen einer GPU abgeschlossen, wird die Schleife verlassen und der Speicher auf der GPU kann freigegeben werden.
7. Um alle Threads der GPUs zu synchronisieren wird gewartet, bis alle GPUs ihre Teilberechnung zu Ende geführt haben und das rekonstruierte Volumen im Host-Speicher vorliegt.
8. Gab es schon eine Rekonstruktion mit einem vorhergehenden A-Scan-Datenblock, werden die beiden Volumen nun aufsummiert. Anschließend wird die Kontrolle an MATLAB zurückgegeben.

4.2.3 Implementierung des SAFT CUDA Kernels für die GPU

Das Ablaufdiagramm der Rekonstruktion mit dem Standard-SAFT-Kernel ist in Abbildung 4.3 dargestellt. Der CUDA-Kernel wird für alle zu rekonstruierenden Voxel aufgerufen und so die SAFT-Rekonstruktion parallel auf der GPU ausgeführt. Die dabei durchlaufenen Schritte sind:

1. Im ersten Schritt wird abhängig von der zugewiesenen *ThreadID* (*Tid*) bestimmt, welcher Voxel \vec{x} gerade berechnet wird. Sind die Koordinaten bekannt, wird eine Schleife über alle N genutzten A-Scans gestartet.
2. Innerhalb der Schleife wird Sender $s[i]$ aus dem *constant memory* geladen und überprüft. Da alle Kernel dieselbe Schleife ausführen, wird jeweils auf dieselbe Speicheradresse zugegriffen. Über den Zugriff auf das *constant memory* werden die Daten dabei innerhalb eines Taktzyklus bereitgestellt. Die A-Scans liegen dabei in einer sortierten Reihenfolge vor, bei der für einen Sender alle genutzten Empfänger nacheinander angeordnet sind. Handelt es sich nicht um den Sender vom letzten Schleifen-Durchlauf $s[i] \neq s[i - 1]$, ist die Position des Senders \vec{s} neu und alle vom Sender abhängigen Daten müssen neu berechnet werden. Die Position des Senders \vec{s} wird ebenfalls aus dem *constant memory* geladen.

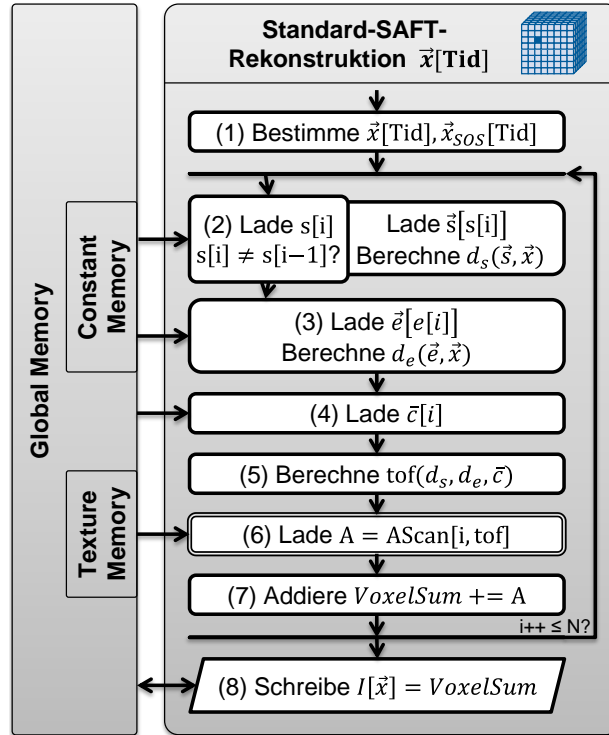


Abbildung 4.3: Ablaufdiagramm des Standard-SAFT-Kernels für einen Voxel.

Der Abstand d_s zwischen Sender \vec{s} und Voxel \vec{x} wird bestimmt. Handelt es sich dagegen um einen bereits genutzten Sender $s[i] = [i - 1]$, liegen die vom Sender abhängigen Werte bereits in den lokalen Registern vor und müssen somit nicht mehr neu geladen bzw. berechnet werden.

3. Durch die Anordnung der A-Scans ist im Normalfall die Empfänger-Position bei jedem weiteren A-Scan neu. Daher wird in jedem Durchlauf der Schleife die aktuellen Empfänger-Koordinaten \vec{e} , abhängig vom aktuellen Empfängerindex $e[i]$, aus dem *constant memory* geladen und der Abstand d_e zwischen Empfänger \vec{e} und Voxel \vec{x} bestimmt.
4. Eine mittlere Schallgeschwindigkeit \bar{c} wird pro A-Scan geladen.
5. Der Zeitpunkt tof wird abhängig von der Gesamtdistanz ($d_s + d_e$) und der mittleren Schallgeschwindigkeit \bar{c} nach Gleichung (3.12) bestimmt und mit der Samplefrequenz $f_s=10$ MHz multipliziert.
6. Die Amplitude A wird aus dem aktuellen A-Scan am Zeitpunkt tof , durch den Zugriff über das *texture memory*, bereits interpoliert gelesen.
7. Die Amplitude A wird in VoxelSum aufsummiert (siehe Gleichung (3.15)). Sind noch nicht alle N A-Scans im A-Scan-Datenblock von MATLAB abgearbeitet, wird die Schleife mit dem nächsten A-Scan neu gestartet.
8. Nachdem alle A-Scans in dem Schleifendurchlauf abgearbeitet wurden, steht die Summe aus allen A-Scan-Beiträgen in VoxelSum und wird in das Ausgabe-Volumen $I[\vec{x}]$ geschrieben.

4.2.4 Performanzanalyse

Rechengeschwindigkeit verschiedener Volumen- und A-Scan-Datenblockgrößen

In Abbildung 4.4 ist die innerhalb der MEX-Funktion erreichte Rechengeschwindigkeit in GVA/s in Abhängigkeit der A-Scan-Datenblockgröße für verschiedene Volumengrößen von 32^3 bis 1024^3 dargestellt. Dabei wurde nur eine der beiden GPUs der NVIDIA GTX 590 Doppel-GPU genutzt. Die maximale Rechengeschwindigkeit von 13,7 GVA/s wird ab einer Volumengröße von 512^3 und einer A-Scan-Datenblockgröße über 10.000 erreicht. Die erreichte Rechengeschwindigkeit ist stark abhängig von der genutzten Volumengröße. Dies kann auf das Zugriffsmuster auf die A-Scan-Messdaten zurückgeführt werden, die bei der parallelen Berechnung von benachbarten Voxel durchgeführt wird. Wird die ROI mit einer höheren Voxelanzahl abgetastet, steigt die Auflösung und benachbarte Voxel liegen enger beieinander als bei einer niedrigen Auflösung. Damit steigt auch die Wahrscheinlichkeit, dass ähnliche Signallaufzeiten (*tof*) berechnet werden. Somit wird auch auf im Speicher nahegelegene Amplitudenwerte im A-Scan zugegriffen, wodurch eine hohe Cache-Ausnutzung erreicht werden kann. Des Weiteren ist die Größe des A-Scan-Datenblocks wesentlich für die maximale Rechengeschwindigkeit der GPU. Höhere Volumengrößen erreichen schneller die maximale Rechengeschwindigkeit als geringe Volumengrößen. Der Grund liegt darin, dass bei einer geringen Voxelanzahl die Berechnung der Voxel so kurz ist, dass der Overhead des Datentransfers für den A-Scan-Datenblock zur GPU nicht vernachlässigt werden kann. Eine A-Scan-Datenblockgröße von 10.000 stellt dabei zudem einen guten praktikablen Kompromiss dar. Die Eingangsdatenmenge beträgt dabei 240 MB, sodass im 1,5 GB großen DDR5 RAM noch genügend Platz für eine große Anzahl Schichten vorhanden ist, die zu rekonstruieren sind. Hieraus folgt, dass die maximale Performanz mit der GPU erreicht wird, wenn Rekonstruktionen

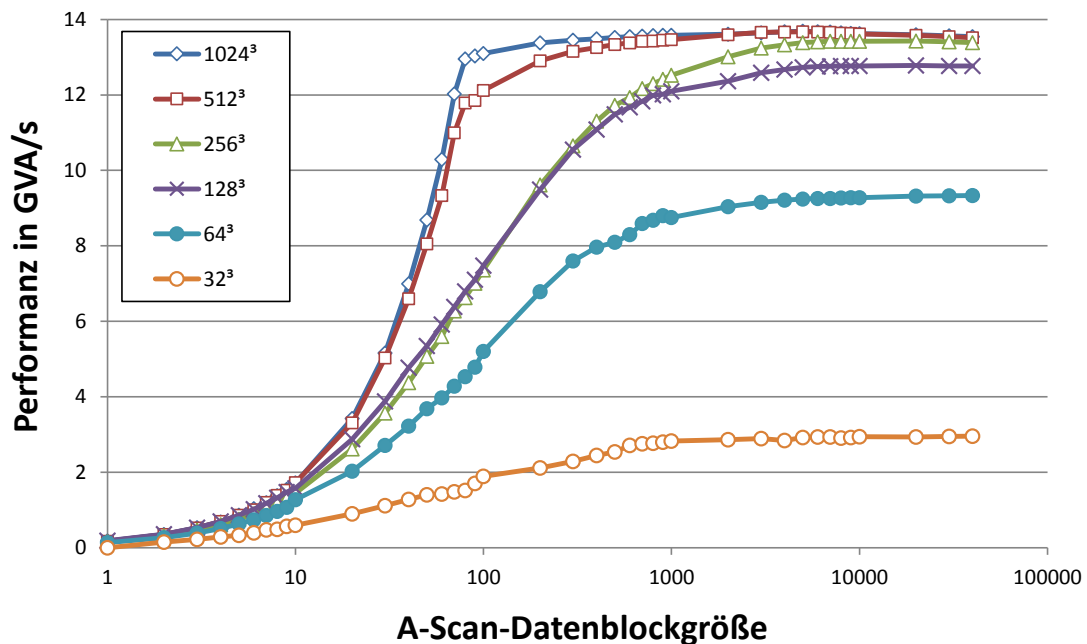


Abbildung 4.4: Erreichte SAFT-Rechengeschwindigkeit mit der Einzel-GPU GTX 590, gemessen in Voxel-Durchsatzrate über der A-Scan-Blockgröße für verschiedene Volumengrößen von 32^3 bis 1024^3 .

mit hohen Volumengrößen und großen A-Scan-Datenblöcken durchgeführt werden. Bei Größen der A-Scan-Datenblöcke über 10.000 kann zudem der Overhead vernachlässigt werden. Die gemessene effektive Rechengeschwindigkeit von MATLAB aus gesehen beträgt 13,6 GVA/s. Dies entspricht 99,2% der in der MEX-Funktion gemessenen maximalen Performanz.

Rechengeschwindigkeit von mehreren GPUs

Für den Einsatz von mehreren GPUs wurden vier NVIDIA GTX 590 Doppel-GPUs in ein externes 2U-Erweiterungssystem von One Stop Systems (OSS-PCIe-2U-ENCL-EXP-4-2) [79] eingebaut. Diese externe GPU-Box stellt vier PCIe Gen 2 Slots zur Verfügung, die mit der maximalen Übertragungsrate von x16 PCIe Lanes arbeiten können. Die damit erreichte Rechengeschwindigkeit ist in Abbildung 4.5 dargestellt und zeigt, dass der SAFT-Algorithmus sehr gut auf mehreren GPUs verteilt skaliert. Mit vier Doppel-GPUs GTX 590 GPUs steigt die Rechengeschwindigkeit auf 106,6 GVA/s. Dies entspricht einem Beschleunigungsfaktor von 7,8 zur Einzel-GPU. Dabei liegt die Gesamt-Beschleunigung leicht unterhalb von Faktor 8, da das Volumen auf mehrere GPUs verteilt wird und somit jede GPU nur noch ein

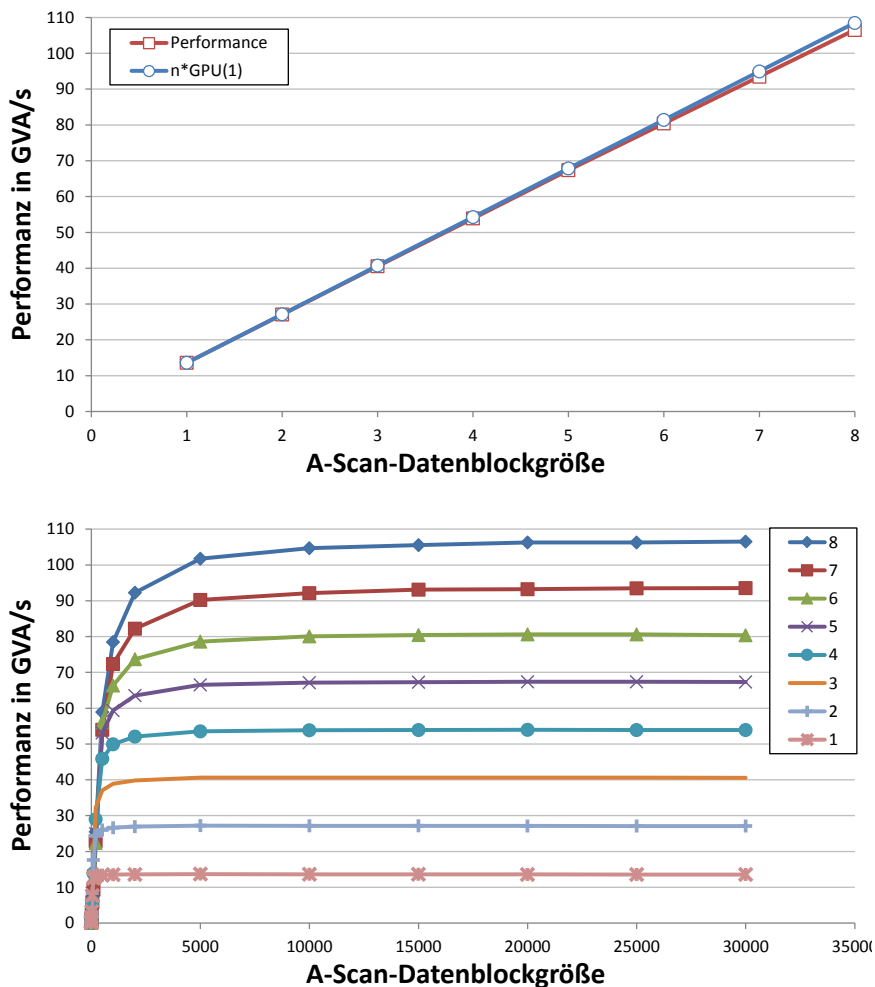


Abbildung 4.5: Erreichte SAFT-Rechengeschwindigkeit mit mehreren GPUs, gemessen für ein Volumen von 1024^3 in Voxel-Durchsatzrate über der Anzahl an GPUs (oben) und der A-Scan-Datenblockgröße (unten).

4. GPU-basierte Beschleunigung der Standard-SAFT

Achtel des Volumens berechnet. Bei einer genügend großen A-Scan-Datenblockgröße von 20.000 und einem genügend großen Volumen von 1024^3 Voxel skaliert die Rechengeschwindigkeit sehr gut mit der Anzahl an eingesetzten GPUs. Die gemessene effektive Rechengeschwindigkeit von MATLAB aus beträgt 103,9 GVA/s (97% von 106,6 GVA/s).

5. Analyse der Bildqualität

5.1 Die Bildqualität der 3D-USCT

Zu der Abbildungsqualität des SAFT-Algorithmus am Beispiel des 3D-USCT-Prototypen am IPE existieren bereits Untersuchungen. So wurde von Schwarzenberg eine detaillierte Untersuchung der Apertur und ihrer Abbildungseigenschaften durchgeführt [38]. Darauf aufbauend wurde eine Apertur vorgeschlagen, die in Bezug auf die räumliche Auflösung, den Kontrast und die Ausleuchtung optimiert ist. Diese optimierte Apertur wurde realisiert und stellt die Grundlage für die aktuellen Messdaten der 3D-USCT am IPE dar. Daher werden in dieser Arbeit genutzte Rekonstruktionsmethoden auf der Grundlage der aktuellen Hardware evaluiert.

Die maximal erreichbare Abbildungsqualität wird durch das Zusammenspiel von genutzter Hardware, ihrer Systemparameter sowie der eingesetzten Rekonstruktionssoftware bestimmt (siehe Abbildung 5.1). Von der Signalaufnahme bis hin zur Rekonstruktion gibt es eine Vielzahl an systembedingten Einflussfaktoren sowie Fehlerquellen, die einen Einfluss auf die Bildqualität nehmen können. Dazu gehören auch externe Einflüsse wie mechanische Toleranzen, Temperaturschwankungen und Patientinnenbewegungen während der Aufnahme. In [38] werden die systembedingten Einflussfaktoren näher erläutert. Die Einflüsse der Hardware lassen sich dabei in die zwei großen Teilbereiche Ultraschallwandler und Geometrie einteilen. Dabei wird zu den Ultraschallwandlern auch die nötige Elektronik zur Ansteuerung und Digitalisierung der Ultraschallsignale sowie der Datentransfer gezählt. Unter der Geometrie wird die geometrische Sensoranordnung verstanden. Zur Software gehören die Signalvorverarbeitung und die Bild-Rekonstruktion mit der SAFT. Durch die Signalvorverarbeitung werden die aufgezeichneten A-Scans für die Rekonstruktion mit SAFT aufbereitet. Dabei wird mit Hilfe des *Adapted Matched Filters* [63] das vom System verursachte Rauschen unterdrückt und die für die Bildrekonstruktion relevanten Informationen der Echosignale hervorgehoben (siehe Kapitel 3.1).

Basierend auf den Annahmen des genutzten Standard-SAFT-Algorithmus (siehe Kapitel 3.3) beträgt die maximale Auflösung bei einer kohärenten Bildgebung $\lambda/2$. Bei einer Annahme einer mittleren Schallgeschwindigkeit von 1500 m/s und der genutzten Mittenfrequenz $f_0 = 2,5$ MHz ergibt sich damit die theoretische Auflösung von

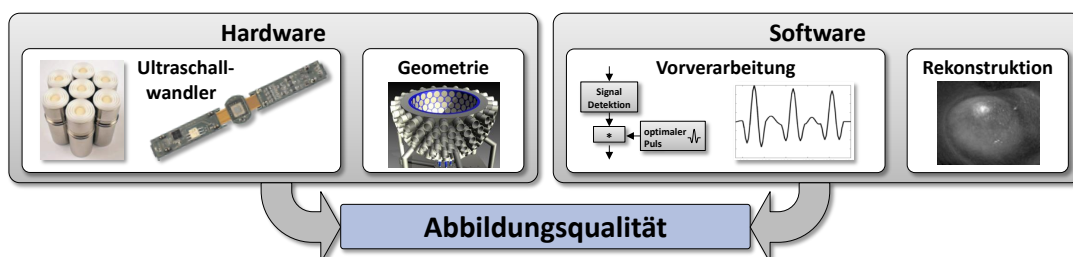


Abbildung 5.1: Einflussfaktoren auf die Abbildungsqualität nach [38].

$$\text{Auflösung} = \frac{\lambda}{2} = \frac{1500 \text{ m/s}}{2 \cdot 2,5 \text{ MHz}} = 0,3 \text{ mm} \quad . \quad (5.1)$$

Diese theoretisch hergeleitete Auflösung wurde von Ruiter et al. [80] experimentell nachgewiesen. Dazu wurde ein Bondingdraht mit dem Durchmesser von 0,07 mm in dem 3D-USCT-Prototypen vermessen. Der Durchmesser liegt dabei unterhalb der Auflösungsgrenze und wirkt daher wie ein Dirac-Impuls auf das gesamte Abbildungssystem. Durch die Impulsantwort kann die Auflösungsgrenze des Abbildungssystems experimentell bestimmt werden. Der Draht wurde vor der Messung verdreht, so dass die Abbildung des Drahts für mehrere Richtungen untersucht werden konnte. Dabei zeigte sich eine durchschnittliche Auflösung in alle Richtungen von 0,24 mm, die damit sogar leicht unterhalb der 0,3 mm liegt. Diese Auflösung wurde jedoch in einem homogenen Medium Wasser mit einem Bondingdraht erreicht. Bei der Aufnahme komplexerer Objekte wie z.B. einer realen Brust wird diese Auflösung zur Zeit noch nicht erreicht. In der Praxis hat sich für komplexere Objekte eine reduzierte Auflösung von $(1,8 \text{ mm})^3$ bewährt. Bei Rekonstruktionen mit einer Korrektur der Schallgeschwindigkeit wird eine Auflösung bis $(0,9 \text{ mm})^3$ genutzt. Dazu wird die Breite des Optimalen Pulses während der Vorverarbeitung wie in Kapitel 3.1 beschrieben angepasst. Diese Auflösungen werden daher auch in dieser Arbeit bei der Rekonstruktion mit realen Messdaten genutzt.

Einen wichtigen Einfluss auf die Bildqualität besitzt auch die *region of interest* (ROI), da sie den Bereich definiert, der scharf abgebildet werden kann. Der Durchmesser und die Höhe der ROI der aktuellen Apertur betragen jeweils 10 cm. Allerdings hat die erste Studie, die im Universitätsklinikum Jena durchgeführt wurde, gezeigt, dass die Brust einen Auftrieb besitzt und sich größere Brüste teilweise außerhalb der aktuellen ROI befinden. Um die ganze Brust aufnehmen zu können, ist eine neue größere Apertur in der Entwicklung, die einen Durchmesser von 20 cm und eine Höhe von 12 cm besitzen soll. Daher wird in dieser Arbeit die größere ROI mit 20 cm Durchmesser genutzt.

In dieser Arbeit sollen Methoden beschrieben und analysiert werden, die es dem SAFT-Algorithmus erlauben über die bisherigen Annahmen hinaus zu gehen, sowie für inhomogene Medien eine ebenso hohe Bildauflösung zu erreichen. Im Fokus der Untersuchungen dieser Arbeit stehen die Methoden und die Software der Rekonstruktionsvarianten. Jedoch gibt die Hardware Randbedingungen vor, die für die Betrachtung der optimalen Bildqualität als Randbedingung mit einbezogen werden.

5.2 Metriken zur Analyse der Bildqualität

Im Folgenden wird auf geeignete Metriken eingegangen, mit denen die Bildqualität der 3D-USCT quantitativ bewertet werden kann. Dabei liegt der Fokus auf dem Verfahren der Reflektivitätsrekonstruktion mittels des SAFT-Algorithmus.

Auflösungsvermögen und Verschiebung

Für die Berechnung des räumlichen Auflösungsvermögens eines bildgebenden Systems werden in [38] zwei Ansätze vorgestellt, die allerdings eine sehr unterschiedliche Rechenkomplexität aufweisen: Der rasterbasierte und der geometrische Ansatz.

Der häufig eingesetzte **rasterbasierte Ansatz** [81, 82, 83] entspricht einer Bildrekonstruktion mittels SAFT im Raum mit anschließender Analyse der dreidimensionalen Abbildungsfunktion. Die Abbildung eines Punktstreuers wird durch die Größe und Form der Punktabbildung beschrieben. Für die Analyse muss daher ein Volumen rekonstruiert werden, das auch die Umgebung des Punktstreuers in ausreichender Größe beinhaltet. Dabei ist zu beachten, dass es bei einer realistischen Annahme einer räumlich inhomogenen Schallgeschwindigkeitsverteilung zu einer Ringabbildung eines Punktstreuers kommen kann. Dadurch vergrößert sich die Ausdehnung der Abbildungsfunktion eines Punktstreuers.

Für die Rekonstruktion wird das Bild in diskrete Bildpunkte bzw. Volumenelemente (Voxel) eingeteilt, die in einem regelmäßigen Raster angeordnet sind. Da die maximal zu erreichende räumliche Auflösung 0,24 mm beträgt, muss eine Bildauflösung von mindestens 0,12 mm oder höher für die Rekonstruktion angewandt werden.

Die Vorteile der Analyse mit dem rasterbasierten Ansatz liegen vor allem in der Flexibilität und Vielfältigkeit der Methode, da dieses Vorgehen direkt das berechnete Volumen der Bildrekonstruktion mittels SAFT nutzt. Der Einfluss beliebiger Fehlerquellen und Parameter, wie Schallgeschwindigkeiten, Position des Punktstreuers und der Ultraschallwandler, kann so mitbestimmt werden. Es können zudem komplexere geometrische Objekte einbezogen und vermessen werden. So kann ebenso der Einfluss von weiteren Objekten wie z.B. der Brust oder Zysten bewertet werden.

Die Komplexität der Auswertung mittels des rasterbasierten Ansatzes ist abhängig von der Anzahl der zu untersuchenden Positionen P , der Anzahl der zu rekonstruierenden Voxel V und der Anzahl der genutzten A-Scans A . Da gewöhnlich $P \ll V$ ist und $P \ll A$, ist der Aufwand insbesondere von der Anzahl der zu rekonstruierenden Voxel V und der A-Scans A abhängig. Dies resultiert in einer Komplexität von $O(V \cdot A)$. Dabei kann je nach gewählter Größe des Datensatzes und umgebenden Volumens, V und A dieselbe Größenordnung aufweisen.

Da der rasterbasierte Ansatz bei einer hohen Auflösung und A-Scan-Datenmenge einen erheblichen Rechenaufwand besitzt, wurde von Schwarzenberg [38] ein **geometrischer Ansatz** entwickelt, der eine effizientere Untersuchung der Punktabbildung ermöglicht.

Hierbei basiert die Idee darauf, dass jeder A-Scan einen Beitrag für einen Voxel liefert. Dieser Beitrag liegt dabei jeweils in einer Ellipsoidschale, mit Sender und Empfänger als Brennpunkte. Nun werden mehrere sog. 3D-Untersuchungslinien durch den Voxel gelegt und die Schnittpunkte mit den Ellipsoidschalen berechnet. Damit kann für jeden A-Scan entsprechend der exakte Anfang und das Ende jeder Rückprojektion auf dieser Linie berechnet werden. Nach einer Sortierung der Schnittpunkte auf diesen Untersuchungslinien kann ein Profil der Summe aller Rückprojektionen ermittelt werden und somit ein exaktes Werteprofil bestimmt werden. Werden mehrere Untersuchungslinien durch den Voxel mit entsprechend hoher Winkelabtastung genutzt, kann so die 3D-Abbildungsfunktion beliebig genau approximiert werden.

Die Komplexität der Auswertung mittels des geometrischen Ansatzes ist ebenfalls abhängig von der Anzahl zu untersuchenden Positionen P und der Anzahl der genutzten A-Scans A . Dagegen wird der Aufwand erheblich reduziert, indem V zu rekonstruierende Voxel durch die Anzahl von Untersuchungslinien L ersetzt wird. Der Aufwand für die Schnittpunktberechnung pro Untersuchungslinie ist dabei zwar

deutlich aufwändiger, jedoch ist die Anzahl der Untersuchungslinien deutlich geringer als die Anzahl der genutzten A-Scans A , $L \ll A$. Damit hängt der Aufwand hauptsächlich von A ab. Allerdings ist zusätzlich eine Sortierung notwendig, wodurch die Komplexität insgesamt auf $O(A \cdot \log A)$ anwächst.

Beide numerische Methoden eignen sich für die Berechnung der Punktabbildung und können beliebige Sensorgeometrien und Punktstreuerpositionen vermessen. Dabei ist für die Signale keine Beschränkung gegeben und beliebige zeitliche wie auch örtliche Fehlerquellen lassen sich mit einberechnen. Kommt es jedoch durch eine fehlerhafte Annahme der Schallgeschwindigkeiten zu einer Verschiebung der Punktstreuerabbildung, wird der falsche Bereich vermessen. Beim rasterbasierten Ansatz kann dagegen zuerst das Maximum in dem rekonstruierten Volumen gesucht werden. Steht das Maximum fest, kann anschließend die Punktabbildung an dieser Stelle bestimmt werden. Die Verschiebung des Maximums kann zudem als weitere Metrik für die Bildqualität genutzt werden. Ein weiterer Vorteil liegt darin, dass direkt und intuitiv die Ausgabedaten der SAFT-Rekonstruktion genutzt werden, die alle Werte und Einflüsse enthalten. Wurde daher das gesamte Volumen einmal rekonstruiert, können weitere beliebige Bildpunkte darin untersucht und für die Messung weiterer Metriken wie z.B. des Kontrasts genutzt werden. Gerade die Berechnung des Kontrasts erfordert in jedem Fall die Rekonstruktion einer genügend großen Umgebung des Punktstreuers. Der steigenden Berechnungsdauer mit rasterbasiertem Ansatz kann durch eine Hardwarebeschleunigung entgegengewirkt werden. So wird in dieser Arbeit insbesondere die Beschleunigung mittels GPUs betrachtet sowie weitergehende Methoden untersucht, um einen optimierten Kompromiss zwischen Berechnungsdauer und Bildqualität zu erreichen. Für die Evaluierung des räumlichen Auflösungsvermögens wird daher die rasterbasierte Methode angewandt.

Die Berechnung des Auflösungsvermögens mittels rasterbasierter Methode wird in drei Schritten durchgeführt:

1. Für den simulierten Messpunkt wird eine **Teilrekonstruktion** durchgeführt, die die Abbildung des Punktstreuers vollständig enthält.
2. **Suchen des Maximas** in dem Teilvolumen und Bestimmung des Abstands zur Position des simulierten Punktstreuers. Die Verschiebung kann über

$$D = \|\vec{x}_{\text{simuliert}} - \vec{x}_{\text{maximum}}\|, \quad (5.2)$$

bestimmt werden, siehe Abbildung 5.2.

3. **Untersuchung von Größe und Form der 3D-Punktabbildung** anhand der drei Schnittbilder in xy -, xz - und yz -Richtung, die durch das Maxima gehen. Dabei wird, wie in Abbildung 5.2 für zwei Beispiele gezeigt, pro Schnittbild eine Analyse der 2D-Halbwertsbreite (FWHM) für N Winkel im Abstand von 5° durchgeführt. In Abbildung 5.2 ist links eine Punktabbildung dargestellt, die bis auf die z -Richtung isotrop ist und die Position des rekonstruierten Maximums mit der realen Position übereinstimmt. Rechts ist eine Punktabbildung dargestellt, die in der yz -Richtung eine Verformung aufzeigt und zudem eine Verschiebung erfahren hat. Dabei wird angenommen, dass die Form und Ausdehnung der Punktabbildung sich nur kontinuierlich verändert und durch die drei Schnittbilder gut abgebildet werden kann.

Für die Evaluierung der 3D-Punktabbildung wird als Maß für die Auflösung die mittlere Halbwertsbreite

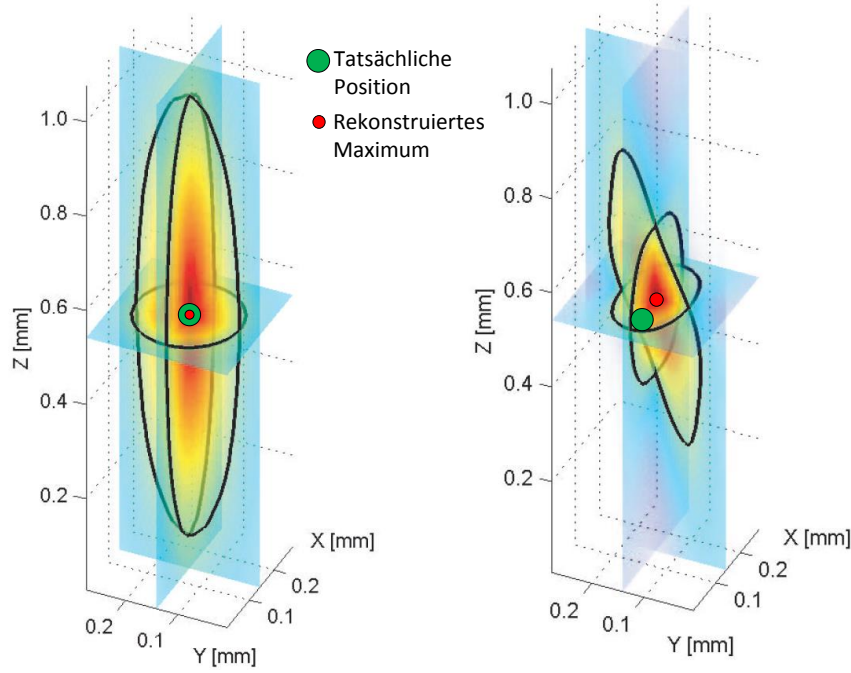


Abbildung 5.2: Zwei Beispiele für die Analyse von rekonstruierten Teilvolumina nach [84]. Die Untersuchung der 3D-Punktabbildung wird mittels Analyse der 2D-FWHM der Schnittbilder in xy -, xz - und yz -Richtung durchgeführt.

$$\overline{FWHM}_{xyz} = \frac{\sum_{i=1}^N FWHM_{xyz}(i)}{N}, \quad (5.3)$$

und als Maß für die Isotropie der Punktabbildung die Standardabweichung

$$\sigma_{FWHM_{xyz}} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N \left(\overline{FWHM}_{xyz} - FWHM_{xyz}(i) \right)^2} \quad (5.4)$$

bestimmt. Über alle Punktstreuer lässt sich für das gesamte Volumen ein Mittelwert \overline{FWHM}_{ges} und eine Standardabweichung $\sigma_{FWHM_{ges}}$ berechnen.

Bildkontrast

Der Kontrast eines Bildes kann über eine Vielzahl von Funktionen ermittelt werden. Für die Bewertung des Kontrasts in der 3D-USCT ist es von Interesse, dass streuende Objekte sich deutlich von nichtstreuenden Objekten abheben, die einen dunklen Hintergrund darstellen sollen. Durch die Rekonstruktion mittels SAFT kommt es jedoch, wie in Kapitel 3.3 beschrieben, prinzipiell zu Rekonstruktionsartefakten (siehe Abbildung 5.3), die sich negativ auf den Kontrast auswirken und schwach streuende Objekte überlagern können.

Die Rekonstruktion kann dabei durch den Einsatz von GPUs beschleunigt werden. Wurde das Volumen bereits für die Untersuchung des Auflösungsvermögens berechnet, kann dieses Volumen direkt genutzt werden.

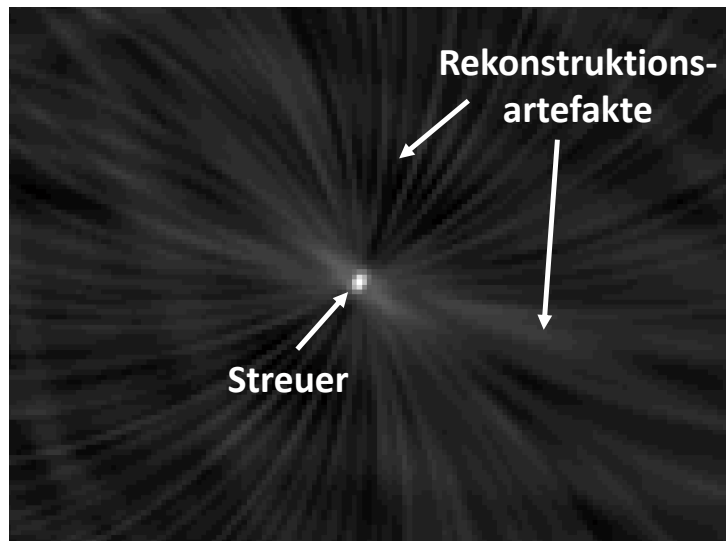


Abbildung 5.3: Exemplarische 2D-Rekonstruktion eines Punktstreuers mit den SAFT-typischen Rekonstruktionsartefakten.

Ein geeignetes Maß, das eine häufige Anwendung im Bereich analoger Systeme und der verlustbehafteten Kompressions-Codecs findet, ist das Spitzen-Signal-zu-Rausch-Verhältnis, oder engl. *peak-signal-to-noise ratio* (PSNR) [85, 86, 87]. Es beschreibt das Verhältnis zwischen einem Originalbild und einem veränderten Vergleichsbild. Für den Fall, dass die Grundwahrheit bekannt ist, kann folgende Definition genutzt werden, die für den 3D-Fall erweitert wurde [88]:

$$PSNR_1 = 10 \cdot \log_{10} \left(\frac{\max(I)^2}{MSE} \right) \quad (5.5)$$

$$MSE = \frac{1}{X \cdot Y \cdot Z} \sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} \sum_{k=0}^{Z-1} [I(x, y, z) - K(x, y, z)]^2 \quad (5.6)$$

Das PSNR ist dabei über das Verhältnis der maximalen Reflektivität des Originals $\max(f(\vec{x}))$ zu der mittleren quadratischen Abweichung (engl. *mean squared error* (MSE)) definiert. Der MSE ist ein Maß, das den Fehler zwischen den beiden $X \cdot Y \cdot Z$ großen Volumen angibt, dem Originalvolumen $f(\vec{x})$ und dem mit SAFT rekonstruierten Volumen $I(\vec{x})$. Das PSNR wird oft auch in logarithmischer Skala, in Dezibel (dB) angegeben, da viele Signale einen breiten Dynamikbereich haben. Zu beachten ist, dass der PSNR-Wert keine absolute Bildqualität feststellt. Daher sind Schlüsse aus Vergleichen mit anderen PSNR-Werten nur gültig, wenn sie sich auf dasselbe Originalbild beziehen.

Für reale Messdaten ist die Grundwahrheit nicht bekannt, da diese erst gemessen werden soll. Daher kann die oben genannte Definition des PSNR nicht angewandt werden. Eine alternative Definition, die insbesondere in der medizinischen Bildgebung Gebrauch findet, nutzt das Verhältnis der maximalen Signalamplitude zur Standardabweichung des Rauschens [73]:

$$PSNR_2 = \frac{\max(I)}{\sigma_{Rauschen}} \quad (5.7)$$

Dabei wird das Rauschen in einem Bereich gemessen, der kein Objekt enthält. Für die 3D-USCT kann z.B. ein Bereich im homogenen Medium Wasser außerhalb der Brust genutzt werden.

Räumlich homogene Objektausleuchtung

Bei der Rekonstruktion der Reflexionsbilder kommt es durch Dämpfungseffekte zu lokalen Abschattungen, die bei der 3D-USCT insbesondere in Richtung Zentrum der Brust zunehmen (siehe Abbildung 2.5). Das Resultat ist eine räumlich inhomogene Amplitude der Reflexion oder auch eine ungleichmäßige Objektausleuchtung. Die Analyse der räumlich homogenen Objektausleuchtung kann mit der Simulation untersucht werden, da dabei die Stärke der Streuung als Grundwahrheit vorgegeben werden kann. Dabei wird jeweils der rekonstruierte Grauwert I aller im Volumen gleichmäßig verteilten N Punktstreuer untersucht und mit den Amplituden der Grundwahrheit I_{GW} verglichen, um den mittleren Fehler

$$\bar{\mathcal{E}}_A(\vec{x}_i) = \frac{\sum_{i=1}^N |I_{GW}(\vec{x}_i) - I(\vec{x}_i)|}{N}, \quad (5.8)$$

zu bestimmen. Als Grundwahrheit wird dabei das Volumen genutzt, das, unter der Annahme, dass keine Dämpfung stattfindet, mit dem Standard-SAFT-Algorithmus rekonstruiert wurde. Das ist nötig, da die Amplitude der Punktstreuer bereits, bedingt durch den Standard-SAFT-Algorithmus, von der räumlichen Position abhängt.

Als Maß für die *räumliche homogene Objektausleuchtung* (RHOA) wird die Standardabweichung des Fehlers \mathcal{E}_A

$$\sigma_{RHOA} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N \left(\bar{\mathcal{E}}_A(\vec{x}_i) - \mathcal{E}_A(\vec{x}_i) \right)^2} \quad (5.9)$$

über alle Punktstreuer bestimmt. Eine niedrige Standardabweichung σ_{RHO} entspricht dabei einer guten homogenen Objektausleuchtung.

5.3 Aufbau der Simulationsumgebung

In Kapitel 5.2 wurden quantitative Metriken eingeführt, die eine Bewertung der erreichbaren Bildqualität des SAFT-Algorithmus ermöglichen. Für die Berechnung dieser Metriken ist eine Grundwahrheit nötig, die dem Original entspricht und mit der das rekonstruierte Bild verglichen werden kann. Diese ist bei realen Experimenten jedoch a priori nicht gegeben. Eine Bewertung der Bildqualität bleibt somit einer subjektiven Einschätzung überlassen. Um dennoch eine quantitative Untersuchung durchführen zu können, wurde eine Simulation erstellt, die den Effekt der zu untersuchenden Einflüsse soweit wie möglich realistisch abdeckt. Dazu gehört insbesondere die Möglichkeit, inhomogene Medien bezüglich Schallgeschwindigkeit und Dämpfung zu simulieren. Über eine definierte Anordnung von Objekten im 3D-USCT kann die

Grundwahrheit vorgegeben werden, aufgrund derer A-Scans simuliert werden (siehe Abbildung 5.4). Nach der Signal-Vorverarbeitung kann anhand der simulierten A-Scans die Rekonstruktion mit SAFT durchgeführt werden. Das rekonstruierte Volumen kann somit mit der Grundwahrheit quantitativ verglichen werden.

Da der SAFT-Algorithmus davon ausgeht, dass jede streuende Struktur mit einer Vielzahl an idealen Punktstreuern approximiert werden kann, werden hauptsächlich Reflexionen an idealen Punktstreuern simuliert. Darüber hinaus soll auch die Möglichkeit bestehen, für die in Kapitel 9 generierten Reflexionscharakteristiken Reflexionen an Oberflächen simulieren zu können. Im Folgenden wird ausgehend von der Simulation der Punktstreuer die Simulation eingeführt und anschließend für Oberflächen erweitert.

Simulation von Reflexionen an Punktstreuern

Abbildung 5.5 stellt eine vereinfachte Anordnung mit nur einem Punktstreuer dar. Dieser befindet sich innerhalb einer simulierten Brust mit einem Radius r von 10 cm. Die mit einer Halbkugel approximierte Brust deckt nahezu den gesamten Bereich der ROI ab. Für die Simulation mit heterogenem Medium kann somit die Schallgeschwindigkeit in der Brust im Vergleich zum Wasser variieren. Die Schallgeschwindigkeit im Wasser wird als $c_{water}=1500$ m/s angenommen. Um mit der Simulation einen weiten Bereich abzudecken, wird in dieser Arbeit eine Schallgeschwindigkeit der Brust von 1450 bis 1550 m/s angenommen. Dieser Wertebereich entspricht realistischen Werten für verschiedene Gewebetypen.

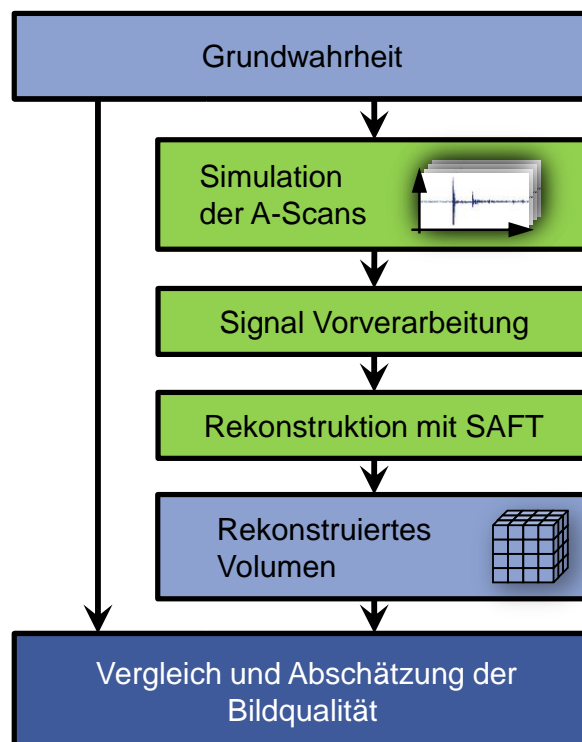


Abbildung 5.4: Überblick der Analyse der Bildqualität mit Simulation.

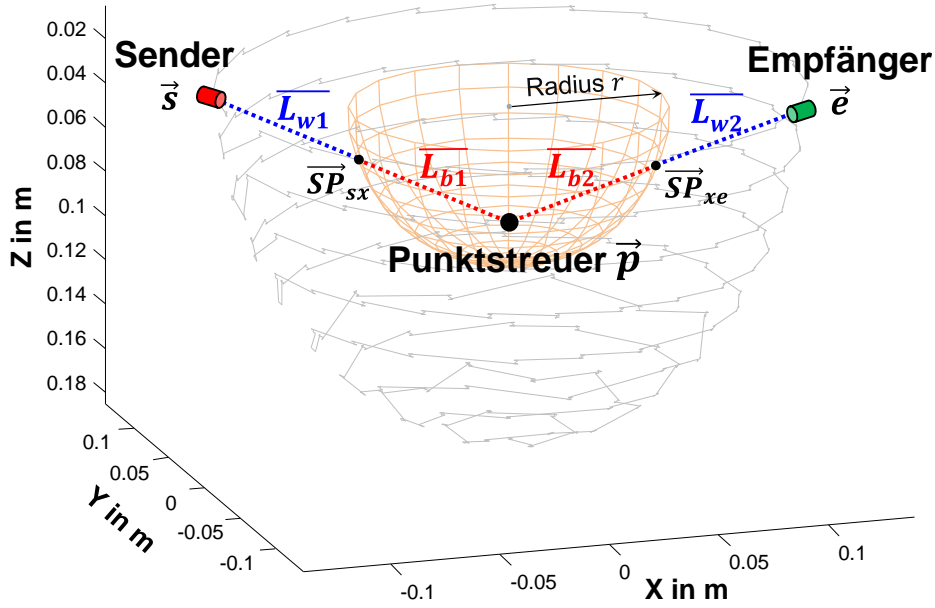


Abbildung 5.5: Simulierter Punktstreuer und die Pfade für ein Sender- und Empfängerpaar.

Die Simulation basiert auf der Strahlverfolgung (Ray-Tracing) sowie der Annahme einer linearen Ausbreitung des Ultraschalls. Effekte wie Brechung und Beugung werden dabei vernachlässigt, da die Rekonstruktion mit SAFT ebenfalls keine Brechung oder Beugung beinhaltet und dieser Einfluss nicht bewertet werden soll. In der Simulation wird eine Reflexion in einem A-Scan jeweils durch die Kombination zweier Werte beschrieben:

1. die Ankunftszeit tof am Empfänger.
2. die Amplitude des gedämpften Ultraschalls A bei der Ankunft am Empfänger.

Mithilfe der Simulation, bei der die geometrischen Positionen der Sender \vec{s} , Empfänger \vec{e} und des Voxels \vec{x} bekannt sind, können nun die Schnittpunkte \vec{SP}_{sx} und \vec{SP}_{xe} mit der Brustoberfläche auf dem Pfad zwischen Sender-Punktstreuer-Empfänger berechnet werden. Aus den Längen der Teilstrecken ergeben sich somit die Distanzen $l_w = \overline{L_{w1}} + \overline{L_{w2}}$ und $l_b = \overline{L_{b1}} + \overline{L_{b2}}$, die der Ultraschall in dem jeweiligen Medium Wasser bzw. Brust durchläuft. Die mittlere Schallgeschwindigkeit auf dem Pfad durch Wasser und Brust kann somit bei $M = 2$ Medien durch das harmonische Mittel nach

$$\bar{c}_{pfad}(\vec{s}, \vec{x}, \vec{e}) = \frac{\sum_{i=1}^M (l_i)}{\sum_{i=1}^M \left(\frac{l_i}{c_i} \right)} = \frac{l_w + l_b}{\frac{l_w}{c_w} + \frac{l_b}{c_b}} \quad (5.10)$$

berechnet werden. Dabei bezeichnet l_i die Länge des Pfads durch das i -te Medium und c_i die Schallgeschwindigkeit innerhalb des i -ten Mediums. Daraus ergibt sich anschließend die Ankunftszeit tof für diese Reflexion mit

$$tof_{pfad}(\vec{s}, \vec{x}, \vec{e}) = \frac{\sum_{i=1}^M (l_i)}{\bar{c}_{pfad}}. \quad (5.11)$$

Für die Simulation der Dämpfung auf dem Pfad wird in der Simulation für die einzelnen Medien ein Dämpfungswert α_i in $dB/cm/MHz$ definiert. Die Amplitude A ergibt sich, bei einer ausgesendeten Amplitude $A_0 = 1$, durch die Dämpfung att , die der Ultraschall auf dem Pfad durch die M Medien mit

$$A_{pfad} = A_0 \cdot att_{pfad}(\vec{s}, \vec{x}, \vec{e}) = 10^{-\frac{1}{20} \cdot \sum_{i=1}^M (\alpha_i \cdot l_i)} \quad (5.12)$$

Simulation von Reflexionen an Oberflächen

Reflexionen an Oberflächen sind deutlich komplexer zu berechnen als Reflexionen an idealen Punktstreuern, die in alle Richtungen isotrop abstrahlen. Je nach rauer oder glatter Oberflächenbeschaffenheit kann die Reflexion diffus, spiegelnd oder eine Mischung aus beidem sein. Die Amplitude des reflektierten Ultraschalls ist dabei von der Oberflächennormalen \vec{N} sowie von der Sender- und Empfängerposition \vec{s}, \vec{e} abhängig. Für die Simulation von Reflexionen an Oberflächen wird nach dem Phong-Beleuchtungsmodell vorgegangen (siehe Kapitel 3.2). Dabei muss die Oberfläche der simulierten Objekte abgetastet werden (siehe Abbildung 5.6 links). Damit erhält man mehrere Reflexionspunkte mit unterschiedlichen Reflexionsstärken (siehe Abbildung 5.6 rechts). Die einzelnen Reflexionspunkte auf den Oberflächen besitzen unterschiedliche Laufwege und Laufzeiten sowie unterschiedliche Ankunftszeiten am Empfänger. Somit muss jeder Strahl einzeln berechnet werden. Damit steigt die Rechenintensität mit der Anzahl R an ausgesendeten Strahlen. Die Berechnung der Ankunftszeit $tof_{pfad}(r)$ für einen Strahl r erfolgt analog zu dem Fall der Reflexion am Punktstreuer, wobei der Schnittpunkt zwischen Strahl und Objektoberfläche den Reflexionspunkt darstellt. Die Berechnung der Amplitude $A_{pfad}(r)$ der Reflexion eines Strahls wird nach Gleichung (9.6) und abhängig von den gewählten Materialparametern $k_{diffus}, k_{spiegelnd}, n$ durchgeführt, mit denen festgelegt wird, wie stark ein Objekt diffus bzw. spiegelnd streut. Durch das Integral über alle am Empfänger auftreffenden Strahlen und ihren Zeitpunkten ergibt sich das Signal für den Empfänger.

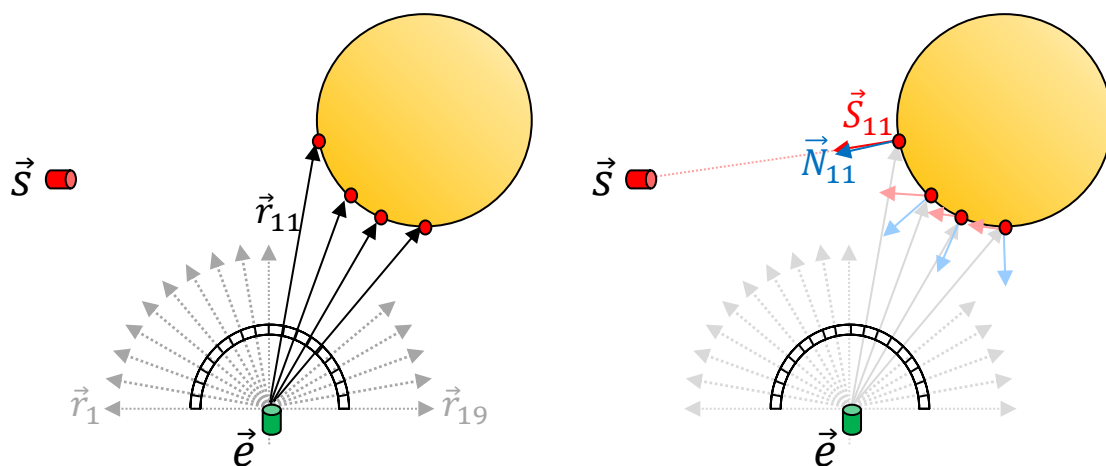


Abbildung 5.6: Schematische Darstellung in 2D der Rasterung des Empfängersichtfeldes bei der Simulation von Reflexionen an Objektflächen.

Simulation von A-Scans

Für ein SEP können sich $J = P + (M \cdot R)$ Reflexionen ergeben, wobei P die Anzahl der Punktstreuer, M die Anzahl der Objekte und R die Anzahl der vom Empfänger ausgesendeten Strahlen ist. Sind die Zeitpunkte $tof(j)$ und die Amplitudenwerte $A(j)$ der J Reflexionen berechnet, kann daraus der A-Scan generiert werden. Der Ablauf der A-Scan-Generierung ist in Abbildung 5.7 vereinfacht für eine Reflexion dargestellt. Dazu wird in dem A-Scan zu den Zeitpunkten $tof(j)$ ein Dirac-Impuls der Größe $A(j)$ eingetragen.

Anschließend wird der letzte Schritt der Signalvorverarbeitung nachgebildet, indem das Signal, bestehend aus Dirac-Impulsen, noch mit dem Optimalen Puls gefaltet wird. Dabei wird die Breite des Optimalen Pulses an die gewählte Auflösung für die Rekonstruktion angepasst (siehe Kapitel 3.1). Durch diesen Schritt wird davon ausgegangen, dass alle auftretenden Reflexionen durch die Signalvorverarbeitung eindeutig erkannt werden können. Dies ist jedoch bei Reflexionen, die zeitlich sehr eng beieinander liegen, abhängig von der genutzten Frequenz und Bandbreite nur begrenzt möglich. Dennoch wird in dieser Arbeit von einer perfekten Signalvorverarbeitung ausgegangen, da die Rekonstruktionsmethoden an sich bewertet werden sollen. Bei der Rekonstruktion mit realen Messdaten ist daher zu beachten, dass die Messsignale durch die Signalvorverarbeitung bereits Fehler enthalten können, die nicht mehr durch die SAFT-Rekonstruktion korrigiert werden können. Eine verbesserte Detektion und Trennbarkeit der einzelnen Reflexionssignale bei limitierter Bandbreite sind aktuelle Forschungsgebiete. So wird untersucht, ob durch den Einsatz von Huffman Codes beim Senden Verbesserungen erreicht werden können [89]. Als letzter Schritt bei der Simulation der A-Scans wird optional ein normalverteiltes Rauschen aufaddiert.

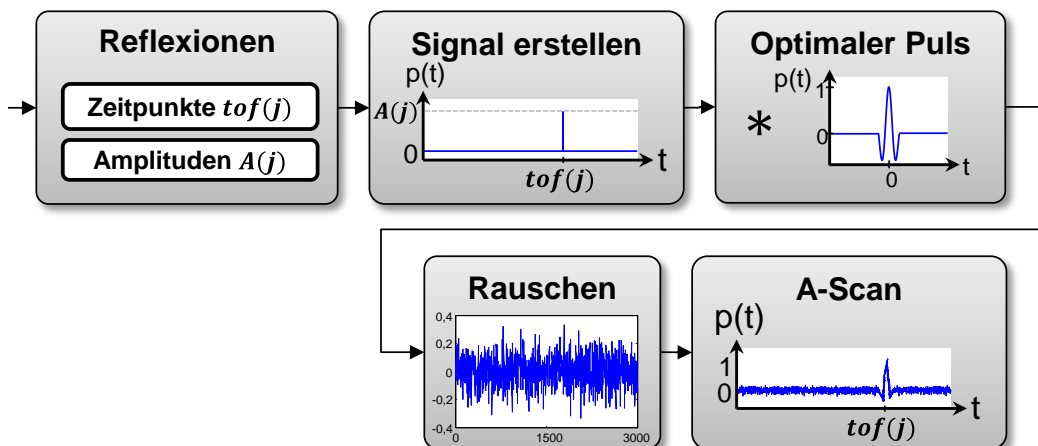


Abbildung 5.7: Überblick über den Ablauf der Simulation der A-Scan-Daten.

6. Reflexionsbildung mit SAFT für heterogene Medien

6.1 Korrektur der Annahme einer konstanten Schallgeschwindigkeit

Die 3D-Ultraschall-Computer-Tomographie (3D-USCT) verspricht hochauflösende Bilder für eine frühzeitige Diagnose von Brustkrebs. Der für die Rekonstruktion der Reflektivität eingesetzte SAFT-Algorithmus und die ihm zugrunde liegenden Annahmen wurden bereits in Kapitel 3.3 näher beschrieben. Dabei ist insbesondere die Annahme einer konstanten Schallgeschwindigkeit von großer Bedeutung, da es bei einer Veränderung der Schallgeschwindigkeit zu einer verschlechterten, bis hin zu gar keiner konstruktiven Überlagerung der ins Bild zurück projizierten Echosignalen kommt. Dieser Effekt verstärkt sich zudem bei kurzen Echosignalen bzw. hohen Frequenzen und verhindert insbesondere die Rekonstruktion von hochauflösenden Bildern.

Im Gegensatz zu der Annahme der Standard-SAFT, bei der die Schallgeschwindigkeit im 3D-USCT als konstant angesehen wird, unterscheidet sich die Schallgeschwindigkeit im Wasser deutlich von den Schallgeschwindigkeiten in der Brust (siehe Table 6.1). Zudem ist die Brust selbst ein heterogenes Medium, das unterschiedliche Gewebearten enthält. Des Weiteren ist die Schallgeschwindigkeit abhängig von der Temperatur und kann einen weiteren Fehler verursachen. Erste Untersuchungen zeigten bereits, dass hochauflösende Bilder nur möglich sind, wenn eine Schallgeschwindigkeitskorrektur durchgeführt wird [91].

Der 3D-USCT-Prototyp am KIT besitzt einen Aperturdurchmesser von 26 cm und eine Höhe von 16 cm. Die von einem Sender ausgesendete US-Welle kann dabei eine Distanz von bis zu maximal 52 cm zurücklegen (siehe Abbildung 6.1). Betrachtet man nur den möglichen Laufweg für die aktuelle ROI des 3D-USCTs, kann die maximale Distanz dagegen auf 46 cm beschränkt werden. Ausgehend von dem maximalen Laufweg kann der noch zulässige Fehler berechnet werden, mit dem die angestrebte

Tabelle 6.1: Verschiedene Schallgeschwindigkeiten für die 3D-USCT nach [30, 90]

Medium	Schallgeschwindigkeit
Wasser (20°)	1482 m/s
Wasser (26°)	1500 m/s
Fett	1460 m/s
Blut (37°)	1570 m/s
Drüsengewebe	ca. 1470 m/s
Krebs	ca. 1510 m/s
Muskel	1580 m/s
Weichgewebe im Mittel	1540 m/s

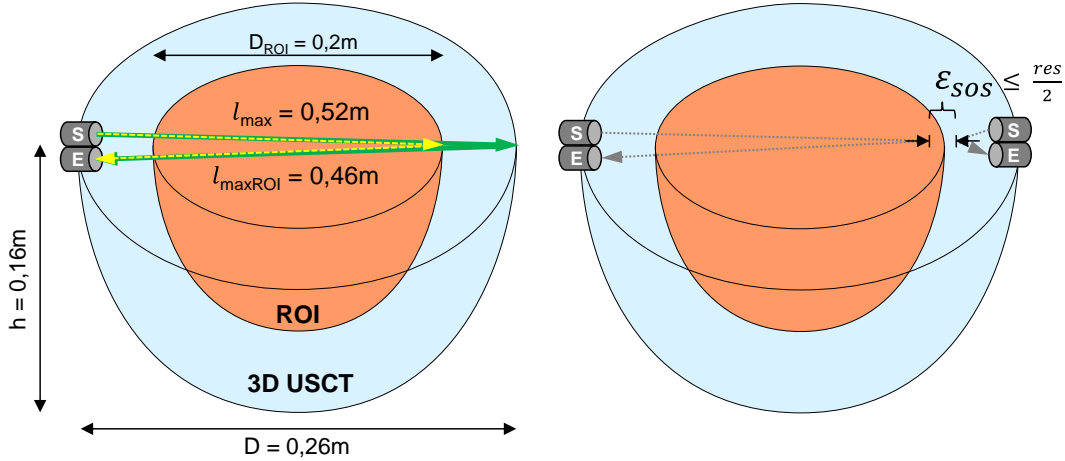


Abbildung 6.1: Der längste Pfad, den die Ultraschallwelle im 3D-USCT-Prototyp (links, grün) und bei Beschränkung auf die Brust (links, gelb) durchlaufen kann. Rechts ist die Berechnung der maximal erlaubten Fehlers \mathcal{E} skizziert.

Auflösung von $res = (0,24\text{ mm})^3$ erreicht werden kann. Die Laufzeit t , die der Ultraschall für einen Laufweg l_{max} bei einer angenommenen Schallgeschwindigkeit im Wasser von $c_{Wasser} = 1500\text{ m/s}$ benötigt, lässt sich durch

$$t = \frac{l_{max}}{c_{Wasser}} \quad (6.1)$$

berechnen. Eine fehlerhaft angenommene Schallgeschwindigkeit $c_2 = c_{Wasser} - \Delta c$ führt in dieser Laufzeit zu einer Abweichung \mathcal{E} , die nicht größer als die Hälfte der angestrebten Auflösung werden darf, damit sich die zurückprojizierten Echosignale im Bild noch konstruktiv überlagern:

$$\begin{aligned} t \cdot c_1 - t \cdot c_2 &\stackrel{!}{\leq} \frac{res}{2} \\ t \cdot c_{Wasser} - t \cdot (c_{Wasser} - \Delta c) &\leq \frac{res}{2} \\ t \cdot \Delta c &\leq \frac{res}{2} \\ \Delta c &\leq \frac{c_{Wasser} \cdot res}{l_{max} \cdot 2} \end{aligned} \quad (6.2)$$

Ausgehend von der maximalen Auflösung von $(0,24\text{ mm})^3$, die erreicht werden kann, bedeutet dies, dass eine maximale Abweichung auf einem Pfad von nur $\Delta c = 0,35\text{ m/s}$ für die gesamte Apertur, bzw. $\Delta c = 0,39\text{ m/s}$ für den Bereich der Brust erlaubt sind. In der 3D-USCT können jedoch Schallgeschwindigkeiten in einem weiten Bereich von ca. 1450 m/s bis 1550 m/s vorkommen, die um mehrere Größenordnungen größer sind als Δc . Damit wird deutlich, dass die Schallgeschwindigkeit einen starken Einfluss auf die Auflösung der Rekonstruktion besitzt. Für hochauflösende Bilder ist daher eine Korrektur der Schallgeschwindigkeit auf einem Pfad von essentieller Bedeutung.

6.1.1 Methodik und Umsetzung

Für die Korrektur der Schallgeschwindigkeit auf einem Pfad kann der Vorteil genutzt werden, den die 3D-USCT bietet. Da es mithilfe der Transmissionssignale möglich ist, neben den Reflexions- auch Schallgeschwindigkeits- und Absorptionsvolumen zu rekonstruieren [92], kann diese Information genutzt werden, um die Bildqualität deutlich zu erhöhen. Dabei ist zu beachten, dass für die 3D-USCT die Auflösung des Reflexionsvolumen mit $(0,24\text{ mm})^3$ ca. um den Faktor 10 höher liegt als die Auflösung der Schallgeschwindigkeitsvolumen (SOS-Volumen) mit $(2\text{ mm})^3$ [91]. Der Grund liegt darin, dass bei der Rekonstruktion der Transmissionsvolumen eine Strahlapproximation genutzt wird und der Wellencharakter des Ultraschalls, d.h. die Beugung und Brechung, vernachlässigt wird. Jedoch können die Schallgeschwindigkeitsvolumen trotz ihrer niedrigeren Auflösung genutzt werden, um genauere Schallgeschwindigkeitswerte für die Rekonstruktion der Reflektivität mit SAFT zu ermitteln.

Abhängig von dem Pfad, den der Ultraschall im Wasser und Messobjekt durchläuft, kann eine mittlere Schallgeschwindigkeit berechnet werden. Dafür muss zuerst der Pfad durch das diskrete SOS-Volumen bestimmt werden. Ein Pfad bezeichnet dabei den Weg vom Sender (\vec{s}) über den zu berechnenden Voxel (\vec{x}) zu dem Empfänger (\vec{e}) und kann mit dem Bresenham-Algorithmus [93] bestimmt werden. Abbildung 6.2 zeigt beispielhaft und schematisch die Berechnung der mittleren Schallgeschwindigkeit für zwei Voxel an den Positionen \vec{x}_1 und \vec{x}_2 innerhalb einer xy-Ebene. Der Pfad, den der Ultraschall von \vec{s} über \vec{x}_1 zu \vec{e} durchläuft, geht nur durch Wasser. Der Pfad über \vec{x}_2 dagegen läuft teilweise durch das Wasser und durch die Brust und ergibt somit eine andere mittlere Schallgeschwindigkeit.

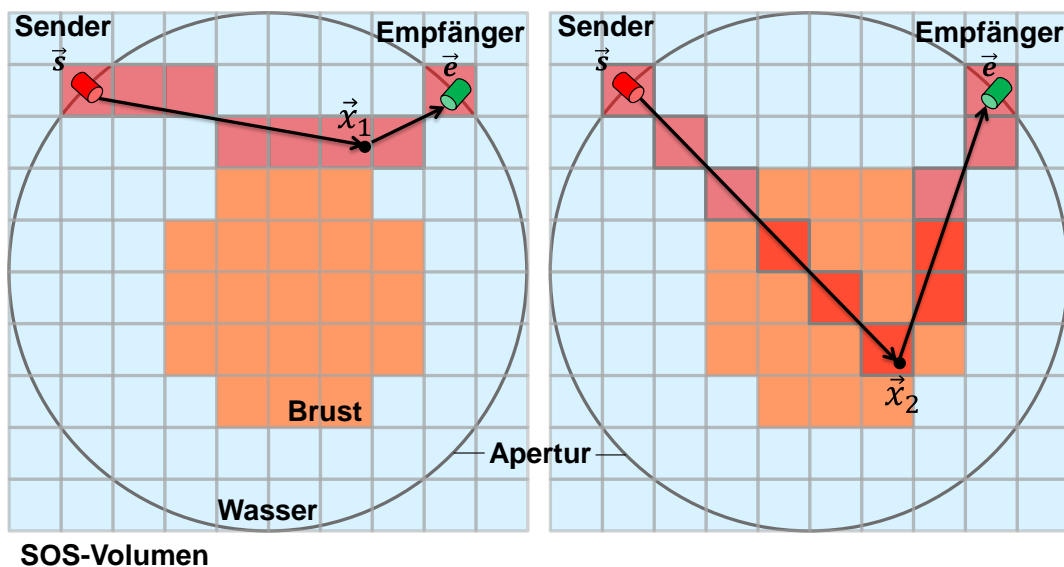


Abbildung 6.2: Illustration der Pfadbestimmung eine z-Schicht durch das SOS-Volumen mithilfe des Bresenham-Algorithmus für zwei Voxel bei \vec{x}_1 und \vec{x}_2 und einem SEP.

Die mittlere Schallgeschwindigkeit \bar{c}_P auf einen Pfad $P(\vec{s}, \vec{x}, \vec{e})$ durch das diskrete SOS-Volumen lässt sich mit dem harmonischen Mittel nach Gleichung (6.3) annähern.

$$\bar{c}_P(\vec{s}, \vec{x}, \vec{e}) = \frac{N}{\sum_{i=1}^N \frac{1}{c(\vec{x}_i)}} \quad (6.3)$$

Dabei bezeichnet N die auf dem Pfad besuchten SOS-Voxel und $c(\vec{x}_i)$ die lokalen Schallgeschwindigkeitswerte im SOS-Volumen. Um die Berechnung zu beschleunigen, wird die Streckenlänge durch die einzelnen SOS-Voxel als konstant angenommen, womit eine Gewichtung über die einzelnen Streckenlänge entfallen kann.

Diese approximierte mittlere Schallgeschwindigkeit $\bar{c}_P(\vec{s}, \vec{x}, \vec{e})$ muss für jede Sender-Voxel-Empfänger Kombination ermittelt werden und ermöglicht mit Gleichung (6.4) eine schallgeschwindigkeitskorrigierte SAFT-Rekonstruktion des Reflexionsvolumens.

$$I(\vec{x}) = \sum_{\forall(s,e)} AScan(s, e, tof_P), \quad tof_P = \frac{\|\vec{s} - \vec{x}\| + \|\vec{x} - \vec{e}\|}{\bar{c}_P(\vec{s}, \vec{x}, \vec{e})} \quad (6.4)$$

6.1.2 Evaluation

Im Folgenden wird untersucht, welchen Einfluss die Korrektur der Schallgeschwindigkeit auf die Qualität des rekonstruierten Volumens hat. Dazu werden in einer Simulation 21 Punktstreuer untersucht, die sich innerhalb einer Brust mit einem Durchmesser von 20 cm befinden (siehe Abbildung 6.3). Aufgrund der Symmetrie der 3D-USCT-Apertur werden nur Punktstreuer innerhalb eines Viertelkreises genutzt, die in einem 5x5 Raster angeordnet sind. Zueinander besitzen sie einen Abstand von 2,2 cm. Die Simulation hat zwei Bereiche mit unterschiedlichen Schallgeschwindigkeiten. Für das Wasser wird eine feste Schallgeschwindigkeit $c_{Wasser} = 1500 \text{ m/s}$

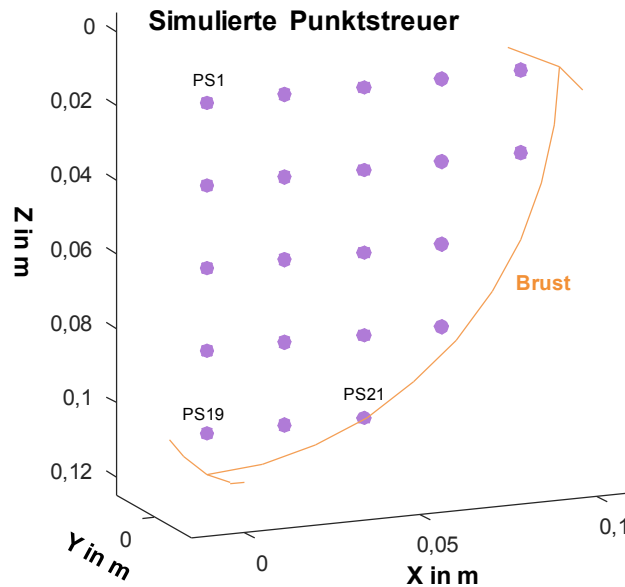


Abbildung 6.3: Anordnung der simulierten 21 Punktstreuer (lila) innerhalb der angedeuteten Brust (hautfarben).

angenommen. Die Schallgeschwindigkeit der Brust wird innerhalb eines realistischen Bereichs von 1460 m/s bis 1540 m/s variiert. Dies entspricht einem Schallgeschwindigkeitsunterschied $\Delta c = \pm 40\text{ m/s}$ zum Wasser. Zudem wird die Auflösung des SOS-Volumens, das für die Korrektur genutzt wird, über den Bereich von 16^3 bis 128^3 Voxeln untersucht. Dieser Bereich entspricht den SOS-Auflösungen von $(16, 2\text{ mm})^3$ bis $(2, 0\text{ mm})^3$. Für die Simulation und anschließende Rekonstruktion werden 10 AP genutzt, was einer Menge von 8873640 A-scans entspricht.

Für jeden Punktstreuer werden folgende drei quantitativen Metriken für die Bildqualität bestimmt, die bereits in Kapitel 5.2 eingeführt wurden:

- **Verschiebung:** Es wird der Abstand zwischen der simulierten Position des Punktstreuers und der Position gemessen, an der die maximale Amplitude im rekonstruierten Bild registriert wird.
- **Kontrast:** Es wird die maximale Amplitude des Voxels an der Position der Punktstreuer genutzt. Sie wird relativ zu der Referenz angegeben. Als Referenz wird dabei die maximale Amplitude verstanden, die mit dem Standard-SAFT rekonstruiert wird, wenn es keinen Unterschied der Schallgeschwindigkeiten in Wasser und Brust gibt.
- **Auflösung:** Für die Beschreibung der *Punktspreizfunktion* (PSF) wird die FWHM-Metrik in 3D berechnet und dargestellt.

Simulationsergebnisse

Abbildung 6.4 zeigt den Verlauf der ermittelten Bildqualitätmetriken, gemittelt über alle simulierten Punktstreuer. Ohne Korrektur ist der negative Einfluss auf die Bildqualität signifikant, sobald ein Unterschied der Schallgeschwindigkeiten zwischen Wasser und Brust auftritt. Dagegen bewirkt eine Korrektur mit Hilfe eines SOS-Volumens von nur 16^3 Voxel bzw. einer SOS-Auflösung von $(16, 2\text{ mm})^3$ bereits eine deutliche Verbesserung des Kontrasts (siehe Abbildung 6.4 oben). Ebenso wird die gemessene Verschiebung durch die SOS-Korrektur signifikant reduziert (siehe Abbildung 6.4 mittig). Die mittlere Verschiebung der Punktstreuer sinkt bei einer SOS-Auflösung von $(96, 2\text{ mm})^3$ auf $0,116\text{ mm}$ und liegt damit unterhalb der halben maximalen Auflösung von $(0, 24\text{ mm})^3$. Für höhere SOS-Auflösungen erreichen die Bildqualitätsmetriken ein Plateau und zeigen keine weitere Verbesserung. Auch die mittlere FWHM in 3D zeigt dasselbe Verhalten. Es fällt jedoch auf, dass die FWHM für einen Schallgeschwindigkeitsunterschied $\Delta c = \pm 40\text{ m/s}$ ohne Korrektur sogar eine bessere Auflösung zeigt als bei einem Schallgeschwindigkeitsunterschied $\Delta c = \pm 20\text{ m/s}$. Um diesen Verlauf zu verstehen, wird im Folgenden eine tiefere Analyse durchgeführt.

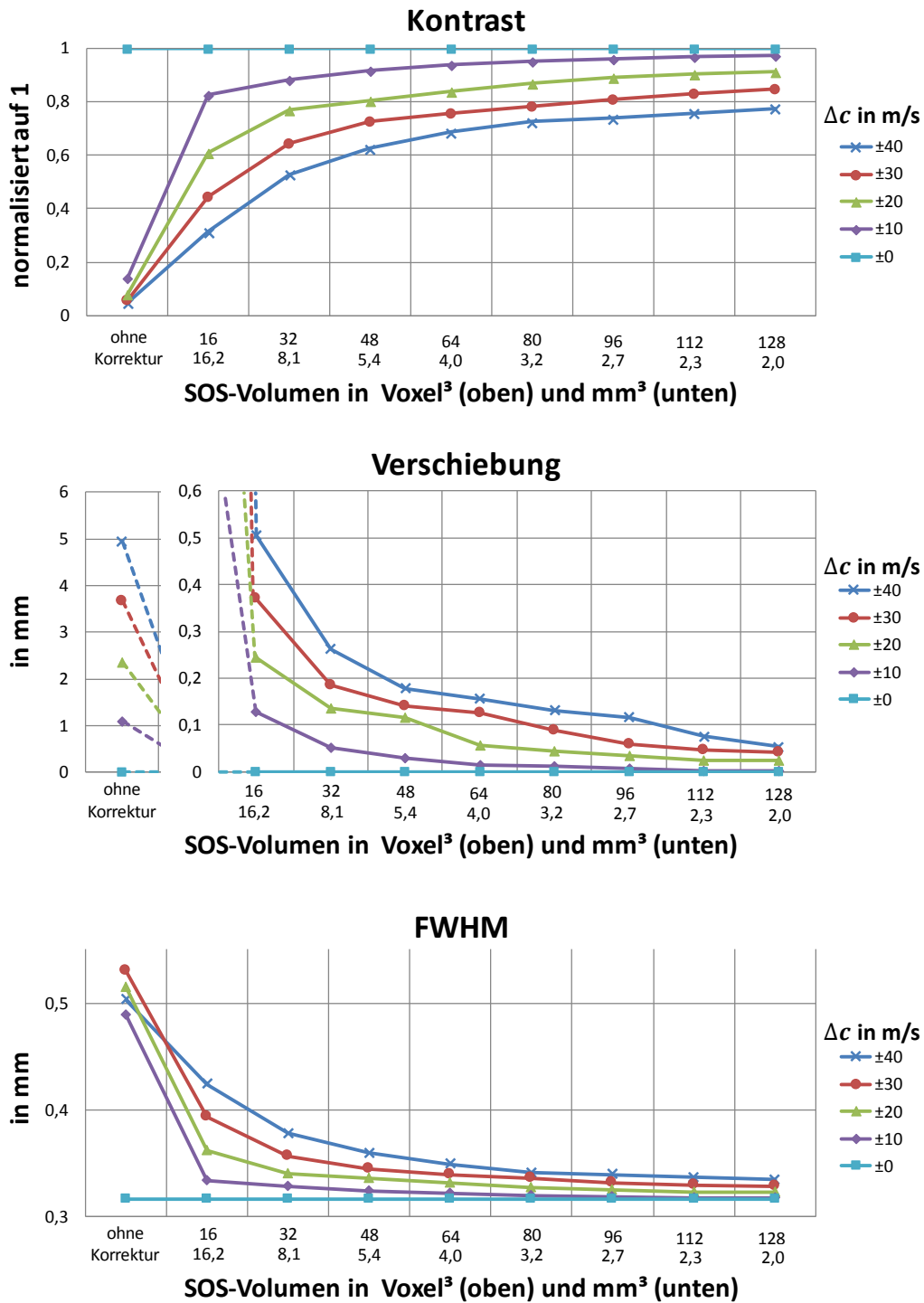


Abbildung 6.4: Ermittelte Werte der Bildqualitätsmetriken über alle 21 Punktstreuer für verschiedene Auflösungen der SOS-Volumen und Schallgeschwindigkeitsunterschiede Δc der Brust zum Wasser.

In Abbildung 6.5 ist die PSF für den Fall dargestellt, bei dem ein Schallgeschwindigkeitsunterschied von $\Delta c = \pm 40 \text{ m/s}$ gegeben ist. Es wird deutlich, dass die PSF ohne Korrektur der Phasenabweichung sehr stark von der Position der Punktstreuer abhängt. Des Weiteren ist die deutliche Verschiebung zu erkennen, bei dem jeweils das Maximum und damit der Punktstreuer detektiert wurde. Die mittlere Verschiebung beträgt $5,0 \text{ mm}$ und ist damit um den Faktor 21 höher als die Auflösung. Bei $\Delta c = \pm 40 \text{ m/s}$ beträgt der mittlere Kontrast $4,5\%$ und die mittlere Auflösung liegt bei $(0,5 \text{ mm})^3$. Im Schlechtesten Fall beträgt die Verschiebung $7,0 \text{ mm}$, der Kontrast sinkt auf nur $0,4\%$ und die PSF weitet sich bis auf $0,8 \text{ mm}$.

Für die Bewertung der PSF wurden die Volumen um den simulierten Punktstreuer rekonstruiert. In Abbildung 6.6 sind beispielhaft die xy - und xz -Schnittbilder jeweils für die Punktstreuer PS1, PS19 und PS21 dargestellt. Die Schnittbilder laufen dabei durch das jeweilige Maximum des rekonstruierten Volumens, an der die Position des Punktstreuers detektiert wird. Bei allen drei Punktstreuern wird deutlich, dass, verursacht durch den Schallgeschwindigkeitsunterschied innerhalb der Brust, sich nicht mehr alle Echosignale in einem Punkt konstruktiv überlagern. Das führt zu breit verstreuten Beiträgen der zurückprojizierten Echosignale im rekonstruierten Volumen, die sich sowohl konstruktiv als auch destruktiv überlagern können. Die Folge davon ist, dass nicht mehr nur ein eindeutiges Maxima im Volumen erzeugt wird, sondern je nach Schallgeschwindigkeit und Position eine Vielzahl von lokalen Maxima entstehen können, wie z.B. bei PS1 und PS21 in der xz -Schicht. Da bei der

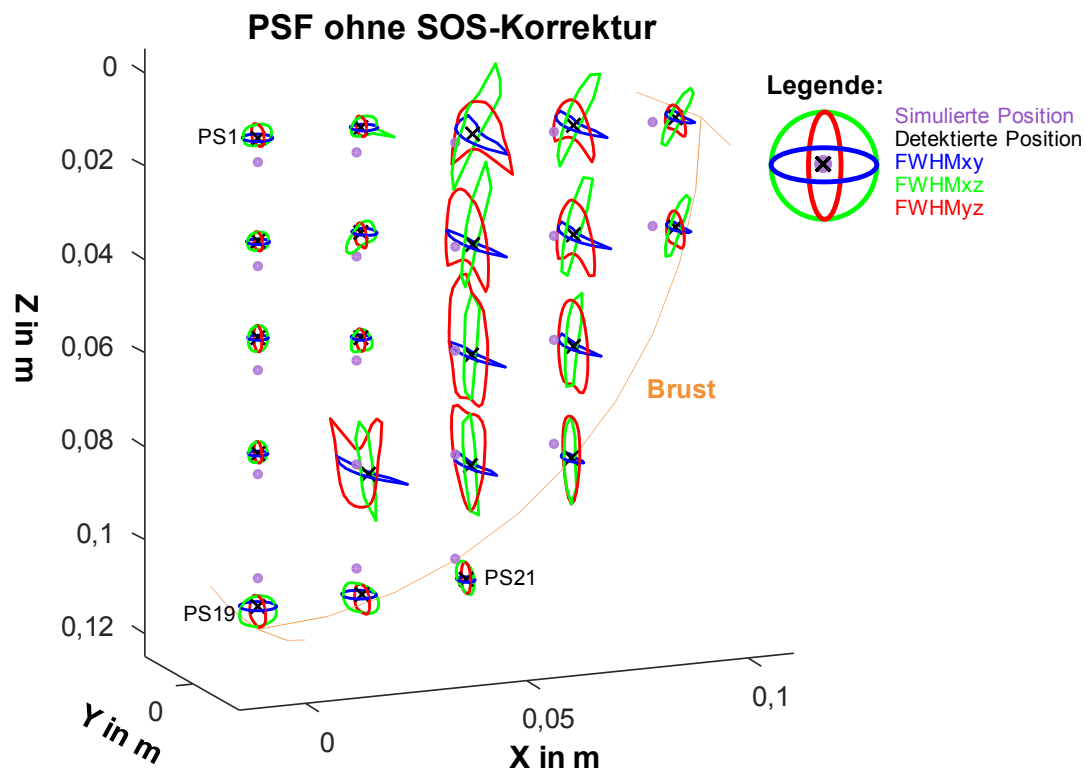


Abbildung 6.5: PSF der simulierten 21 Punktstreuer in 3D und Verschiebung von simulierter Position (lila) und detektierter Position (schwarz) ohne SOS-Korrektur. Die PSF ist bei $\Delta c = \pm 40 \text{ m/s}$ und um einen Faktor 30 vergrößert dargestellt.

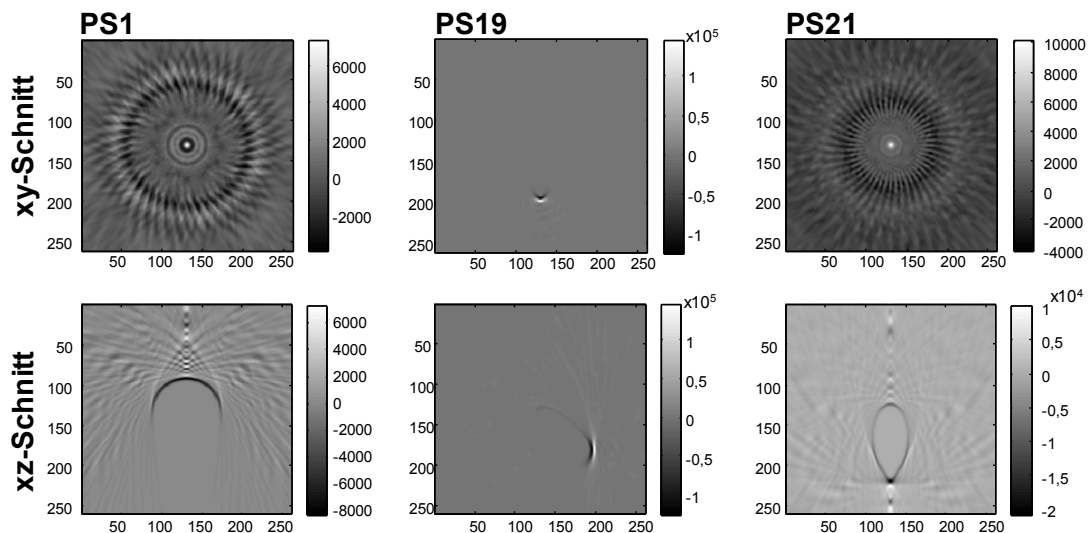


Abbildung 6.6: xy- und xz-Schnittbilder bei $\Delta c = \pm 40 \text{ m/s}$ ohne SOS-Korrektur für die simulierten Punktstreuer PS1, PS19 und PS21 (siehe Abbildung 6.5).

Bestimmung der PSF zuerst nach dem Maximum gesucht wird, wird die PSF nur an der Position des lokalen Maximums berechnet. Dies verdeutlicht, dass neben der PSF auch die Verschiebung und die maximale Amplitude mit in die Bewertung der Bildqualität und Auflösung einbezogen werden müssen. Dies gilt insbesondere für den Fall von größeren Unterschieden in den Schallgeschwindigkeiten. Die maximale Amplitude ist dabei ein guter Indikator dafür, wie viele Echosignale sich konstruktiv in einem Punkt überlagern. Die Punktstreuer PS19 und PS21 zeigen eine knapp doppelt so große Amplitude wie PS1. Da beide sehr nah an der Brustkante positioniert sind, wird für die meisten Echosignale nur ein geringer Fehler bezüglich der Schallgeschwindigkeit gemacht und es kommt zu einer größtenteils konstruktiven Überlagerung. Ein weiterer Effekt wird durch Anordnung der Punktstreuer bei PS1 und PS19 sichtbar. Da sie sich in der Mitte der Apertur befinden, kommt es zu symmetrischen Rekonstruktionsartefakten, die insbesondere im xy-Schnitt hohe Amplituden erreichen.

Wird dagegen die SOS-Korrektur mit einer SOS-Auflösung von 128^3 Voxeln bzw. einer SOS-Auflösung von $(2,0 \text{ mm})^3$ bei der Rekonstruktion eingesetzt, verbessert sich die Bildqualität signifikant. In Abbildung 6.7 ist die PSF wiederum bei einem Schallgeschwindigkeitsunterschied von $\pm 40 \text{ m/s}$ dargestellt. Eine Verschiebung der detektierten Punktstreuer ist bei einer mittleren Verschiebung von $0,05 \text{ mm}$ nicht mehr auszumachen und liegt unterhalb der Auflösungsgrenze. Dies bedeutet, dass es lediglich ein eindeutiges Maximum für jeden Punktstreuer gibt. Dies ist auch in den rekonstruierten Volumen für die drei Punktstreuer PS1, PS19 und PS21 deutlich erkennbar (siehe Abbildung 6.8). Bei $\Delta c = \pm 40 \text{ m/s}$ liegt der mittlere Kontrast bei $77,5\%$ und die mittlere Auflösung bei $(0,33 \text{ mm})^3$. Mit SOS-Korrektur beträgt die Verschiebung im schlechtesten Fall nur noch $0,2 \text{ mm}$ und liegt damit noch unter der maximalen Auflösung von $(0,24 \text{ mm})^3$, der Kontrast sinkt auf nur 29% und die PSF weitet sich bis auf $0,4 \text{ mm}$.

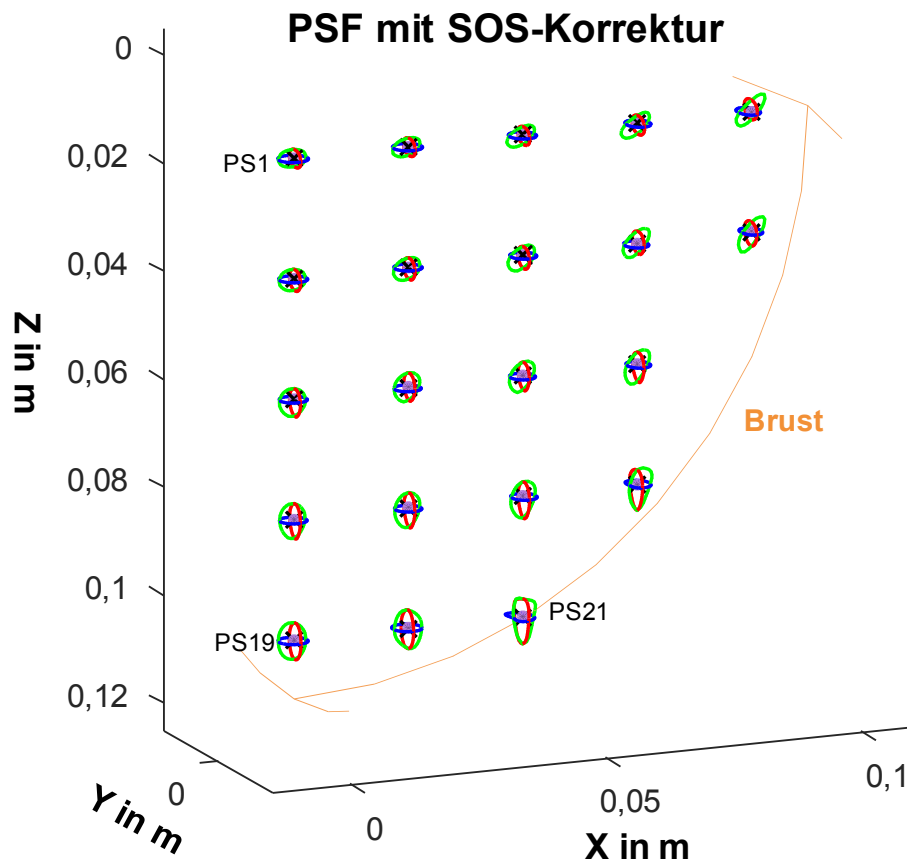


Abbildung 6.7: PSF der simulierten 21 Punktstreuer in 3D mit SOS-Korrektur. Die PSF ist bei $\Delta c = \pm 40 \text{ m/s}$ und um einen Faktor 30 vergrößert dargestellt.

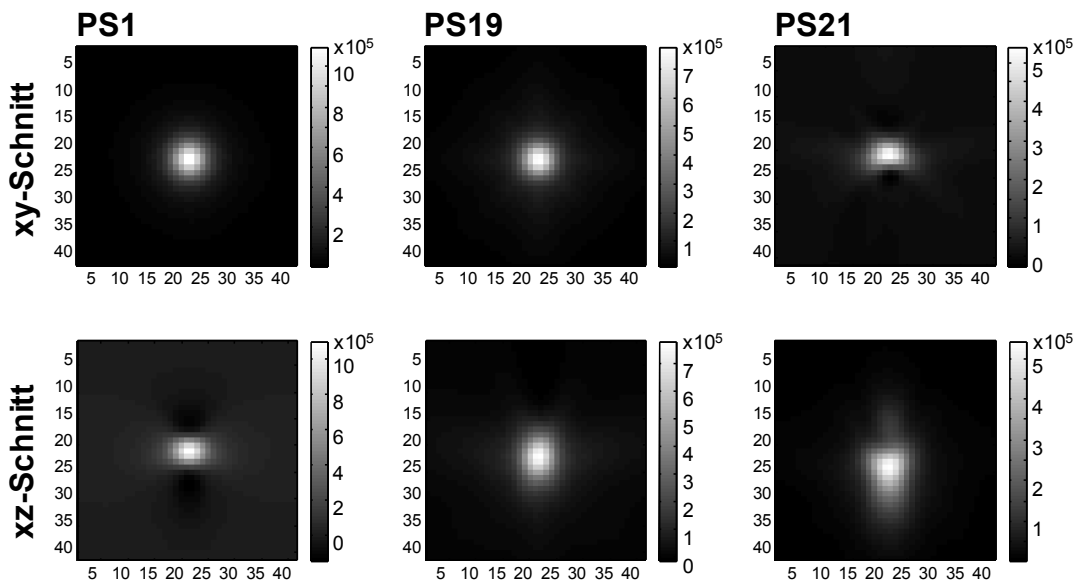


Abbildung 6.8: xy- und xz-Schnittbilder bei $\Delta c = \pm 40 \text{ m/s}$ mit SOS-Korrektur für die simulierten Punktstreuer PS1, PS19 und PS21 (siehe Abbildung 6.7).

Die Untersuchung der Bildqualität wurde bisher mit Rekonstruktionen durchgeführt, für die alle A-Scans der 10AP voll genutzt wurden. Wird dagegen wie in Kapitel 3.3 beschrieben nur die Hälfte der A-Scans genutzt, kann die Auflösung weiter verbessert werden. Die Untersuchung wurde daher für den Winkelbereich von 45° - 90° wiederholt (siehe Abbildung 3.14). Die mittlere Verschiebung erhöht sich dabei leicht von 0,05 mm auf 0,06 mm, der mittlere Kontrast liegt bei 46,8% und die mittlere Auflösung bei $(0,28 \text{ mm})^3$. Die geringe Verschlechterung in der Verschiebung begründet sich durch die geringere Anzahl an A-Scans, die bei der Rekonstruktion genutzt werden. Da für den mittleren Kontrast bei halber Menge an A-scans noch ein Faktor 2 für einen fairen Vergleich eingerechnet werden muss, liegt dieser bei 93,6% und erreicht hier ein um 16,1% besseres Ergebnis.

Ergebnisse mit experimentellen Messdaten

Neben der Untersuchung mit Simulationsdaten wurde die Schallgeschwindigkeitskorrektur qualitativ auch an experimentellen Messdaten untersucht. Dafür wurde ein 0,2 mm dicker Bondingdraht in einen vertikalen Gelatinezylinder eingebettet (siehe Abbildung 6.9, links). In Abbildung 6.9 rechts sind entsprechend die rekonstruierten xy-Schnittbilder der Standard-SAFT und der SAFT mit SOS-Korrektur dargestellt. Zudem wird ein Bereich, in dem sich der Draht und die Kante des Gelatinephantoms befinden, vergrößert dargestellt.

Bei der Rekonstruktion mit der Standard-SAFT ohne SOS-Korrektur erscheint der Draht kreisförmig statt punktförmig (siehe Abbildung 6.9, oben rechts). Ebenso wird der Rand des Gelatinephantoms doppelt dargestellt. Beide Abbildungsfehler sind Folgen einer unfokussierten Abbildung aufgrund der falschen Annahme, dass es keine Schallgeschwindigkeitsunterschiede im Medium gibt. Wird dagegen die SOS-Korrektur genutzt, können diese Abbildungsfehler, wie in Abbildung 6.9 unten rechts dargestellt, korrigiert werden. Der Draht wird korrekt, als punkt- bzw. zylinderförmiger Draht, abgebildet. Ebenso erscheint die Oberfläche des Gelatine-Phantoms nun als eine Oberfläche und wird nicht mehr doppelt dargestellt. Für die Rekonstruktion mit SOS-Korrektur wurde ein SOS-Volumen der Auflösung von 128^3 Voxel genutzt. Als Kontrastmaß wurde das PSNR nach Gleichung (5.7) berechnet. Der Kontrast verdreifacht sich dabei nahezu von 24,3 auf 70,1 im Vergleich zu dem rekonstruierten Bild ohne Korrektur.

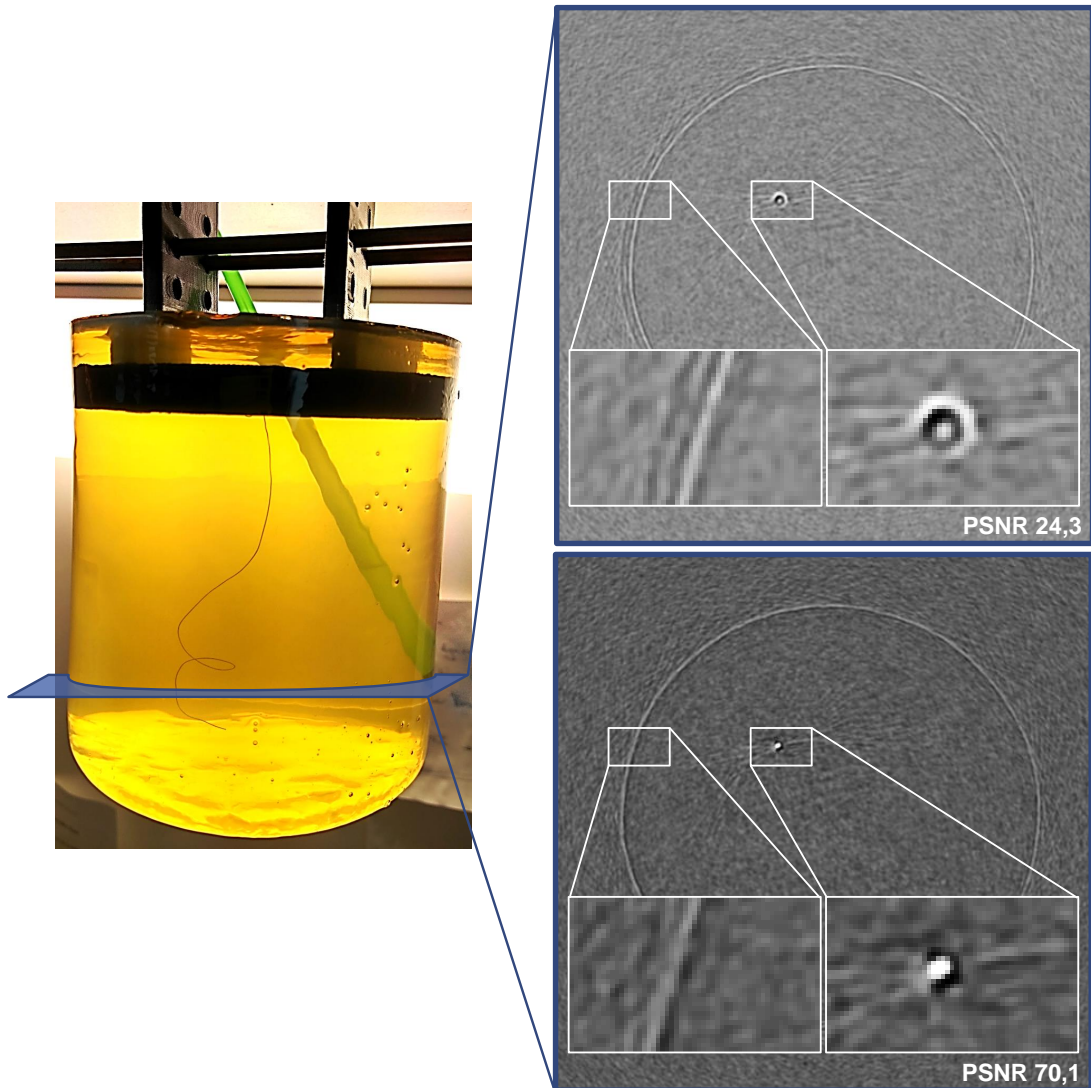


Abbildung 6.9: Bondingdraht mit Durchmesser 0,2 mm in einem Gelatinezylinder (links). Rekonstruktion einer Schicht ohne (oben rechts) und mit SOS-Korrektur (unten rechts).

6.2 Korrektur der Annahme einer konstanten Dämpfung

Neben der Annahme einer konstanten Schallgeschwindigkeit wird beim einfachen Standard-SAFT-Algorithmus vorausgesetzt, dass durch das Medium eine konstante Dämpfung oder Absorption verursacht wird. Tatsächlich hat Wasser eine relativ geringe Dämpfung ($0,002 \text{ dB/cm/MHz}$). Die Brust dagegen besitzt eine deutlich höhere Dämpfung ($0,63 \text{ dB/cm/MHz}$) und kann daher zu einer lokalen Abschattung von inneren Regionen der Brust führen (siehe Abbildung 2.5). Zudem wird an der Grenzschicht von Wasser zur Brust ein Teil der US-Welle reflektiert. Dies entspricht praktisch einer hohen Dämpfung, da nur noch ein Teil der US-Welle in die Brust eindringen kann. Um ein gleichmäßiges Bild insbesondere der inneren Regionen der Brust bezüglich Ausleuchtung und Kontrast zu bekommen, wird eine Dämpfungskorrektur benötigt.

Die Dämpfung beschreibt die Abnahme der Amplitude bzw. der Leistung der US-Welle infolge von Absorption, Reflexionen, Streuung und Beugung, während sich eine US-Welle im Medium ausbreitet [67]. Der stärkste Dämpfungseffekt ist die Umwandlung der Wellenenergie in Wärme. Sie wird während der Wellenausbreitung aufgrund von Viskositäts- und Relaxationsverlusten im Medium verursacht. Ebenso reduzieren Streuungen im Gewebe an kleineren Objekten und Reflexionen an den Grenzen zweier Gewebearten mit unterschiedlichen Schallwiderständen die Energie der US-Welle, die bei einem Empfänger zu messen ist. Dabei kann es zu einer teilweisen Reflexion oder auch einer totalen Reflexion der US-Welle kommen. Dieser Effekt trägt zur Dämpfung ca. mit einem Anteil von 10-15% bei [67]. Die Stärke der Dämpfung wird daher von den Eigenschaften des Mediums bzw. Gewebes bestimmt, das die US-Welle durchwandert. Die Abnahme der Amplitude durch die Dämpfung kann allgemein nach dem Lambert-Beerschen Gesetz als eine Funktion in der Form

$$A = A_0 e^{-\alpha l} \quad (6.5)$$

beschrieben werden. Wobei l die Distanz ist, die der US durch das Medium mit dem Dämpfungskoeffizient α zurückgelegt hat. Empirische Untersuchungen in [94, 95] zeigen, dass der Dämpfungskoeffizient α von der Frequenz f abhängig ist und mit steigender Frequenz ebenfalls zunimmt. Es gilt daher:

$$\alpha = \alpha_0 f^n, \quad 1 \leq n \leq 2 \quad (6.6)$$

Für die üblicherweise in der medizinischen Ultraschallbildgebung genutzten Frequenzen konnte gezeigt werden, dass sich der Dämpfungskoeffizient für die meisten Gewebearten linear proportional zur genutzten Frequenz verhält [26]. Daher wird häufig bei Weichgewebe von $n = 1$ ausgegangen. Der Dämpfungskoeffizient wird in der Literatur meist in der Einheit dB/cm angegeben. Daher wird noch eine Umrechnung von der Einheit Neper benötigt. Mit $1 \text{ dB} = 1 \text{ Np} \cdot \ln(10)/20$ kann Gleichung (6.5) umgeschrieben werden zu:

$$A = A_0 10^{-\frac{1}{20} \alpha l} \quad (6.7)$$

Eine Übersicht der Dämpfungskoeffizienten für verschiedene Gewebearten sind in Tabelle 6.2 nach [26, 67, 96] angegeben.

Tabelle 6.2: Dämpfungskoeffizient für verschiedene Gewebetypen für die 3D-USCT nach [26, 67, 96]

Gewebetyp	Dämpfungskoeffizient α_0 bei 1MHz
Wasser	0,0022 dB/cm
Blut	0,18 dB/cm
Fett	0,63 dB/cm
Muskel	1,3-3,3 dB/cm
Weichgewebe	0,3-0,8 dB/cm

Bei der Rekonstruktion mit der Standard-SAFT wurde bisher davon ausgegangen, dass die Dämpfung des US im gesamten 3D-USCT konstant bzw. keine Dämpfung vorhanden ist. Gegensätzlich zu der Annahme der Standard-SAFT, ist das zu untersuchende Medium heterogen. Z.B. unterscheidet sich der Dämpfungskoeffizient im Wasser mit $0,002 \text{ dB/cm/MHz}$ deutlich von dem Dämpfungskoeffizient in der Brust, der gut mit der Dämpfung in Fett von ca. $0,63 \text{ dB/cm/MHz}$ angenähert werden kann. Eine einfache Abschätzung macht den Einfluss der Dämpfung durch ein heterogenes Medium deutlich. Die Öffnung des 3D-USCT-Prototyps hat einen Durchmesser von 26 cm. Wird die Brust mit dem Radius $r = 10 \text{ cm}$ angenommen, läuft der US bei der Strecke von einem Sender bis zur Mitte der Brust und zurück zu einem Empfänger durch 12 cm Wasser und 20 cm Brust. Die Amplitude des US erfährt dabei eine Dämpfung auf 16% der ausgesendeten Amplitude und führt ohne Kompensation zu einer verringerten Bildqualität in den Reflexionsbildern. Dies verdeutlicht die Notwendigkeit einer Dämpfungskorrektur für heterogene Medien, um ein gleichmäßiges Bild insbesondere der tieferen Regionen der Brust bezüglich Ausleuchtung und Kontrast zu erreichen.

6.2.1 Methodik und Umsetzung

Die 3D-USCT bietet neben der Rekonstruktion von SOS-Volumen auch die Möglichkeit Dämpfungsvolumen der Brust zu rekonstruieren [97]. Die Rekonstruktion der Dämpfungsvolumen geschieht dabei ebenfalls durch die Transmissionstomographie. Jedoch wird hier anstatt der benötigten Laufzeit, wie bei der Schallgeschwindigkeitsrekonstruktion, die Amplitude ausgewertet. Ist die Information über die Dämpfungseigenschaften des Volumen bekannt, kann diese für die Korrektur der Reflexionsvolumen in 3D angewandt werden.

Die Dämpfung, die der US erfährt, ist von dem Pfad abhängig, den der US durch das Objekt durchläuft. Für eine Dämpfungskorrektur wird daher, wie bei der Schallgeschwindigkeitskorrektur, für jeden Voxel und jedes SEP ein eigener Korrekturfaktor benötigt. Ist die Dämpfung att_P auf dem Pfad $P(\vec{s}, \vec{x}, \vec{e})$ bekannt, kann diese als Korrekturfaktor folgendermaßen in den SAFT-Algorithmus nach Gleichung (3.15) eingesetzt werden:

$$I(\vec{x}) = \sum_{\forall(s,e)} \left(\frac{1}{att_P} \cdot AScan(s, e, tof_P) \right) \quad (6.8)$$

Für die Schallgeschwindigkeitskorrektur werden die Pfade durch das SOS-Volumen bereits mit dem Bresenham-Algorithmus [93] bestimmt. Das Dämpfungsvolumen

(ATT-Volumen) besitzt aufgrund der gleichen Rekonstruktionsmethode die gleiche Auflösung wie das SOS-Volumen. Daher kann neben der mittleren Schallgeschwindigkeit zusätzlich die Dämpfung auf dem Pfad direkt mitbestimmt werden und steht somit als Korrekturfaktor für die Amplitude $A_{Scan}(s, e, tof_P)$ zur Verfügung. Die Dämpfung auf dem Pfad wird somit durch

$$att_P = att_{bres}(\vec{s}, \vec{x}, \vec{e}) = 10 \cdot \frac{1}{20} \cdot \frac{l_{Ppath}}{N} \cdot \sum_{i=1}^N \alpha_i \quad (6.9)$$

approximiert. Dazu wird der Mittelwert über die lokalen Dämpfungskoeffizienten α_i von N Voxeln des ATT-Volumens auf dem Pfad berechnet und mit der jeweiligen Pfad-Länge l_P gewichtet. Um die Berechnung zu beschleunigen, werden die Strecken, die der US in den einzelnen ATT-Voxel durchläuft, als konstant angenommen, womit eine Gewichtung über die einzelnen Strecken entfallen kann. Die aus dem A-Scan geladene Amplitude $A_{Scan}(s, e, tof_P)$ kann somit korrigiert werden, bevor sie mit der SAFT für einen Voxel aufsummiert wird.

6.2.2 Evaluation

Im Folgenden wird der Einfluss untersucht, den die Dämpfungskorrektur auf die Qualität des rekonstruierten Volumens hat. Für eine quantitative Bewertung der Dämpfungskorrektur wird eine Simulation in Kombination mit drei Metriken genutzt. Es werden die Metriken der räumlich homogenen Objektausleuchtung (RHOA) und des Kontrasts genutzt, die bereits in Kapitel 5.2 eingeführt wurden. Zusätzlich wird der absolute Fehler pro Voxel berechnet, der bei der Berechnung des Korrekturfaktors für die Dämpfung mit dem Bresenham-Algorithmus auftritt. Darüber hinaus wird die Dämpfungskorrektur anschließend anhand von in-vivo Daten evaluiert.

Der Gewinn der Dämpfungskorrektur für die Bildqualität wird mit einer Simulation quantifiziert. Um den Einfluss von Rauschen abschätzen zu können, wird den simulierten A-Scans ein normalverteiltes Rauschen mit Mittelwert 0 und Standardabweichungen in unterschiedlicher Höhe hinzugefügt. Die Anordnung der durchgeführten Simulation ist in Abbildung 6.10 links dargestellt. Sie besteht aus mehreren Objekten mit unterschiedlichen Eigenschaften (Größe, Dämpfungskoeffizient, Position), um einen realistischen Fall darzustellen. Die Dämpfung im Wasser wird mit $\alpha_{Wasser}=0,002$ dB/cm/MHz angenommen. Die Brust wird durch eine Halbkugel mit 10 cm Durchmesser approximiert. Für die Dämpfung der Brust wird der Wert für fetthaltiges Gewebe $\alpha_{Brust}=0,5$ dB/cm/MHz genutzt. Zusätzlich sind drei Objekte mit unterschiedlicher Größe und Dämpfungskoeffizienten eingefügt. Zwei Läsionen besitzen mit $\alpha_{Läsion1}=0,8$ dB/cm/MHz und $\alpha_{Läsion2}=0,7$ dB/cm/MHz höhere Dämpfungskoeffizienten als die Brust. Eine Zyste dagegen ist mit Flüssigkeit gefüllt und besitzt mit $\alpha_{Zyste}=0,009$ dB/cm/MHz einen ähnlich geringen Dämpfungskoeffizienten wie das Wasser. Um die Homogenität der Reflektivität abschätzen zu können, sind mehrere Punktstreuer gleichmäßig in einem Raster von 15^3 Punktstreuer angeordnet. Die US-Wandler der KIT-3D-USCT besitzen eine Mittenfrequenz f_0 von 2,5 MHz und ca. 50% Bandbreite. Daher wird die Dämpfung mit $\alpha = \alpha_0 \cdot f_0$ in dB/cm mit der Mittenfrequenz $f_0 = 2,5$ MHz approximiert. In Abbildung 6.10, rechts ist der gedämpfte Amplitudenverlauf über einen exemplarischen Pfad für eine SEP und einen Punktstreuer aufgezeigt. Dieser ist auch in der Simulationsanordnung Abbildung 6.10 links dargestellt. Der Ultraschall wird vom Sender ins Wasser ausgestrahlt

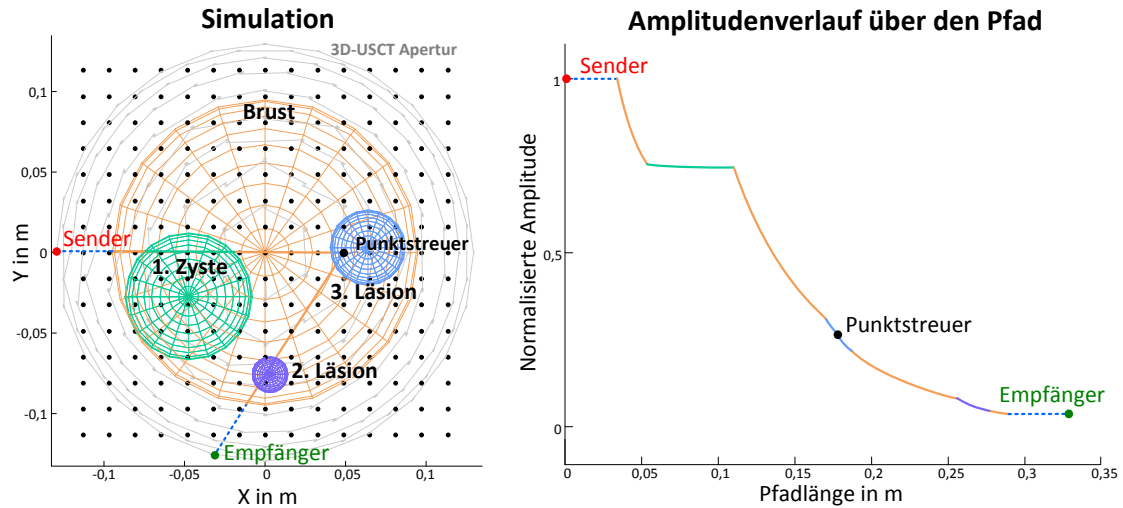


Abbildung 6.10: Simulation mit Läsionen innerhalb der Brust (links) und die Dämpfung auf einem exemplarischen Pfad (rechts).

und erfährt dort eine sehr geringe Dämpfung. Die Brust besitzt einen höheren Dämpfungskoeffizienten, wodurch die Ultraschallamplitude deutlich abnimmt. Durchläuft der Ultraschall anschließend die Zyste, erfährt dieser wiederum eine geringe Dämpfung. Der weitere Pfad geht wiederum durch die Brust und eine Läsion, wo er von dem Punktstreuer in Richtung Empfänger reflektiert wird. Ist der Ultraschall am Empfänger angekommen, besitzt er somit eine reduzierte Amplitude, die der Dämpfung auf dem Pfad entspricht.

Kontrasterhaltende Dämpfungskorrektur

Das Ziel der Dämpfungskorrektur ist eine räumlich homogene Reflektivität über das gesamte Volumen. Dies wird, wie beschrieben, durch das Verstärken der Amplitude erreicht, indem abhängig vom durchlaufenen Pfad des US ein eigener Korrekturfaktor genutzt wird. Durch diese ortsabhängige Verstärkung wird allerdings neben dem Signal auch das Hintergrundrauschen mit angehoben. Dadurch kann eine mehr oder weniger inhomogene Verteilung des PSNR über das Volumen auftreten. Um ein hohes PSNR zu gewährleisten, wurde eine einfache Schwellwerttechnik eingeführt, die den Korrekturfaktor für die Dämpfung auf einen maximalen Korrekturfaktor beschränkt: $att_P = \min(att_P, att_{Schwellwert})$. Der geeignete Schwellwert für den Korrekturfaktor hängt von der Größe und der Dämpfung des zu messenden Objekts ab und ist daher unbekannt. Durch eine Analyse der A-Scans kann jedoch für jedes Objekt individuell vor der Rekonstruktion ein geeigneter Schwellwert bestimmt werden. Für einen geeigneten Schwellwert wurde das Verhältnis aus der maximalen Amplitude des Signals zu der maximalen Amplitude des Rauschens getestet. Dieses Spitzen-Signal-zu-Spitzen-Rausch-Verhältnis wird im Folgenden mit *peak-signal-to-peak-noise ratio* (PSPNR) bezeichnet. Das PSPNR wird für jeden A-Scan nach der Vorverarbeitung gemessen. Dazu wird das Signal im Bereich vor dem Transmissionspuls untersucht, das nur Rauschen enthalten kann, und so der maximale Amplitudenwert des Rauschens bestimmt. Das PSPNR wird somit durch das Verhältnis zwischen maximalem Amplitudenwert im Signal nach dem Transmissionspuls zum maximalen Amplitudenwert des Rauschens bestimmt. Abbildung 6.11 zeigt die Verteilung des PSPNR-Verhältnisses über alle A-Scans für eine Beispielpatientin. Das

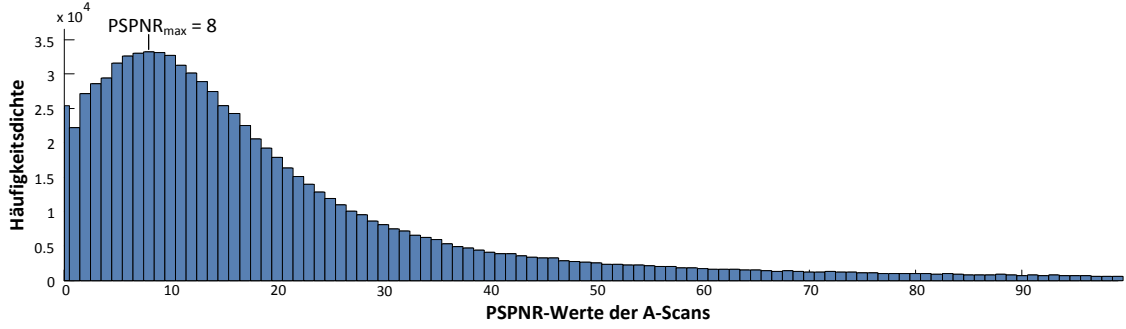


Abbildung 6.11: Histogramm der PSPNR-Werte über alle A-Scans der in-vivo Daten für eine Beispielpatientin. Das Histogramm ist auf den Bereich bis 100 begrenzt und enthält 95 % der Daten. Der beobachtete maximale PSPNR liegt bei $1,9 \cdot 10^5$.

Histogramm ist wegen der besseren Lesbarkeit bis zu einem PSPNR-Wert von 100 dargestellt und enthält 95% der Daten. Der maximale PSPNR-Wert beträgt $1,9 \cdot 10^5$. Es ist erkennbar, dass die meisten A-Scans einen $\text{PSPNR} = 8$ besitzen. Dieser Wert wird im folgenden PSPNR_{max} bezeichnet. Dabei besitzen 24,4 % der A-scans ein $\text{PSPNR} \leq \text{PSPNR}_{max}$ und 49,1 % der A-scans ein $\text{PSPNR} \leq 2 \cdot \text{PSPNR}_{max}$. Wird dieser PSPNR_{max} als Schwellwert für die Dämpfungskorrektur genutzt, wird auch für die in-vivo Daten ein guter Kontrast erreicht. Darüber hinaus wurden die Daten der ersten klinischen Studie analysiert, um den Bereich der PSPNR_{max} Werte von allen gemessenen Brüsten zu bestimmen. Der Mittelwert der PSPNR_{max} Werte liegt bei 7,6 mit einer Standardabweichung von 2,1. Der PSPNR_{max} liegt dabei für alle Messungen in dem Bereich von 5,0 bis 13,0.

Ergebnisse

Absoluter Fehler:

Mit der Simulation kann pro Pfad die Dämpfung analytisch mit Gleichung (6.10) berechnet werden und als Grundwahrheit genutzt werden. Dazu werden zuvor die Längen l_i der M Teilpfade auf dem Pfad berechnet (siehe Abbildung 6.10 rechts).

$$\text{att}_{analytic}(\vec{e}, \vec{r}, \vec{x}) = 10^{-\frac{1}{20} \cdot \sum_{i=1}^M (\alpha_i \cdot l_i)} \quad (6.10)$$

Die analytisch bestimmte Dämpfung $\text{att}_{analytic}(\vec{s}, \vec{x}, \vec{e})$ wird anschließend mit der Dämpfung $\text{att}_{bres}(\vec{s}, \vec{x}, \vec{e})$ verglichen, die mit Bresenham-Linien-Algorithmus nach Gleichung (6.9) berechnet wird. Dieser Vergleich wird für jeden Punktstreuer über alle SEP der Anzahl N durchgeführt und der mittlere absolute Fehler

$$\bar{\mathcal{E}}_{att} = \frac{\sum_{\forall(e,r)} |\text{att}_{analytic}(\vec{s}, \vec{x}, \vec{e}) - \text{att}_{bres}(\vec{s}, \vec{x}, \vec{e})|}{N} \quad (6.11)$$

über alle 3375 Punktstreuer bestimmt.

Der mittlere Fehler der Dämpfungskorrektur ist in Tabelle 6.3 für zwei Fälle dargestellt. Im ersten Fall sind nur zwei Bereiche mit unterschiedlicher Dämpfung angegeben. Im zweiten Fall wurden zusätzlich noch eine Zyste und zwei Läsionen simuliert.

Tabelle 6.3: Mittlerer Fehler der Bresenham-Approximation über alle simulierten Punktstreuer

Simulation	Korrektur	Mittelwert	Min	Max	STD
Wasser und Brust allein	nicht genutzt	29,8	17,9	43,0	7,1
Wasser und Brust allein	genutzt	1,4	0,4	3,3	0,6
zusätzlich Zyste und Läsionen	nicht genutzt	23,2	8,1	42,	8,3
zusätzlich Zyste und Läsionen	genutzt	1,3	0,3	2,7	0,6

Beide Fälle wurden mit und ohne Dämpfungskorrektur durchgeführt. Da für diese Messung kein Rauschen simuliert wurde, wurde auch kein Schwellwert für die Dämpfungskorrektur festgelegt. Ohne Dämpfungskorrektur ist der Fehler im Mittel 30 bzw. 23 mal höher als mit Dämpfungskorrektur. Ebenso ist eine deutlich erhöhte Standardabweichung zu beobachten. dIES entspricht einer räumlich starken inhomogenen Reflektivität. Mit Dämpfungskorrektur ist der Fehler im Mittel in beiden Fällen deutlich geringer. Ebenso wird die Standardabweichung mit 0,6 deutlich verringert. Dies entspricht einer homogenen Reflektivität.

Räumlich homogene Reflektivität:

Für die Bewertung der räumlich homogenen Reflektivität werden 3375 Punktstreuer simuliert, die gleichmäßig verteilt in einem Raster von 15^3 Punktstreuern angeordnet sind (siehe Abbildung 6.10). Dabei wird als Maß für die *räumliche homogene Objektausleuchtung* (RHOA) die Standardabweichung σ_{RHOA} über die maximalen Amplituden der rekonstruierten Punktstreuer nach Gleichung (5.9) berechnet. Als Referenz für eine Grundwahrheit werden dabei die maximalen Amplituden der Punktstreuer genutzt, die ohne Dämpfung simuliert und mit dem Standard-SAF-T rekonstruiert wurden. Damit wird sicher gestellt, dass nur der Einfluss der Dämpfung und der Dämpfungskorrektur bewertet wird. Das ist nötig, da die Amplituden der Punktstreuer bereits, bedingt durch die *Punktspreizfunktion* (PSF), bei der Rekonstruktion von ihrer räumlichen Position abhängig sind.

Abbildung 6.12 zeigt die Rekonstruktion einer xy-Schicht bei $z = 0\text{ m}$. Zusätzlich ist der Verlauf der Amplitude für den Schnitt bei $y = 0\text{ m}$ dargestellt. Die Amplituden werden, für den ohne Dämpfung simulierten Fall, mit dem Standard-SAF-T nahezu gleichmäßig rekonstruiert. Nur am äußeren Rand sinkt die Amplitude bedingt durch die räumliche Varianz der PSF. Wird die Dämpfung mit simuliert, kommt es zu einer Abschattung der mittleren Bereiche in der Brust. Wird die Dämpfungskorrektur dagegen genutzt, zeigt der Verlauf der Amplitude, dass die Korrektur eine räumlich homogene Reflektivität wiederherstellen kann.

Als quantitatives Maß um die räumlich homogene Reflektivität zu bewerten, wurde die Standardabweichung σ_{RHOA} über alle Punktstreuer berechnet und mit der Standard-SAF-T-Rekonstruktion ohne Dämpfung verglichen. σ_{RHOA} ist dabei normalisiert und relativ in % zu der maximalen Amplitude der rekonstruierten Punktstreuer der Referenz angegeben. Ein niedriger σ_{RHOA} -Wert entspricht einer guten Homogenität. Abbildung 6.13 (oben) zeigt die Verteilung von σ_{RHOA} über alle Punktstreuer für verschiedene Rauschlevel und genutzte Schwellwerte. Der Einsatz der Dämpfungskorrektur führt in nahezu allen Fällen zu einem 25% niedrigeren σ_{RHOA} . Das zeigt, dass die Dämpfungskorrektur die Homogenität wieder herstellen kann.

6. Reflexionsbildgebung mit SAFT für heterogene Medien

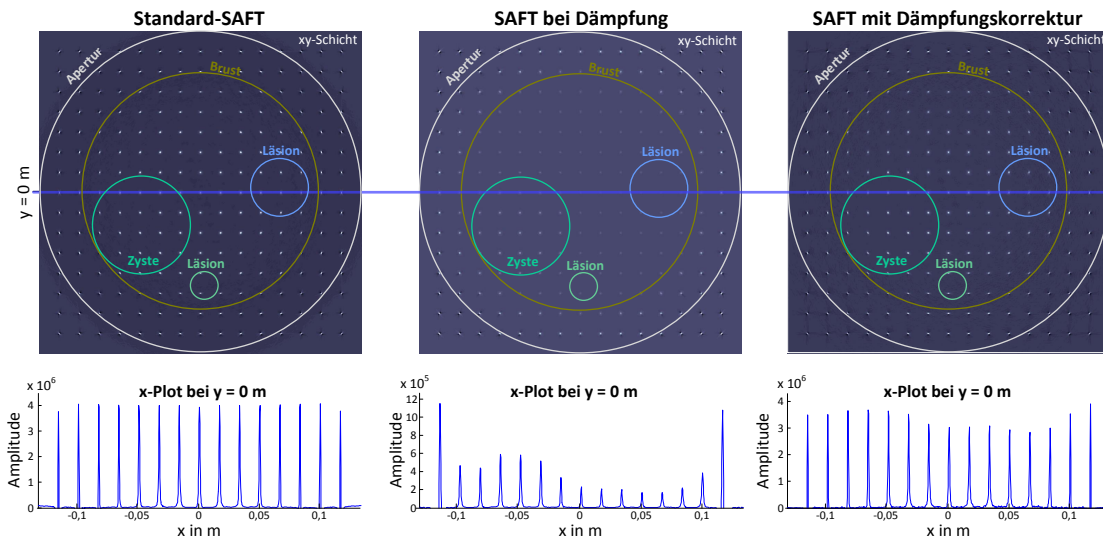


Abbildung 6.12: Rekonstruierte Bilder der simulierten Punktstreuer ohne zusätzliches Rauschen: ohne Dämpfung (links), mit simulierter Dämpfung (Mitte) und mit Dämpfungskorrektur (rechts). Die SAFT-Rekonstruktion einer xy-Schicht (oben) und ein Schnitt in x-Richtung (bei $y = 0 \text{ m}$) (unten).

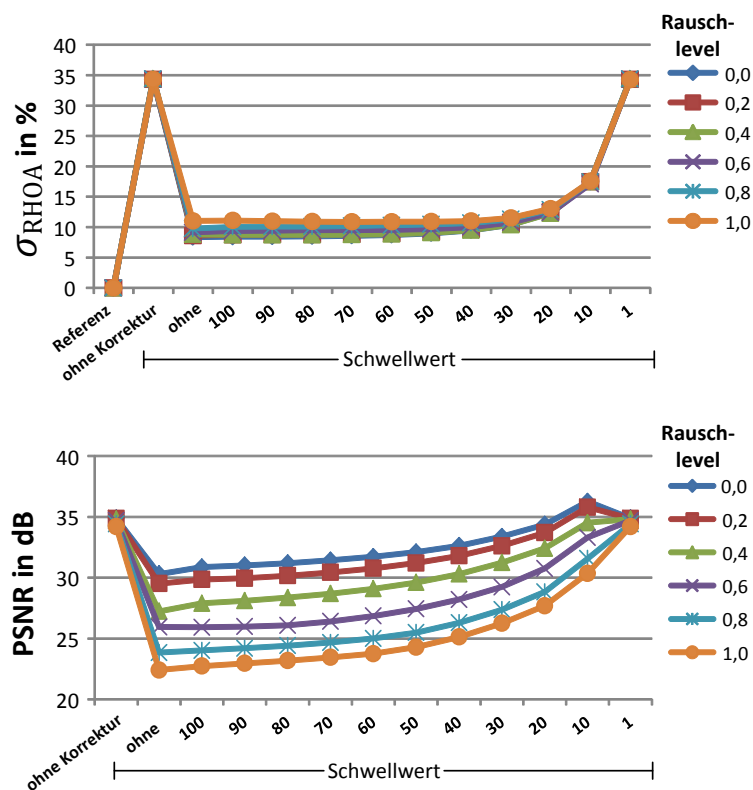


Abbildung 6.13: Standardabweichung (oben) und PSNR (unten) der simulierten Punktstreuer für den Fall mit Zyste und Läsionen für verschiedene Rauschlevel und Korrektur-Schwellwerte.

Wird der Schwellwert allerdings zu gering gewählt (<30), wird die Korrektur beschränkt und die Homogenität kann nicht wieder erreicht werden. Interessant ist, dass das Rauschen nur einen sehr geringen Einfluss auf σ_{RHOA} hat. Selbst bei einem Rauschlevel in der Höhe des Signals (Rauschlevel = 1) erreicht die Korrektur eine gute Homogenität, wenn ein genügend großer Schwellwert genutzt wird. Dieser Effekt lässt sich durch den Einfluss der großen Anzahl an genutzten A-Scans erklären, bei dem das Rauschen bei einem Mittelwert = 0 durch das Aufsummieren stark unterdrückt wird.

Kontrast:

Für die Bewertung des Kontrasts wird das PSNR der rekonstruierten Punktstreuer nach Gleichung (5.7) bestimmt. Abbildung 6.13 unten zeigt den Verlauf des PSNR für verschiedene Rauschlevel und Korrektur-Schwellwerte. Während σ_{RHOA} nur sehr schwach vom Rauschen beeinflusst wird, zeigt das PSNR eine starke Abhängigkeit. Das bestätigt, dass die Dämpfungskorrektur ebenso das Rauschen verstärkt, aber auch die Notwendigkeit, die Dämpfung zu begrenzen, um einen hohen Kontrast zu gewährleisten. Bei der Dämpfungskorrektur muss demnach ein Kompromiss zwischen homogener Ausleuchtung und hohem Kontrast gefunden werden.

Ein ähnliches Ergebnis ist für die Kontrastwerte bei den rekonstruierten in-vivo Daten zu beobachten. Da die Grundwahrheiten für die in-vivo Daten nicht bekannt sind, wird das PSNR durch das Verhältnis von maximaler Amplitude zu der Standardabweichung des Rauschens außerhalb der Brust bestimmt. Die mit Schallgeschwindigkeitskorrektur rekonstruierten Bilder sind in Abbildung 6.14 dargestellt. In Abbildung 6.14 oben links ist die Rekonstruktion ohne Dämpfungskorrektur gezeigt, die einen PSNR von 26 besitzt. Die genutzte Dämpfungskorrektur erhöht das PSNR auf 52,2 (siehe Abbildung 6.14 oben rechts). Es werden jedoch auch neue Rauschartefakte sichtbar. Der beste Kompromiss zwischen Dämpfungskorrektur und Verstärkung des Rauschens wird mit einem Schwellwert von $PSNR_{max} = 8$ erreicht (siehe Abbildung 6.14 unten), der mit der PSPNR-Analyse bestimmt wurde (siehe Abbildung 6.11). Das PSNR liegt hier bei 72,2 ohne an Dynamikumfang zu verlieren. Die Rekonstruktionsartefakte außerhalb der Brust werden deutlich reduziert und die Strukturen innerhalb der Brust werden deutlicher. Die Rausch-Artefakte, die erst durch die Dämpfungskorrektur aufgetreten sind, sind nun nicht mehr zu erkennen.

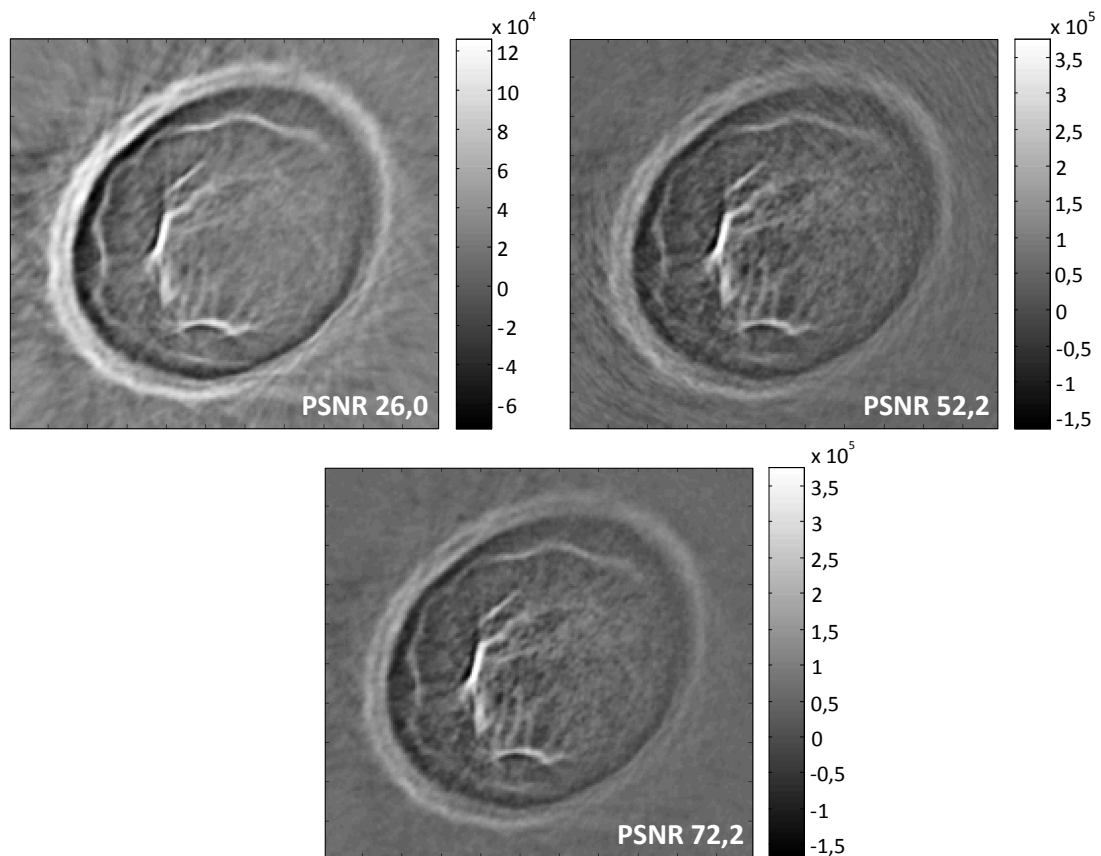


Abbildung 6.14: Vergleich der rekonstruierten Bilder von in-vivo Daten einer Beispielpatientin ohne (oben links) und mit Dämpfungskorrektur ohne Schwellwert (oben rechts). Rekonstruktion mit Schwellwert 8 (unten).

7. Beschleunigung der SAFT für heterogene Medien

7.1 Beschleunigung der korrigierten SAFT-Algorithmen

Der Standard-SAFT-Algorithmus, der für die Rekonstruktion der Reflektivität in der 3D-USCT eingesetzt wird, ist aufgrund der hohen Datenmenge sehr rechenintensiv. Da jedoch jeder Voxel parallel und unabhängig von einander berechnet werden kann, ist die SAFT sehr gut für die parallele Berechnung geeignet und wurde auf GPUs beschleunigt ausgeführt. In diesem Kapitel wird untersucht, wie die Berechnung der schallgeschwindigkeits- und dämpfungskorrigierten Versionen des SAFT-Algorithmus, die in Kapitel 6.1 und 6.2 für heterogene Medien eingeführt wurden, optimiert und in den aktuellen Standard-SAFT-Algorithmus integriert werden kann.

Die Rekonstruktion mit dem SAFT-Algorithmus basiert darauf, dass sich die zurückprojizierten Echosignale an den Orten konstruktiv überlagern, an denen eine Reflexion stattgefunden hat. Ob es zu einer konstruktiven oder destruktiven Überlagerung kommt, hängt dabei wesentlich von der Genauigkeit der mittleren Schallgeschwindigkeit ab. Ein Fehler der mittleren Schallgeschwindigkeit besitzt, wie in Kapitel 6.1 untersucht, einen signifikanten Einfluss auf die Bildqualität. Es wurde eine Grenze für die maximal erlaubte Abweichung von 0,39 m/s hergeleitet, da die genutzte Schallgeschwindigkeit über die Korrektheit der Information in dem rekonstruierten Volumen entscheidet. Bei der Dämpfungskorrektur wird dagegen eine verbesserte homogene Reflektivität erreicht, mit dem Ziel den Kontrast zu erhöhen. Die Korrektur wird jedoch bei realen Daten von dem SNR der vorverarbeiteten A-Scans beschränkt. In der Praxis wird daher die Dämpfungskorrektur mit dem Einsatz eines Korrekturschwellwerts bestmöglich durchgeführt.

In diesem Kapitel werden zuerst die Berechnungsstrategien anhand der schallgeschwindigkeitskorrigierten Version eingeführt, beschleunigt und evaluiert. Anschließend wird auf die Version mit Dämpfungskorrektur näher eingegangen und der Programmablauf sowie die Performanz dargestellt.

Für die Schallgeschwindigkeitskorrektur wird im SAFT-Algorithmus vor allem ein eigener Schallgeschwindigkeitswert $\bar{c}_P(\vec{e}, \vec{x}, \vec{r})$ pro SEP und pro zu rekonstruierendem Voxel benötigt. Dieser kann mit Hilfe der zuvor berechneten SOS-Volumen bestimmt werden, indem die mittlere Schallgeschwindigkeit auf dem jeweiligen Pfad (Sender, Voxel, Empfänger) berechnet wird.

Dabei ist zu beachten, dass die Auflösung des Reflexionsvolumen mit $(0,24 \text{ mm})^3$ ca. um den Faktor 50 höher liegt als die Auflösung der SOS-Volumen mit $(1,2 \text{ cm})^3$. Dabei stellt die Auflösung von $(1,2 \text{ cm})^3$ den schlechtesten Fall dar. In der Praxis werden SOS-Volumen bis zu einer Auflösung von $(4,0 \text{ mm})^3$ genutzt [92]. Diese Auflösungen entsprechen dabei den Volumen von $2167^2 \times 1333$ Voxel für das gesamte Reflexionsvolumen und $128^2 \times 79$ Voxel für das SOS-Volumen. Die zu berechnenden Voxel können dabei, auf Grund der kleineren ROI, auf eine Anzahl von $1667^2 \times 1000$ für das Reflexionsvolumen und $100^2 \times 60$ für das SOS-Volumen reduziert werden. Von

diesen Volumengrößen ausgehend, nimmt die Komplexität der SAFT-Berechnung mit Schallgeschwindigkeitskorrektur auf $O(V \cdot A \cdot P) = O(V^2 \cdot A^2)$ weiter zu. Wobei V die Anzahl der zu rekonstruierenden Voxel im Reflexionsvolumen, A die Anzahl der genutzten A-Scans und P die Anzahl der Pfade bezeichnet. Da $P = V \cdot A$ steigt die Komplexität nun auf das Quadratische an. Um diese Komplexität zu meistern, ist eine effiziente Implementierung notwendig, die eine hohe Rechengeschwindigkeit erreicht. Daher werden im Folgenden verschiedene Berechnungsstrategien für den SAFT-Algorithmus mit Schallgeschwindigkeitskorrektur analysiert und evaluiert. Dabei muss ein Kompromiss zwischen einem begrenzten Speicherbedarf, hoher Rechengeschwindigkeit und hoher Bildqualität gefunden werden.

7.1.1 Berechnungsstrategien

Der zur Verfügung stehende Speicher auf den GPUs ist limitiert. Die genutzten GPUs besitzen aktuell je nach Ausführung eine maximale Speichergröße von 1,5 GB (GTX 590 GPU) bis maximal 6 GB (GTX TITAN). Daher wird eine geeignete Berechnungsstrategie benötigt, die der Größe des jeweiligen GPU-Speichers angepasst ist. Tabelle 7.1 gibt einen Überblick über vier Berechnungsstrategien der Pfade, die jeweilige Anzahl der zu speichernden Werte und dem entsprechenden Speicherbedarf. Da bei mehreren genutzten *Aperturpositionen* (AP) die Sender und Empfänger neu positioniert werden und die Rekonstruktion sequentiell abgearbeitet wird, beziehen sich die Werte nur auf eine AP.

Akronym	Berechnungsstrategie	zu speichernde Daten	Speicherbedarf
Kernel	1. Berechnung im SAFT-Kernel	4 Byte · 1 Voxel	4 B
PV	2a. Vorberechnung aller Pfade für alle Z-Schichten	4 Byte · 1667 ² · x1000 Voxel · (628 · 1413) Transducer	9,0 PB
PV1	2b. Vorberechnung aller Pfade für 1 Z-Schicht	4 Byte · 1667 ² · x1 Voxel · (628 · 1413) Transducer	9,0 TB
TPV1	3. Vorberechnung Sender & Empfänger Pfade für eine Z-Schicht	(4+1) Byte · (628+1413) Transducer · 1667 ² · x1 Voxel	26,4 GB
SOSTPV1	4a. Vorberechnung Sender & Empfänger Pfade für eine SOS-Z-Schicht	(4+1) Byte · (628+1413) Transducer · 100 ² · x1 SOS-Voxel	97,3 MB
SOSTPV	4b. Vorberechnung Sender & Empfänger Pfade für alle SOS-Z-Schichten	(4+1) Byte · (628+1413) Transducer · 100 ² · x60 SOS-Voxel	5,7 GB

Tabelle 7.1: Übersicht der Berechnungsstrategien und ihr jeweiliger Speicherbedarf.

Strategie 1 - Kernel: Die erste Berechnungsstrategie der mittleren SOS auf den Pfaden ist die direkte Berechnung im GPU-SAFT-Kernel für jeden Voxel. Aufgrund der direkten Berechnung wird die geringste Speichermenge von nur einem mittleren SOS-Wert benötigt ($float=4$ Byte). Der Nachteil dieser Strategie ist allerdings, dass die Teil-Pfade von einem Sender zu dem Voxel und von dem Voxel zu einem Empfänger vielfach berechnet werden müssen.

Strategie 2 - Pfadvorbereitung (PV): Werden alle mittleren SOS-Werte im Voraus berechnet und bei Bedarf im SAFT-Kernel geladen, führt das zu dem größten Speicherverbrauch (PV). Die Werte für jeden Voxel und jeden A-Scan müssen gespeichert werden. Das ist mit einer Datenmenge von 9,0 Petabyte um viele Größenordnungen zu groß für den GPU-Speicher. Auch die sequentielle Berechnung von einzelnen Z-Schichten hat einen Speicherbedarf von 9,0 TB und ist ebenfalls deutlich zu groß (PV1).

Strategie 3 - Teilpfadvorbereitung (TPV): Um die Datenmenge zu reduzieren und die Rechengeschwindigkeit zu erhöhen, können die Pfade, die berechnet werden sollen, in zwei Gruppen eingeteilt werden: In die Pfade von Sender zum Voxel und die Pfade vom Voxel zum Empfänger. Jeder Pfad wird dabei im Voraus nur einmal berechnet und kann anschließend mehrfach geladen werden. Es muss pro Pfad nur noch eine Summe der Schallgeschwindigkeiten über jeden besuchten SOS-Voxel ($float=4$ Byte), sowie die Anzahl auf dem Pfad besuchten SOS-Voxel ($char=1$ Byte) in eine LookUp-Tabelle gespeichert werden. Die mittlere Schallgeschwindigkeit wird anschließend im SAFT-Kernel mit dem harmonischen Mittel der zwei Teilpfade berechnet. Die Anzahl der möglichen Sender- und Empfänger-Kombinationen reduziert sich um einen Faktor 435 von $628 \cdot 1413 = 887364$ auf $628 + 1413 = 2041$. Der Speicherbedarf von 26,4 GB ist dennoch zu hoch für den GPU-Speicher.

Strategie 4 - Teilpfadvorbereitung für SOS-Voxel (SOSTPV): Eine weitere Optimierung kann durchgeführt werden, indem alle Voxel des Reflexionsvolumens zusammengefasst werden, die sich in einem SOS Voxel befinden. Damit müssen nur noch die zwei Teilpfade von Sender zum SOS-Voxel und vom SOS-Voxel zum Empfänger für eine SOS-Z-Schicht berechnet und gespeichert werden (SOSTPV1). Mit dieser Strategie sinkt der Speicherbedarf für eine SOS-Z-Schicht mit 97,3 MB auf eine für den GPU-Speicher geeignete Größe.

Der Speicherbedarf für die Vorberechnung und Speicherung aller 100 SOS-Z-Schichten (SOSTPV) ist jedoch mit 5,7 GB zu groß. Daher muss die Vorberechnung der SOS-Pfade und die SAFT-Berechnung sequentiell in mehreren Durchläufen erfolgen. Zusammenfassend sind aufgrund des Speicherbedarfs nur die **Strategien Kernel und SOSTPV1** für aktuelle GPUs geeignet und werden im Folgenden weiter untersucht.

Bildqualität der Berechnungsstrategien

Neben der Limitierung, die durch die Speichergröße der GPU-Hardware gegeben ist, spielt die Bildqualität eine wichtige Rolle. Um die beste Bildqualität mit der Schallgeschwindigkeitskorrektur zu erreichen, muss die genutzte mittlere Schallgeschwindigkeit c_P in Gleichung (6.4) so genau wie möglich sein. In Kapitel 6.1 wurde für den Bereich der ROI eine maximal erlaubte Abweichung der Schallgeschwindigkeit auf den Pfaden von $\Delta c = 0,39 m/s$ hergeleitet, um eine kohärente Bildrekonstruktion zu ermöglichen. Mit **Strategie Kernel** wird für jeden Voxel eine mittlere Schallgeschwindigkeit berechnet. Dabei wird die mittlere Schallgeschwindigkeit für den

Pfad von und zu der genauen Voxelposition bestimmt (siehe Abbildung 7.1 oben links). Dies ist genauer als bei der optimierten **Strategie SOSTPV1**. Hier wird die mittlere Schallgeschwindigkeit für den Pfad von und zu der SOS-Voxelposition bestimmt, in dem sich der Voxel befindet (siehe Abbildung 7.1 oben rechts). Alle Voxel des Reflexionsbildes, die sich innerhalb eines SOS-Voxels befinden, werden durch die Position des jeweiligen SOS-Voxel approximiert. Der dabei gemachte Fehler wird größer, je weiter sich der Voxel an der Kante des SOS-Voxels befindet. Beim Übergang zwischen zwei SOS-Voxel ist der Fehler der mittleren SOS am höchsten, da angenommen wird, dass sich der Voxel in der Mitte des SOS-Voxel befindet. Dadurch können Eckartefakte in dem rekonstruierten Volumen auftreten. Dieser Effekt ist besonders stark, wenn zwei benachbarte SOS-Voxel im SOS-Volumen einen hohen Unterschied ihrer SOS-Werte aufweisen (siehe Abbildung 7.1 unten). Die Eckartefakte sind daher abhängig von dem zu messenden Objekt innerhalb des 3D-USCTs. Um den Unterschied der Schallgeschwindigkeit benachbarter SOS-Voxel zu untersuchen, wurde eine Analyse aller Brustdaten durchgeführt, die aus der ersten klinischen Studie gewonnen wurden. Abbildung 7.2 zeigt die Häufigkeit der auftre-

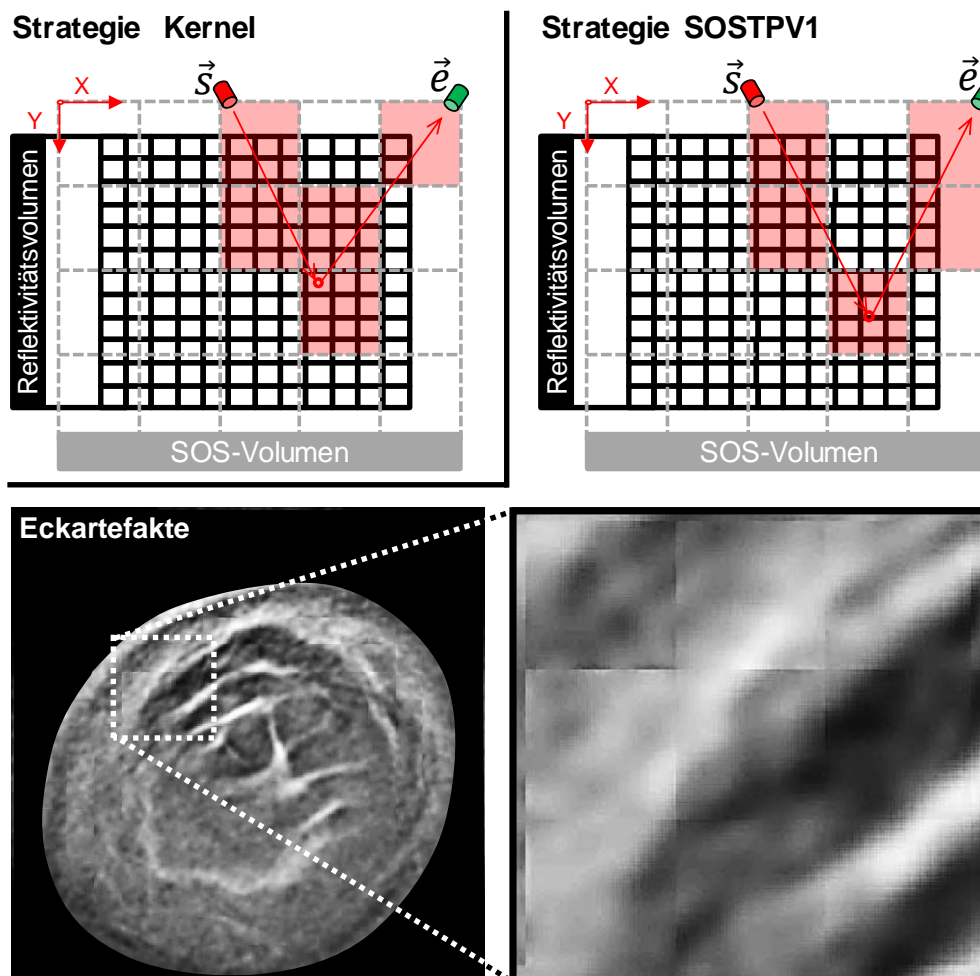


Abbildung 7.1: Schematischer Vergleich zwischen Strategie Kernel (links) und Strategie SOSTPV1 (rechts oben). Bei Strategie SOSTPV1 treten Eckartefakte an der Grenze zwischen zwei benachbarten SOS-Voxel auf, die einen hohen Unterschied der Schallgeschwindigkeitswerte aufweisen (unten).

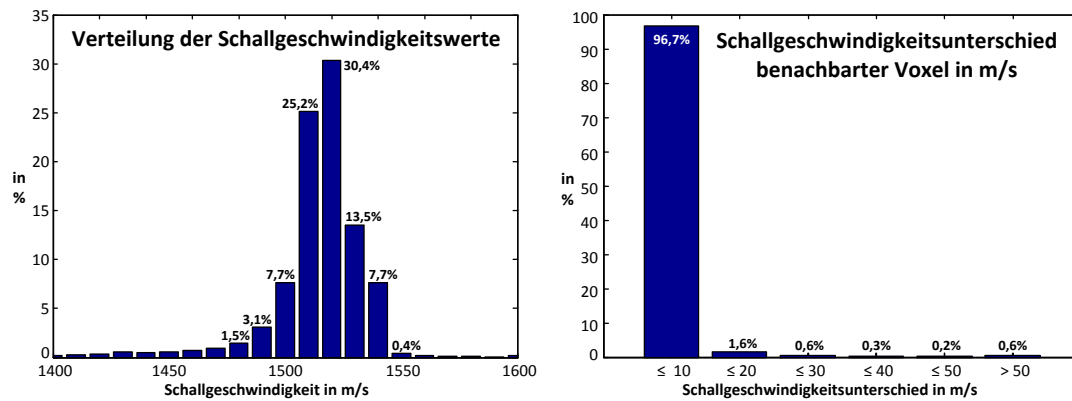


Abbildung 7.2: Histogramm der SOS-Werte (links) und der Unterschied zwischen benachbarten SOS-Voxel (rechts) auf Grundlage der rekonstruierten SOS-Volumen der klinischen Studie.

tenden Schallgeschwindigkeit und die Verteilung der gemessenen Unterschiede zu den benachbarten SOS-Voxel. Über alle Brustdaten zeigen 96,7% aller SOS-Voxel zu ihren benachbarten SOS-Voxel einen Unterschied der Schallgeschwindigkeit von weniger als 10 m/s und 98,3% einen Unterschied von weniger als 20 m/s. Dennoch besitzen einige SOS-Voxel einen höheren Schallgeschwindigkeitsunterschied zu ihren benachbarten SOS-Voxel, bei denen ungewollte Eckartefakte im rekonstruierten Bild sichtbar werden können. Zudem ist das Auftreten von Eckartefakten nicht allein abhängig von dem Unterschied zum benachbarten SOS-Voxel. Der insgesamt durchlaufene Pfad und die jeweilig lokalen Schallgeschwindigkeiten können sich je nach Pfad relativ zueinander zu einem größeren Schallgeschwindigkeitsunterschied aufsummieren und ebenfalls Eckartefakte verursachen.

Um diese Artefakte bei der auf geringen Speicherbedarf optimierten Rekonstruktionsstrategie SOSTPV1 zu vermeiden, stehen zwei Möglichkeiten zur Verfügung: Der Einsatz einer höheren Auflösung von SOS-Volumen erhöht die Genauigkeit der mittleren Schallgeschwindigkeit durch eine höhere Abtastung des Volumens. Dabei wird auch der Speicherbedarf erhöht, da mehr Werte für mehr SOS-Voxel gespeichert werden. Ebenso wird die Rechendauer erhöht. Für jeden weiteren Voxel muss pro SEP ein neuer Pfad berechnet werden. Zudem steigt die Länge aller Pfade, die mit dem Bresenham-Algorithmus durchwandert werden müssen, deutlich an. Da bei einer höheren SOS-Volumen-Auflösung weniger Voxel innerhalb eines SOS-Voxel liegen, wird seltener auf die zuvor berechneten Daten der Teilpfade zugegriffen. Für die Berechnung des SAFT-Kernels ergibt sich damit ein negativeres Cache-Verhalten. Besteht allerdings trotz höherer SOS-Auflösung weiterhin ein hoher Unterschied der Schallgeschwindigkeiten für zwei benachbarte SOS-Voxel, treten die Eckartefakte nach wie vor auf. Der zweite Ansatz ist die Interpolation zwischen den vorberechneten mittleren SOS-Werten. Die Interpolation führt zu einem kontinuierlichen Übergang der mittleren SOS-Werte, die Eckartefakte werden vermieden und die Bildqualität erhöht (siehe Abbildung 7.3). Da keine neuen Pfade berechnet werden müssen, bleibt die Dauer und die Datenmenge für die Vorberechnung gleich. Ein weiterer Vorteil ergibt sich zudem bei der Berechnung auf GPUs. Werden die SOS-Daten in den *textur memory* der GPU geschrieben, kann eine lineare Interpolation direkt in der Hardware, bei einem nur geringen Performanzverlust, durchgeführt werden. Stan-

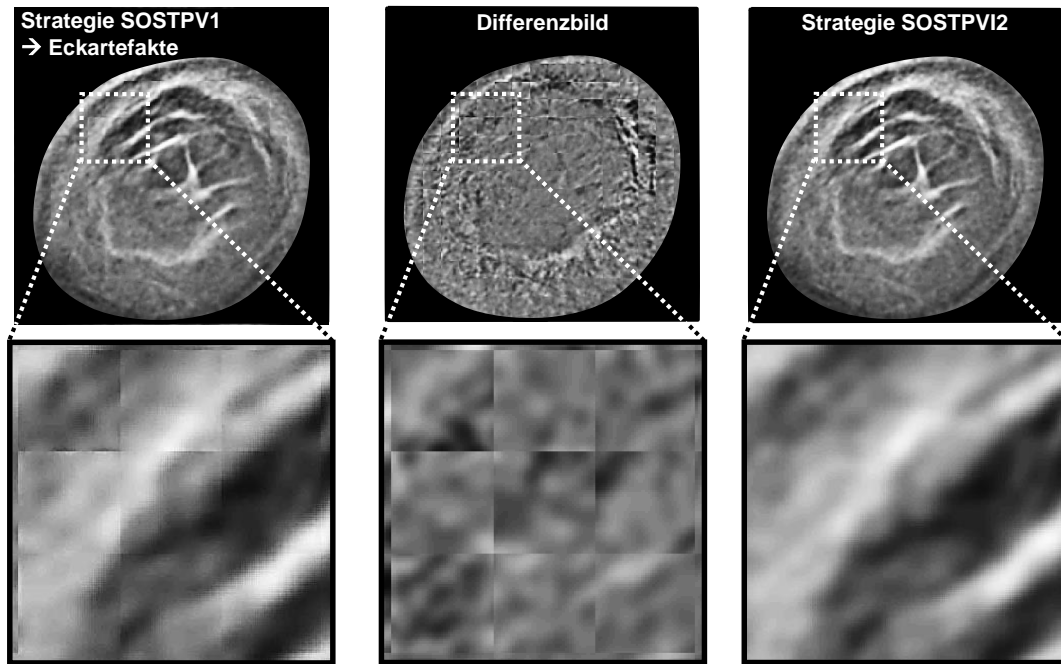


Abbildung 7.3: Eckartefakte an der Position von benachbarten SOS-Voxel mit hohem Unterschied der SOS-Voxel bei der optimierten Strategie SOSTPV1 (links). Die Eckartefakte können mit der Interpolation zwischen den mittleren SOS-Werten bei Strategie SOSTPVI2 vermieden werden (rechts). Die Differenz beider Bilder (mitig).

dardmäßig ist der *textur memory* nicht beschreibbar und es wird ein zusätzlicher Speicher und Kopiervorgang benötigt. CUDA stellt jedoch die Möglichkeit zur Verfügung, den Speicher zuvor als *Surface* zu definieren [39]. So können die SOS-Werte direkt vom GPU-Kernel der Pfadberechnung in den *surface memory* geschrieben werden. Im SAFT-Kernel wird anschließend, wie bei dem *textur memory* gewohnt, auf die Daten zugegriffen. Für eine Interpolation in 3D werden mindestens zwei SOS-Z-Schichten benötigt. Ein weiterer Nachteil ist, dass die Interpolation nur bei Gleitkommazahl-Datentypen unterstützt wird [39]. Das bedeutet, dass die Anzahl der besuchten SOS-Voxel auf dem Pfad nicht mehr als *char* mit 1 Byte gespeichert werden können. Es muss ebenfalls als Gleitkommazahl in *float* mit 4 Bytes gespeichert werden. Dies stellt jedoch auch einen Vorteil dar, da kombinierte Datentypen wie *float2* oder *float4* von CUDA unterstützt werden. Somit kann direkt mit einem Lesezugriff von nur einer Speicheradresse die mittlere Schallgeschwindigkeit und die Anzahl an besuchten SOS-Voxeln interpoliert ausgelesen werden. Für den Speicherbedarf bedeuten diese Änderungen einen 3,2 mal höheren Speicherbedarf. Diese Variante der Strategie SOSTPV1 mit zusätzlicher Interpolation wird im Folgenden als **Strategie SOSTPVI2** bezeichnet (Tabelle 7.2).

Akronym	Berechnungsstrategie	zu speichernde Daten	benötigter Speicher
SOSTPVI2	5. Vorberechnung Sender & Empfänger Pfade für zwei SOS-Z-Schichten	(4+4) Byte · (628+1413) Transducer · $100^2 \times 2$ SOS-Voxel	311,4 MB

Tabelle 7.2: Übersicht der Berechnungsstrategie SOSTPVI2 und dem Speicherbedarf.

7.1.2 Evaluation

Für eine quantitative Evaluierung der Schallgeschwindigkeitskorrektur wird eine Grundwahrheit benötigt. Dazu wird die in Kapitel 5.3 beschriebene Simulation genutzt. Der genutzte Aufbau der Simulationsumgebung ist in Abbildung 7.4 dargestellt. Es werden zwei Bereiche mit unterschiedlichen Schallgeschwindigkeiten simuliert. Die Apertur des KIT 3D-USCT ist mit Wasser gefüllt und wird mit einer Schallgeschwindigkeit von $c_{Wasser}=1500$ m/s angenommen. Eine Brust wird mit einer Halbkugel mit einem Radius von 10 cm approximiert. Die Schallgeschwindigkeitskorrektur beruht auf der Berechnung der mittleren Schallgeschwindigkeit auf den Pfad, den der Ultraschall durchläuft. Der Fehler, der bei dieser Berechnung gemacht wird, stellt somit eine geeignete Metrik zur Bewertung der Korrektur da. Mit Hilfe der Simulation wird analytisch eine mittlere Schallgeschwindigkeit als Grundwahrheit für jeden Pfad berechnet, die anschließend mit dem Wert verglichen wird, der bei der Korrektur berechnet wird. Da anzunehmen ist, dass der Fehler von der Höhe des Unterschieds der Schallgeschwindigkeiten in den zwei Bereichen abhängt, werden mehrere Simulationen mit unterschiedlichen Werten für c_{Brust} durchgeführt. Die Schallgeschwindigkeiten in dem Bereich 1460 - 1540 m/s stellen dabei einen realistischen Wertebereich für verschiedene Brustgewebe dar (siehe Tabelle 6.1).

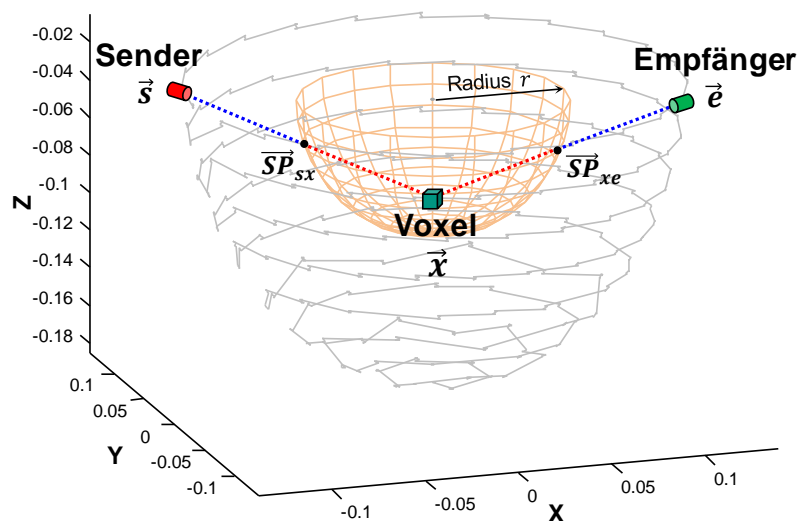


Abbildung 7.4: Aufbau der Simulation und ein Pfad durch das heterogene Medium.

Die mittlere Schallgeschwindigkeit $\bar{c}_{analytisch}(\vec{s}, \vec{x}, \vec{e})$ für einen Voxel wird dabei analytisch bestimmt, indem zuerst die Schnittpunkte der Strecken von Sender-Voxel und Voxel-Empfänger mit der Halbkugel ($\overrightarrow{SP}_{sx}, \overrightarrow{SP}_{xe}$) berechnet werden. Anschließend werden die Längen der Strecken innerhalb des Wassers (l_{Wasser}) und der Brust (l_{Brust}) bestimmt. Die mittlere Schallgeschwindigkeit auf einem Pfad ergibt sich mit dem harmonischen Mittel nach

$$\bar{c}_{analytisch} = \frac{l_{Wasser} + l_{Brust}}{\frac{l_{Wasser}}{c_{Wasser}} + \frac{l_{Brust}}{c_{Brust}}}. \quad (7.1)$$

Die analytisch bestimmte mittlere Schallgeschwindigkeit wird anschließend mit der mittleren Schallgeschwindigkeit verglichen, der mit dem Bresenham-Algorithmus bestimmt wird. Der mittlere Fehler $\bar{\mathcal{E}}$ pro Voxel über alle SEP, mit der Anzahl N , wird mit

$$\bar{\mathcal{E}}(\vec{x}) = \frac{\sum_{\forall(e,r)} |\bar{c}_{analytisch}(\vec{s}, \vec{x}, \vec{e}) - \bar{c}_{bres}(\vec{s}, \vec{x}, \vec{e})|}{N} \quad (7.2)$$

berechnet. Zusätzlich wird die Standardabweichung pro Voxel über alle SEP mit

$$\sigma_{\mathcal{E}}(\vec{x}) = \sqrt{\frac{1}{N-1} \sum_{\forall(e,r)} (\bar{\mathcal{E}}(\vec{x}) - (\bar{c}_{analytisch}(\vec{s}, \vec{x}, \vec{e}) - \bar{c}_{bres}(\vec{s}, \vec{x}, \vec{e})))^2} \quad (7.3)$$

bestimmt. Damit steht neben dem mittleren Fehler auch die Standardabweichung zur Verfügung. Unter der Annahme eines gaußverteilten mittleren Fehlers kann so mit $(\bar{\mathcal{E}} + \sigma_{\mathcal{E}}) < 0,39 \text{ m/s}$ eine strengere Randbedingung untersucht werden, bei der für mindestens 68,3% der Voxel der Fehler geringer als 0,39 m/s ist.

1. Fehler ohne Schallgeschwindigkeitskorrektur

Es wurde der Fehler bei der Berechnung der mittleren Schallgeschwindigkeit über alle Voxel \vec{x} des 3D-USCT-Volumens gemessen. In Abbildung 7.5 ist der mittlere Fehler nach Gleichung (7.2) sowie der mittlere Fehler plus der Standardabweichung aus Gleichung (7.3) dargestellt. Für den Fall, dass keine Schallgeschwindigkeitskorrektur durchgeführt wird, d.h. von einer konstanten Schallgeschwindigkeit ausgegangen wird, steigt der Fehler sehr schnell über die Grenze von 0,39 m/s, die für eine kohärente Bildrekonstruktion erforderlich ist. Eine Schallgeschwindigkeitskorrektur ist bereits bei einem Schallgeschwindigkeitsunterschied von $\Delta c > 0,7 \text{ m/s}$ nötig.

2. Fehler mit Schallgeschwindigkeitskorrektur

Zum Vergleich des Fehlers, der bei heterogenen Medien ohne eine Schallgeschwindigkeitskorrektur gemacht wird, wurde der Fehler für die drei verbliebenen Berechnungsstrategien berechnet. Es wurde ebenfalls über den SOS-Bereich von 1460 – 1540 m/s simuliert. Dies entspricht einem Schallgeschwindigkeitsunterschied Δc zwischen den Schallgeschwindigkeiten im Wasser und der Brust von 0 m/s bis $\pm 40 \text{ m/s}$. Da die Genauigkeit der Korrektur von der Auflösung der genutzten SOS-Volumen abhängt, wurde zudem über mehrere Auflösungen (16^3 bis 256^3 Voxel bzw. 3,3 cm bis 4,1 cm) simuliert und entsprechend der absolute Fehler gemessen. In Abbildung 7.6

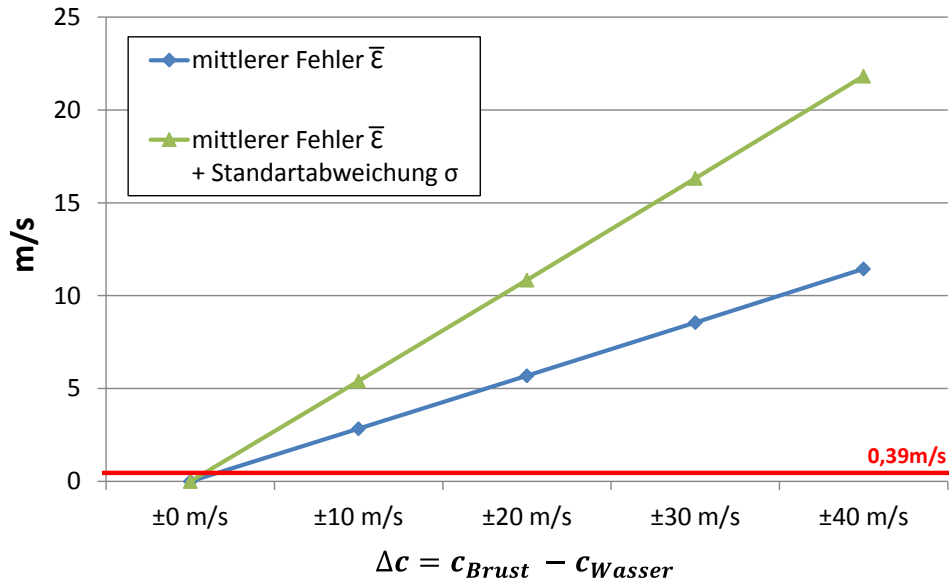


Abbildung 7.5: Mittlerer Fehler $\bar{\mathcal{E}}$ und Standardabweichung $\sigma_{\mathcal{E}}$ der mittleren Schallgeschwindigkeit ohne Schallgeschwindigkeitskorrektur über mehrere Schallgeschwindigkeits-Unterschiede zwischen c_{Brust} und c_{Wasser} .

sind die Ergebnisse der Berechnungsstrategien dargestellt. Für alle Berechnungsstrategien sinkt der Fehler bei Einsatz höherer SOS-Auflösungen. Eine Schallgeschwindigkeitskorrektur mit einer SOS-Auflösung $>128^3$ Voxeln erhöht zwar den Berechnungsaufwand, zeigt dagegen nur noch eine geringe Verbesserung. Die Korrektur erreicht somit eine Grenze. Der mittlere Fehler liegt daher bei allen drei Strategien für $\Delta c = \pm 40 \text{ m/s}$ bei etwa $0,28 \text{ m/s}$. Tabelle 7.3 gibt die genauen Werte des Fehlers für den Fall an, dass die Differenz der Schallgeschwindigkeiten $\pm 10 \text{ m/s}$ beträgt.

Auflösung SOS Strategie	$\bar{\mathcal{E}}$ in m/s		$\sigma_{\mathcal{E}}$ in m/s		minimale SOS Auflösung mit $(\bar{\mathcal{E}} + \sigma_{\mathcal{E}}) < 0,39 \text{ m/s}$
	16^3	128^3	16^3	128^3	
1. Ohne Korrektur	2,82		2,57		/
Im Kernel (Strategie Kernel)	0,32	0,08	0,35	0,10	64^3
Vorbereitung (Strategie SOSTPV1)	0,35	0,09	0,38	0,11	64^3
Vorbereitung mit Interpolation (Strategie SOSTPVI2)	0,29	0,08	0,29	0,10	48^3

Tabelle 7.3: Mittlerer Fehler $\bar{\mathcal{E}}$ und mittlerer Fehler plus Standardabweichung $\sigma_{\mathcal{E}}$ der Berechnungsstrategien bei einem Schallgeschwindigkeits-Unterschied von $\Delta c = \pm 10 \text{ m/s}$. Die Werte sind für die Auflösungen des Schallgeschwindigkeits-Volumens von 16^3 und 128^3 Voxeln angegeben. Zudem ist die minimal nötige Auflösung gegeben, um unterhalb des Fehlers für eine kohärente Rekonstruktion zu bleiben.

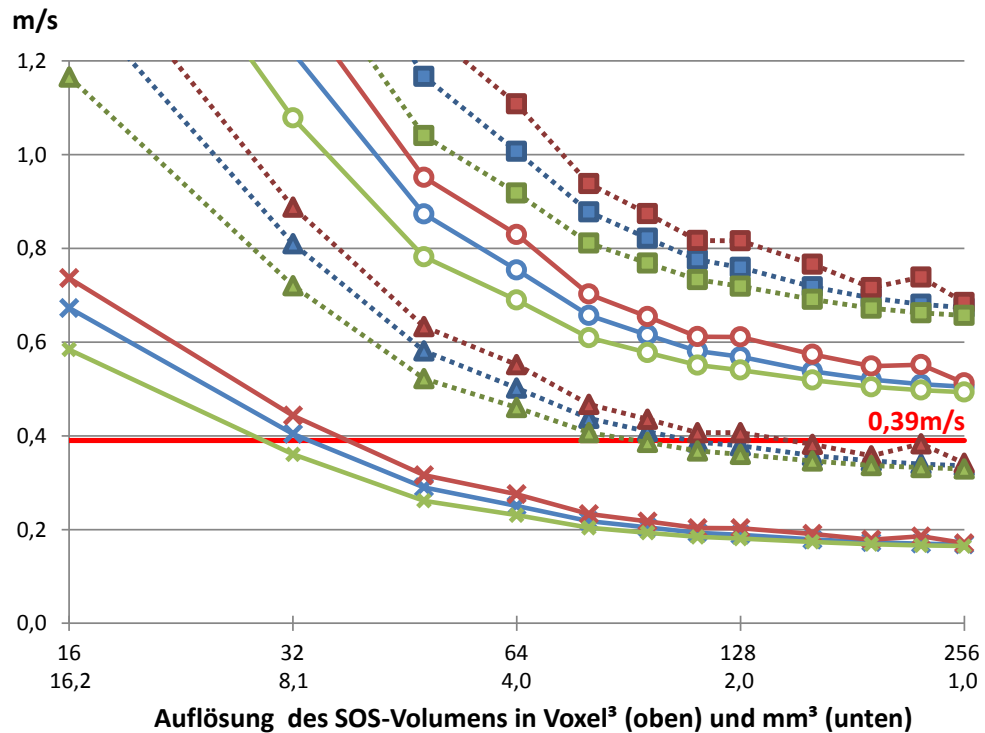
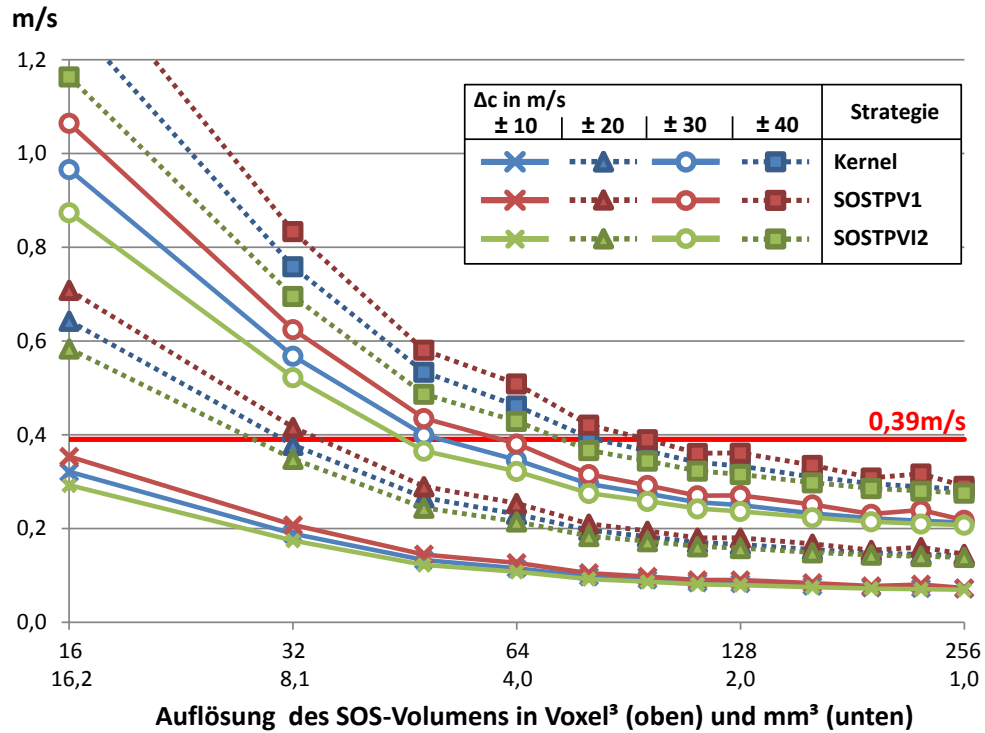


Abbildung 7.6: Mittlerer Fehler $\bar{\varepsilon}$ (oben) und mittlerer Fehler plus Standardabweichung $\bar{\varepsilon} + \sigma_{\varepsilon}$ (unten) der mittleren Schallgeschwindigkeit mit Schallgeschwindigkeitskorrektur für mehrere Auflösungen der SOS-Volumen und bei verschiedenen Schallgeschwindigkeitsunterschieden $\Delta c = 0 \text{ m/s}$ bis $\pm 40 \text{ m/s}$ zwischen c_{Brust} und c_{Wasser} .

Aus den Ergebnissen des mittleren Fehlers $\bar{\mathcal{E}}$ wird der positive Einfluss einer Schallgeschwindigkeitskorrektur deutlich. Werden die Strategien Kernel bzw. SOSTPV1 eingesetzt, kann bei einem Unterschied von $\Delta c = \pm 10 \text{ m/s}$ bereits mit einer Auflösung des SOS-Volumens von 64^3 Voxel die erforderliche Genauigkeit erreicht werden. Strategie SOSTPVI2 erreicht die erforderliche Genauigkeit bei einem Volumen von 48^3 Voxel bereits etwas früher. Der kleinere Fehler von Strategie SOSTPVI2 ist bei geringeren Auflösungen deutlicher ausgeprägt. Da Strategie SOSTPVI2 den geringsten Fehler besitzt, stellt sie von den untersuchten Berechnungsstrategien die beste Approximation für die mittlere Schallgeschwindigkeit dar.

Neben dem mittleren Fehler über alle Voxel ist zudem die räumliche Verteilung des Fehlers von Bedeutung, da der mittlere Fehler $\bar{\mathcal{E}}(\vec{x})$ von der Position des Voxels \vec{x} abhängt. In Abbildung 7.7 ist für die verschiedenen Berechnungsstrategien die Verteilung des Fehlers und die Standardabweichung für eine xy-Schicht dargestellt. Es wird deutlich, dass der Fehler innerhalb der simulierten Brust gering ist und sich die größten Fehler am Rand der simulierten Brust befinden. Reflektierende Strukturen innerhalb der Brust lassen sich somit gut fokussiert darstellen.

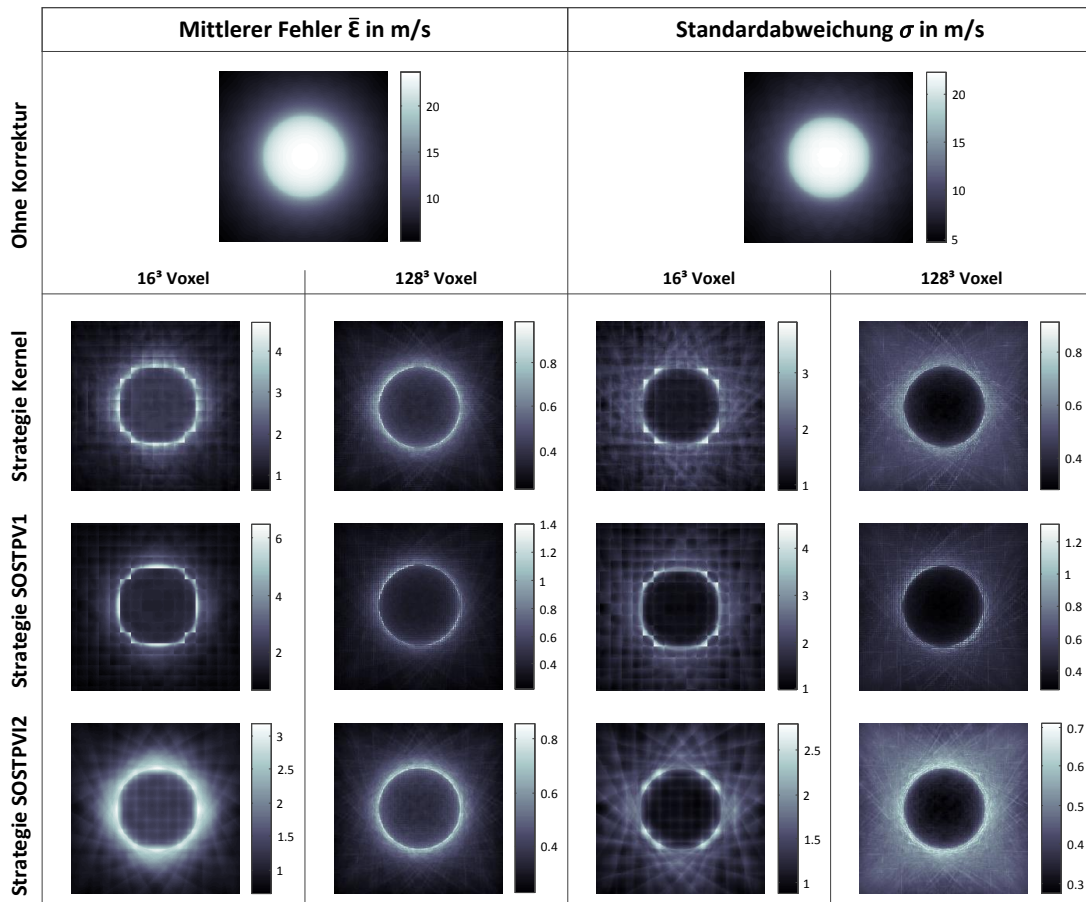


Abbildung 7.7: Räumliche Darstellung des mittleren Fehlers und der Standardabweichung pro Voxel einer XY-Schicht für einen Unterschied von $\Delta c = \pm 10 \text{ m/s}$ und für die Auflösungen von 16^3 und 128^3 Voxel für das SOS-Volumen.

7.1.3 Implementierungen der korrigierten SAFT-Algorithmen

Die Implementierung der Rekonstruktion mit SOS-Korrektur bildet die Grundlage für die Implementierung der Rekonstruktion mit SOS- und Dämpfungskorrektur. Da die Abläufe sehr ähnlich sind, werden in diesem Abschnitt die Implementierungen der Schallgeschwindigkeits- und Dämpfungskorrektur gemeinsam beschrieben und an entsprechenden Stellen auf die geringen Abweichungen hingewiesen. Da dieselbe Auflösung für SOS- und ATT-Volumen genutzt wird, sind die SOS- und ATT-Volumen-Koordinaten identisch ($\vec{x}_{ATT} = \vec{x}_{SOS}$). Daher werden die ATT-Volumen-Koordinaten ebenfalls mit dem Index für die SOS-Volumen-Koordinaten angegeben. Das Ablaufdiagramm des SAFT-Kerns für die Rekonstruktion mit SOS- und Dämpfungskorrektur ist in Abbildung 7.8 dargestellt. Die SAFT-Rekonstruktion mit Korrektur verläuft in zwei Schritten. In der Pfad-Vorbereitung werden die mittleren Schallgeschwindigkeiten und die Dämpfung auf den einzelnen Pfaden von und zu jedem SOS-Voxel berechnet. Die SAFT-Rekonstruktion wird im zweiten Schritt für jeden Voxel durchgeführt. Die Vorbereitung wird zuerst für die Teil-Pfade Sender zum Voxel durchgeführt. Anschließend werden die Werte für den zweiten Teil-Pfad von Voxel zum Empfänger bestimmt. Da der Programmablauf für die Berechnungen bis auf die genutzten Sender bzw. Empfänger gleich ist, wird in Abbildung 7.8 links nur der Ablauf für die Sender-Pfade detailliert dargestellt.

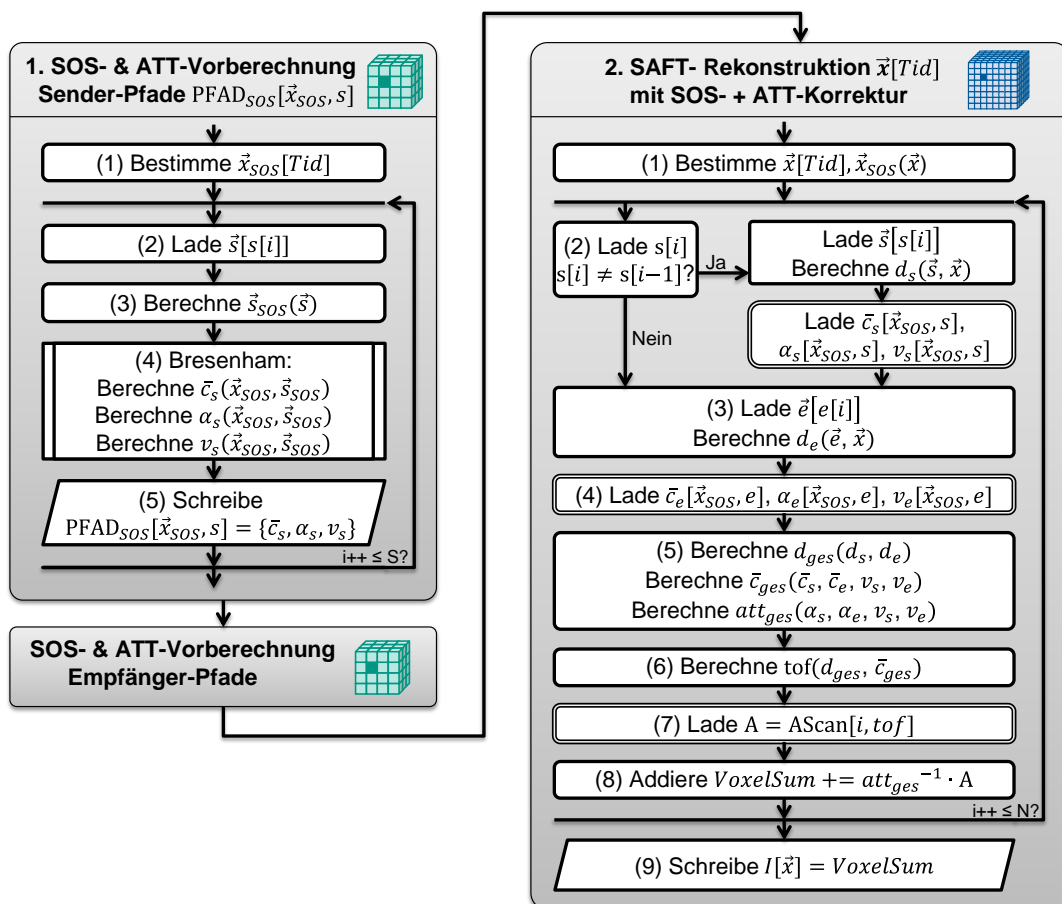


Abbildung 7.8: Ablaufdiagramm der SAFT-Rekonstruktion mit SOS- und ATT-Korrektur für einen Voxel.

1. Vorberechnung der mittleren Schallgeschwindigkeit und Dämpfung

Der CUDA-Kernel für die Vorberechnung wird für alle SOS-Voxel aufgerufen und auf der GPU parallel berechnet.

1. Über die eindeutige *ThreadID* (*Tid*), die jeder CUDA-Kernel bei seinem Aufruf zugewiesen bekommt, wird zu Beginn bestimmt, für welchen SOS-Voxel \vec{x}_{SOS} der Teilpfad berechnet werden soll. Ist die Position des SOS-Voxels \vec{x}_{SOS} bekannt, wird eine Schleife über alle S genutzten Sender gestartet.
2. Innerhalb der Schleife wird die Position des Senders \vec{s} aus dem *constant memory* geladen. Die Rekonstruktion wird sequentiell mit den aus MATLAB übergebenen A-Scan-Datenblöcken durchgeführt (siehe Kapitel 4.2.1). Daher muss, bevor die Position des Senders gelesen werden kann, zuvor durch eine Lookup-Tabelle bestimmt werden, welcher Sender $s[i]$ in dem aktuellen Durchlauf genutzt wird.
3. Ist die Position des Senders bekannt, wird die Position des Senders \vec{s}_{SOS} in SOS-Volumen-Koordinaten berechnet.
4. Da der Pfad beim Sender \vec{s}_{SOS} beginnt und beim SOS-Voxel \vec{x}_{SOS} endet, sind Start- und Endpunkt des zu berechnenden Pfads in SOS-Volumen-Koordinaten bekannt. Nun wird mit dem Bresenham-Algorithmus der Pfad voxelweise durchlaufen. Mit jedem SOS-Voxel wird ein neuer Schallgeschwindigkeitswert in c_s aufaddiert. Die Anzahl der besuchten SOS-Voxel v_s ist durch die größte Distanz bereits bekannt und muss nicht extra aufsummiert werden. Am Ende von Schritt 4 stehen somit die Summen der Schallgeschwindigkeitswerte in c_s sowie die Anzahl der besuchten SOS-Voxel in v_s zur Verfügung. Wird die Dämpfungskorrektur genutzt, werden die Dämpfungskoeffizienten in α_s pro SOS-Voxel aufsummiert.
5. Im letzten Schritt werden die ermittelten Werte zu der *float4*-Variable $PFAD_{SOS}$ kombiniert und in die für den SOS-Voxel entsprechende Speicherstelle geschrieben. Wird nur die Schallgeschwindigkeitskorrektur genutzt, genügt eine *float2*-Variable für $PFAD_{SOS}$. Damit ist die Berechnung für einen Sender-Pfad beendet und es wird, soweit noch nicht alle S Sender-Pfade abgearbeitet wurden, der nächste Sender-Pfad berechnet.

2. SAFT-Rekonstruktion mit Schallgeschwindigkeits- und Dämpfungskorrektur

Sind alle Sender- und Empfänger-Pfade zu allen SOS-Voxel vorberechnet, kann in einem zweiten Schritt die SAFT-Rekonstruktion mit Schallgeschwindigkeits- und Dämpfungskorrektur durchgeführt werden. Dafür wird der CUDA-Kernel für die SAFT-Rekonstruktion für alle Voxel aufgerufen und auf der GPU parallel berechnet.

1. Abhängig von der zugewiesenen *Tid* wird bestimmt, welcher Voxel \vec{x} berechnet wird. Für den Zugriff auf die vorberechneten Schallgeschwindigkeits- und Dämpfungswerte in (2) und (4) werden auch die SOS-Volumen-Koordinaten des Voxels \vec{x}_{SOS} benötigt. Sind die Koordinaten bekannt, wird eine Schleife über alle N genutzten A-Scans gestartet.
2. Innerhalb der Schleife wird Sender $s[i]$ aus dem *constant memory* geladen und überprüft. Die A-Scans liegen dabei in einer sortierten Reihenfolge vor, bei der für einen Sender alle genutzten Empfänger nacheinander angeordnet sind. Handelt es sich nicht um den Sender vom letzten Schleifen-Durchlauf $s[i] \neq s[i-1]$, ist die Position des Senders \vec{s} neu und alle vom Sender abhängigen

Daten müssen bestimmt werden. Die Position des Senders \vec{s} wird aus dem *constant memory* geladen. Der Abstand d_s zwischen Sender \vec{s} und Voxel \vec{x} wird bestimmt. Abhängig von den SOS-Volumen-Koordinaten des Voxels \vec{x}_{SOS} wird auf die für den Pfad vorberechneten Werte c_s und v_s zugegriffen. Wird die Dämpfungskorrektur genutzt, wird zusätzlich α_s gelesen. Dabei werden die Werte durch die Textureinheit der GPU bereits interpoliert zurückgegeben. Handelt es sich dagegen um keinen neuen Sender $s[i] = [i - 1]$, liegen die vom Sender abhängigen Werte bereits in den lokalen Registern vor und müssen nicht mehr neu geladen bzw. berechnet werden.

3. Durch die Reihenfolge, in der die A-Scans abgearbeitet werden, ändert sich die Empfängerposition bei jedem A-Scan. Daher werden in jedem Durchlauf der Schleife die aktuellen Empfängerkoordinaten \vec{e} , abhängig vom aktuellen Empfängerindex $e[i]$ aus dem *constant memory* geladen und der Abstand d_e zwischen Empfänger \vec{e} und Voxel \vec{x} bestimmt.
4. Abhängig von den SOS-Volumen-Koordinaten des Voxels \vec{x}_{SOS} und dem Empfänger e wird auf die für den Pfad vorberechneten Werte c_e und v_e zugegriffen. Wird die Dämpfungskorrektur genutzt, wird zusätzlich α_e gelesen. Diese können durch die GPU bereits interpoliert gelesen werden.
5. Die Informationen über die Teilpfade werden nun zu einer Gesamtdistanz d_{ges} , einer mittleren Schallgeschwindigkeit \bar{c}_{ges} und gegebenenfalls einer Gesamtdämpfung att_{ges} auf dem gesamten Pfad von Sender über den Voxel zum Empfänger verrechnet.
6. Der Zeitpunkt tof wird abhängig von der Gesamtdistanz d_{ges} und der mittleren Schallgeschwindigkeit \bar{c}_{ges} nach Gleichung (3.12) bestimmt.
7. Die Amplitude A wird aus dem aktuellen A-Scan am Zeitpunkt tof , durch den Zugriff über das *texture memory*, bereits interpoliert gelesen.
8. Die Amplitude A wird nach Gleichung (6.8) gegebenenfalls mittels des reziproken Werts der Gesamtdämpfung att_{ges} korrigiert und in *VoxelSum* aufsummiert. Sind noch nicht alle N A-Scans im A-Scandatenblock von MATLAB abgearbeitet, wird die Schleife mit dem nächsten A-Scan neu gestartet.
9. Nach dem Schleifendurchlauf steht die Summe aller A-Scanbeiträge in der Variablen *VoxelSum* und wird in das Ausgabevolumen $I[\vec{x}]$ geschrieben.

7.1.4 Performanz

Tabelle 7.4 zeigt die Rechengeschwindigkeit der Berechnungsstrategien für die Bild-Rekonstruktion mit Korrektur der Annahme einer konstanten Schallgeschwindigkeit im Vergleich zur Rechengeschwindigkeit ohne Korrektur. Die Standard-SAFT-Rekonstruktion besitzt mit 13,6 GVA/s die höchste Rechengeschwindigkeit. Die Rekonstruktion mit Korrektur der Annahme einer konstanten Schallgeschwindigkeit nach Strategie Kernel, die direkt im SAFT-Kernel stattfindet, reduziert die Rechengeschwindigkeit um einen Faktor 34 auf 0,4 GVA/s. Dabei besitzt die Berechnung der mittleren Schallgeschwindigkeit den größten Anteil von 95% und die SAFT-Rekonstruktion nur 5% der Berechnungsdauer. Die Rekonstruktion mit der optimierten Vorberechnung der mittleren Rechengeschwindigkeiten nach Strategie SOSTPV1 ist dagegen um das 18,5-fache schneller und erreicht etwas mehr als die halbe Rechengeschwindigkeit der Version ohne Korrektur. Dabei beträgt die Vorberechnung der mittleren Schallgeschwindigkeiten auf den Pfaden nur noch 0,5% der gesamten Berechnungsdauer. Die SAFT-Rekonstruktion bleibt die Hauptaufgabe mit 99,5% der

Berechnungsstrategie	Vorbereitung der SOS-Pfade	SAFT Algorithmus	Performanz in GVA/s
Ohne Korrektur	-	2 h 26 min 100 %	13,6
im SAFT Kernel (Strategie Kernel)	78 h 34 min 95%	4 h 8 min 5 %	0,4
Vorbereitung (Strategie SOSTPV1)	1 min 0,5 %	4 h 24 min 99,5 %	7,4
Vorbereitung mit Interpolation (Strategie SOSTPVI2)	4 min 1,5 %	4 h 26 min 98,5 %	7,3

Tabelle 7.4: Vergleich der Rechengeschwindigkeit der Berechnungsstrategien auf einer Einzel-GPU der GTX 590 Doppel-GPU. Für ein Volumen von $1024^2 \times 64$ Voxel und SOS-Volumen von 128^3 Voxeln.

Berechnungsdauer. Es treten jedoch Eckartefakte bei größeren Unterschieden zwischen benachbarten SOS-Voxel auf. Strategie SOSTPVI2 erweitert die Berechnung nach Strategie SOSTPV1 mit einer Interpolation und vermeidet so die unerwünschten Eckartefakte. Trotz zusätzlicher Interpolation erreicht Strategie SOSTPVI2 nahezu dieselbe Rechengeschwindigkeit, indem der *textur memory* der GPU ausgenutzt wird, der die 3D Interpolation direkt in Hardware ermöglicht. Daher ist Strategie SOSTPVI2 die am besten geeignete Berechnungsstrategie für diese Anwendung. Wird dieses Strategie genutzt um zudem eine Dämpfungskorrektur durchzuführen, kann eine maximale Performanz von 4,2 GVA/s erreicht werden. Diese liegt damit etwas über der halben Rechengeschwindigkeit der Schallgeschwindigkeitskorrektur.

7.2 Weitere Beschleunigung der SAFT für heterogene Medien

Für die klinische Anwendbarkeit von 3D-USCT sollen Bilder innerhalb einer Zeitspanne von 15 min rekonstruiert werden können. Bei Rekonstruktionen mit Korrektur der Schallgeschwindigkeit wird, wie in der ersten klinischen Studie [7], aktuell eine Auflösung bis $(0,9 \text{ mm})^3$ genutzt (siehe Kapitel 5.1). Dies entspricht einem Volumen innerhalb der ROI von $444^2 \times 266$ Voxel. Es werden Messdaten von zehn Aperturpositionen verwendet. Dabei beschränkt sich die Rekonstruktion für die Reflexionstomographie auf etwa die Hälfte der Messdaten, um nur die A-Scans zu nutzen, die eine gute Zeitauflösung der Reflexionen erlauben (siehe Abbildung 3.14). Dadurch ergibt sich eine genutzte Menge von ca. 4,5 Millionen A-Scans. Abbildung 7.9 gibt einen Überblick über die benötigten Laufzeiten für die Standard-SAFT und der SAFT mit Schallgeschwindigkeits- und Dämpfungskorrektur bei dem Einsatz von acht GPUs. Da zum Zeitpunkt dieser Arbeit nur eine GTX TITAN zur Verfügung stand, werden die Performanzwerte bei acht GTX TITAN GPUs extrapoliert angegeben. Dabei wird die Annahme getroffen, dass der Overhead, wie bei den GTX 590 GPUs in Kapitel 4.2.4, bei einer genügend großen Auslastung vernachlässigbar ist. Während mit der GTX 590 für das Volumen mit der Standard-SAFT 37 Minuten nötig sind, werden knapp zwei Stunden für die SAFT-Version mit Korrekturen benötigt. Werden acht GTX TITAN genutzt, kann die Rechendauer auf etwa eine Stunde reduziert werden. Dennoch liegt die Berechnungsdauer deutlich über den geforderten 15 min. Um die Rekonstruktion innerhalb der gewünschten Zeit durchführen zu können, wird eine Rechengeschwindigkeit von mindestens 32,3 GVA/s pro GPU benötigt. Eine weitere Beschleunigung kann durch geeignete Approximationen durchgeführt werden. Dies kann jedoch zu einem Verlust an Bildqualität führen. Im Folgenden wird eine Approximation vorgestellt, die eine deutliche Beschleunigung erreicht, obwohl sie eine Schallgeschwindigkeits- und Dämpfungskorrektur beinhaltet. Durch eine dreistufige Berechnung können sowohl die Speicher-Zugriffe als auch die Rechenoperationen deutlich reduziert werden. Der durch die Approximation ent-

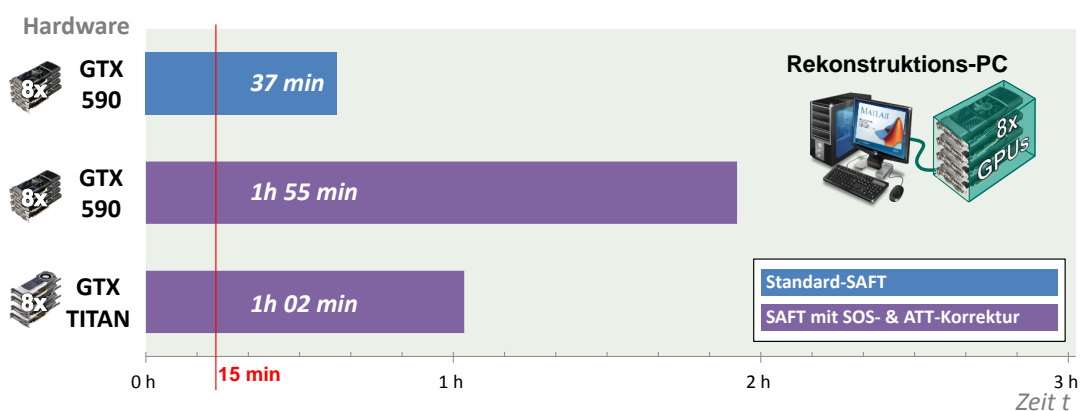


Abbildung 7.9: Überblick der Rechendauer für die Anforderungen der klinischen Anwendbarkeit ($444^2 \times 266$ Voxel, 10 Aperturpositionen, 15 min) über mehrere SAFT-Versionen und GPUs.

stehende Fehler ist dabei innerhalb der Brust, dem für die 3D-USCT interessanten Bereich (ROI), deutlich geringer als außerhalb der Brust.

7.2.1 Methodik und Umsetzung

Bisher wurde in der beschleunigten Implementierung der SAFT die mittlere Schallgeschwindigkeit und die Dämpfung auf den Pfaden im Speicher abgelegt. Die Berechnung erfolgte dabei für die Positionen der SOS- und ATT-Voxel. Erst in einem zweiten Schritt wurden die SOS- und ATT-Werte, abhängig von der Position des Reflexionsvoxels, interpoliert und für die Berechnung des *tof*-Zeitpunktes genutzt, der den Zugriff auf die A-Scans bestimmt. Dieses Vorgehen ermöglicht die volle Ausnutzung der Information der SOS- und ATT-Volumen, da für jeden SOS- und ATT-Voxel die exakten Werte im Voraus berechnet wurden. Bei diesem Vorgehen werden eine Menge an Variablen benötigt, um den *tof*-Zeitpunkt zu berechnen. So wird die Sender- und Empfängerposition abhängig vom A-Scan bestimmt und in SOS-Volumenkoordinaten umgerechnet. Für die mittlere Schallgeschwindigkeit werden die interpolierten Schallgeschwindigkeiten und die Anzahl der besuchten SOS-Voxel auf den Teilpfaden gelesen. Die Länge des Pfades muss pro Voxel und SEP bestimmt und durch die mittlere Schallgeschwindigkeit geteilt werden, um den *tof*-Zeitpunkt zu ermitteln. Ist dieser bestimmt, kann die Amplitude aus dem A-Scan gelesen, abhängig von der Dämpfung korrigiert und aufsummiert werden. Dieses Vorgehen beschreibt letztendlich den SAFT-Algorithmus als eine Zuordnung vom A-Scansignal zum Voxel.

Die Idee einer neuen Art der Beschleunigung beruht nun darauf, den *tof*-Zeitpunkt nur noch für die SOS-Voxel zu berechnen. Anschließend wird der entsprechende *tof*-Zeitpunkt in dem SAFT-Kernel für die Reflexionsvoxel interpoliert gelesen und direkt für den Zugriff auf den A-Scan genutzt. Es wird so ein weiterer Berechnungsschritt eingeführt, der jedoch nur für die SOS-Voxel vorberechnet werden muss und bereits die SOS-Korrektur beinhaltet. Der *tof*-Zeitpunkt hängt jedoch nicht von der Dämpfung *att* auf dem Pfad ab. Soll die Dämpfungskorrektur durchgeführt werden, muss die Dämpfung *att* neben dem *tof*-Zeitpunkt ebenfalls in einer *float2*-Variablen gespeichert werden. Diese dreistufige Berechnungsstrategie mit Approximation der *tof*-Zeitpunkte wird im Folgenden als **Strategie TOFI** bezeichnet.

Implementierung des weiter beschleunigten SAFT-Algorithmus

Abbildung 7.10 stellt den Programmablauf für den weiter beschleunigten SAFT-Algorithmus nach Strategie TOFI mit Schallgeschwindigkeits- und Dämpfungskorrektur dar. Der Programmablauf der beiden Vorberechnungsschritte entspricht nahezu dem bereits beschriebenen Ablauf in Kapitel 7.1.3 für die beschleunigte Berechnung nach Strategie SOSTPVI2. Daher werden im Folgenden die Vor- bzw. Nachteile der Änderungen beschrieben.

Vorteile

- Mit dem zweiten Vorberechnungsschritt wird der größte Teil der SAFT-Berechnung vorgezogen und nur noch für die geringere Anzahl an SOS-Voxel durchgeführt. Bei einer Auflösung von $1667^2 \times 1000$ Voxel für das Reflexionsvolumen und $100^2 \times 60$ SOS-Voxel für das SOS-Volumen (siehe Kapitel 7.1) ist die SOS-Auflösung um einen Faktor 16,7 kleiner. Das bedeutet, dass nur noch für 0,02% der Voxel die Berechnung des *tof*-Zeitpunktes durchgeführt werden muss.

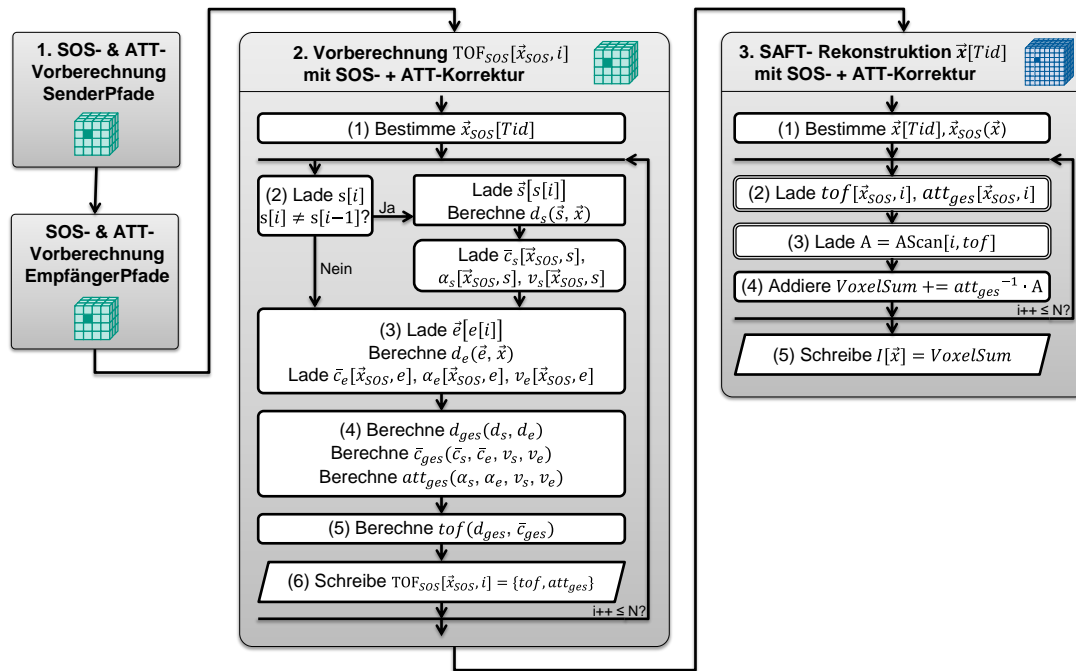


Abbildung 7.10: Ablaufdiagramm für den weiter beschleunigten SAFT-Algorithmus nach Strategie TOFI mit Schallgeschwindigkeits- und Dämpfungskorrektur für einen Voxel.

- Der Vorteil durch die Pfad-Vorbereitung der Teil-Pfade bleibt erhalten. D.h. die Teil-Pfade müssen nach wie vor nur einmal vorberechnet werden.
- In den Vorbereitungsschritten kann direkt in den Koordinaten der SOS-Voxel gerechnet werden. D.h. es wird für die Bestimmung der mittleren Schallgeschwindigkeit, Dämpfung und Anzahl der SOS-Voxel auf den Teilpfaden keine Interpolationen mehr benötigt. Statt den, für die lineare Interpolation in 3D notwendigen acht Werten, muss jeweils nur ein Wert gelesen werden.
- Die Berechnung im SAFT-Kern für die Reflexionsvoxel kann minimiert werden. Es wird eine Schleife über alle A-Scans benötigt, die unabhängig von Sender- und Empfänger-Koordinaten durchlaufen werden kann. Über einen Lesezugriff auf das *Surface* liegen direkt der *tof*-Zeitpunkt und die Dämpfung att_{ges} interpoliert vor und können für die SAFT-Berechnung genutzt werden.

Nachteile

- Die in dem zweiten Vorbereitungsschritt berechneten *tof*- und *att*-Werte müssen zwischengespeichert werden, wodurch der Speicherverbrauch weiter erhöht wird. Da im SAFT-Kern bei Zugriff auf die Daten interpoliert werden soll, werden sie in einem weiteren *Surface* gespeichert. Die Größe des benötigten Speichers ist dabei abhängig von der Anzahl N der genutzten A-Scans in einem A-Scandatenblock und der Anzahl V_{SOS} an SOS-Voxel. Für die Interpolation in 3D müssen wiederum mindestens zwei SOS-Z-Schichten vorberechnet und gespeichert werden. Insgesamt bedeutet das eine benötigte Speichergröße von zusätzlich 160 kB pro A-Scan bei einem SOS-Volumen von $100^2 \times 2$ SOS-Voxel. In der GPU-Hardware sind jedoch die maximalen Größen

der *Textures* und *Surfaces* beschränkt [39]. Daher kann in einer *Textur* eine maximale Anzahl von $N=1024$ A-Scans genutzt werden. Dies entspricht einem Speicherbedarf von 163,8 MB.

Durch die Nutzung mehrerer *Surfaces* bzw. *Textures* der Anzahl T kann diese Beschränkung überwunden werden. Jedoch wird für die Zuordnung zu den T *Surfaces* eine zusätzliche Abfrage im Vorberechnungs- und SAFT-Kernel benötigt. Wie in Kapitel 4.2.4 beschrieben, werden üblicherweise A-Scandatenblöcke der Größe von 10.000 A-Scans oder größer genutzt. Ein A-Scandatenblock dieser Größe würde mit einem Speicherbedarf von 1,6 GB die GPU-Speichergröße der GTX 590 überschreiten. Die Vorberechnung der *tof*-Zeitpunkte und die Rekonstruktion wird in diesem Fall sequentiell mit einer geringeren Anzahl an A-Scans mehrfach durchgeführt. Tabelle 7.5 fasst den Speicherbedarf zusammen.

- Ein weiterer Nachteil ist der Bildqualitätsverlust, der durch die Approximation der Wurzel in der Abstandsberchnung durch die lineare Interpolation auftritt. Der gemachte Fehler ist abhängig von dem Verhältnis der Auflösungen, dem Abstand zwischen Sender und Empfänger und der Position des Voxel \vec{x} . Dieser Fehler wird im Folgenden weiter untersucht.

7.2.2 Fehleranalyse der Approximation

Der durch die Approximation der *tof*-Zeitpunkte bei linearer Interpolation gemachte Fehler soll im Folgenden abgeschätzt werden. Dabei wird der *tof*-Zeitpunkt in Abtastpunkten (Samples) bei einer Abtastrate der A-Scans von $f_s = 10$ MHz angegeben. Unter der vereinfachenden Annahme einer konstanten Schallgeschwindigkeit wird für jeden *tof*-Zeitpunkt eine Ellipse in das zu rekonstruierende Volumen projiziert (siehe Abbildung 7.11). Der Laufweg vom Sender zum Empfänger über einen beliebigen Punkt auf der Ellipse besitzt immer dieselbe Länge $l = \frac{c \cdot tof}{f_s}$. Dagegen ist die Distanz d von einem beliebigen Punkt auf der Ellipse zum Ellipsenmittelpunkt abhängig von dem Abstand \overline{SE} zwischen Sender und Empfänger und dem Winkel α . Die Größe des Abstands liegt dabei zwischen d_2 bei 90° und d_1 bei 180° . Während für d_1 als oberer Grenzfall mit

$$d_1 = \frac{l}{2} = \frac{c}{2f_s} \cdot tof \quad (7.4)$$

Akronym	Berechnungsstrategie	zu speichernde Daten	benötigter Speicher
TOFI	Vorbereitung der - Sender- und Empfängerpfade - <i>tof</i> -Indizes und <i>att</i> für zwei SOS-Z-Schichten	(4+4) Byte · 1024 A-Scans · T genutzte Texturen · $100^2 \times 2$ SOS-Voxel	163,8 MB · T

Tabelle 7.5: Übersicht der Berechnungsstrategie TOFI und ihr Speicherbedarf.

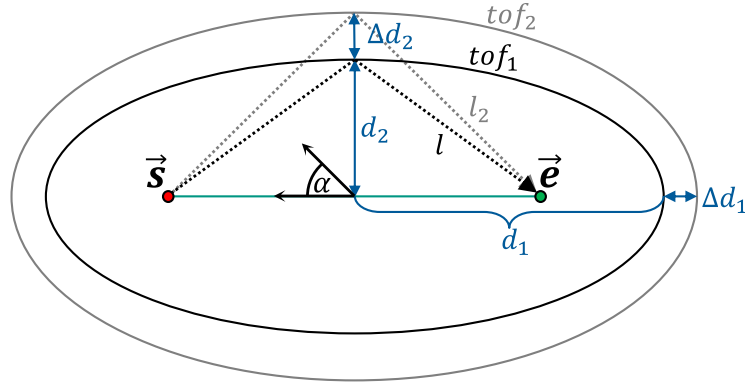


Abbildung 7.11: Schematische Darstellung von zurückprojizierten Ellipsen für zwei *tof*-Samples und entsprechenden Schrittweiten Δd_1 und Δd_2 .

ein linearer Zusammenhang zwischen d und tof besteht, hängt d_2 als unterer Grenzfall zusätzlich von dem Abstand \overline{SE} ab:

$$\begin{aligned} d_2 &= \sqrt{\left(\frac{l}{2}\right)^2 - \left(\frac{\overline{SE}}{2}\right)^2} \\ &= \frac{1}{2} \sqrt{\left(\frac{c \cdot tof}{f_s}\right)^2 - (\overline{SE})^2} \end{aligned} \quad (7.5)$$

wobei d_2 für $l \gg \overline{SE}$ in d_1 übergeht.

Dieser Verlauf der *tof*-Samples wird nun an den Positionen der SOS-Voxel abgetastet. *tof* steht in einem linearen Zusammenhang in die Richtung von d_1 und kann daher über eine lineare Interpolation ohne Fehler approximiert werden. Da das Volumen an den SOS-Voxel abgetastet wird, kann es nur an den Positionen der Sender und Empfänger zu Fehlern kommen, da diese gewöhnlich nicht direkt an der Position der SOS-Voxel liegen.

Die Brust liegt bei der 3D-USCT jedoch normalerweise in der Richtung von d_2 . Der Fehler durch die Approximation des *tof*-Verlaufs mittels linearer Interpolation ist bei geringem Abstand d_2 groß. Für einen zunehmenden Abstand d_2 zu Sender und Empfänger geht der Verlauf von *tof* in Richtung d_2 ebenfalls in einen linearen Verlauf über. Liegen Sender und Empfänger nah beieinander wird dieser Übergang in einen linearen Verlauf früher erreicht. Dieser Fall ist insbesondere interessant, da für die Reflexionstomographie die SEP mit geringem Abstand voneinander genutzt werden. Ein weiterer interessanter Aspekt ist, dass die genutzte ROI in einer gewissen Entfernung zu den SEP liegt. Da die größten Fehler bei einem geringen Abstand d_2 auftreten, liegen diese hauptsächlich außerhalb der Brust.

Soll die maximale Auflösung von $res = (0,24\text{ mm})^3$ erreicht werden, ist die Abweichung begrenzt, die durch die *tof*-Approximation entstehen darf. Damit sich die zurückprojizierten Echosignale noch konstruktiv überlagern, darf keine größere Abweichung Δd als die Hälfte der Auflösung verursacht werden:

$$\Delta d \stackrel{!}{\leq} \frac{res}{2} \quad (7.6)$$

Für d_1 lässt sich somit die maximal erlaubte Abweichung Δtof_{max} über

$$\begin{aligned}\Delta d_1 &= d_1(tof_2) - d_1(tof_1) \stackrel{!}{\leq} \frac{res}{2} \\ \frac{tof_2 \cdot c}{2f_s} - \frac{tof_1 \cdot c}{2f_s} &\leq \frac{res}{2} \\ \Delta tof_{max} &\leq \frac{res \cdot f_s}{c}\end{aligned}\quad (7.7)$$

unabhängig vom Abstand \overline{SE} und dem Laufweg l ermitteln. Unter der Annahme einer Schallgeschwindigkeit von $c = 1500 \text{ m/s}$ ergibt dies eine maximale Abweichung Δtof_{max} von 1,6 Samples bzw. $0,16 \mu\text{s}$.

Die maximale Abweichung für den zweiten Grenzfall besitzt dagegen einen komplexeren Zusammenhang

$$\begin{aligned}\Delta d_2 &= d_2(l_2) - d_2(l_1) \stackrel{!}{\leq} \frac{res}{2} \\ \sqrt{\left(\frac{l_1 + \Delta l}{2}\right)^2 - \left(\frac{\overline{SE}}{2}\right)^2} - \sqrt{\left(\frac{l_1}{2}\right)^2 - \left(\frac{\overline{SE}}{2}\right)^2} &\leq \frac{res}{2},\end{aligned}\quad (7.8)$$

der sich umschreiben lässt zu

$$\begin{aligned}\sqrt{(l_1 + \Delta l)^2 - \overline{SE}^2} &\leq res + \sqrt{l_1^2 - \overline{SE}^2} \\ (l_1 + \Delta l)^2 &\leq \left(res + \sqrt{l_1^2 - \overline{SE}^2}\right)^2 + \overline{SE}^2 \\ \Delta l &\leq \sqrt{\left(res + \sqrt{l_1^2 - \overline{SE}^2}\right)^2 + \overline{SE}^2} - l_1.\end{aligned}\quad (7.9)$$

Mit $l = \frac{tof \cdot c}{f_s}$ kann der maximal erlaubte Fehler Δtof bestimmt werden zu

$$\Delta tof_{max} \leq \frac{f_s}{c} \cdot \sqrt{\left(res + \sqrt{\left(\frac{tof_1 \cdot c}{f_s}\right)^2 - \overline{SE}^2}\right)^2 + \overline{SE}^2} - tof_1. \quad (7.10)$$

Dabei ist Δtof_{max} vom Abstand \overline{SE} und dem Wert des tof_1 -Samples abhängig. Der tof_1 -Sample ist wiederum von der Position des Voxels \vec{x} abhängig, der rekonstruiert werden soll. Für bestimmte SEP und Voxel kann der maximale Fehler somit ermittelt werden. Der kleinstmögliche Sample $tof_1 = \overline{SE} \cdot f_s / c$ entspricht dabei der Laufzeit von Sender zum Empfänger. Kleinere Samples sind zwar im A-Scan vorhanden, können jedoch nur Rauschen enthalten, da der ausgesendete Ultraschall noch nicht beim Empfänger angekommen ist. Auf diese Samples wird bei der Rekonstruktion daher nicht zugegriffen.

Die Approximation soll auch in Abhängigkeit von dem Abstand \overline{SE} und der Position des Voxel \vec{x} weiter abgeschätzt werden. Daher wird die Größe des gemachten Fehlers anhand von drei Fällen abgeschätzt, welche die beiden Grenzfälle sowie einen mittleren Fall abdecken (siehe Abbildung 7.12 links). Im ersten Fall sind der Sender

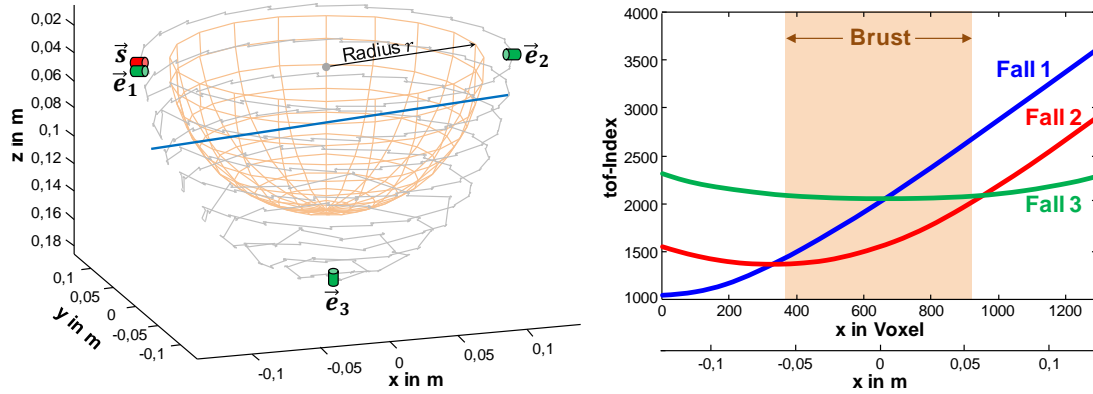


Abbildung 7.12: Anordnung der Sender und Empfänger der drei untersuchten Fälle mit Ortslinie in x -Richtung bei $y=0$ m und $z=0,13$ m in blau (links). Verläufe des tof -Index für die drei Fälle entlang der Ortslinie (rechts).

\vec{s} und Empfänger \vec{e}_1 so nah wie möglich positioniert und liegen auf einem *transducer array system* (TAS). In dem zweiten Fall sind der Sender \vec{s} und Empfänger \vec{e}_2 so weit wie möglich entfernt, d.h. sie liegen sich gegenüber. Dieser Fall stellt ein typischerweise genutztes SEP für die Transmissionstomographie dar. Im dritten Fall ist der Sender \vec{s} auf der höchst möglichen Position und der Empfänger \vec{e}_3 ist auf dem untersten TAS positioniert. Dieser Fall stellt den maximal genutzten Abstand \overline{SE} in der Reflexionstomographie dar (siehe auch Abbildung 3.14).

In Abbildung 7.12 rechts sind die Verläufe des tof -Index entlang der Ortslinie in x -Richtung beispielhaft bei $y=0$ m und $z=0,13$ m für alle drei Fälle dargestellt. Es ist zu erkennen, dass sich abhängig von der Position der Sender und Empfänger sowie der Position eines Voxels \vec{x} der Verlauf ändert. Der Verlauf ist dabei immer kontinuierlich. In Abbildung 7.13 rechts sind die Verläufe des tof -Index zudem in xy -Richtung für alle drei Fälle dargestellt, jeweils für die z -Schicht mit dem größten Fehler. Zusätzlich ist in Abbildung 7.13 links die Verteilung des Fehlers über das gesamte Volumen jeweils in einem Histogramm dargestellt.

Die statistischen Fehlermaße über das gesamte Volumen sind mit Δtof_{ges} in Tabelle 7.6 angegeben. Eine besondere Fehlerbetrachtung wird hier für den Voxel \vec{b} durchgeführt, der den Wort Case darstellt. Dabei ist \vec{b} als der Voxel definiert, der den Brustanfang für das jeweilige SEP darstellt. Voxel \vec{b} befindet sich somit auf der Kante der ROI mit minimaler tof -Laufzeit und besitzt dadurch den größten Fehler durch die Approximation. $\Delta tof_{max}(\vec{b})$ stellt den maximal erlaubten Fehler für den Voxel \vec{b} dar. $\Delta tof(\vec{b})$ ist der gemessene Fehler für den Voxel \vec{b} . Die Fehlerwerte werden für die Fälle einer linearen Interpolation als auch einer Spline-Interpolation angegeben. Für den ersten Fall liegt der Fehler für den Brustanfang $\Delta tof(\vec{b})$ um den Faktor 7,5 deutlich unterhalb des maximal erlaubten Fehlers, der mittlere Fehler ist um den Faktor 26,5 niedriger. Für den dritten Fall liegt der Fehler für den Brustanfang um den Faktor 4,8 unterhalb des maximal erlaubten Fehlers, der mittlere Fehler ist um den Faktor 8,1 niedriger. Für den zweiten Fall wird die Laufzeit zum Brustanfang \vec{b} gleich der Laufzeit zum Empfänger genutzt. Für den Fall $tof(\vec{b}) = \overline{SE} \cdot f_s / c$ beträgt der maximal erlaubte Fehler 0,00074 Samples, der mit einem gemessenen Fehler von 0,15 Samples um das 214-fache überschritten wird. Allerdings wird dieser

SEP	\overline{SE} in m	$\Delta tof_{max}(\vec{b})$ in Samples	Inter- polation	$\Delta tof(\vec{b})$ in Samples	Δtof_{ges} in Samples			
					Mean	STD	Min	Max
1. \vec{s}, \vec{e}_1	0,001	1,59	linear spline	0,21 0,0006	0,06 0,0001	0,04 0,002	0,0 0,0	9,2 4,4
2. \vec{s}, \vec{e}_2	0,26	0,0007	linear spline	0,15 0,0001	0,06 0,0001	0,03 0,0007	0,0 0,0	4,3 1,3
3. \vec{s}, \vec{e}_3	0,21	0,57	linear spline	0,12 0,00009	0,07 0,0001	0,04 0,002	0,0 0,0	10,7 5,2

Tabelle 7.6: Überblick der ermittelten *tof*-Werte für die drei Fälle.

Fall bei der Reflektivitätsrekonstruktion nicht genutzt. Für alle drei Fälle liegt die Summe aus mittlerem Fehler und der Standardabweichung (STD) über das gesamte Volumen unter 0,11 Samples. Es treten mit maximal 10,7 Samples zwar deutlich größere Fehler auf, diese liegen wie in Abbildung 7.13 rechts gezeigt, jedoch außerhalb der Brust. Die lineare Interpolation kann hoch performant in der GPU-Hardware durchgeführt werden. Eine Interpolation höherer Ordnung kann durch eine kubische Spline-Interpolation laut [98] mit nur acht linearen Interpolationen ebenfalls effizient auf der GPU ausgeführt werden. Eine Approximation mit Spline-Interpolation ist daher etwas rechenaufwändiger durchzuführen, zeigt jedoch im Mittel einen um den Faktor 600 geringeren Fehler. Das Potential einer Spline-Interpolation eignet sich daher insbesondere für geringere SOS-Auflösungen, bei denen die lineare Interpolation, bedingt durch die zu geringe Abtastung, zu großen Fehlern führen würde.

7.2.3 Evaluation

Bisher wurde der Fehler jeweils nur für einen A-Scan betrachtet. Da sich der Fehler jedoch über verschiedene SEP verändert, ist am Ende die bei der SAFT-Rekonstruktion gebildete Summe für die Bildqualität entscheidend. Daher wurde die PSF-Untersuchung aus Kapitel 6.1.2 ebenfalls mit der beschleunigten Berechnung nach Strategie TOFI durchgeführt. Da die Bildqualität der Strategie TOFI vom Abstand zu den Sendern und Empfängern abhängig ist, wurden mehrere Punktstreuer auf einem Viertel der Brust simuliert. Diese sind zudem in den Ecken der SOS-Voxel positioniert um den Worst-Case abzubilden. Die Schallgeschwindigkeiten wurden mit $c_{Wasser} = 1500 \text{ m/s}$ und $c_{Brust} = 1460 \text{ m/s}$ angenommen. Das Ergebnis ist für drei Blickrichtungen in Abbildung 7.14 rechts dargestellt. Für einen Vergleich ist zudem die Untersuchung der PSF für den korrigierten SAFT nach Strategie SOSTPVI2 dargestellt (siehe Abbildung 7.14 links). Tabelle 7.7 gibt dabei die Bildqualitätsmetriken gemittelt über alle simulierten Punktstreuer an. Der Kontrast wird dabei relativ zu dem Ergebnis der SOSTPVI2-Rekonstruktion berechnet. Es ist zu erkennen, dass der TOFI-Algorithmus nahezu dieselbe Auflösung, Kontrast und Verschiebung erreicht.

Neben den Simulationsdaten wird die Methode mit experimentellen Daten evaluiert, die in Kapitel 8.2 eingeführt werden. Dabei werden die Rekonstruktionsvarianten Standard-SAFT, Schallgeschwindigkeitskorrigierter SAFT und beschleunigter TOFI-SAFT miteinander verglichen. Es werden jeweils drei interessante Schnitte in x-, y- und z-Richtung dargestellt.

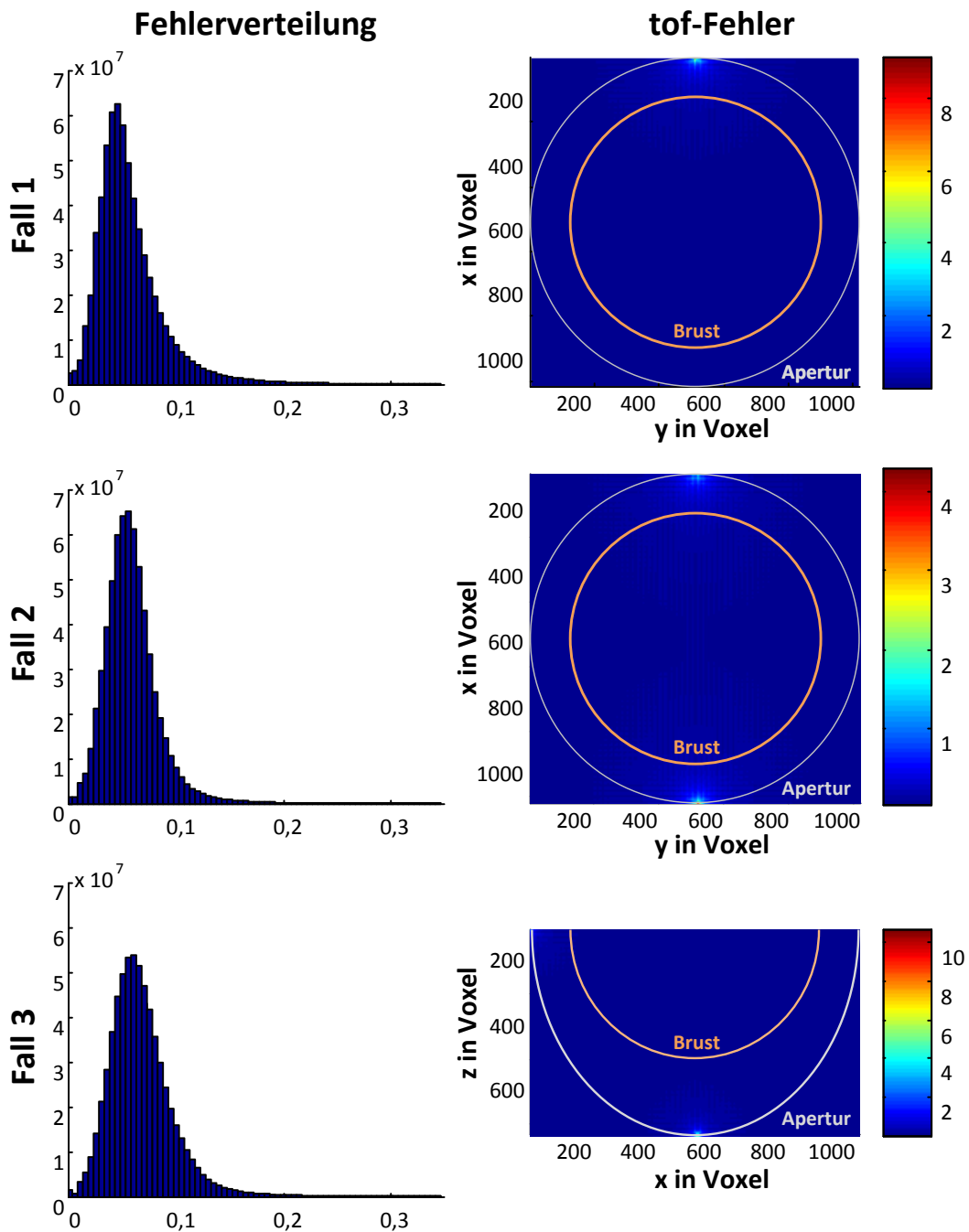


Abbildung 7.13: Verteilung des *tof*-Fehlers als Histogramm über das gesamte Volumen für die drei untersuchten Fälle (links). *tof*-Fehler für die beiden ersten Fälle (rechts oben, mittig) und für den dritten Fall (rechts unten) jeweils für die Ebene mit dem größten Fehler dargestellt.

Tabelle 7.7: Vergleich der genutzten Bildqualitätsmetriken für die Berechnungsstrategie SOSTPVI2 und TOFI.

Strategie	\overline{FWHM}_{ges}	rel. Kontrast	Verschiebung
SOSTPVI2	0,33 mm	1	0,097 mm
TOFI	0,34 mm	0,98	0,104 mm

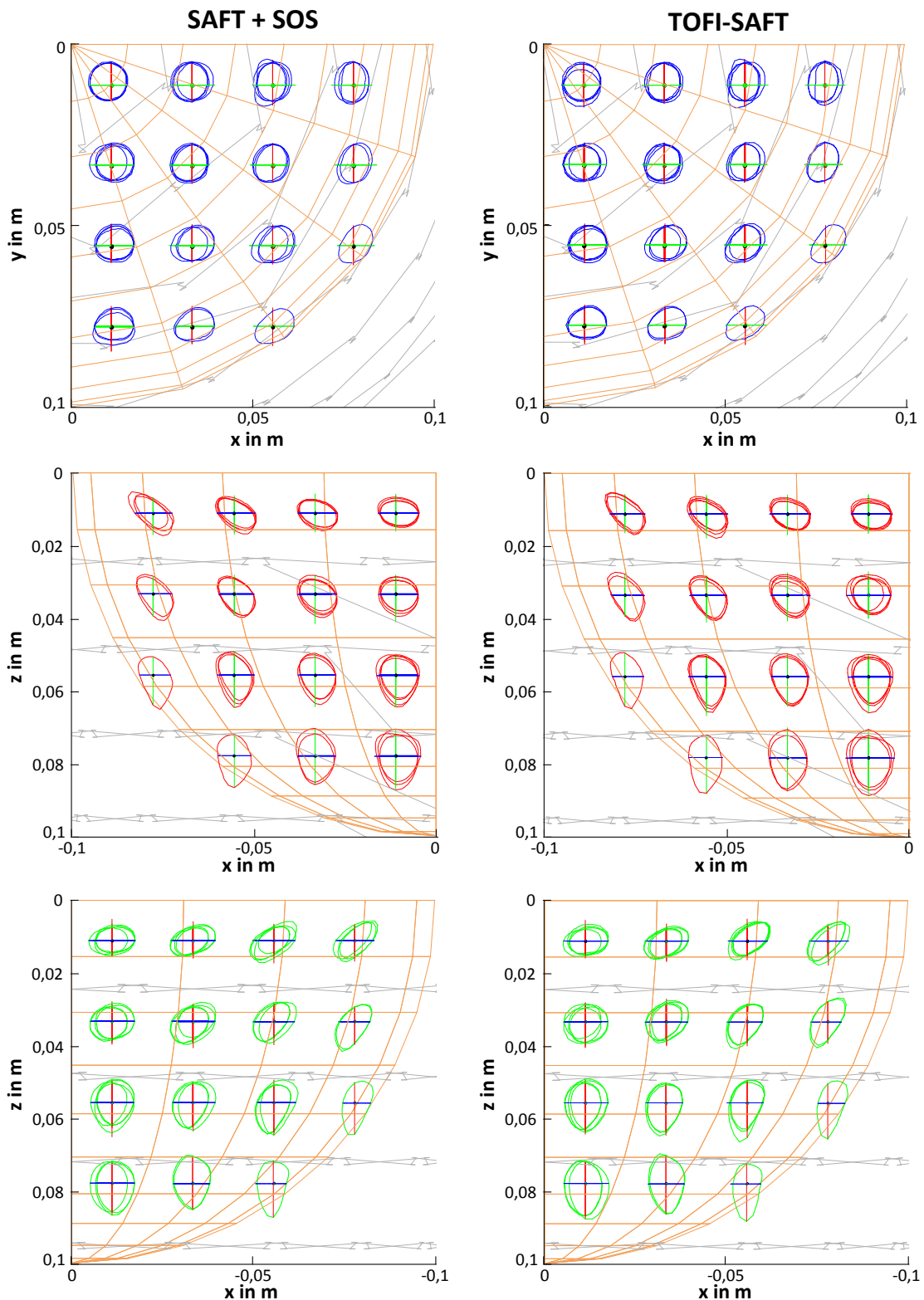


Abbildung 7.14: PSF der Punktstreuer bei Rekonstruktion mit dem korrigierten SAFT-Algorithmus (links) und dem beschleunigten TOFI-SAFT-Algorithmus (rechts) aus allen drei Blickrichtungen. Für die Darstellung der PSF wurde ein Vergrößerungsfaktor von 60 verwendet.

Schraube

Bei der Rekonstruktion der Schraube ergibt sich ein nahezu identisches Bild sowohl was den Dynamikbereich als auch den PSNR-Wert betrifft (siehe Abbildung 7.15). Hieraus lässt sich schließen, dass sich insbesondere Messobjekte gut rekonstruieren lassen, die weniger komplexe Streumuster enthalten.

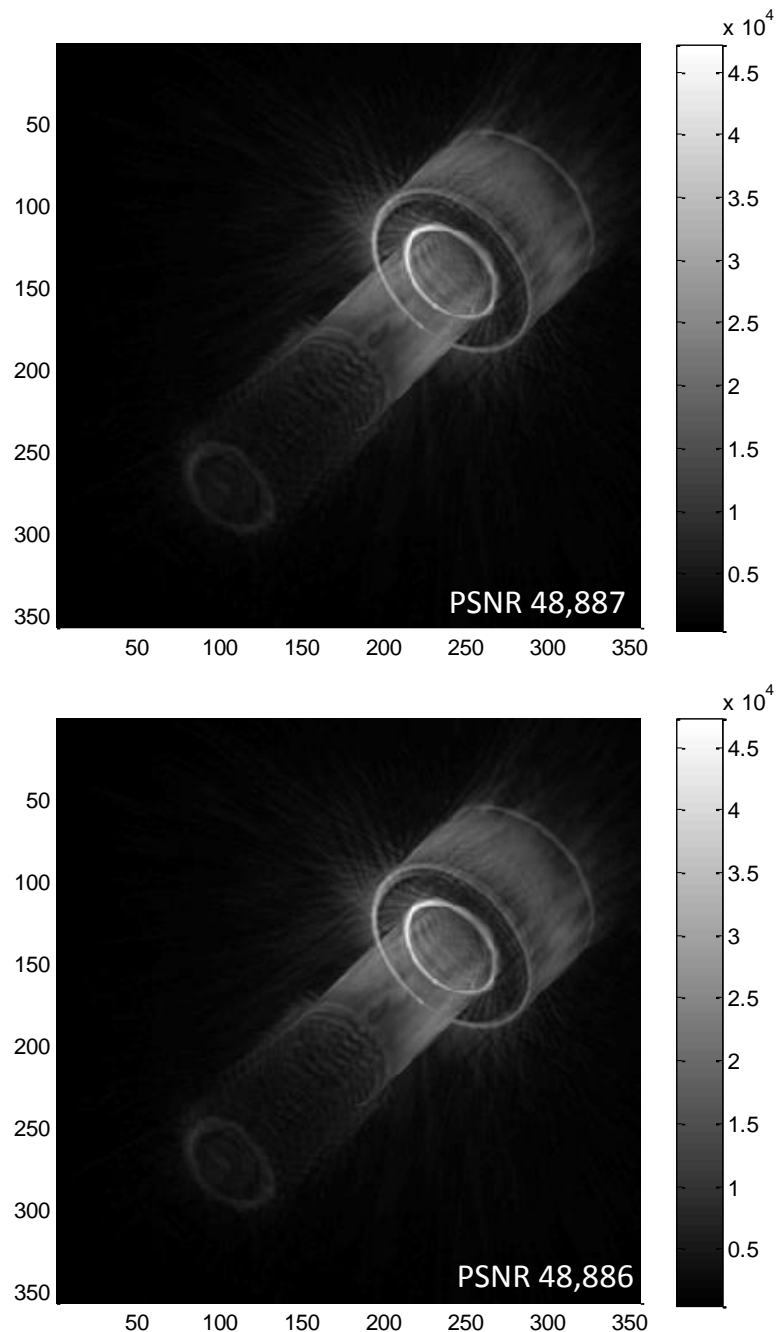


Abbildung 7.15: Die Rekonstruktion der Schraube bei *maximum-intensity* Darstellung mit der Standard-SAFT (oben) und mit dem TOFI-SAFT (unten).

Gelatine-Phantom

Bei der Rekonstruktion mit dem Standard-SAFT ergibt sich die in Abbildung 7.16 oben gezeigte doppelte Gelatineoberfläche und ein unfokussierter Bondingdraht (siehe Kapitel 6.1.2). Mit SOS-Korrektur bildet sich der Draht dagegen scharfkantig und fokussiert ab. Die doppelte Gelatineoberfläche wird als eine Oberfläche abgebildet. Im xz-Schnittbild lassen sich die vier Schnitte durch den Draht eindeutig unterscheiden und lokalisieren. Der Kontrast wird erhöht. Insbesondere im xy-Schnitt steigt der Kontrast um einen Faktor 2,4 an. Wird mit dem beschleunigten TOFI-Algorithmus rekonstruiert, zeigt sich ebenfalls eine gute Fokussierung des Drahts und der Gelatineoberfläche durch die SOS-Korrektur. Ein Unterschied zwischen TOFI und SAFT mit SOS-Korrektur ist mit dem Auge nicht zu erkennen, zeigt sich aber in einem geringen Unterschied des PSNR-Wertes von maximal 0.6.

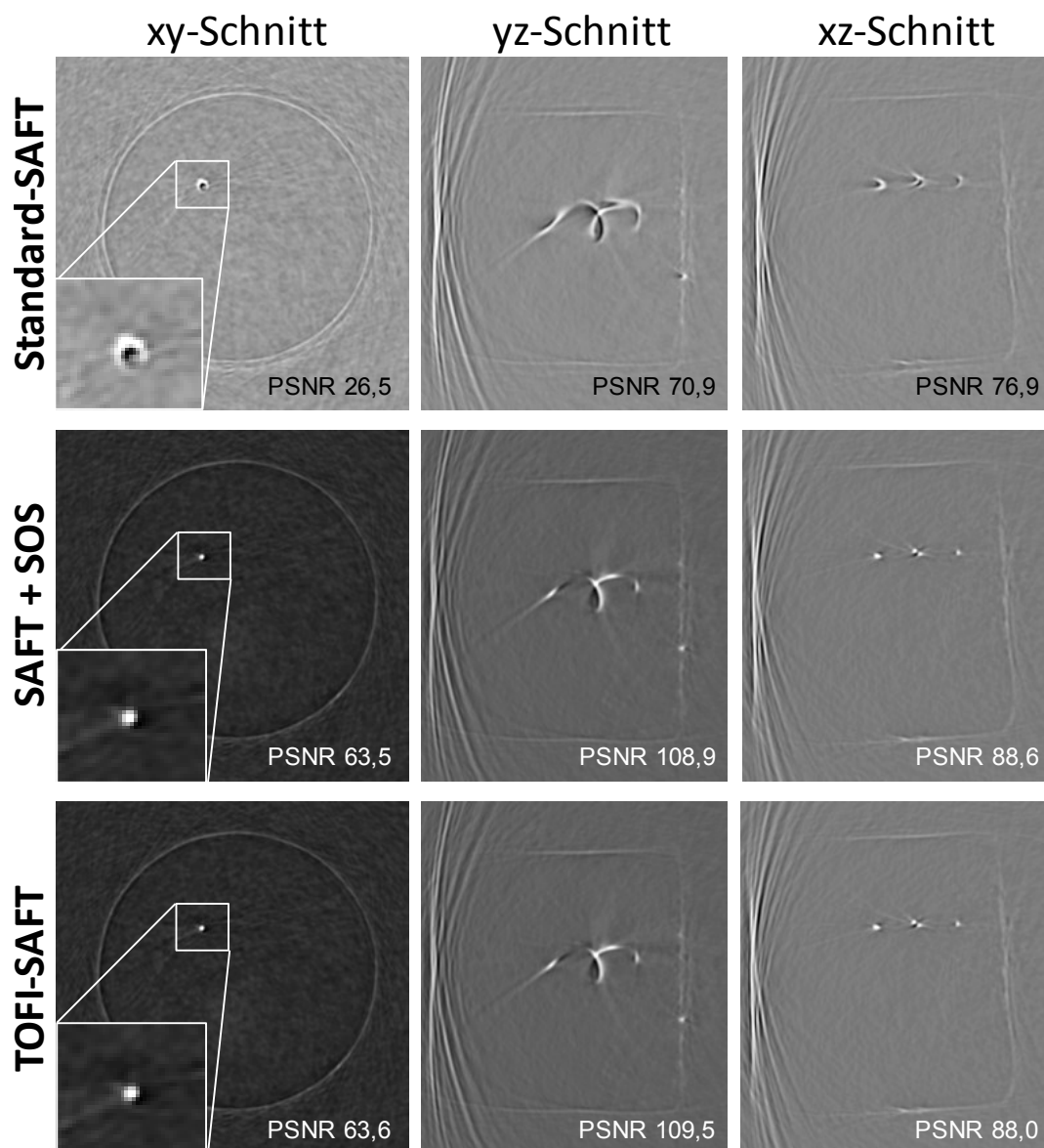


Abbildung 7.16: Vergleich des Gelatine-Phantoms (siehe Abbildung 6.9) bei drei der Rekonstruktionsvarianten für jeweils einen Schnitt in jede Richtung.

Brust-Phantom

Bedingt durch die unterschiedlichen Schallgeschwindigkeiten im Phantom und Wasser kommt es bei der Rekonstruktion mit dem Standard-SAFT zu Artefakten (siehe Abbildung 7.17 oben). Es sind Verschiebungen bzw. Größenänderungen der Einschlüsse sowie eine nicht geschlossene Oberfläche erkennbar (siehe Pfeile). Diese Artefakte können bei der Rekonstruktion mit SOS-Korrektur vermieden werden. Dennoch weist die Standard-SAFT etwas höhere PSNR-Werte auf. Die Ursache liegt dabei in der fehlenden SOS-Korrektur. Die reflektierenden runden Oberflächen der Einschlüsse werden auf Oberflächen mit kleinerem Radius zurückprojiziert. Dies führt bei der Rekonstruktion zu erhöhten Reflektivitäten I , die direkt in die PSNR-Metrik eingehen (siehe Gleichung (5.7)). Der beschleunigte TOFI-Algorithmus erreicht auch hier nahezu dieselben PSNR-Werte wie die SAFT mit SOS-Korrektur.

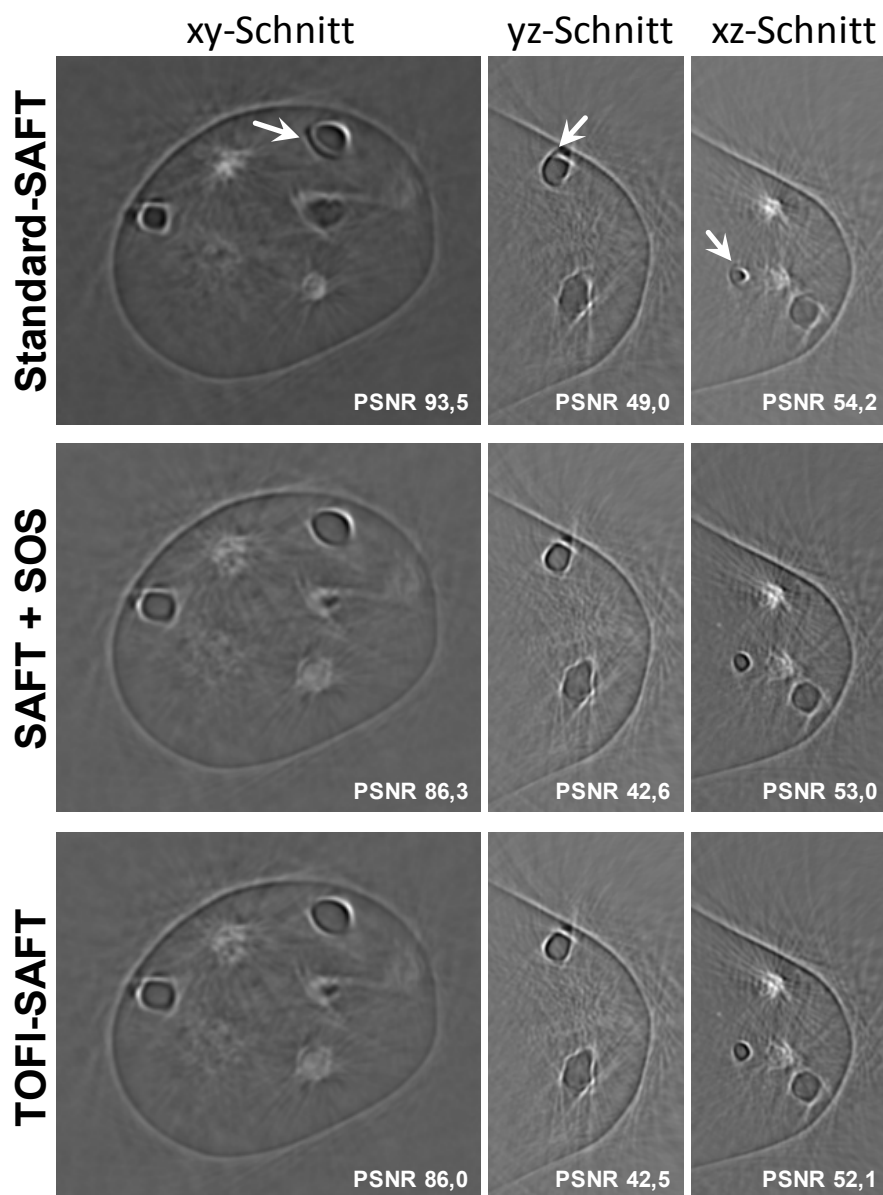


Abbildung 7.17: Vergleich des Brust-Phantoms bei drei der Rekonstruktionsvarianten für jeweils einen Schnitt in jede Richtung.

Reale Brust

Wie in Abbildung 7.18 dargestellt, zeigt sich auch bei dem komplexesten Messobjekt, der realen Brust, dasselbe gute Ergebnis der TOFI-SAFT. Verglichen zur Standard-SAFT beinhaltet die TOFI-SAFT die SOS-Korrektur und führt zu leicht verbesserten PSNR-Werten, um bis zu 2,4. Die PSNR-Werte der SOS-korrigierten SAFT und der TOFI-SAFT sind bis auf den yz-Schnitt, der einen geringen Unterschied von 0,1 aufweist, nicht zu unterscheiden.

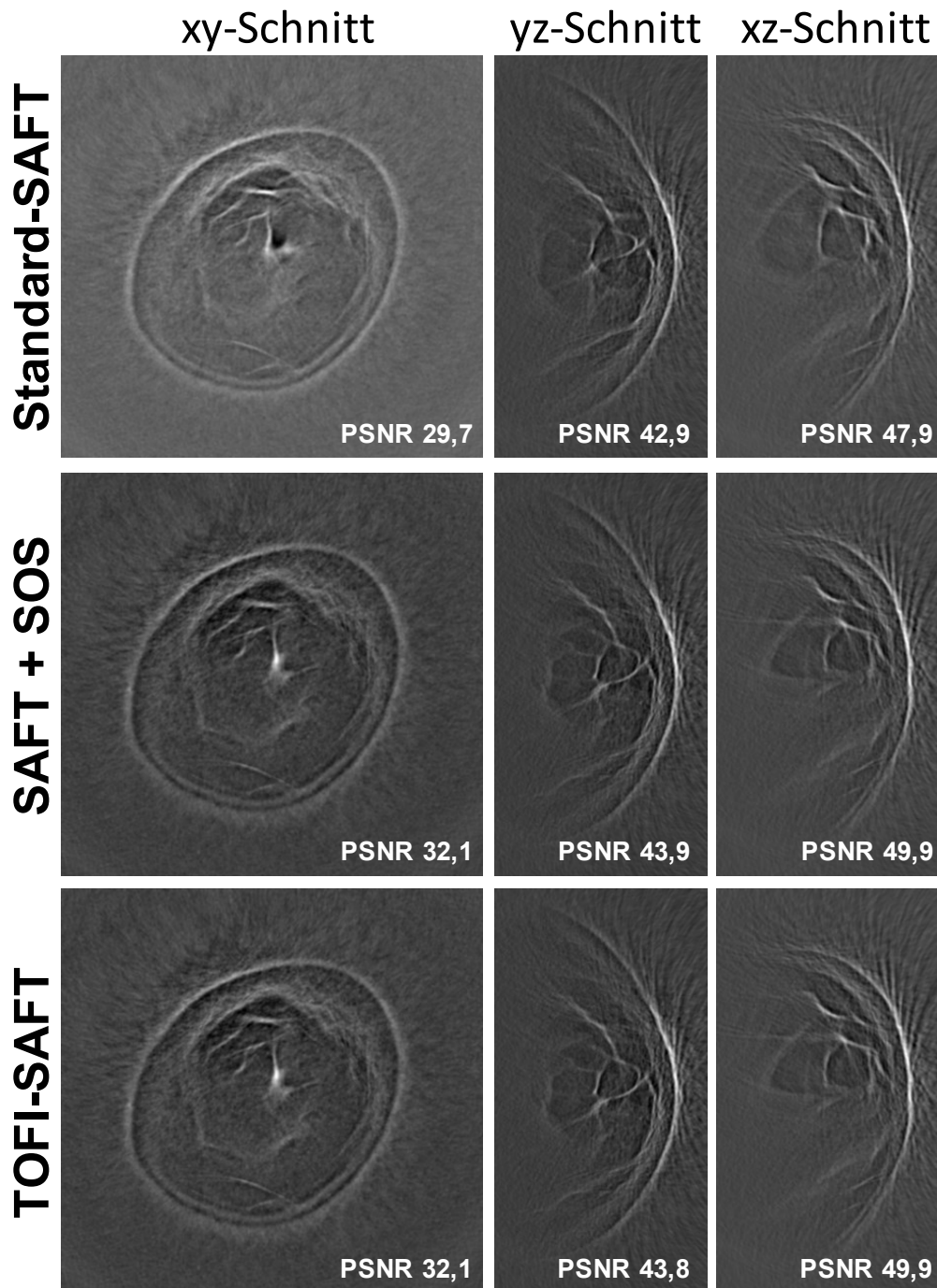


Abbildung 7.18: Vergleich der realen Brust bei drei der Rekonstruktionsvarianten für jeweils einen Schnitt in jede Richtung.

7.2.4 Performanz

Der TOFI-Algorithmus wurde mit dem Ziel entwickelt, den korrigierten SAFT-Algorithmus zu beschleunigen, um eine Rechengeschwindigkeit zu ermöglichen, die für eine klinische Anwendbarkeit nötig ist. Die erreichte Rechengeschwindigkeit ist in Abbildung 7.19 für die verschiedenen Rekonstruktionsvarianten dargestellt. Der TOFI-Algorithmus erreicht auf einer GTX 590 mit 13,0 GVA/s eine Rechengeschwindigkeit, die nahezu der Standard-SAFT ohne Korrekturen entspricht. Auf einer GTX Titan wurde dagegen eine weitere Beschleunigung um den Faktor 2,1 gegenüber der Standard-SAFT erreicht. Die GTX Titan ist insbesondere für den TOFI-Algorithmus geeignet, da sie 3,5 mal mehr Textureinheiten besitzt als die GTX 590. Über diesen Faktor lässt sich auch die weitere Beschleunigung um den Faktor 3,4 erklären. Die Kombination aus TOFI-Algorithmus und GTX Titan GPU stellt somit eine Rechengeschwindigkeit zur Verfügung, die die für eine klinische Anwendbarkeit geforderte Rechenleistung von 32,3 GVA/s pro GPU sogar um den Faktor 1,7 überschreitet.

Zusammenfassend lässt sich der beschleunigte TOFI-Algorithmus qualitativ für heterogene Medien zwischen Bildqualität und Performanz, wie in Abbildung 7.20 folgendermaßen einsortieren: Der Standard-SAFT besitzt zwar eine hohe Rechenleistung, ist jedoch aufgrund seiner Annahmen für heterogene Medien ungeeignet. Eine essentielle Verbesserung bringt die Schallgeschwindigkeitskorrektur, da sie für eine bessere konstruktive Überlagerung der zurückprojizierten Ellipsen sorgt und damit die Basis einer SAFT-Rekonstruktion für heterogene Medien darstellt. Die Dämpfungskorrektur ermöglicht darüber hinaus eine Kontrastanhebung für innere Strukturen und ermöglicht so eine Rekonstruktion mit verbesserter gleichmäßiger Ausleuchtung der Brust. Allerdings besitzt sie dabei die längste Rechendauer. Der TOFI-Algorithmus beinhaltet dagegen beide Korrekturen und liefert zudem eine signifikante Beschleunigung. Durch die genutzte Approximation kommt es zu Fehlern. Anhand unterschiedlich komplexen Messobjekten konnte gezeigt werden, dass sich die Bildqualität dabei nur geringfügig verschlechtert. Mit einer Spline-Interpolation lässt sich die Approximation weiter verbessern, erhöht aber gleichzeitig den Rechenaufwand. Der TOFI-SAFT ist daher als Rekonstruktionsalgorithmus für die 3D-USCT-Rekonstruktion von heterogenen Medien mit GPUs besonders gut geeignet.

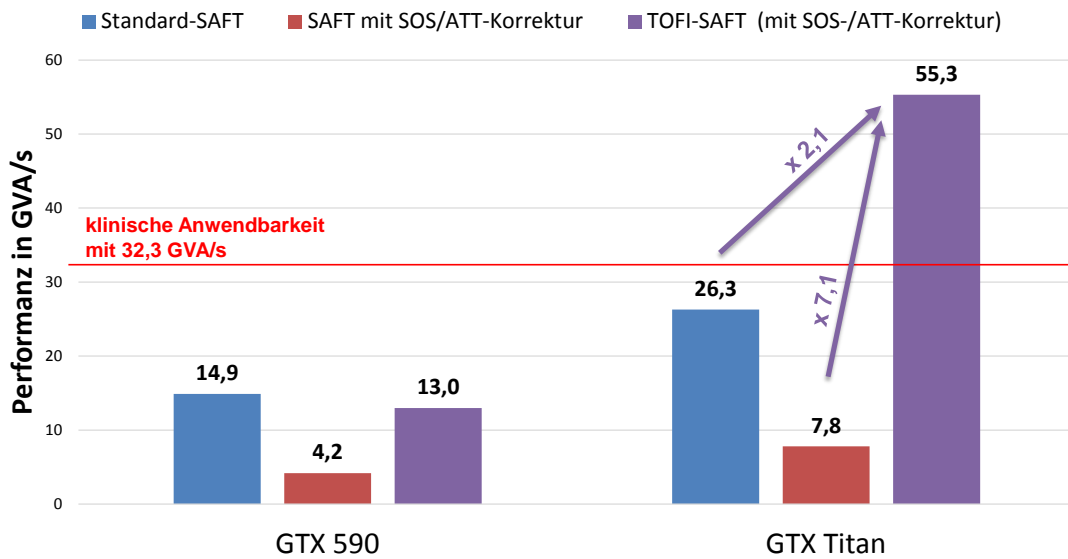


Abbildung 7.19: Gegenüberstellung der erreichten Rechengeschwindigkeit bei den verschiedenen SAFT-Varianten und GPUs.

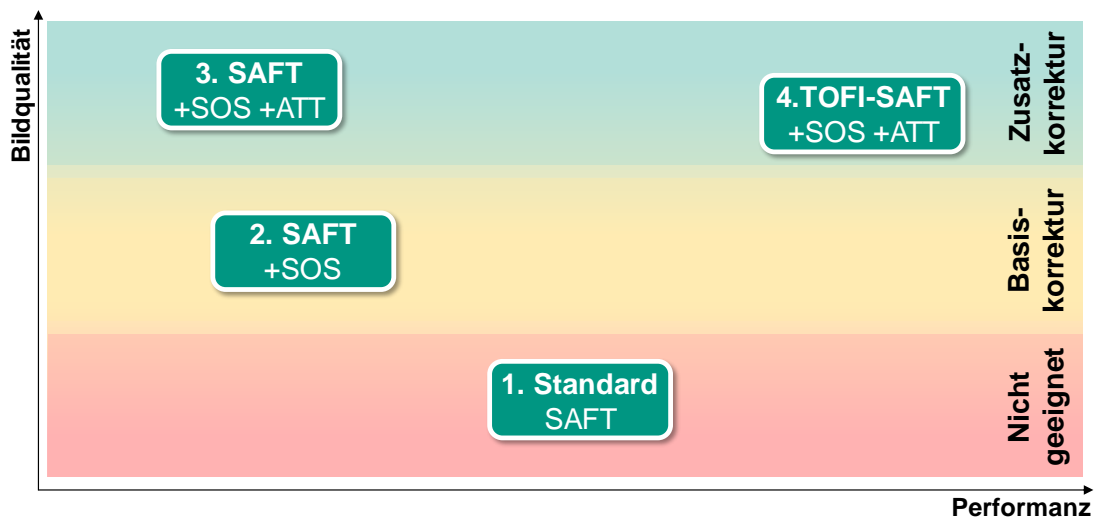


Abbildung 7.20: Überblick der verschiedenen SAFT-Varianten und Einteilung in Bildqualität und Performanz für heterogene Medien.

8. Erweiterung der SAFT zur Reduktion der Gratinglobe-Artefakte

8.1 Methodik und Umsetzung

Bei der Rekonstruktion mit SAFT kommt es, methodisch bedingt durch die Rückprojektionen, zu unerwünschten Rekonstruktionsartefakten. In diesem Kapitel wird eine Methode aufgezeigt und untersucht, mit der die Rekonstruktionsartefakte unterdrückt und die tatsächlichen Streuer hervorgehoben werden können. In Abbildung 8.1 links ist der einfache Fall eines isotropen Punktstreuers mit einer sehr geringen Anzahl von zehn A-Scans dargestellt. Abbildung 8.1 rechts stellt den Verlauf über die Beiträge der SEP dar, die bei der Rekonstruktion für drei beispielhafte Voxel mit der SAFT aufsummiert werden. Die Stärke der Artefakte hängt dabei von der Position der zu messenden Objekte, ihren Reflexionseigenschaften und den Positionen der genutzten Ultraschallwandler ab. Ist das Objekt ein isotroper Punktstreuer wie in Abbildung 8.1 an der Position \vec{x}_1 , der in alle Richtung gleichmäßig abstrahlt, so detektieren alle Ultraschallwandler ein Ultraschallsignal. Bei der Rekonstruktion werden diese Echosignale als Ellipsoide ins Bild zurück projiziert. Jeder zu rekonstruierende Voxel \vec{x} erhält auf diese Weise pro A-Scan einen Beitrag $AScan(\vec{s}, \vec{e}, tof_P)$. Diese Beträge werden mit der SAFT nach Gleichung (3.15) zu einem Wert für die Reflektivität $I(\vec{x})$ aufsummiert. Das Ziel dabei ist es, die Verteilung der Reflektivität bzw. der akustischen Impedanzunterschiede $f(\vec{x})$ bestmöglich zu bestimmen.

Der Wert $AScan(\vec{s}, \vec{e}, tof_P)$ an einer Stelle eines A-Scans setzt sich dabei jedoch nicht nur aus dem Beitrag eines Echosignals von $f(\vec{x})$ zusammen, sondern aus dem

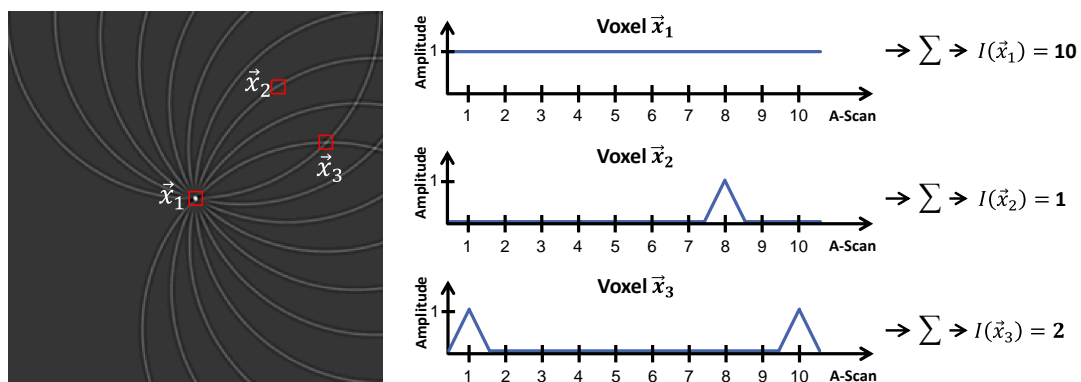


Abbildung 8.1: Bildrekonstruktion eines Punktstreuers mit zehn A-Scans (links). Für drei Positionen \vec{x}_1 , \vec{x}_2 und \vec{x}_3 ist rechts der Verlauf über die A-Scan-Beiträge gegeben, über den mit dem SAFT-Algorithmus eine Summe gebildet wird.

Integral aller Echosignale, welche dieselbe Laufzeit (tof) besitzen. Sie bilden für den dreidimensionalen Fall eine Fläche \mathcal{G} , für die gilt

$$\mathcal{G}_{tof}(\vec{x}) = \left\{ \vec{x} : tof(\vec{s}, \vec{x}, \vec{e}) = \frac{\|\vec{s} - \vec{x}\| + \|\vec{x} - \vec{e}\|}{c_P(\vec{s}, \vec{x}, \vec{e})} \right\}. \quad (8.1)$$

D.h. die Messung zu dem Zeitpunkt tof entspricht dem Integral der Impedanzunterschiede $f(\vec{x})$ über dem Gebiet \mathcal{G}_{tof} [38]

$$AScan(s, e, tof) = \int_{\mathcal{G}_{tof}} f(\vec{x}) d\mathcal{G}_{tof}(\vec{x}). \quad (8.2)$$

Die Gleichung (3.15) für die SAFT-Rekonstruktion lässt sich somit umschreiben zu

$$\begin{aligned} I(\vec{x}) &= \sum_{\forall(s,e)} AScan(s, e, tof) \\ &= \sum_{\forall(s,e)} \int_{\mathcal{G}_{tof}} f(\vec{x}) d\mathcal{G}_{tof}(\vec{x}). \end{aligned} \quad (8.3)$$

Definiert man $\bar{\mathcal{G}}_{tof}$ als das Gebiet \mathcal{G}_{tof} ohne den Punkt \vec{x} , kann nach [99] weiter geschrieben werden

$$\begin{aligned} I(\vec{x}) &= S \cdot E \cdot f(\vec{x}) + \sum_{\forall(s,e)} \int_{\bar{\mathcal{G}}_{s,e,tof}} f(\vec{x}) d\mathcal{G}_{s,e,tof}(\vec{x}) \\ &= S \cdot E \cdot f(\vec{x}) + \mathcal{E}(\vec{x}). \end{aligned} \quad (8.4)$$

Dabei gibt der erste Term die fehlerfreie Rekonstruktion der Impedanzunterschiede $f(\vec{x})$ für den Voxel \vec{x} an. Da eine Anzahl von $S \cdot E$ A-Scans genutzt wird, ist die Reflektivität $I(\vec{x})$ um den Faktor $S \cdot E$ verstärkt. Der zweite Term stellt den Rekonstruktionsfehler $\mathcal{E}(\vec{x})$ dar. Er wird durch die Echosignale anderer Streuer verursacht, die sich auf den Gebieten $\bar{\mathcal{G}}_{s,e,tof}$ befinden. Im Vergleich zu dem Echosignal von $f(\vec{x})$ werden diese Signale nicht verstärkt und daher relativ um den Faktor $1/(S \cdot E)$ unterdrückt. Je mehr A-Scans für die Rekonstruktion genutzt werden, desto höher wird das Signal-zu-Fehler-Verhältnis und der Fehler wird unterdrückt.

Es ist zu beachten, dass der Fehler nicht ausschließlich von der Anzahl der genutzten A-Scans abhängt, sondern auch von der Anzahl, Position und Stärke der streuenden Objekte, die gemessen werden. Ist z.B. das Echosignal eines wahren Streuers bei \vec{x} viel schwächer verglichen zu den Echosignalen der Summe über $\bar{\mathcal{G}}_{s,e,tof}$, kann der Fehlerterm $\mathcal{E}(\vec{x})$ überwiegen und zu einem Kontrastverlust führen.

Eine zusätzliche Reduktion des Fehlers kann zu einer deutlichen Kontrastverbesserung des Bildes führen. Die Beiträge $AScan(s, e, tof_P)$ pro A-Scan können für einen Bildpunkt $I(\vec{x})$, wie in Abbildung 8.1 rechts, hintereinander aufgereiht betrachtet werden. Ist, wie im Voxel a), ein isotroper Punktstreuer enthalten, besitzen die Beiträge über alle A-Scans die gleiche Amplitude. Dagegen befindet sich an der Position b) und c) kein Streuer. Dennoch sind durch den Streuer in Position a) Beiträge vorhanden, die je nach Position zu unterschiedlich starken Artefakten im Bild führen.

Unter der Annahme eines isotropen Punktstreuers müsste der Verlauf der Amplitudenbeiträge wie in a) konstant sein. In der Realität besitzt allerdings jedes Objekt oder jede Oberfläche eine gewisse Reflexionscharakteristik. Zudem kommt es je nach Position des Streuers zu einer unterschiedlichen Dämpfung der Echosignale, so dass eher eine kontinuierliche Änderung der Amplitudenbeiträge beobachtbar sein sollte, wie beispielhaft in Abbildung 8.2 oben dargestellt. Ebenfalls ist die Reihenfolge der SEP-Beiträge mitentscheidend für den kontinuierlichen Verlauf der Amplitudenbeiträge eines Voxels. Wird für einen Streuer nun eine kontinuierliche Änderung der Amplitudenbeiträge angenommen, so können Ausreißer im Verlauf der Amplitudenbeiträge, die durch weitere Streuer bzw. starke Beiträge aus $\overline{\mathcal{G}}_{s,e,tof}$ verursacht werden, als fehlerhafter Beitrag betrachtet und durch einen Medianfilter herausgefiltert werden.

Umsetzung

Für die konkrete Umsetzung der Reduktion der Rekonstruktionsartefakte wurde der vorhandene SAFT-Code um einen 1D-Medianfilter erweitert. Die hier genutzte Implementierung wurde im Wesentlichen in der Studienarbeit von Christian Jülg [100] erstellt. Der GPU-Kernel-Code für den Standard-SAFT-Algorithmus berechnet das Volumen voxel-parallel. Pro *Thread* durchläuft der Algorithmus alle A-Scans und berechnet jeweils einen Amplitudenwert, der zu einer Gesamtsumme addiert wird. Dieses Vorgehen entspricht der Darstellung in Abbildung 8.1 rechts. Dadurch ist es möglich, den jeweiligen Median direkt in der Abarbeitung der A-Scans während der Rekonstruktion zu berechnen. Die dazu nötigen drei Schritte sind in Abbildung 8.2 dargestellt. Der SAFT-Algorithmus arbeitet die A-Scans nacheinander ab. Dadurch ergibt sich ein Zugriff auf die A-Scanmesswerte wie in Abbildung 8.2 oben dargestellt. Durch die Reihenfolge, in der die A-Scans vorliegen, wird pro Sender s_i auf die A-Scans von mehreren Empfängern $e = [0..1412]$ zugegriffen. Für die Berechnung des laufenden Medians wird ein Zwischenspeicher in der Größe der gewählten Fensterbreite W des Median-Filters benötigt (siehe Abbildung 8.2 unten links). Dieser

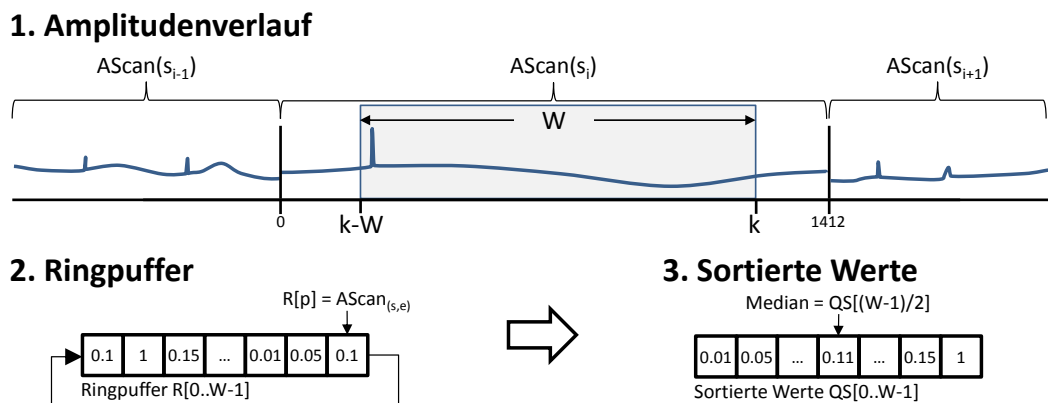


Abbildung 8.2: Ermittlung des laufenden Medians mit der Fensterbreite W aus dem Amplitudenverlauf (1.) der A-Scans. Die Werte im Fenster werden in einem Ringpuffer $R[0..W-1]$ zwischengespeichert (2.) und in $QS[0..W-1]$ sortiert (3.). Der Median kann laufend aus der Speicherstelle $QS[(W-1)/2]$ gelesen werden.

wird als Ringpuffer $R[0..W-1]$ realisiert, um jeweils die aktuellen Amplitudenwerte im Speicher vorzuhalten. Mit jedem Schritt k wird ein neuer Wert $A\text{Scan}(s, e)$ an die aktuelle Stelle $p = k \bmod W$ in den Ringpuffer geschrieben. Ist der Puffer gefüllt, wird erneut an der Stelle $R[0]$ begonnen und somit jeweils der älteste Wert überschrieben. Die Werte des aktuellen Median-Fensters im Ringpuffer werden in einen zweiten Speicher $QS[0..W-1]$ übernommen und dort sortiert (siehe Abbildung 8.2 unten rechts). Der Median kann somit laufend an der Stelle $QS[(W-1)/2]$ gelesen und in $I(\vec{x})$ aufsummiert werden.

Wird bei der fortlaufenden Abarbeitung der A-Scans der Sender s gewechselt, springt der Index des Empfängers e von dem letzten Empfängerindex zurück auf den Index des ersten Empfängers. Werden alle Empfänger für die Rekonstruktion genutzt, geschieht das alle $E=1413$ Schritte. Da hierbei ein Sprung der Amplitudenwerte zu erwarten ist, wird bei jedem Wechsel des Senders der Ringpuffer wieder neu gefüllt.

Die Berechnung des fortlaufenden Medians kann dabei mit verzweigungsfreien Instruktionen durchgeführt werden [101]. Eine verzweigungsfreie Instruktionsabarbeitung ist insbesondere für die Berechnung mit GPUs vorteilhaft, da sie nach dem SIMT-Prinzip (*single instruction multiple threads*) arbeiten. Bei einem verzweigungsfreien Vorgehen werden mögliche Kontrollflussdivergenzen vermieden, die zu einer sequentiellen Ausführung der einzelnen Kontrollflusspfade und Performanzverlust führen würden.

8.2 Evaluation

Die Reduktion der Rekonstruktionsartefakte verspricht einen höheren Kontrast zwischen reflektierenden Strukturen und dem nicht-reflektierenden Hintergrund. Für die Evaluation wird daher die PSNR-Metrik als Maß für den Kontrast genutzt. Sie eignet sich, da sie den maximalen Wert ins Verhältnis zum Hintergrundrauschen stellt. Für Simulationsdaten kann die PSNR-Metrik nach Gleichung (5.5) genutzt werden, da die Grundwahrheit bekannt ist. Für reale Messdaten kommt dagegen die PSNR-Metrik nach Gleichung (5.7) zum Einsatz, bei der für das Rauschen ein Bereich außerhalb des Objekts gewählt wird, der nur Rauschen enthalten kann.

Da die Entstehung der Rekonstruktionsartefakte vom Objekt abhängig ist, wird die Methode an vier Experimenten mit verschiedenartigen Streucharakteristiken evaluiert.

1. Bei den **Simulationsdaten** werden ideale Punktstreuer untersucht, die infinitesimal klein sind und in alle Richtungen gleich streuen.
2. Für Messdaten aus dem Bereich des NDT wird eine **Schraube** untersucht, die sehr starke Oberflächenreflexionen besitzt. Da sie aus Metall gefertigt wurde, stellt sie ein sehr homogenes Medium dar. Es wird von einer sehr guten Fokussierung an der Oberfläche ausgegangen, da Schallgeschwindigkeitsunterschiede im Wasser vernachlässigt werden können. Da die Schraube während der Messung schief innerhalb der 3D-USCT angeordnet wurde, wird hier zur besseren Darstellung eine *maximum intensity*-Projektion genutzt. Dabei wird das gesamte rekonstruierte Volumen auf die xy-Ebene projiziert, indem jeweils nur der maximale Wert in z-Richtung übernommen wird.

$$I(\vec{x}, \vec{y}) = \max_j (I(\vec{x}, \vec{y}, j)) \quad (8.5)$$

3. Das **Brust-Phantom** besitzt ein relativ homogenes Medium, allerdings mit verschiedenen Einschlüssen. Damit ergeben sich klare Oberflächenreflexionen. Die Schallgeschwindigkeitsunterschiede dürfen nicht vernachlässigt werden.
4. Die **reale Brust** ist ein inhomogenes Medium aus Fett, Drüsengewebe und Bindegewebe. Dies ergibt eine Mischung aus vielen verschiedenen schwachen und stärkeren Streuern. Damit stellt die reale Brust ein komplexes Medium für die Rekonstruktion dar.

Ein entscheidender Faktor für die Wirksamkeit der Methode ist die genutzte Fensterbreite und die Reihenfolge der genutzten SEP für den Median-Filter. Neben der Standardreihenfolge der Empfänger (Abbildung 8.3 links) wird mit einer veränderten Reihenfolge untersucht, ob ein verbessertes Filterergebnis erreicht werden kann, indem Empfänger von weiteren TAS-Ringen mit einbezogen werden. Dadurch gehören die SEP-Beiträge über eine größere Filterlänge zu derselben Region, wodurch die Lokalität erhöht wird (Abbildung 8.3 rechts).

Da der Medianfilter alle E Schritte neu gefüllt wird, könnten Fensterbreiten von 1 bis E untersucht werden. Allerdings ist, bis auf einen idealen Punktstreuer, nicht davon auszugehen, dass ein Objekt in alle Richtungen gleich abstrahlt. Neun Empfänger befinden sich in Abstand von 2 mm auf einem TAS. Der durchschnittliche Abstand zwischen zwei benachbarten TAS liegt bei 33,5 cm.

Wird eine Filterbreite kleiner oder gleich neun gewählt, werden hauptsächlich Beiträge herausgefiltert, die eine starke Veränderung innerhalb eines TAS bewirken. Bei einer größeren Filterbreite können die Informationen benachbarter TAS miteinbezogen werden. Das bedeutet jedoch auch, dass Strukturen, die sehr gerichtet reflektieren, als Fehlersignal herausgefiltert werden können.

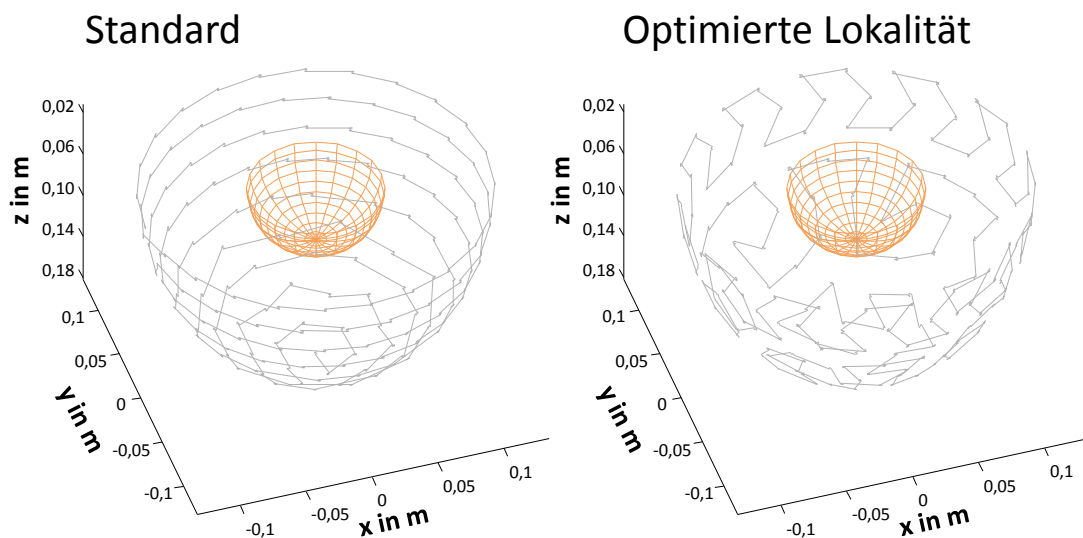


Abbildung 8.3: Standardreihenfolge der Empfänger (links). Die veränderte Reihenfolge, um für den 1D-Median eine verbesserte Lokalität zu erreichen (rechts).

Simulationsdaten

In Abbildung 8.4 sind die Rekonstruktionen mehrerer simulierter Punktstreuer über verschiedene Fensterbreiten dargestellt. Bei Fensterbreiten bis neun ist kaum eine Veränderung beobachtbar, so dass die Artefaktreduktion nur in einem weiten Abstand vom Punktstreuer überhaupt eine Wirkung zeigt. Dieses Verhalten begründet sich darin, dass die ins Bild projizierten Ellipsen sehr eng beieinander liegen und über einen weiten Bereich für dieselben Voxel Beiträge liefern. Dabei befinden sich die entsprechenden SEP alle auf einem TAS. Erst bei einer Fensterbreite größer neun, bei der SEP von mehreren TAS genutzt werden, kommt es zu einer Artefaktreduktion. Dieser Effekt lässt sich nutzen um die Rechengeschwindigkeit zu erhöhen. Statt alle neun Empfänger von einem TAS nacheinander zu nutzen, kann die Reihenfolge im Voraus so abgeändert werden, dass jeweils nur einer der neun Empfänger eines TAS in Reihe genutzt wird. Dadurch lässt sich mit einem Neuntel der Medianfilterlänge nahezu dasselbe Ergebnis erreichen.

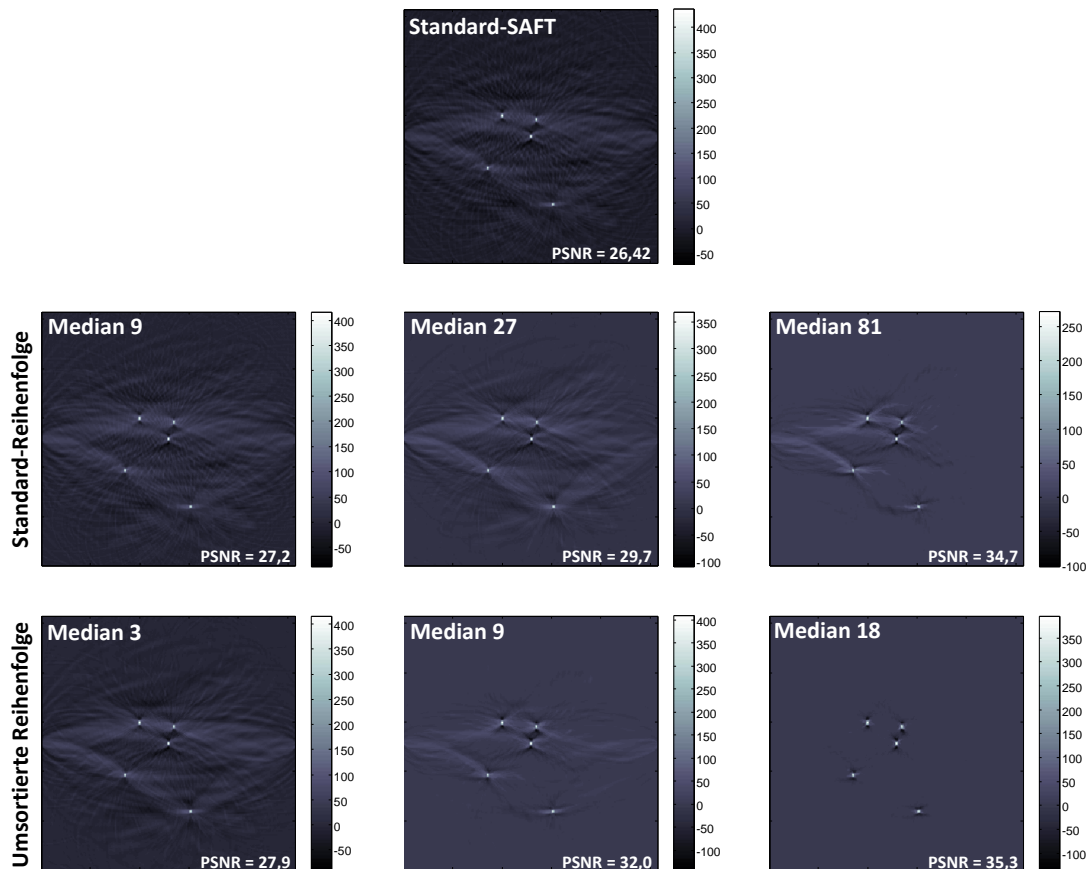


Abbildung 8.4: Rekonstruktionsvergleich von fünf simulierten Punktstreuern mit der Standard-SAFT (oben) und bei zusätzlicher Artefaktereduktion mit unterschiedlichen Filterlängen (mittig und unten). Mittig sind die Rekonstruktionen bei Standardreihenfolge und unten bei umsortierter Reihenfolge dargestellt.

Schraube

Die Rekonstruktion der Schraube ist in Abbildung 8.5 mit der *maximum intensity*-Projektion dargestellt. Ist die Medianfilterlänge sehr gering, kommt es zu Verschmierungen der Rekonstruktionsartefakte und damit zu einem geringeren Kontrast. Der Grund liegt darin, dass die Artefakte als Struktur erkannt werden und die Löcher zwischen ihnen herausgefiltert werden. Mit steigender Filterlänge ab 27 werden die Rekonstruktionsartefakte immer besser unterdrückt. Ab einer Filterlänge von 45 ist kaum noch ein Unterschied auszumachen. Da die Schraube eine klare und stark streuende Oberfläche besitzt, funktioniert die Artefaktreduktion an diesem Beispiel sehr gut.

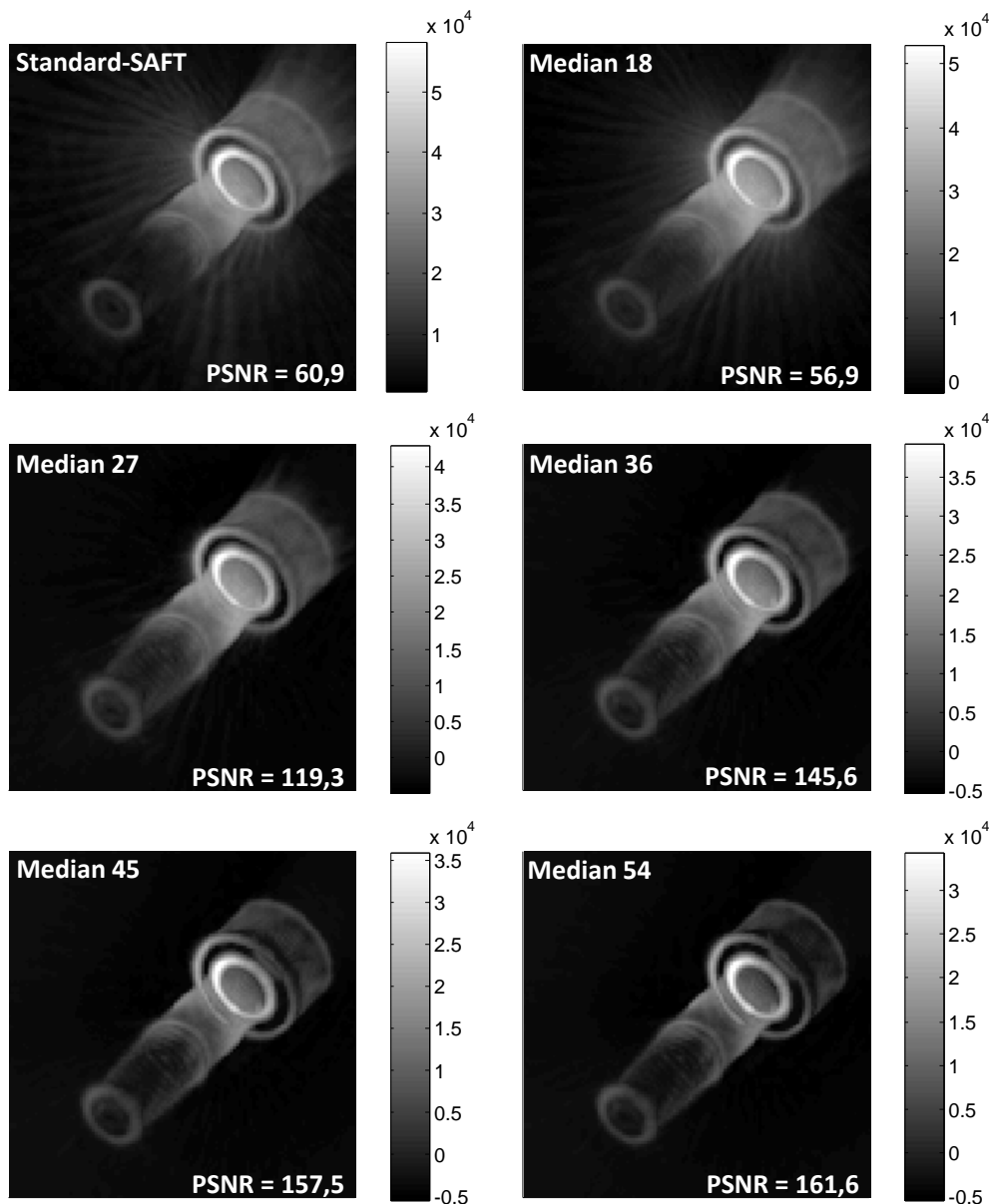


Abbildung 8.5: Rekonstruktion der Schraube über mehrere Medianfilterlängen bei vier AP und einer Auflösung von $(1,8\text{ mm})^3$.

Brustphantom

In Abbildung 8.6 ist das Brustphantom für vier verschiedene Schnittbilder dargestellt. Dabei wird der Einfluss der beiden Reihenfolgen untersucht (siehe Abbildung 8.3). Es wurde pro TAS nur einer der neun Empfänger genutzt und eine feste Filterlänge von 10 gewählt. Ohne Artefaktreduktion sind mehrere deutliche Artefakte erkennbar, die oben mit Pfeilen markiert sind. Diese lassen sich mit dem Einsatz der Artefaktreduktion deutlich verringern (siehe Abbildung 8.3 mittig). Jedoch werden auch einzelne Strukturen unterdrückt. Diese sind mit Pfeilen markiert. Wird die Reihenfolge mit optimierter Lokalität (LO) genutzt, werden diese Strukturen deutlich weniger herausgefiltert (siehe Abbildung 8.3 mittig). Insbesondere die Brustkante wird schärfer abgebildet. Dabei sinkt der gemessene Kontrastwert geringfügig ab.

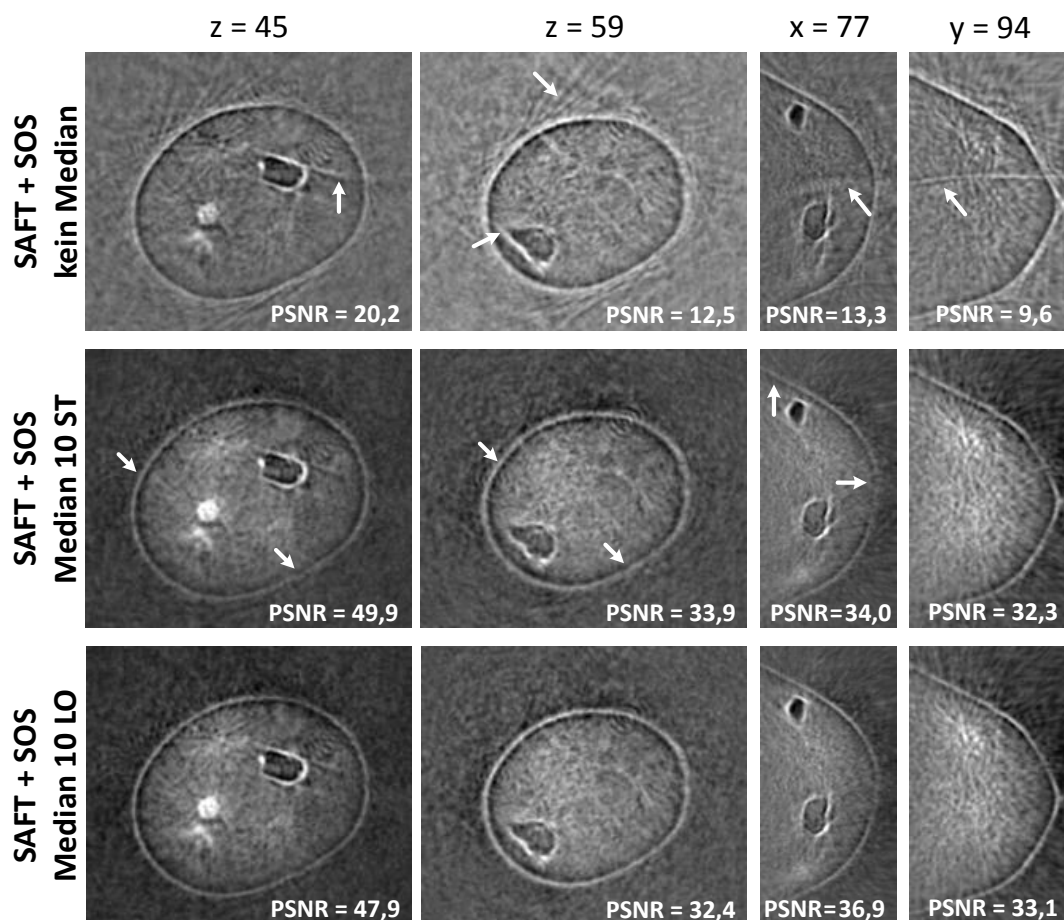


Abbildung 8.6: Rekonstruktion einzelner Schichten des Brustphantoms bei einer Medianfilterlänge von 10. Rekonstruktion mit schallgeschwindigkeitskorrigierter SAFT (oben). Mittig, die Rekonstruktion mit Standardreihenfolge (ST) und unten mit der veränderten Reihenfolge für eine optimierte Lokalität (LO). Es wurden vier AP und eine Auflösung von $(0,9 \text{ mm})^3$ genutzt.

Reale Brust

Die Rekonstruktion der realen Brust mit Schallgeschwindigkeitskorrektur ist in Abbildung 8.7 für sechs Filterlängen dargestellt. Wird eine zu geringe Filterlänge genutzt, kommt es zu einer Aufsummierung und Verstärkung von negativen Werten. Diese lassen sich durch die negativen Werte des Optimalen Pulses erklären, die bei der Signalvorverarbeitung genutzt werden. Da die Brust viele komplexe Streuer enthält, kommt es daher vermehrt zu negativen Signalen, die der Medianfilter somit verstärkt. Die Medianfilterlänge 18 besitzt für diese Brustmessung den höchsten Kontrastwert, der um den Faktor 2 höher liegt als bei der Standard-SAFT. Der Hintergrund wird deutlich reduziert und der Dynamikbereich besitzt ausschließlich positive Werte. Wird die Filterlänge weiter erhöht, wirkt sich dies wiederum negativ auf den Kontrast aus und der Dynamikbereich bekommt negative Werte. Dies bedeutet, dass die positiven Anteile des Optimalen Pulses verstärkt herausgefiltert und die negativen Werte verstärkt aufsummiert werden.

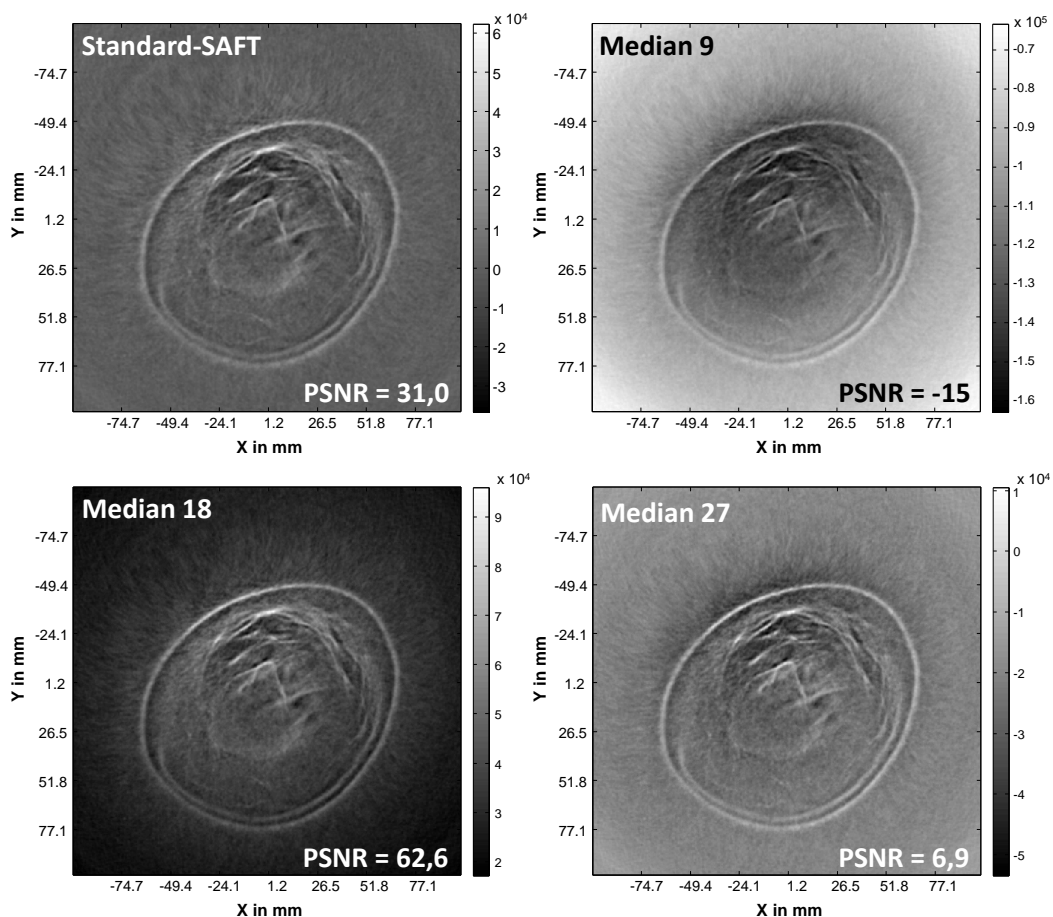


Abbildung 8.7: Rekonstruktion der realen Brust über mehrere Medianfilterlängen bei zehn AP und einer Auflösung von $(0,9 \text{ mm})^3$.

Die SAFT-Rekonstruktion mit Artefaktunterdrückung zeigt bei weniger komplexen Objekten sehr gute Ergebnisse. Insbesondere Rekonstruktionsartefakte, die durch Oberflächen wie bei der Schraube verursacht werden, können gut herausgefiltert werden. Beim Brustphantom konnten deutliche Artefakte reduziert und somit eben-

falls gute Ergebnisse erzielt werden. Für den komplexen Fall der Brust zeigt sich, dass eine geeignete Filterlänge genutzt werden muss, um einen höheren Kontrast zu erreichen. Interessant ist, dass der Dynamikbereich nahezu bei allen Filterlängen gleich bleibt, jedoch bei einer zu geringen oder zu hohen Filterlänge eine Verschiebung in den negativen Bereich erfährt. Diese Abhängigkeit von der Filterlänge lässt sich am geringsten bei der Simulation und der Schraube beobachten. Daher liegt die Vermutung nahe, dass sich auf Grund einer ungenauen Schallgeschwindigkeit die ins Bild projizierten Ellipsen nicht kohärent überlagern. Durch den Optimalen Puls ergeben sich vermehrt negative Werte, die der Medianfilter hervorhebt.

Performanz

Die gemessene Rechengeschwindigkeit ist in Tabelle 8.1 über mehrere Fensterbreiten angegeben. Der Rechenaufwand steigt dabei nahezu linear mit der Filterlänge an. Über eine Umsortierung, wie bei den Simulationsdaten gezeigt, kann auch eine geringere Filterlänge genutzt werden, um nahezu identische Ergebnisse zu erreichen.

Fensterbreite	1	9	18	27	54
GVA/s	10,161	1,438	0,613	0,446	0,232
Verlangsamung	1	7,07	16,58	22,78	43,80

Tabelle 8.1: Übersicht der Rechenleistung bei verschiedenen Medianfilterlängen auf der Nvidia GTX Titan [100].

9. Erweiterung der SAFT zur Gewinnung von Reflexionscharakteristiken

Alle bisher genutzten SAFT-Rekonstruktionsvarianten in der 3D-USCT-Reflexionsbildung berechnen aus den aufgenommenen A-Scans einen Wert, der die Reflektivität an der Position der Voxel möglichst realitätsgetreu wiedergeben soll. Dabei wird jedoch die Richtungsinformation der Reflexionen, die durch den Zusammenhang der drei Positionen von Sender, Empfänger und Voxel aus den Daten gewonnen werden kann, nicht ausgenutzt.

In diesem Kapitel werden Methoden beschrieben, die eine Berechnung von Reflexionsprofilen ermöglichen. Das bisherige SAFT-Rekonstruktionsverfahren kann dazu erweitert werden. Die Reflexionsprofile werden pro Voxel berechnet und stellen somit eine neuartige Bildgebungsmodalität für die 3D-USCT dar. Für die Evaluation der neuen richtungsabhängigen Daten werden verschiedene Methoden untersucht. Zum einen werden Methoden zur Datenrepräsentation analysiert.

Zum einen werden Methoden analysiert, die aus den berechneten Reflexionsprofilen neue Informationen extrahieren können. Dies kann z.B. zur Verbesserung der Bildqualität genutzt werden. Zum anderen wird evaluiert, ob die Daten für eine verbesserte Segmentierung oder Gewebeklassifizierung genutzt werden können. Durch die Berechnung der Reflexionsprofile pro Voxel steigt der Speicherbedarf signifikant an. Daher wird auf die Optimierung der Rechengeschwindigkeit ein wichtiges Augenmerk gelegt. Die in diesem Kapitel vorgestellten Methoden wurden größtenteils im Rahmen der Masterarbeit von Patrick Hucker [102] untersucht. Für diese Arbeit wurde jedoch die genutzte Datenmenge von einer AP auf zehn AP erweitert, um eine genauere Aussage treffen zu können.

9.1 Das Reflexionsmodell nach Phong

In Kapitel 3.2 wurden die verschiedenen Reflexionsarten eingeführt, die in der 3D-USCT auftreten können. In der 3D-USCT werden reflektierende Oberflächenstrukturen abgebildet, an denen es in der Realität meist zu einer Mischform aus einer spiegelnden und einer diffusen Streuung kommt. Für die Simulation und Analyse der Reflexionscharakteristik von Oberflächen wird daher ein geeignetes Reflexionsmodell benötigt. Vollständige und einfach zu berechnende richtungsabhängige Reflexionsmodelle sind zum Zeitpunkt dieser Arbeit nicht bekannt. Daher wird im Folgenden auf einfache und bewährte Modelle der Computergrafik zurückgegriffen und für die 3D-USCT angewendet. Gebhardt gibt in [68] einen Überblick über die verschiedenen Modelle, die in der Computergrafik genutzt werden.

Das Beleuchtungsmodell nach Phong [103] beruht auf empirischen Beobachtungen und ist physikalisch nicht korrekt. Es kann z.B. vorkommen, dass mehr Licht von einem Objekt reflektiert wird, als auf das Objekt eingestrahlt wurde [104]. Aufgrund seiner geringen Rechenintensität hat es sich jedoch in der Computergrafik

etabliert und stellt einen guten Kompromiss zwischen Genauigkeit und Rechengeschwindigkeit dar. Auch neuere und korrektere Modelle basieren weiterhin auf dem Phong-Modell.

Beim Beleuchtungsmodell nach Phong werden Vereinfachungen genutzt, die sich auch für die Anwendung in der 3D-USCT eignen [102]. Es wird angenommen, dass es nur punktförmige Lichtquellen gibt. In der 3D-USCT entspricht dies einem Sender. Er sendet eine Ultraschallwelle aus, die sich kugelförmig ausbreitet. Das Phong-Modell ist ein Beleuchtungsmodell für Oberflächen, wird jedoch lokal für einzelne Voxel angewandt. Eine Oberfläche wird so über den Verlauf mehrerer Voxel repräsentiert. Ebenso werden in der 3D-USCT einzelne Voxel unabhängig voneinander rekonstruiert, können jedoch zusammengenommen eine Oberflächenstruktur darstellen.

Für eine Anwendung des Phong-Beleuchtungsmodells in der 3D-USCT können die benötigten Vektoren, wie in Abbildung 9.1 skizziert, genutzt werden. Dabei wird die Lichtquelle durch den Ultraschallsender ersetzt. Die Funktion des Beobachters wird vom Ultraschallempfänger übernommen. Dieser misst den bei ihm auftreffenden Schalldruck der Echosignale und wandelt sie in ein Spannungssignal der Amplitude A_r um.

Die **Amplitude** A_r einer Reflexion wird nach dem Phong-Modell, wie bereits in Kapitel 3.2 eingeführt, durch die Summe aus den einzelnen Komponenten

$$A_r = A_{diffus} + A_{spiegelnd} \quad (9.1)$$

der diffusen Reflexion A_{diffus} und der spiegelnden Reflexion $A_{spiegelnd}$ beschrieben.

Der **diffuse** Anteil berechnet sich basierend auf dem Lambertischen Kosinusetz (siehe Gleichung 3.9) mittels

$$\begin{aligned} A_{diffus} &= A_e \cdot k_{diffus} \cdot \cos \theta \\ &= A_e \cdot k_{diffus} \cdot (\vec{S} \cdot \vec{N}). \end{aligned} \quad (9.2)$$

Dabei ist A_e die Amplitude des einfallenden Ultraschalls und θ der Einfallswinkel zwischen dem Vektor \vec{S} zum Sender und der Oberflächennormalen \vec{N} . Der materialabhängige Parameter k_{diffus} wird diffuser Reflexionskoeffizient genannt. Er liegt

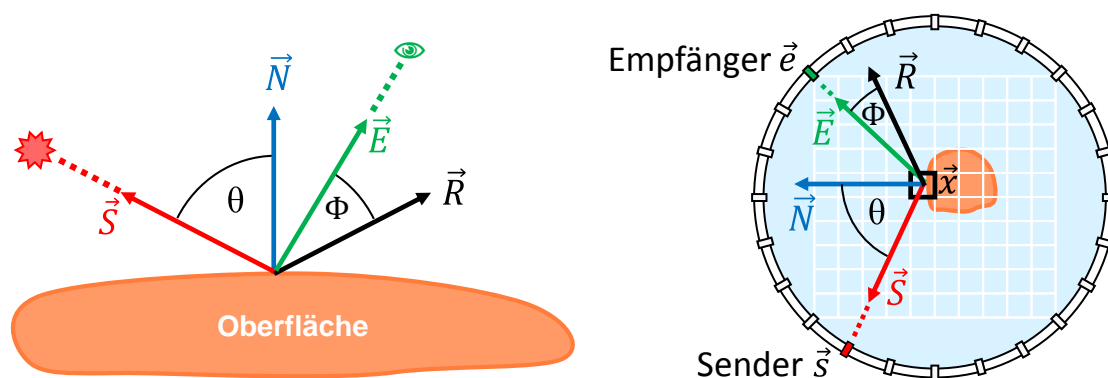


Abbildung 9.1: Schematische Darstellung der relevanten Vektoren beim Phong-Beleuchtungsmodell (links) und in der 3D-USCT (rechts) nach [102].

zwischen 0 und 1 und beschreibt den Anteil, der diffus von der Oberfläche reflektiert wird. Bei der diffusen Reflexion ist die Amplitude abhängig vom Winkel θ . Je kleiner der Winkel zwischen dem Vektor zum Sender \vec{S} und der Oberflächennormalen \vec{N} ist, desto größer ist die reflektierte Amplitude.

Zu einer ideal **spiegelnden** Reflexion kommt es selten in der Realität. Normalerweise weist auch die spiegelnde Reflexion einen gewissen Streuanteil auf, der von der Beschaffenheit der Oberfläche abhängt (siehe Abbildung 9.2). Die Amplitude der gespiegelten Reflexion berechnet sich beim Phong-Modell nach

$$\begin{aligned} A_{\text{spiegelnd}} &= A_e \cdot k_{\text{spiegelnd}} \cdot \cos^n \phi \\ &= A_e \cdot k_{\text{spiegelnd}} \cdot \left(\vec{R} \cdot \vec{E} \right)^n, \end{aligned} \quad (9.3)$$

mit ϕ als Winkel zwischen dem Vektor \vec{E} zum Empfänger und der Hauptreflexionsrichtung \vec{R} , um die herum eine Streuung stattfindet. Der materialabhängige Parameter $k_{\text{spiegelnd}}$ wird Glanzlicht-Reflexionskoeffizient genannt und liegt zwischen 0 und 1. Über den materialabhängigen Parameter n kann der Streuanteil der Reflexion definiert werden. Raue Oberflächen wirken matt und werden mit kleinen Werten von n bis 1 modelliert. Glatte Oberflächen spiegeln dagegen stärker und werden mit einem größeren n von z.B. 100 berechnet. Für eine ideal spiegelnde Reflexion wäre n unendlich. Für die Berechnung von $A_{\text{spiegelnd}}$ wird zusätzlich der Reflexionsvektor \vec{R} benötigt. Dieser kann über das Reflexionsgesetz Einfallswinkel=Ausfallswinkel hergeleitet und über

$$\vec{R} = 2\vec{N}(\vec{S} \cdot \vec{N}) - \vec{S} \quad (9.4)$$

berechnet werden. Dabei ist \vec{S} der Vektor zum Sender und \vec{N} der Normalenvektor.

Für die 3D-USCT ergibt sich somit aus Gleichung (9.1) für die reflektierte Amplitude

$$A_r = A_e \cdot \left(k_{\text{diffus}} \cdot (\vec{N} \cdot \vec{S}) + k_{\text{spiegelnd}} \cdot (\vec{R} \cdot \vec{E})^n \right). \quad (9.5)$$

A_r und A_e werden als bekannt angenommen, da die Amplituden an Empfänger und Sender bekannt sind. Das Verhältnis von A_r zu A_e gibt das Maß der Reflektivität

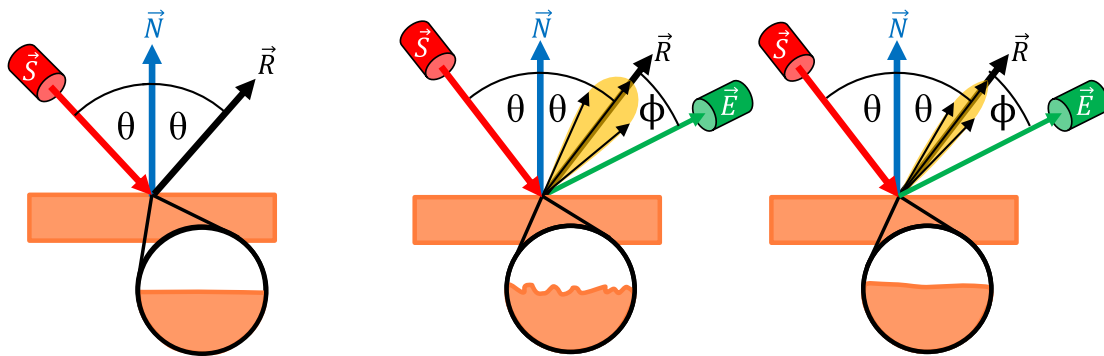


Abbildung 9.2: Schematische Darstellung von spiegelnden Reflexionen. Ideal (links), für einen kleinen Parameter n (mittig) und einem großen Parameter n (rechts). Die Oberflächenbeschaffenheit ist vergrößert dargestellt.

der Struktur innerhalb des Voxels wieder und kann mit Gleichung (9.4) geschrieben werden:

$$\begin{aligned} \frac{A_r}{A_e} &= k_{diffus} \cdot (\vec{N} \cdot \vec{S}) + k_{spiegelnd} \cdot (\vec{R} \cdot \vec{E})^n \\ &= k_{diffus} \cdot (\vec{N} \cdot \vec{S}) + k_{spiegelnd} \cdot \left[(2\vec{N}(\vec{S} \cdot \vec{N}) - \vec{S}) \cdot \vec{E} \right]^n \end{aligned} \quad (9.6)$$

Der Zusammenhang in Gleichung 9.6 kann für ein Voxel in der 3D-USCT folgendermaßen interpretiert werden: Befindet sich innerhalb des Voxels eine stark heterogene bzw. raue Oberflächenstruktur, ist der diffuse Anteil (k_{diffus}) groß. Beinhaltet das Voxel dagegen eine stark homogene bzw. glatte Oberfläche, wie z.B. einen Teil der Brustoberfläche, so ist der Anteil der spiegelnden Reflexion ($k_{spiegelnd}$) größer.

Soll das Phong-Modell genutzt werden, um aus den Messdaten zu ermitteln wie stark eine Struktur diffus oder spiegelnd reflektiert, sind folgende Herausforderungen zu meistern: Da das Phong-Modell die Oberflächennormale benötigt, muss diese bei einer gewollten Parameterabschätzung von k_{diffus} , $k_{spiegelnd}$ und n zuerst geschätzt oder bestimmt werden. In der Computergrafik wird gewöhnlich von undurchlässigen Oberflächen ausgegangen. Objekte wie Glas oder Nebel sind besondere Fälle, die zu einer erhöhten Rechenlast führen. In der 3D-USCT kann dagegen im Normalfall davon ausgegangen werden, dass das untersuchte Medium für den Ultraschall durchlässig ist und den Ultraschall nur in einem gewissen Maße dämpft. Daraus folgt, dass der Ultraschall sowohl an der äußeren Oberfläche mit der Normalen \vec{N} , als auch an der inneren Oberfläche mit der Normalen $-\vec{N}$ reflektiert werden kann. Es muss daher beachtet werden, dass es zu zwei Oberflächennormen kommen kann, die entgegengesetzt sind. Durch die n -te Potenz besitzt die Gleichung (9.6) einen nicht-linearen Term, der eine Parameterschätzung erschwert.

9.2 Berechnung der Reflexionscharakteristik

In diesem Kapitel soll darauf eingegangen werden, wie eine richtungsabhängige Charakteristik von einem Objekt in der 3D-USCT ermittelt werden kann. Dabei wird in dieser Arbeit der Begriff Reflexionscharakteristik für eine richtungsabhängige Verteilung der Reflektivität verwendet, die für ein einzelnes Voxel gilt. Die Reflexionscharakteristik gibt daher Auskunft darüber, wie stark und in welche Richtung eine Struktur reflektiert, die sich innerhalb des Voxels befindet.

Die Idee, um diese Reflexionscharakteristik pro Voxel zu ermitteln, besteht darin, die Beiträge der jeweiligen *Sender-Empfänger-Paare* (SEP) jeweils einer Richtung $R_{\vec{j}}$ zuzuordnen. In Abbildung 9.3 wird der Unterschied im Programmablauf zum Standard-SAFT-Algorithmus vereinfacht dargestellt. Beim Standard-SAFT wird eine Gesamtsumme über alle Beiträge der SEP berechnet und die Richtungsinformation vernachlässigt (Abbildung 9.3 mittig). Im Gegensatz dazu werden für die Berechnung der Reflexionscharakteristik sogenannte Richtungsvektoren \vec{V} berechnet, mithilfe derer die Zuordnung zu einer von J Richtungen durchgeführt werden kann (Abbildung 9.3 rechts). Ein Richtungsvektor $\vec{V}(\vec{s}, \vec{x}, \vec{e})$ wird dabei pro SEP jeweils aus den drei Positionen von Sender \vec{s} , Empfänger \vec{e} und Voxel \vec{x} bestimmt. Für die Bestimmung eines Richtungsvektors aus drei Koordinaten gibt es mehrere Möglichkeiten. Daher werden in dieser Arbeit Annahmen getroffen, auf die im Folgenden näher eingegangen wird.

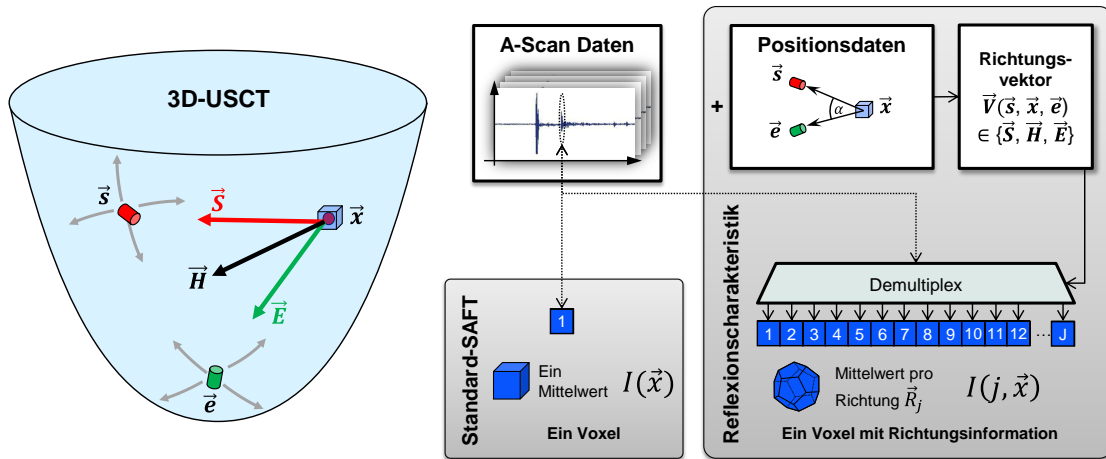


Abbildung 9.3: Gegenüberstellung des Programmablaufs beim Standard-SAFT und beim SAFT-Algorithmus mit Berechnung der Reflexionscharakteristik.

Richtungsvektoren

In dieser Arbeit werden für die Zuordnung folgende drei Richtungsvektoren genutzt (siehe Abbildung 9.3 links):

1. $\vec{V}(\vec{s}, \vec{x}, \vec{e}) = \vec{S}(\vec{s}, \vec{x}, \vec{e})$ als Vektor vom Voxel \vec{x} zum Sender \vec{s} :
 Durch diese Zuordnung wird ermittelt, wie hoch die noch messbare Amplitude über alle Empfänger ist, wenn das Ultraschallsignal aus der Richtung des Senders \vec{s} gesendet und innerhalb des Voxels \vec{x} reflektiert wurde. Oder einfacher ausgedrückt: Wie stark reflektiert das Medium innerhalb des Voxels den Ultraschall, der aus der Richtung des Senders \vec{s} kommt. Der resultierende Reflektivitätswert unter dieser Annahme wird maximal, wenn es zu einer diffusen Reflexion kommt. Daher wird diese Annahme auch als *Annahme einer diffusen Reflexion* bezeichnet.
2. $\vec{V}(\vec{s}, \vec{x}, \vec{e}) = \vec{E}(\vec{s}, \vec{x}, \vec{e})$ als Vektor vom Voxel \vec{x} zum Empfänger \vec{e} :
 Es wird ermittelt, wie hoch die gemessene Amplitude am Empfänger \vec{e} ist, wenn alle Sender \vec{s} ein Ultraschallsignal ausgesendet haben, welches jeweils innerhalb des Voxels \vec{x} reflektiert wurde. Oder einfacher ausgedrückt: Wie stark reflektiert das Medium innerhalb des Voxels den Ultraschall, unabhängig von der Einfallrichtung, in die Richtung des Empfängers \vec{e} . Diese Annahme hat keine direkte Beziehung zu einer Reflexionsart und wird daher nach der Richtung des Vektors *Voxel-Empfänger* bezeichnet.
3. $\vec{V}(\vec{s}, \vec{x}, \vec{e}) = \vec{H}(\vec{s}, \vec{x}, \vec{e})$ als Winkelhalbierende zwischen Sendervektor \vec{S} und Empfängervektor \vec{E} :
 Es wird ermittelt, wie groß die gemessene Amplitude am Empfänger \vec{e} ist, wenn der Sender \vec{s} das Ultraschallsignal ausgesendet hat und wenn sich innerhalb des Voxels \vec{x} eine Oberfläche mit der Normalen $\vec{N}(\vec{x}) = \vec{H}$ befindet. Oder einfacher ausgedrückt: Wie stark kommt es zu einer spiegelartigen Reflexion unter der Annahme einer Oberflächennormalen. Der resultierende Reflektivitätswert unter dieser Annahme wird maximal, wenn es zu einer spiegelnden Reflexion kommt. Daher wird diese Annahme auch als *Annahme einer spiegelnden Reflexion* bezeichnet.

Geometrien

Neben der Bestimmung eines Richtungsvektors spielt die Einteilung der möglichen Richtungen eine große Rolle. Das Ziel der Reflexionscharakteristik ist es, eine richtungsabhängige Verteilung der Reflektivitätswerte für jedes einzelne Voxel zu ermitteln. Dazu wird ein Richtungsvektor \vec{V} genutzt, der vom Voxelmittelpunkt \vec{x} aus definiert ist und in kartesischen Koordinaten über die drei Vektoren in x-, y- und z-Richtung beschrieben werden kann (siehe Abbildung 9.4). Für den Richtungsvektor ist ausschließlich die Richtung und nicht die Länge wichtig. Geht man von einer Einheitskugel mit Radius $r = 1$ aus, kann der Richtungsvektor mit zwei Winkeln in Kugelkoordinaten angegeben werden. Dabei stellt θ den Polar- oder Elevationswinkel und ϕ den Azimutwinkel dar (siehe Abbildung 9.4).

Die richtungsabhängige Verteilung der Reflektivitätswerte kann in unterschiedlicher Auflösung für eine größere oder geringere Anzahl an Richtungen durchgeführt werden. In Abbildung 9.3 werden die Beiträge beispielhaft in zwölf verschiedene Richtungen eingeteilt. Wollte man für jede mögliche Winkelkombination die Reflektivität speichern, so müsste für jedes Voxel und jede SEP der Elevationswinkel, Azimutwinkel und die Amplitude gespeichert werden. Dies würde bei dem Volumen von $1667^2 \times 1000$ Voxel für die ROI, 628 Sendern und 1413 Empfängern und 10 AP bereits zu einer Datenmenge von über 296 PB im *float*-Format führen. Aus diesem Grund wird eine Methode benötigt, um die Anzahl der gespeicherten Reflexionsdaten zu reduzieren. Dabei wird die Oberfläche einer fiktiven Einheitskugel um den Voxelmittelpunkt \vec{x} herum in Flächenabschnitte eingeteilt. Diese Einteilung ist beispielhaft für den 2D-Fall in Abbildung 9.5 dargestellt. Der Kreis wird dabei in $J=12$ Kreisabschnitte eingeteilt und kann durch ein Zwölfeck angenähert werden. Die Kreisabschnittsvektoren \vec{R}_j gehen jeweils durch die Mitte der Kreisabschnitte und geben die Hauptrichtung der Kreisabschnitte an. Der für ein SEP beispielhaft eingezeichnete Richtungsvektor $\vec{V}(\vec{s}, \vec{x}, \vec{e})$ kann nun mit den Kreisabschnittsvektoren \vec{R}_j verglichen werden und wird dem Kreisabschnitt 10 zugeordnet, dessen Kreisabschnittsvektor die kleinste Winkeldifferenz zum Richtungsvektor \vec{V} besitzt. Für den 3D-Fall werden dreidimensionale Körper genutzt, die eine Kugel approximieren. Vorteilhaft sind

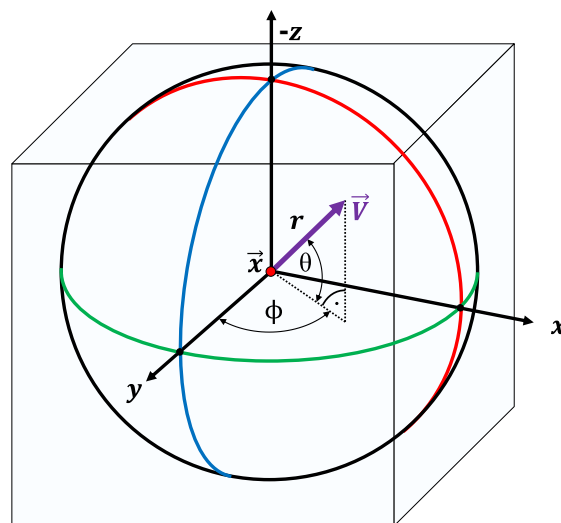


Abbildung 9.4: Ein Richtungsvektor \vec{V} mit Ursprung im Voxelmittelpunkt \vec{x} im kartesischen Koordinatensystem und im Kugelkoordinatensystem.

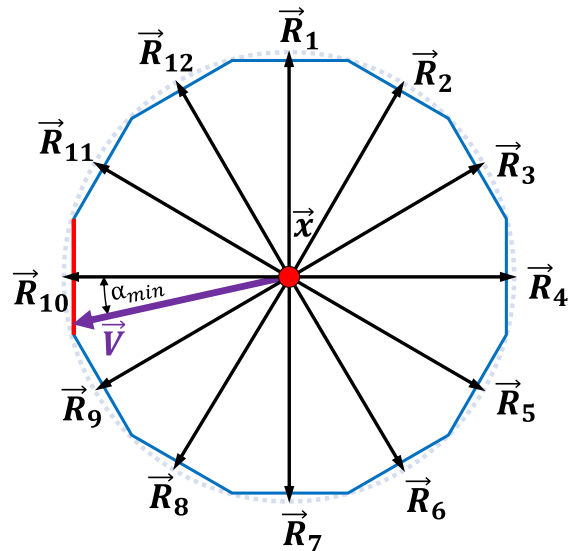


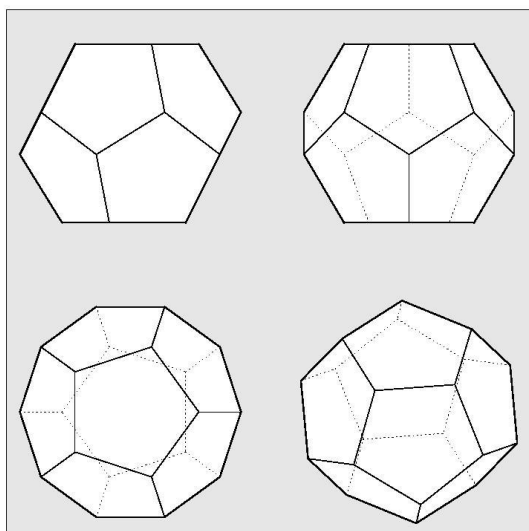
Abbildung 9.5: Beispiel für eine Einteilung der Vektoren im 2D-Fall. Zuordnung des Richtungsvektors \vec{V} zu Kreisabschnitt 10 (rot).

dabei Körper, die eine gleichmäßige Einteilung der Kugeloberfläche in formgleiche Teilabschnitte erlauben. Diese Kugeloberflächenteilabschnitte der genutzten Geometrien werden im Folgenden auch vereinfacht Flächen genannt. In dieser Arbeit werden für die Berechnung der Reflexionscharakteristiken drei Körper genutzt [102]:

Der **Dodekaeder** (griech. Zwölfblächner) gehört zu der Klasse der platonischen Körper (vollkommen regelmäßige konvexe Polyeder) [105]. Er besteht aus zwölf gleichmäßigen Fünfecken, besitzt 20 Ecken und 30 gleich lange Kanten. Die Normalen der zwölf Oberflächen verteilen sich gleichmäßig im Raum. Abbildung 9.6 links stellt den Dodekaeder aus mehreren Blickwinkeln dar.

Der **Ikosaeder** (griech. Zwanzigblächner) ist wie der Dodekaeder ein platonischer Körper. Er besteht aus 20 gleichmäßigen Dreiecken, besitzt 12 Ecken und 30 gleich

Dodekaeder



Ikosaeder

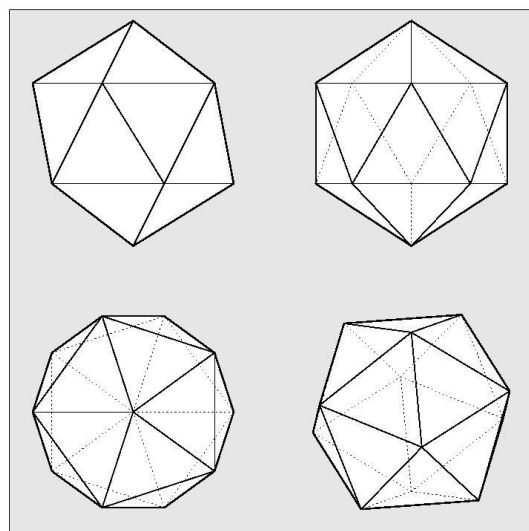


Abbildung 9.6: Geometrie des Dodekaeders (links) und Ikosaeders (rechts) [106].

lange Kanten. Die Normalen der 20 Oberflächen verteilen sich ebenfalls gleichmäßig im Raum. Abbildung 9.6 rechts zeigt den Ikosaeder aus mehreren Blickwinkeln.

Für eine höhere Abtastung der Kugeloberfläche in beliebiger Größe wurde zudem die Geometrie des **Gradnetz** untersucht. Das Gradnetz bietet im Vergleich zu den Platonischen Körpern eine regelmäßige Einteilung des Elevationswinkels θ und des Azimutwinkels ϕ . Der Nachteil hierbei ist allerdings, dass die gleichmäßige Winkelabdeckung zu ungleichmäßigen und unterschiedlich großen Teilabschnitten der Einheitskugel-Oberfläche führt (siehe Abbildung 9.7 links). Dennoch bietet das Gradnetz eine für den Menschen leicht zu interpretierende Darstellung der Reflexionscharakteristik, da sie auf eine ebene Fläche projiziert werden kann, bei der der Verlauf der Reflexionscharakteristik ersichtlich ist. Für die Evaluation wurden zwei Gradnetze untersucht. Für ein *niedriger aufgelöstes Gradnetz* wurde eine Diskretisierung des Azimutwinkels ϕ in 30° -Schritten gewählt. Das ergibt zwölf Werte, da der Azimutwinkel von -180° dem Azimutwinkel $+180^\circ$ entspricht. Für den Elevationswinkels θ wurde eine Diskretisierung in 18° -Schritten gewählt mit insgesamt 9 Werten. Die Pole ($\theta = \pm 90^\circ$) werden als rotationssymmetrisch angenommen. Damit ergeben sich mit dieser Diskretisierung insgesamt $12 \cdot 9 + 2 = 110$ Winkelkombinationen. Dies lässt sich mit einer zweidimensionalen Matrix darstellen (siehe Abbildung 9.7 rechts). Die Umrechnung der Richtungsvektoren und Oberflächennormalen vom kartesischen Koordinatensystem in das Kugelkoordinatensystem erfolgt mit

$$\text{Azimutwinkel } \phi = \arctan2(y, x) \quad (9.7)$$

$$\text{Elevationswinkel } \theta = \arctan2(z, \sqrt{x^2 + y^2}) . \quad (9.8)$$

Gradnetz

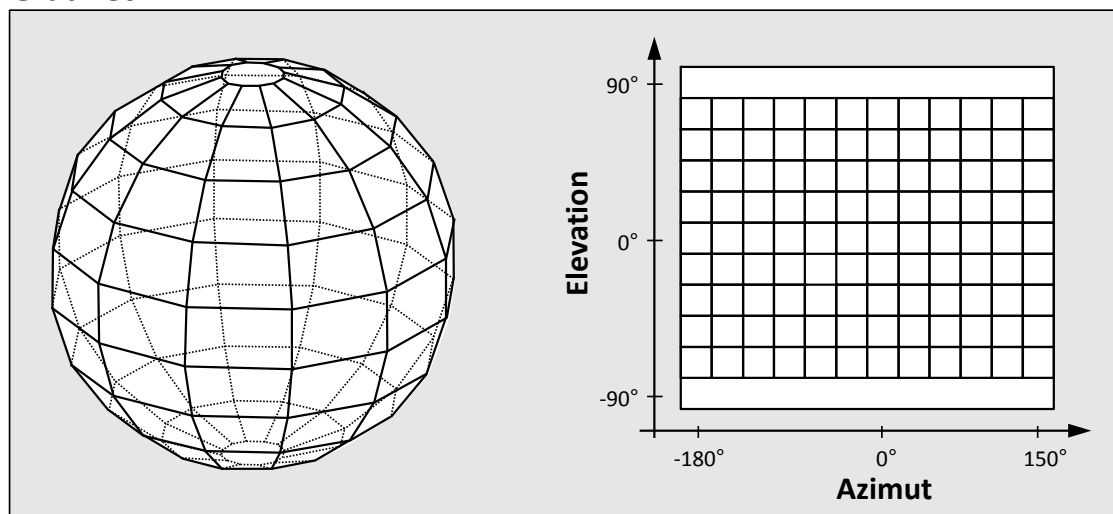


Abbildung 9.7: Geometrie des Gradnetzes (30° , 18°) (links) und seine 2D-Matrixdarstellung (rechts).

Wobei $\arctan2(a,b)$ wie folgt definiert ist

$$\arctan2(a,b) = \begin{cases} \arctan\frac{a}{b}, & b > 0 \\ \arctan\frac{a}{b} + \pi, & a \geq 0, b < 0 \\ \arctan\frac{a}{b} - \pi, & a < 0, b < 0 \\ +\pi/2, & a > 0, b = 0 \\ -\pi/2, & a < 0, b = 0 \\ \text{undefiniert}, & a = 0, b = 0 \end{cases} \quad (9.9)$$

Zudem wurde ein *hochaufgelöstes Gradnetz* implementiert. Die Diskretisierung liegt dabei für die Azimut- und Elevationswinkel jeweils bei 5° woraus sich 2522 Winkelkombinationen ergeben.

Werden diese Geometrien für die Berechnung der Reflexionscharakteristik genutzt, werden die Beiträge der einzelnen SEP jeweils einer Fläche zugeordnet. Für den Vergleich des Richtungsvektors \vec{V} stehen die Oberflächennormalen der Flächen \vec{R}_j zur Verfügung, die durch die genutzten Geometrien bekannt sind. Damit können die Reflexionsdaten z.B. mit dem Dodekaeder auf zwölf Flächen pro Voxel reduziert werden. Zum Speichern der Reflexionscharakteristik benötigt man den Index der Fläche und die jeweilige Amplitude. Da die Flächen beim Gradnetz nicht gleich groß sind und es zu unterschiedlich vielen Beiträgen der Flächen kommen kann, wird zudem die Anzahl der genutzten SEP-Beträge pro Fläche gespeichert. Diese Anzahl steht für eine anschließende Normierung zur Verfügung. Tabelle 9.1 fasst die genutzten Geometrien und ihren Speicherbedarf zusammen. Bei einem vollständigen Volumen in höchster Auflösung von $1667^2 \times 1000$ Voxel für die ROI steigt der Speicherbedarf im Vergleich zum Standard-SAFT mit 11 GB je nach Geometrie auf 267 GB bis 56 TB. Daher werden für diese Arbeit entweder nur einzelne Schichten rekonstruiert oder die Auflösung des Volumens reduziert. In den ersten Experimenten hat sich für reale Daten eine Auflösung von $(0,9 \text{ mm})^3$ bewährt, die einer Breite des Optimalen Pulses von 48 Sampeln entspricht. Diese Auflösung wird auch in diesem Kapitel verwendet.

Geometrie	Flächen	Speicherbedarf	
		pro Voxel	Volumen
Standard-SAFT	1	$1 \cdot 4 \text{ Byte}$	11 GB
Dodekaeder	12	$12 \cdot (2 \text{ Werte} \cdot 4 \text{ Byte})$ $= 96 \text{ Byte}$	267 GB
Ikosaeder	20	$20 \cdot (2 \text{ Werte} \cdot 4 \text{ Byte})$ $= 160 \text{ Byte}$	445 GB
Gradnetz ($30^\circ, 18^\circ$)	110	$(12 \cdot 9 + 2) \cdot (2 \text{ Werte} \cdot 4 \text{ Byte})$ $= 880 \text{ Byte}$	2,4 TB
Gradnetz ($5^\circ, 5^\circ$)	2522	$(72 \cdot 35 + 2) \cdot (2 \text{ Werte} \cdot 4 \text{ Byte})$ $= 20,2 \text{ kByte}$	56 TB

Tabelle 9.1: Überblick der Geometrien mit jeweiligem Speicherbedarf für ein Voxel und für ein Volumen von $1667^2 \times 1000$ Voxel.

Implementierung

Für die Implementierung der Berechnung der Reflexionscharakteristik wird nach dem Programmablauf in Abbildung 9.3 vorgegangen. Die Berechnung der Reflexionscharakteristik erweitert dabei den bereits in Kapitel 4.2.3 beschriebenen Ablauf des SAFT-Algorithmus. Die einzelnen Schritte des Programmablaufs sind in Abbildung 9.8 als Flussdiagramm dargestellt und werden im Folgenden beschrieben.

1. Im ersten Schritt wird abhängig von der zugewiesenen *ThreadId*(Tid) bestimmt, welcher Voxel \vec{x} gerade berechnet wird. Sind die Koordinaten bekannt, wird eine Schleife über alle N genutzten A-Scans gestartet.
2. Im zweiten Schritt werden, abhängig von der genutzten SAFT-Implementierung, alle nötigen Schritte durchgeführt um den *tof*-Index zu berechnen.
3. Ist der *tof*-Index bekannt, kann die Amplitude A interpoliert geladen werden.
4. Dieser Schritt ist der Kern der Berechnung der Reflexionscharakteristik. Aus den Positionsdaten von Sender, Voxel und Empfänger wird der Richtungswinkel \vec{V} unter der gewählten Annahme (\vec{S} , \vec{H} oder \vec{E}) berechnet. Abhängig von der genutzten Geometrie wird anschließend die Fläche j bestimmt, zu der der Beitrag des aktuellen SEP zugeordnet wird. Für den Dodekaeder und den Ikosaeder wird hier der gewählte Richtungsvektor \vec{V} mit den Oberflächennormalen \vec{R}_j der genutzten Geometrien verglichen. Für die Zuordnung wird die Fläche j mit dem geringsten Winkel zwischen Oberflächennormalen und Richtungswinkel \vec{V} ausgewählt. Für das Gradnetz müssen keine Vergleiche mit den Oberflächennormalen \vec{R}_j durchgeführt werden. Der Winkel lässt sich in Kugelkoordinaten umrechnen und kann entsprechend der genutzten Winkeldiskretisierung direkt der Fläche j zugeordnet werden (Nearest-Neighbor-Interpolation).
5. Die Amplitude A , als Beitrag des aktuellen SEP, wird für die Fläche j in *VoxelSum*[j , 1] aufaddiert. In *VoxelSum*[j , 2] wird die Anzahl der genutzten Beiträge für die Fläche j mitgezählt, um eine Normierung durchführen zu können. Sind noch nicht alle A-Scans abgearbeitet, wird die Schleife mit dem nächsten A-Scan fortgeführt.
6. Nachdem alle A-Scans in der Schleife abgearbeitet wurden, stehen die Summen aus allen A-Scan-Beiträgen für alle J Flächen in *VoxelSum* und werden in das Ausgabe-Volumen $I[\vec{x}, 1..J]$ geschrieben.

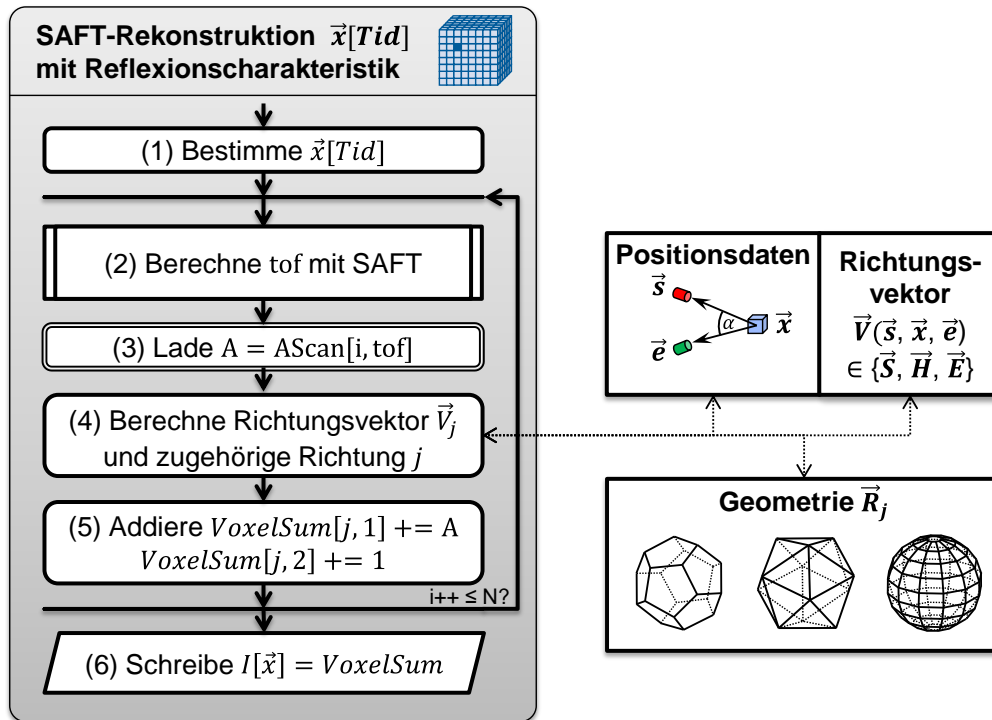


Abbildung 9.8: Erweiterung des Programmablaufs der Standard-SAFT für die Berechnung der Reflexionscharakteristik.

9.3 Evaluation der Reflexionscharakteristik

Nachdem die Reflexionscharakteristiken für die einzelnen Voxel berechnet wurden, liegen die Daten in einer fünfdimensionalen Matrix C der Größe $2 \times J \times X \times Y \times Z$ vor:

$$C[2, j, \vec{x}] = C[2, j, x, y, z] \quad (9.10)$$

$$A[j, \vec{x}] = C[1, j, \vec{x}] \quad (9.11)$$

$$N[j, \vec{x}] = C[2, j, \vec{x}] \quad (9.12)$$

Die erste Dimension enthält zwei Werte: Die aufsummierten Amplituden A und die Anzahl der aufsummierten Amplituden N . Die zweite Dimension hat J Elemente, entsprechend der genutzten Geometrie mit einer Anzahl von J Flächen bzw. Oberflächennormalen \vec{R}_j . Die letzten drei Dimensionen ergeben sich durch die x-, y-, und z-Koordinaten des Voxels im Volumen. Diese Daten stehen nun zu einer weiteren Verarbeitung einer Vielzahl an Möglichkeiten zur Verfügung. Um das Potential der Reflexionscharakteristik zu bewerten, wird im Folgenden eine Evaluation anhand zweier Fragestellungen durchgeführt:

1. Extraktion neuer Informationen

Können im Vergleich zur Standard-SAFT neue Informationen über das Medium bzw. Gewebe gewonnen werden?

2. Klassifikation

Kann anhand der pro Voxel berechneten Reflexionscharakteristik eine Klassifikation der reflektierenden Gewebestrukturen durchgeführt werden?

9.3.1 Extraktion neuer Informationen

Gewichtung der rekonstruierten Volumen

Ein Vorteil ist, dass für eine Rekonstruktion mit der Standard-SAFT kein neuer Durchlauf benötigt wird. Da die Summen für mehrere Richtungen einzeln aufsummiert wurden, kann mit

$$I_{St}(\vec{x}) = \sum_{j=1}^J A(j, \vec{x}) \quad (9.13)$$

durch eine einfache Addition über alle Richtungen das Volumen der Standard-SAFT erzeugt werden.

Werden die Amplituden A pro Richtung jeweils mit der Anzahl der aufsummierten Amplituden N gewichtet, kann ein gewichtetes Bild I_G berechnet werden.

$$I_G(\vec{x}) = \sum_{j=1}^J \frac{A(j, \vec{x})}{N(j, \vec{x})}. \quad (9.14)$$

Abbildung 9.9 vergleicht ein mit der Standard-SAFT rekonstruiertes Bild (links) mit dem gewichteten SAFT (rechts). Wird die Gewichtung durchgeführt, kann der Kontrast von Strukturen, die eine größere Distanz zum Mittelpunkt der Apertur besitzen, angehoben und verbessert werden. So wird insbesondere die Brustkante und die reflektierenden Strukturen im linken unteren Bereich von Abbildung 9.9 rechts deutlich angehoben (siehe Pfeil). Da dabei allerdings auch das Rauschlevel außerhalb der Brust mit angehoben wird, ist der gemessene PSNR-Wert im Vergleich zum Bild der nicht gewichteten Standard-SAFT um das 2,4-fache verringert.

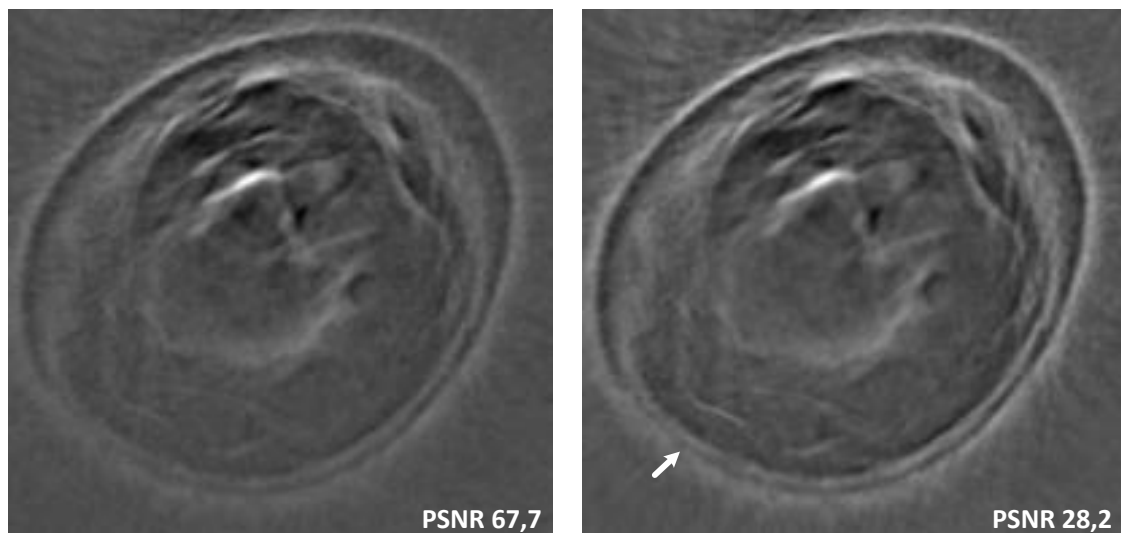


Abbildung 9.9: xy-Schnittbild einer realen Brust bei einer Rekonstruktion mit dem Standard-SAFT (links) und mit dem gewichteten SAFT (rechts).

Bestimmung der Oberflächennormalen

Da die Reflexionscharakteristik sowohl bei einem diffusen als auch bei einem spiegelnden Streuer von der Oberflächennormalen abhängt, wird untersucht, ob die Oberflächennormale $\vec{N}(\vec{x})$ geschätzt werden kann. Ausgehend von dem Reflexionsmodell nach Phong kann über die Richtungsvektoren \vec{S} und \vec{H} die Oberflächennormale $\vec{N}(\vec{x})$ ermittelt werden. Beinhaltet ein Voxel eine diffus reflektierende Oberfläche, wird die Amplitude einer diffusen Reflexion maximal, wenn der Winkel zwischen Sendervektor \vec{S} und Oberflächennormale $\vec{N}(\vec{x})$ gegen 0 geht. Für den diffusen Fall muss daher der Kugelabschnittsvektor (\vec{R}_j) in der Reflexionscharakteristik ermittelt werden, der den maximalen Reflexionswert besitzt. Beinhaltet ein Voxel dagegen eine eher spiegelartig reflektierende Oberfläche kann diese über die Zuordnung mit der Winkelhalbierenden \vec{H} als Richtungsvektor ermittelt werden. Da die Amplitude einer spiegelnden Reflexion maximal wird, wenn der Winkel zwischen der Winkelhalbierenden \vec{H} und Oberflächennormalen $\vec{N}(\vec{x})$ gegen 0 geht, wird auch hier der maximale Reflexionswert der Reflexionscharakteristik gesucht. Ein weiterer Ansatz in [102] geht von der Annahme aus, dass die Oberflächennormalen $\vec{N}(\vec{x})$ auch insgesamt die primäre Reflexionsrichtung darstellen. Als Maß für eine Gesamtreflexionsrichtung werden die Kugelabschnittsvektoren (\vec{R}_j) mit der aufsummierten Amplitude A multipliziert und mit der Anzahl der aufsummierten Amplituden N gewichtet, bevor sie vektoriell aufsummiert werden.

$$\vec{N}_{ges}(\vec{x}) = \sum_{j=1}^J \frac{\vec{R}_j \cdot A(j, \vec{x})}{N(j, \vec{x})} \quad (9.15)$$

Für jedes Voxel im Volumen kann so ein Vektor berechnet werden, der die Hauptreflexionsrichtung wiedergibt und die gesuchte Oberflächennormalen $\vec{N}(\vec{x})$ approximiert. Es muss jedoch beachtet werden, dass eine reflektierende Oberfläche zwei sich gegensätzliche Normalen $\pm N$ besitzt. Wenn sich eine Oberflächenstruktur innerhalb des Voxels befindet, finden starke Reflexionen in beide Normalenrichtungen $\pm N$ statt. Werden die Vektoren \vec{R}_j vektoriell miteinander addiert, ergibt sich demnach nur die Differenz der Reflektivitäten an den beiden sich gegenüberliegenden Flächen. Jedoch kann aufgrund der Dämpfung normalerweise davon ausgegangen werden, dass die Reflexionen an der Seite, die zur Brustoberfläche hin gerichtet ist immer stärker ausfallen als die Reflexionen ins Körperinnere. Abbildung 9.10 zeigt die nach Gleichung (9.15) ermittelten Vektorfelder für einen Ausschnitt an einer Brustkante unter der Zuordnung für die drei Richtungsvektoren \vec{S} , \vec{E} und \vec{H} . Dabei werden zwei Informationen der einzelnen Vektoren dargestellt. Die Richtung wird über die Richtung der Einheitsvektoren und die Stärke der Reflektivität farbcodiert pro Voxel angegeben. Für alle drei Annahmen für die genutzten Richtungsvektoren zeigt sich, dass die Oberflächennormale der Brustkante gut ermittelt werden kann. An der Innenseite der Brustoberfläche kommt es wie erwartet zu einer Umkehrung der Normalen, da eine Oberfläche immer zwei Normalen $\pm \vec{N}$ besitzt. Im Vergleich zwischen den genutzten Annahmen für die Richtungsvektoren zeigt sich teilweise eine andere detektierte Richtung der Oberflächennormalen und eine niedrigere maximale Reflektivität für die Annahme einer spiegelnden Reflexion.

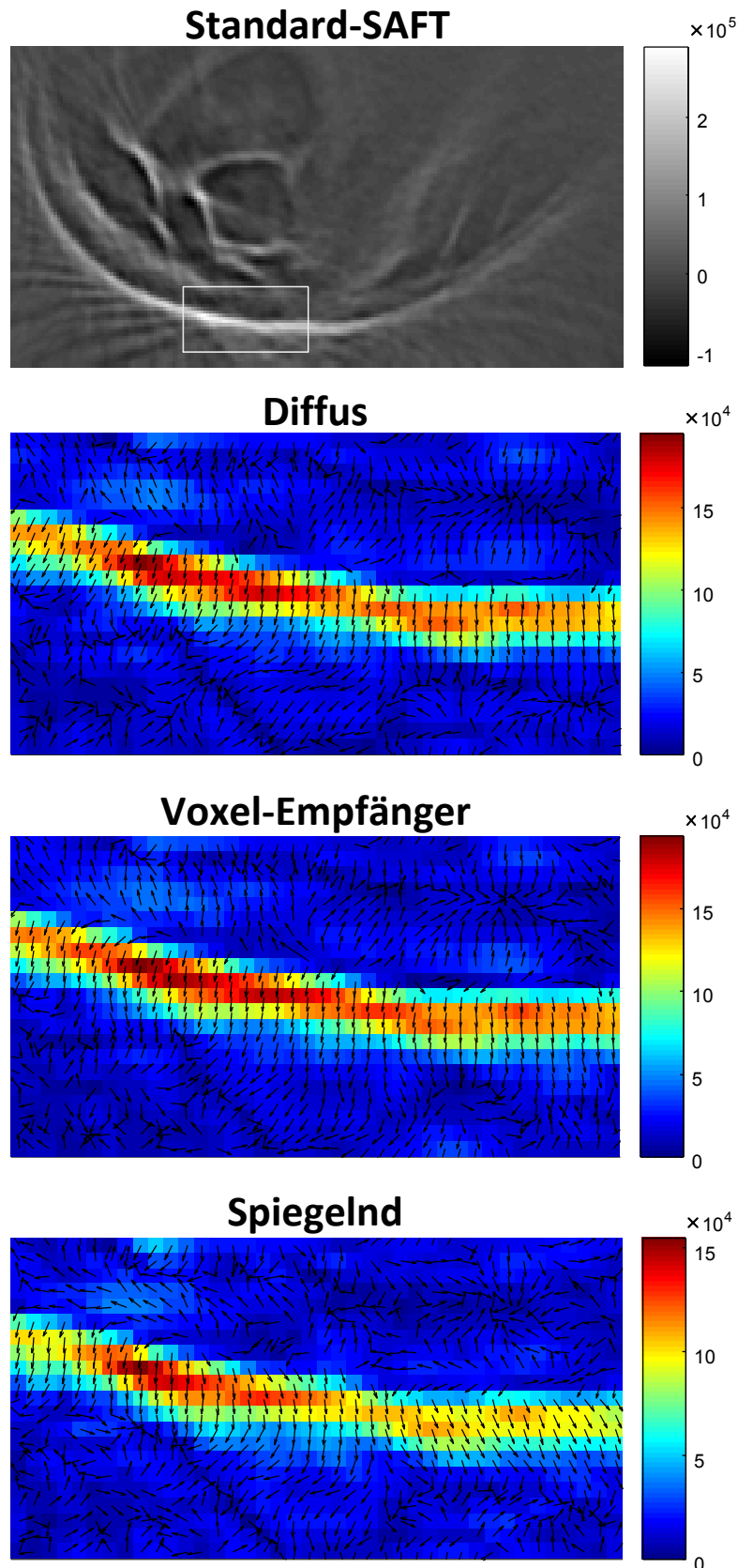


Abbildung 9.10: Ermittelte Vektorfelder für einen Ausschnitt an der Brustkante, der im Bild der Standard-SAFT weiß markiert ist. Bei Annahme einer diffusen sowie spiegelnden Reflexion und mit dem Richtungsvektor \vec{E} . Dargestellt ist das xz-Schnittbild einer realen Brust (siehe Kapitel 8.2).

Richtungsabhängige Betrachtungen

Es können Volumen für die einzelnen Richtungen berechnet werden, um zu sehen welche Beiträge aus den einzelnen Richtungen kommen. Die fünf-dimensionalen Daten C werden dabei auf eine Fläche bzw. Kugelabschnittsvektor \vec{R}_j zu einem dreidimensionalen Volumen reduziert. Für die j -te Fläche ergibt sich das richtungsabhängige 3D-Volumen mit

$$I_j(\vec{x}) = \frac{A(j, \vec{x})}{N(j, \vec{x})}. \quad (9.16)$$

Die Gewichtung der Reflektivitätswerte ist wichtig, da für die einzelnen Richtungen und Voxel, bedingt durch die Apertur, eine unterschiedliche Anzahl an SEP-Beiträgen genutzt wurden.

In Abbildung 9.11 sind die XY-Schnittbilder der einzelnen zwölf Flächen dargestellt, die mit der Geometrie des Dodekaeders mit der Annahme $\vec{V} = \vec{S}$ berechnet wurden. Zudem ist das Bild der Standard-SAFT dargestellt, das der Summe aus allen Richtungen nach Gleichung (9.13) entspricht. Im Vergleich der Bilder der einzelnen Richtungen bzw. Flächen zeigt sich, dass die Strukturen im Bild eine sehr starke richtungsabhängige Reflektivität besitzen. Den Flächen 3 und 6 werden am wenigsten Beiträge zugewiesen. In Abbildung 9.12 werden YZ-Schnittbilder der Flächen 1 und 5 für die drei Reflexionsannahmen gegenübergestellt. Auch für diese Schnittbilder zeigen die Strukturen eine richtungsabhängige Reflektivität. Die Fläche 1 enthält andere Informationen als Fläche 5. Im Vergleich der unterschiedlichen Annahmen für die Richtungsvektoren zeigt sich eine unterschiedliche Hervorhebung von Strukturen, die für die jeweilige Richtung bevorzugt diffus oder spiegelnd reflektieren (siehe Pfeile in Abbildung 9.12). Die Bilder für die diffuse Annahme und der Voxel-Empfänger Annahme weisen insgesamt eine starke Ähnlichkeit auf. Dies könnte darauf hinweisen, dass es für die Reflexion an diesen Gewebegrenzen egal ist von welcher Richtung der Ultraschall ausgesendet wurde. Sender und Empfänger könnten in ihrer Funktion ausgetauscht werden, woraus sich auch für die Voxel-Empfänger Annahme ein diffuses Reflexionsmodell ergeben würde. Ein Unterschied ist für Fläche 1 jedoch an der Brustkante beobachtbar, die in einem weiteren Bereich abgebildet ist.

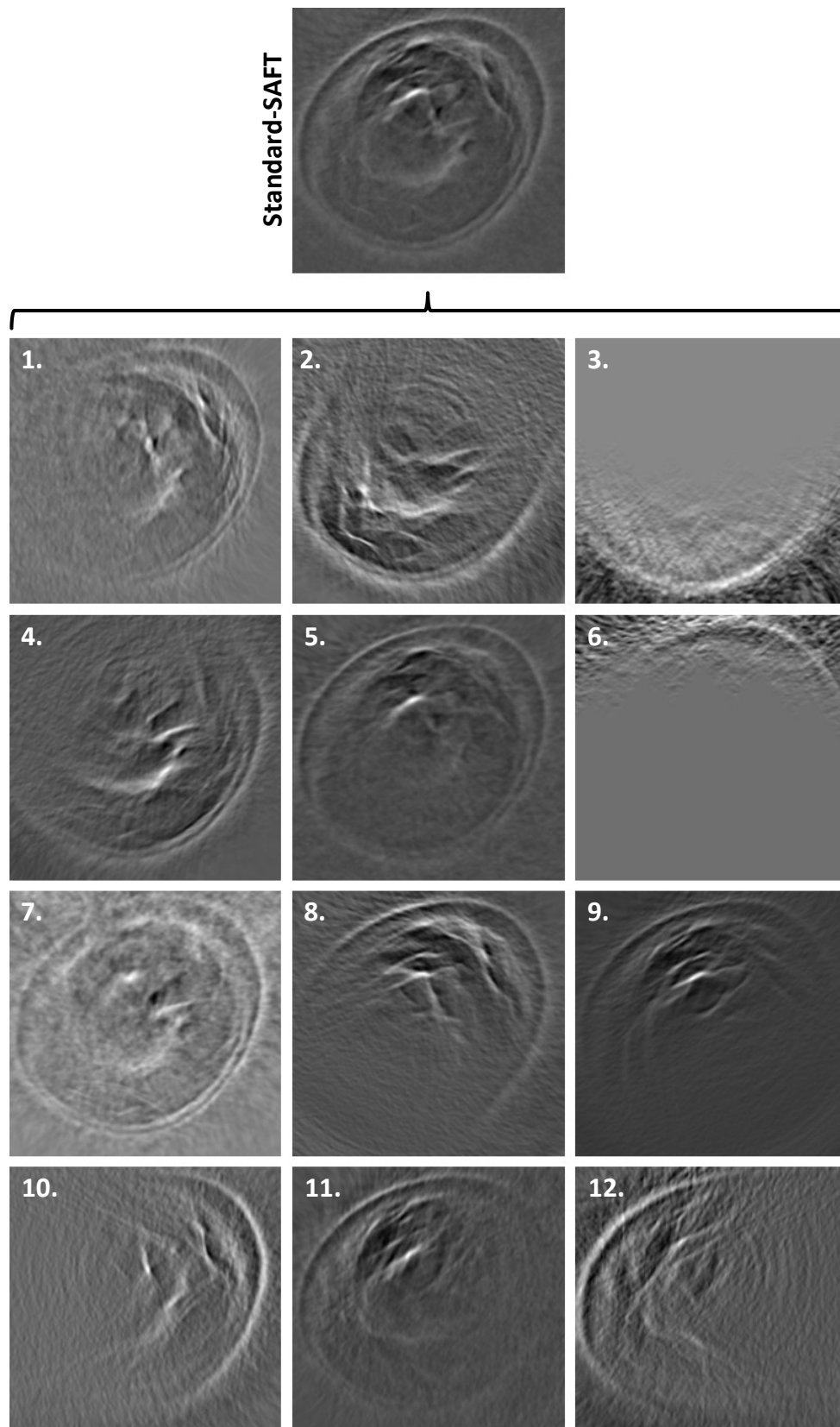


Abbildung 9.11: xy-Schnittbilder einer realen Brust von verschiedenen Kugelabschnittsvektoren $\vec{R}_{j=1..12}$ eines Dodekaeders, berechnet mit der Annahme $\vec{V} = \vec{S}$. Der Wertebereich liegt für alle zwölf Bilder zwischen 0 und 1.

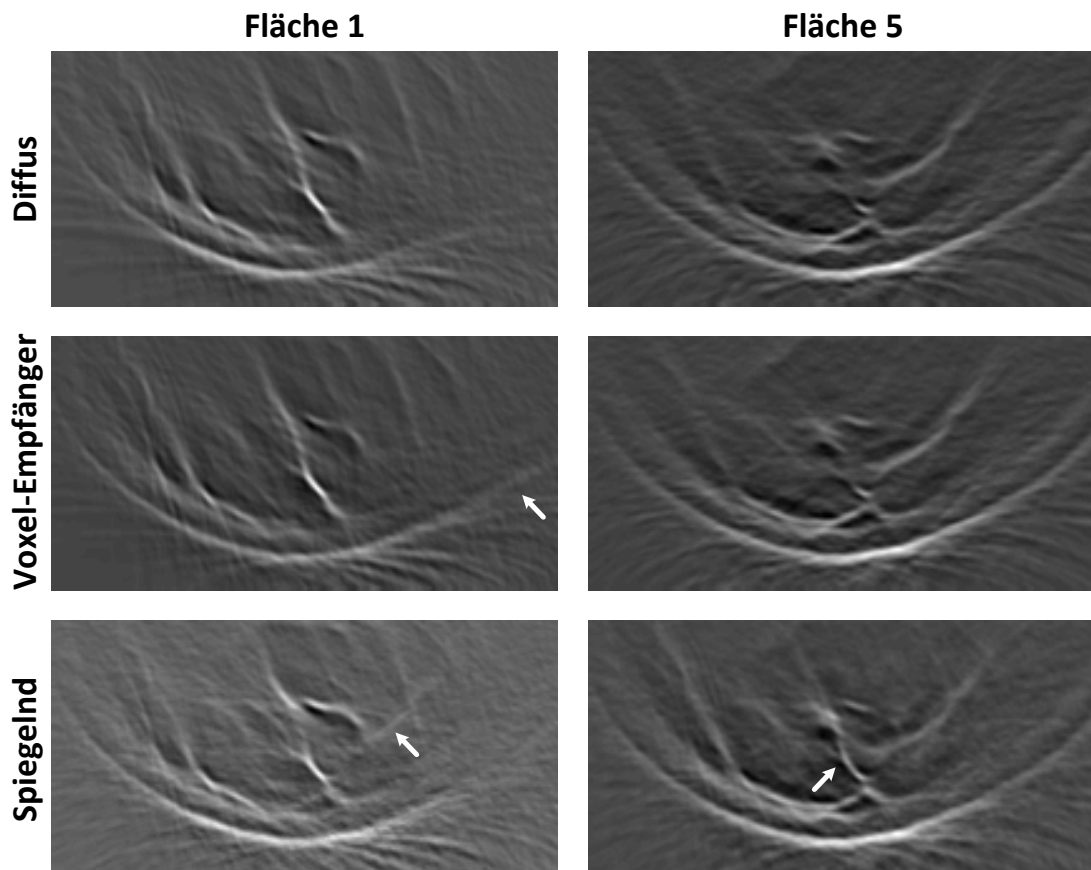


Abbildung 9.12: yz -Schnittbilder einer realen Brust von verschiedenen Kugelabschnittsvektoren $\vec{R}_{j=1,5}$ eines Dodekaeders, berechnet mit verschiedenen Annahmen für \vec{V} . Der Wertebereich liegt für alle Bilder zwischen 0 und 1.

Erhöhung des Kontrasts mit der energiereichsten Fläche

Unter der Annahme, dass die Flächen mit dem höchsten Reflektivitätswert auch die besten Echosignale liefern, können die Reflexionscharakteristiken der einzelnen Voxel jeweils auf die Fläche mit den größten Reflexionswert reduziert werden [102]. Das Maximum kann sowohl vor (I_{maxj}) als nach der Gewichtung (I_{Gmaxj}) bestimmt werden.

$$I_{maxj}(\vec{x}) = \max_j (A(j, \vec{x})) \quad (9.17)$$

$$I_{Gmaxj}(\vec{x}) = \max_j \left(\frac{A(j, \vec{x})}{N(j, \vec{x})} \right) \quad (9.18)$$

In Abbildung 9.13 wird das Bild der Standard-SAFT (oben) den Bildern bei Reduzierung auf die energiereichste Fläche (unten) für alle drei Richtungsvektoren $\vec{V} = \vec{S}, \vec{E}$ und \vec{H} gegenübergestellt. Da, relativ gesehen, für einige Flächen eine nur geringe Anzahl an Beiträgen aufsummiert wurden, kommt es bei der Gewichtung für den spiegelnden Fall zu sehr hohen Werten im oberen Bereich die das Bild insgesamt dominieren. Um diesen Effekt zu begrenzen, wurde die Anzahl der Beiträge N experimentell auf ein Minimum von 50.000 bzw. 200.000 Beiträgen begrenzt. Während es im Bild bei minimal 50.000 Beiträgen noch zu einigen dominierenden Voxel außerhalb der Brust kommt, können diese mit der Grenze von 200.000 Beiträgen herausgefiltert werden. Auch hier kann durch die Gewichtung ein verbesserter und gleichmäßiger Kontrast für Strukturen erreicht werden, die sich tiefer innerhalb der Brust befinden. Wie bereits in Abbildung 9.12 beobachtbar, ähneln sich die Bilder für den Fall einer diffusen und der Voxel-Empfänger Annahme stark. Die Bilder bei der Annahme einer spiegelnden Reflexion zeigen sehr ähnliche Strukturen, heben allerdings Strukturen im Inneren der Brust stärker hervor. Der Dynamikbereich liegt für alle drei Annahmen in derselben Größenordnung. Dies kann darauf hindeuten, dass die Brustoberfläche wie auch die innersten Strukturen sowohl diffus als auch spiegelnd reflektieren. In Abbildung 9.13 (unten) wird die Information aus allen drei Annahmen kombiniert dargestellt, indem die Bilder der drei Annahmen aufsummiert wurden. Dabei wurde für den spiegelnden Fall die Grenze von 200.000 Beiträgen genutzt. Als Maß für den Kontrast wurde das PSNR ermittelt. Der Kontrast steigt im Vergleich zur Standard-SAFT durch die Maximumdetektion um 53% für den ungewichteten und um 31% für den gewichteten Fall.

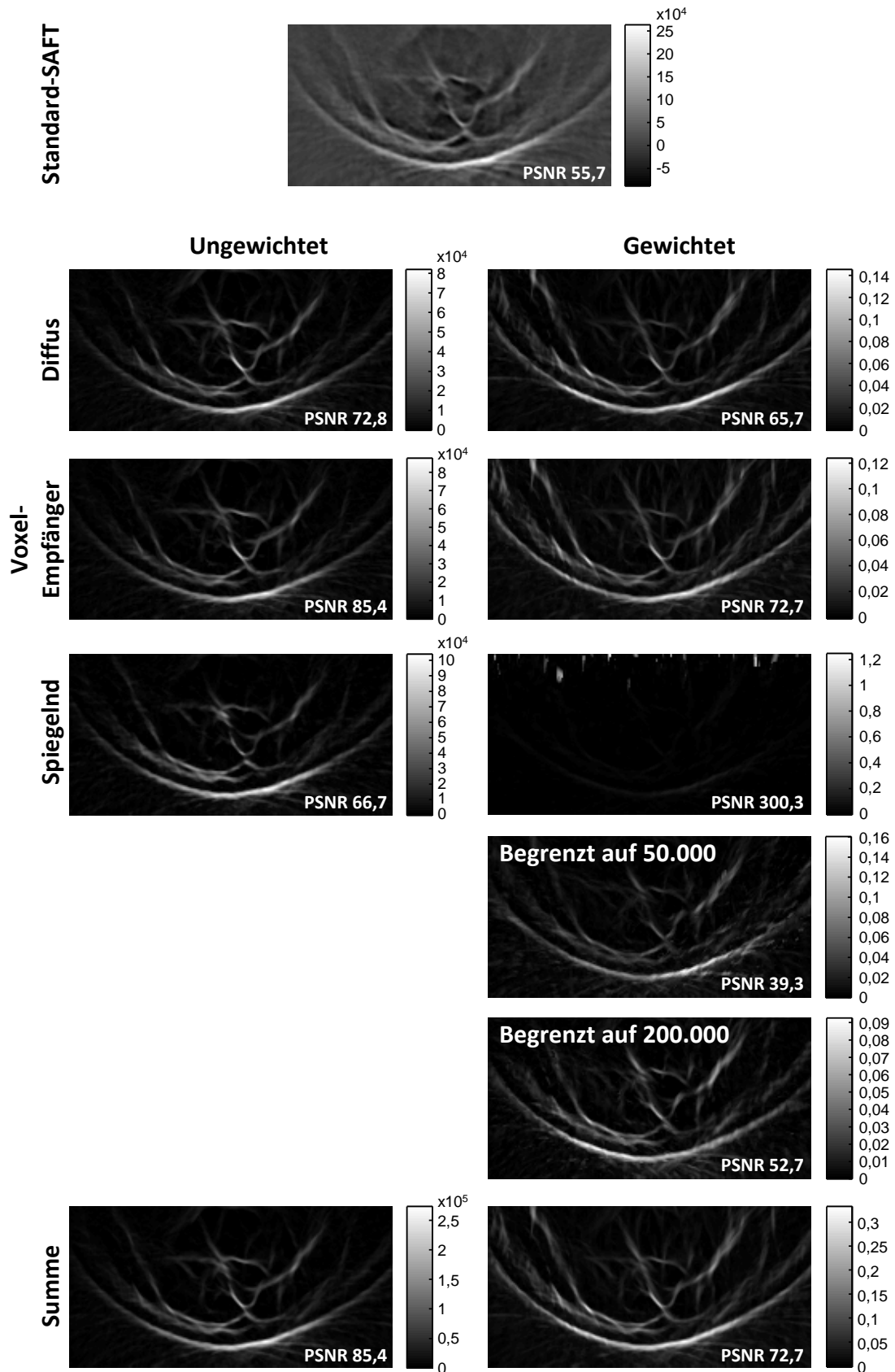


Abbildung 9.13: yz-Schnittbilder einer realen Brust mit der Standard-SAFT (oben) im Vergleich der energiereichsten Fläche eines Dodekaeders (mittig) für verschiedene Annahmen, sowie die Summe aus allen drei Annahmen. Links die ungewichteten und rechts die gewichteten Bilder.

Standardabweichung über alle Flächen

Ein weiterer Parameter für das Medium kann über die Standardabweichung der Reflektivitätswerte der Flächen ermittelt werden [102]. Stellt die Struktur innerhalb des Voxels einen isotropen Streuer dar, ist die Standardabweichung gering. Befindet sich eine Oberflächenstruktur innerhalb des Voxels, die zu einer gerichteten Reflexion (diffusen oder spiegelnden) führt, ist die Standardabweichung dagegen groß. Die Standardabweichung ohne (I_{stdj}) und mit Gewichtung (I_{Gstdj}) bestimmt sich durch

$$I_{stdj}(\vec{x}) = std_j(A(j, \vec{x})) \quad (9.19)$$

$$I_{Gstdj}(\vec{x}) = std_j \left(\frac{A(j, \vec{x})}{N(j, \vec{x})} \right). \quad (9.20)$$

Abbildung 9.14 zeigt den Vergleich zwischen dem Standard-SAFT (oben links) und der Darstellung der Standardabweichung über die Reflektivitätswerte der Flächen (unten). Zudem sind aus zwei Bildern Bereiche vergrößert dargestellt (oben rechts).

Es ist zu erkennen, dass die einzelnen Äste des Drüsengewebes sowie die Brustkante eine hohe Standardabweichung und damit eine starke gerichtete Reflexion aufweisen (Bereich A). Um diese Strukturen herum zeichnen sich Bereiche mit einer sehr niedrigen Standardabweichung scharf ab (Bereich B). Da sich im Bild der Standard-SAFT an diesen Stellen keine hohen Reflektivitätswerte befinden, deutet dies auf Bereiche ohne Reflexionen hin.

Es sind jedoch auch Bereiche mit hoher Standardabweichung zu beobachten, insbesondere in Bereichen, bei denen das Bild der Standard-SAFT hohe negative Reflektivitätswerte aufweist (Bereich C). Die hohen negativen Werte deuten dabei auf die Aufsummierung der negativen Anteile des optimalen Pulses hin und können als eine Abschattung hinter streuenden Strukturen verstanden werden. Als eine weitere Ursache kann auch eine fehlerhafte Annahme der Schallgeschwindigkeit bei der Rekonstruktion zu diesen Abbildungen führen. Eine fehlerhaft angenommene Schallgeschwindigkeit kann dazu führen, dass es nur teilweise zu einer konstruktiven Überlagerung der Ellipsen an den richtigen Positionen kommt. Somit würde eine niedrige Standardabweichung auf eine korrekte Überlagerung, eine hohe Standardabweichung dagegen auf eine nur sporadische konstruktive Überlagerung hindeuten.

Die Darstellung der Standardabweichung hebt reflektierende Strukturen scharfkantig von nichtreflektierendem Gewebe ab und ermöglicht somit eine verbesserte Abgrenzung der Strukturen im Bild.

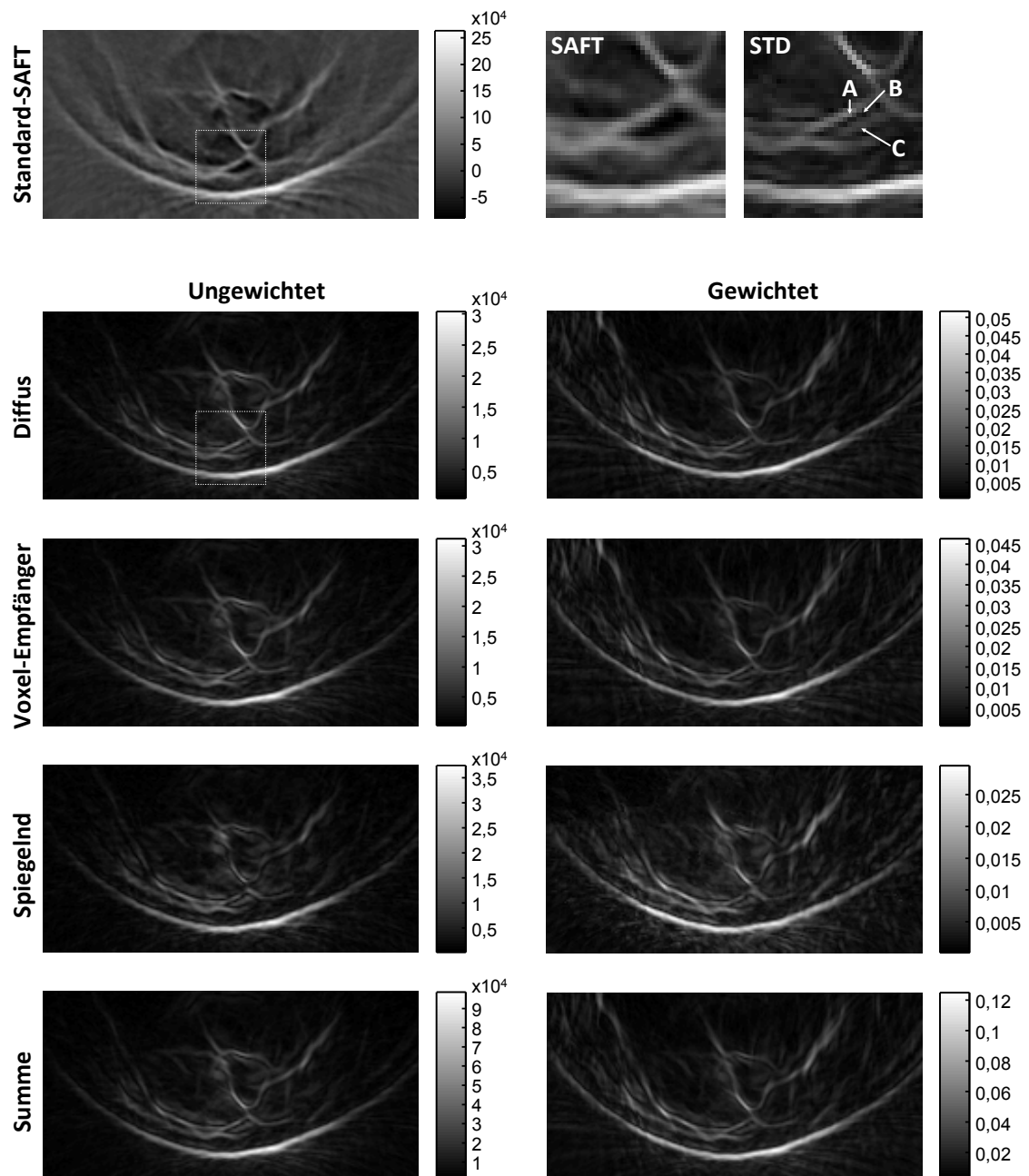


Abbildung 9.14: yz-Schnittbilder einer realen Brust mit der Standard-SAFT (oben links) im Vergleich zur Standardabweichung über die Reflektivitätswerte der Flächen eines Dodekaeders (unten) für verschiedene Annahmen, sowie die Summe aus allen drei Annahmen. Links die ungewichteten und rechts die gewichteten Bilder. Zwei Bereiche sind vergrößert dargestellt (oben rechts).

9.3.2 Klassifikation

Simulation von diffusen und spiegelnden Reflexionen

Mit dem Reflexionsmodell nach Phong (siehe Kapitel 9.1) können zwei Arten der Reflexion modelliert werden: Diffuse und spiegelnde Reflexionen. Damit stellt sich die Frage, ob sich diese beiden Reflexionsarten unterscheiden lassen, wenn die Zuordnung mit den Richtungsvektoren \vec{S} bei einer diffusen Annahme, oder aber mit dem Richtungsvektor \vec{H} bei einer spiegelnden Annahme durchgeführt wird.

Für eine Evaluation wurde in [102] eine reflektierende Struktur innerhalb eines Voxels mit Gleichung (9.6) nach Phong simuliert, und die Reflexionscharakteristik mit der Geometrie des niedriger aufgelösten Gradnetzes (30° , 18°) berechnet. Für die Oberflächennormale \vec{N} wurde der Azimut- und Elevationswinkel von $\phi = \theta = 0^\circ$ gewählt. Die Parameter k_{diffus} , $k_{spiegelnd}$ und n wurden für drei Fälle variiert. Es wurde eine rein diffuse, eine rein spiegelnde und eine gemischte Reflexion simuliert. Um die Richtcharakteristik über alle Richtungen bewerten zu können, wurde eine kugelförmige Apertur simuliert, auf der 10.000 Sender und Empfänger nach [107] zufällig, jedoch gleichmäßig positioniert wurden. Dabei wurde die Richtcharakteristik nach Gleichung (9.14) gewichtet berechnet. Zur Evaluation der 3D-USCT wurde die Simulation zudem mit einer halbkugelförmigen Apertur durchgeführt und so die halbellipsoide Apertur der 3D-USCT angenähert. Für diese Arbeit wurde die Simulation auf 10 AP mit insgesamt 8.873.640 SEP erweitert, um eine genauere Aussage zu ermöglichen.

Abbildung 9.15 zeigt die berechneten Reflexionscharakteristiken für eine kugelförmige Apertur. In der linken Abbildungsspalte wurde eine rein diffuse Reflexion simuliert und in der rechten Spalte eine rein spiegelnde Reflexion. Abbildung 9.15 links oben zeigt eine Reflexionscharakteristik, die unter der Annahme einer diffusen Reflexion simuliert und berechnet wurde. Die gewählten Parameter waren $k_{spiegelnd} = 0$, $k_{diffus} = 1$. Gut erkennbar ist der Kosinusverlauf, der durch den diffusen Anteil des Phong-Modells in Gleichung (9.6) erzeugt wird. In Abbildung 9.15 rechts mittig wird eine Reflexionscharakteristik dargestellt, die unter der Annahme einer spiegelnden Reflexion simuliert und berechnet wurde. Dabei wurde für den Exponenten $n = 2$ und für die Parameter $k_{spiegelnd} = 1$, $k_{diffus} = 0$ gewählt (siehe Gleichung (9.3)). Hier ist der quadratische Kosinusverlauf durch den Exponenten n sichtbar, der durch den spiegelnden Anteil des Phong-Modells erzeugt wird.

Die berechneten Reflexionscharakteristiken bei falscher Annahme sind in Abbildung 9.15 rechts oben und links mittig dargestellt. Dabei wird der Kosinusverlauf stark gedämpft und mit einem Offset abgebildet. Dabei lässt sich beobachten, dass sich die Beiträge unter einer falschen Annahme gleichmäßig über alle Winkel verteilen und zu einem Offset führen. So bildet sich die diffuse Reflexionscharakteristik mit einem Mittelwert von 0,39 bei einer spiegelnden Annahme als Offset von 0,42 in etwa ab. Die spiegelnde Reflexionscharakteristik mit einem Mittelwert von 0,12 bildet sich dagegen bei einer diffusen Annahme als ein Offset von 0,09 in etwa ab. Bei einer Berechnung mit dem Richtungsvektor \vec{E} ergeben sich ebenfalls stark gedämpfte Kosinusverläufe. Die simulierte diffuse Reflexion führt zu einem Offset von 0,48 (Abbildung 9.15 unten links). Eine simulierte spiegelnde Reflexion führt dagegen zu einem Offset von 0,08 (Abbildung 9.15 unten rechts).

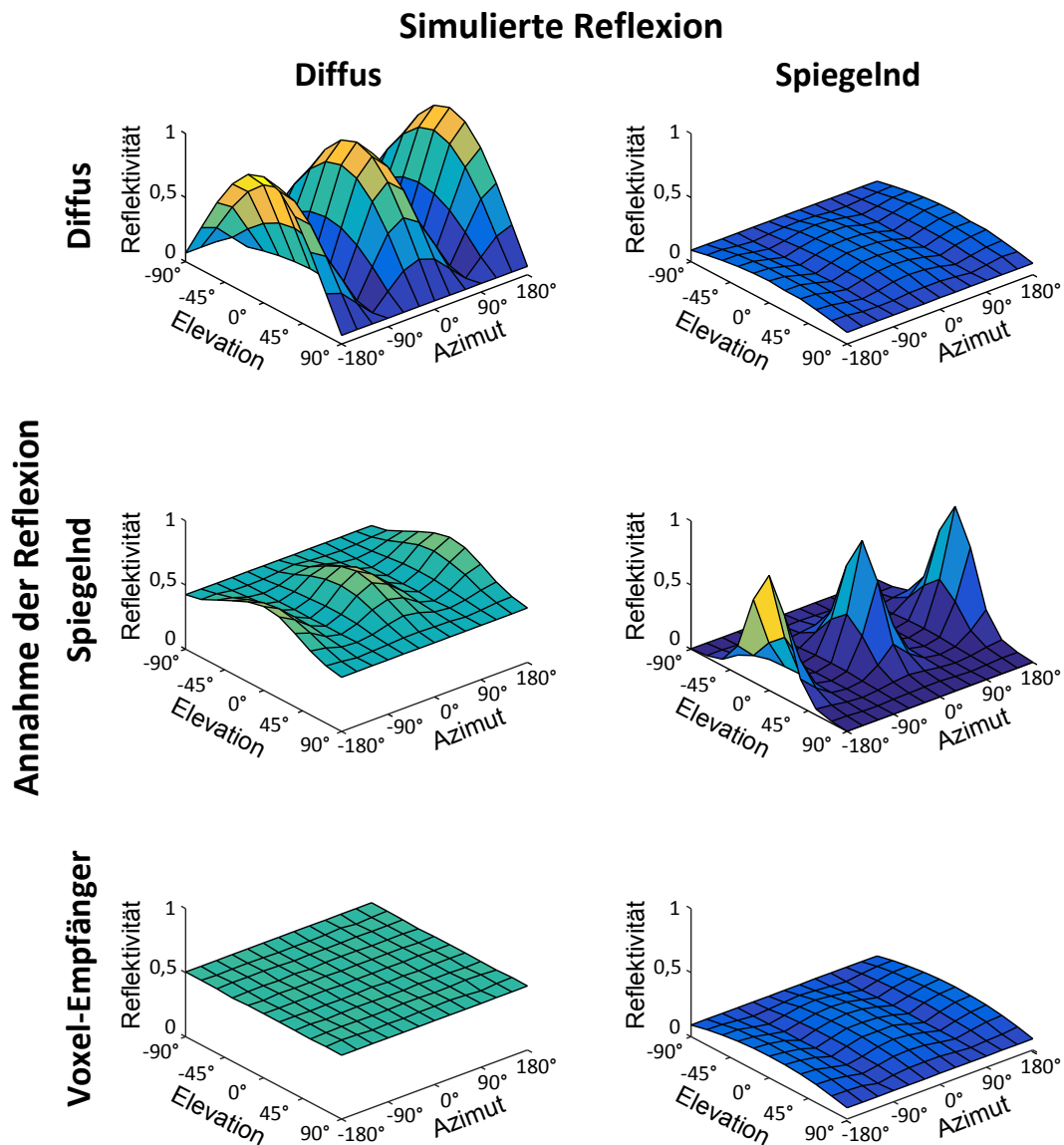


Abbildung 9.15: Berechnete Reflexionscharakteristiken unter verschiedenen Reflexionsannahmen eines Voxels innerhalb einer kugelförmigen Apertur bei Simulation einer rein diffusen Reflexion (links) und einer rein spiegelnde Reflexion (rechts).

Wird eine gemischte Reflexion mit den Parametern $k_{diffus} = 0,5$, $k_{spiegelnd} = 0,5$ und Exponenten $n = 2$ simuliert, ergeben sich, abhängig von der genutzten Annahme, die in Abbildung 9.16 dargestellten Reflexionscharakteristiken. Die ermittelten Reflexionscharakteristiken bei einer kugelförmigen Apertur sind in Abbildung 9.16 links dargestellt. Nach Entfernen des jeweiligen Offsets besitzen die maximalen Amplituden eine Reflektivität von 0,51 und 0,56 für die diffuse bzw. spiegelnde Annahme. Dies entspricht nahezu den, für die Simulation gewählten, Parametern k_{diffus} und $k_{spiegelnd}$. Im Fall einer halbkugelförmigen Apertur werden die Beiträge nur einem Teil der Winkelbereiche zugeordnet, da nicht in allen Richtungen Sender bzw. Empfänger vorhanden sind (siehe Abbildung 9.16 rechts). Jedoch weisen die Richtungen, für die Sender bzw. Empfänger vorhanden sind, nahezu dieselben Werte auf, die bei einer kugelförmigen Apertur berechnet wurden. Eine Ausnahme bildet die Re-

flexionscharakteristik bei einer spiegelnden Annahme (siehe Abbildung 9.16 mittig). Hier kommt es für einen Elevationswinkel $\phi > 45^\circ$ zu deutlich erhöhten Reflektivitäten. Eine genauere Untersuchung der Ursache konnte aufzeigen, dass für diese Richtungen nur eine geringe Anzahl an Beiträgen vorhanden waren und ihren Ursprung im diffusen Anteil der Reflexion haben. Durch die genutzte Gewichtung nach Gleichung (9.14), kommt es hier zu einer deutlichen Übergewichtung der diffusen Beiträge.

Die Ergebnisse der Simulation zeigen, dass sich Reflexionsarten und selbst Mischformen aus den zurückprojizierten Echosignalen extrahieren lassen. Die Beiträge der jeweils nicht angenommenen Reflexionsarten bilden sich dabei als ein Offset ab, der über eine Berechnung des Mittelwerts entfernt werden kann.

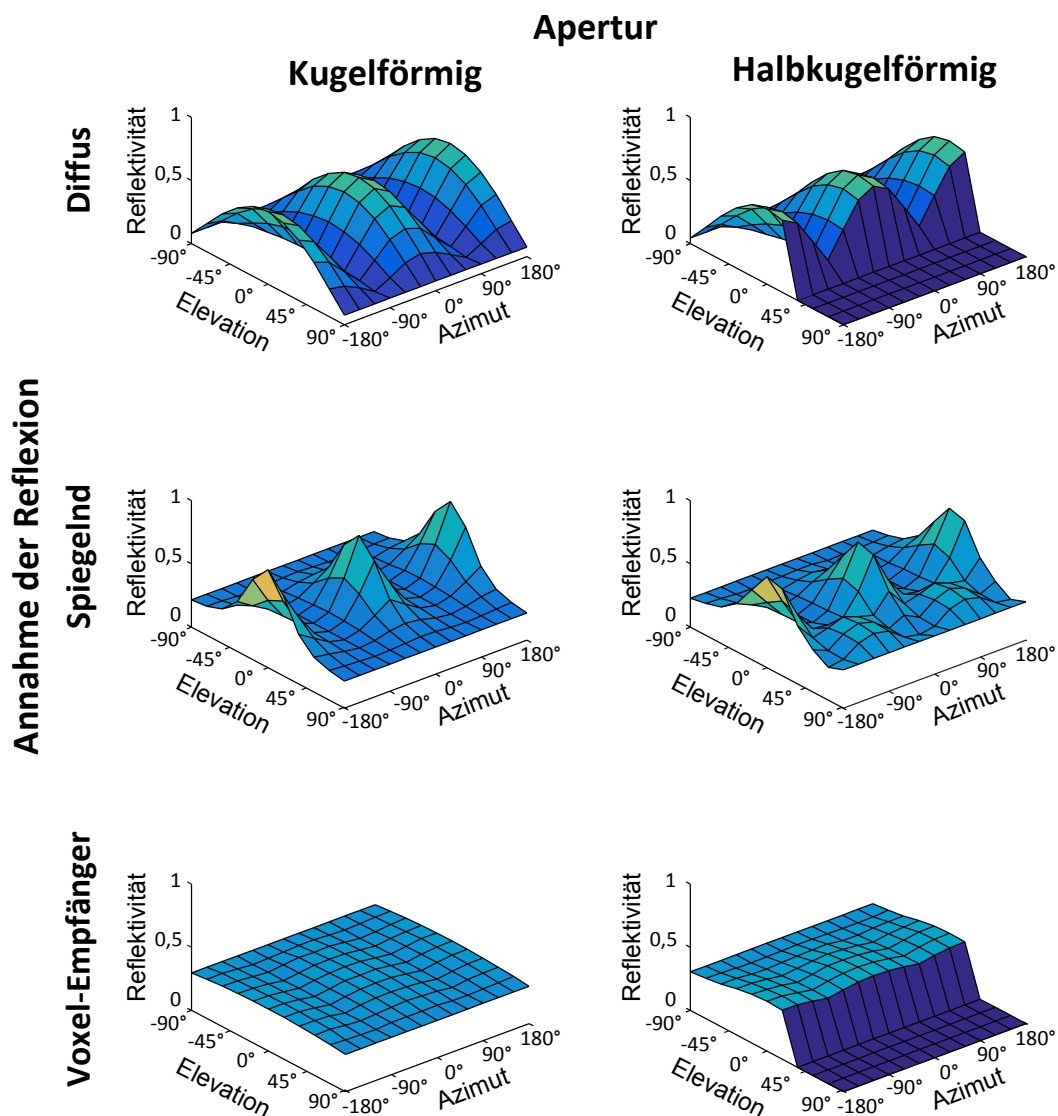


Abbildung 9.16: Berechnete Reflexionscharakteristiken unter verschiedenen Reflexionsannahmen eines Voxels innerhalb einer kugelförmigen (links) und halbkugelförmigen Apertur (rechts) bei Simulation einer gemischten Reflexion.

Schätzung der Parameter für einen Voxel

Aus den gewonnenen Reflexionscharakteristiken können die Parameter über das Phong-Modell geschätzt werden. Dafür wurde in [102] ein Modellschätzer, basierend auf dem Levenberg-Marquardt Algorithmus [108], eingesetzt. Dabei wird ein beliebiges Schätzmodell, das aus zu schätzenden Parametern b besteht, mithilfe des quadratischen Fehlers optimiert (siehe Tabelle 9.2).

Die Simulation und Schätzung der Parameter wurde 100 mal durchgeführt. Für jeden Durchlauf wurden die Parameter zufällig neu gewählt. Die Sender und Empfänger von 500 SEP wurden dabei zufällig auf der Halbkugel verteilt. Die Oberflächennormale \vec{N} sowie die Reflexionsparameter k_{diffus} , $k_{spiegelnd}$ und n wurden folgendermaßen zufällig erzeugt: $k_{diffus} = \text{rand}[0..1]$ liegt zwischen 0 und 1, $k_{spiegelnd} = 1 - k_{diffus}$ und $n = 1 + \text{rand}[0..1]$.

Die Untersuchung wurde mit einer diffusen Reflexionsannahme und drei unterschiedlich komplexen Schätzmodellen durchgeführt. Für die Berechnung der Daten wurde das Gradnetz mit einer Abtastung von 30° bzw. 18° eingesetzt. Die Ergebnisse sind in Tabelle 9.2 dargestellt. Das erste Schätzmodell enthält zwei zu schätzende Parameter. $b_{k,diffus}$ für den Reflexionsparameter und \vec{b}_N für die Oberflächennormale. Das zweite Schätzmodell wird um den Exponenten b_n erweitert. Das dritte Schätzmodell berücksichtigt zudem einen Offset b_{offset} . Für die Schätzung der diffusen Parameter geht der Anteil der spiegelnden Reflexion als Störgröße mit ein. Die Größe des mittleren absoluten Fehlers in Tabelle 9.2 zeigt die Genauigkeit der Schätzmodelle für die Normale von unter 12° . Dies liegt bereits unterhalb der Abtastung des Gradnetzes von 30° bzw. 18° . $b_{k,diffus}$ wird auf etwa 0,14 genau geschätzt. Das zweite Schätzmodell mit b_n als Exponenten bringt keinen Vorteil. Wird dagegen noch ein Offset modelliert, verdoppelt sich die Schätzgenauigkeit der Oberflächennormalen \vec{b}_N .

Vergleich der Reflexionscharakteristiken pro Voxel

Eine weitere Möglichkeit Gewebe zu klassifizieren besteht darin, die Reflexionscharakteristiken einzelner Voxel miteinander zu vergleichen und so ihre Ähnlichkeit zu bestimmen. In [102] wurden die in Abbildung 9.17 dargestellten Schritte für eine erste Evaluation der Methode genutzt. Diese Untersuchung wird im Folgenden durchgeführt, wobei im Gegensatz zu [102] eine höhere Datenmenge und das niedriger aufgelöste Gradnetz eingesetzt wird.

Schätzmodell	mittlerer Fehler für	
	$b_{k,diffus}$	$\angle(\vec{N}, \vec{b}_N)$
1. $I = \left b_{k,diffus} \left(\vec{S} \cdot \vec{b}_N \right) \right $	0,14	$11,1^\circ$
2. $I = \left b_{k,diffus} \left(\vec{S} \cdot \vec{b}_N \right)^{b_n} \right $	0,09	$11,7^\circ$
3. $I = \left b_{k,diffus} \left(\vec{S} \cdot \vec{b}_N \right)^{b_n} + b_{offset} \right $	0,1	$5,9^\circ$

Tabelle 9.2: Überblick des mittleren Fehlers der Parameterschätzung bei drei unterschiedlich komplexen Schätzmodellen.

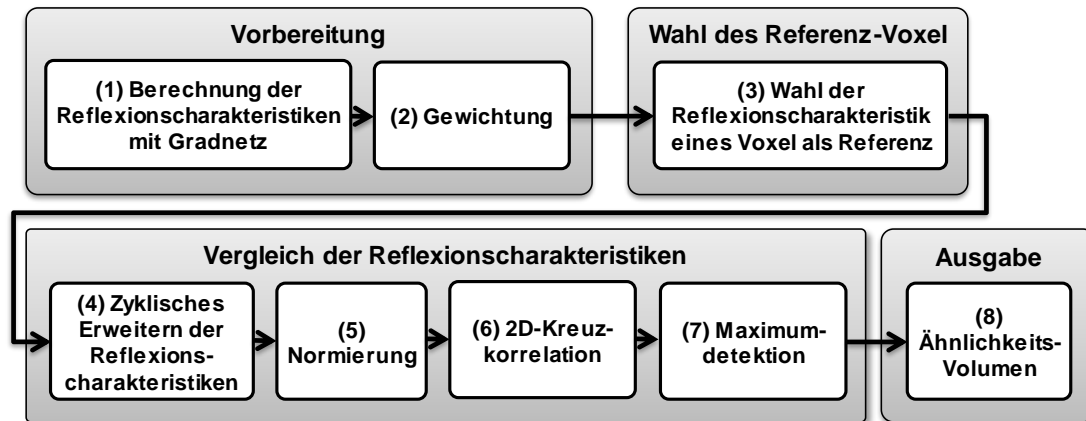


Abbildung 9.17: Ablauf für die Bestimmung der Ähnlichkeit von Reflexionscharakteristiken nach [102].

1. In [102] kam es zu deutlichen Lücken in den Reflexionscharakteristiken, die mit einem Gaußfilter aufgefüllt werden mussten. Für die Berechnung der Reflexionscharakteristiken wird daher die Datenmenge von 10 AP statt nur einer AP genutzt, um mehr Beiträge für die Winkel zu erhalten. Zudem wird das Gradnetz (30° , 18°) statt dem höher aufgelösten Gradnetz (5° , 5°) eingesetzt. Die Berechnung wird unter der Annahme einer diffusen Reflexion durchgeführt.
2. Im zweiten Schritt wird eine Gewichtung durch die Anzahl der SEP-Beiträge durchgeführt.
3. Ist die Vorbereitung der Daten abgeschlossen, stehen die Reflexionscharakteristiken für alle Voxel zur Verfügung. Es wird ein Voxel gewählt, dessen Reflexionscharakteristik als Referenz für den Vergleich genutzt wird (siehe Abbildung 9.18 oben links). Hier wird die Reflexionscharakteristik eines einzelnen Voxels genutzt, wobei auch die Möglichkeit besteht über mehrere Reflexionscharakteristiken von Voxel zu mitteln, die zu einem gleichem Gewebetyp gehören.
4. Um Probleme bei der 2D-Korrelation für den Vergleich an den Rändern zu vermeiden, wird die Symmetrie der Kugel ausgenutzt und die Reflexionscharakteristik nach außen hin zyklisch erweitert. Dies geschieht unter Beachtung der entsprechenden Winkel (siehe Abbildung 9.18 oben rechts). Durch die zyklische Erweiterung wird ein Vergleich ermöglicht, der auch Rotationen der Normalen zulässt. Allerdings ist der Vergleich bei einer Rotation in Elevationsrichtung problematisch, da die Kugelabschnittsflächen bei dem Gradnetz eine unterschiedliche Größe in Elevationsrichtung aufweisen. Eine Rotation der Normalen in Azimutrichtung ist dagegen unproblematisch.
5. In Schritt fünf wird eine Normierung durchgeführt, bei der die Reflektivitätswerte nach dem Entfernen eines Offsets durch das Maximum dividiert werden (siehe Abbildung 9.18 oben rechts). Diese Normierung wird durchgeführt, um die Ähnlichkeit des Verlaufs der Reflexionscharakteristiken, unabhängig von der maximalen Amplitude, zu bewerten.
6. Nun wird eine 2D-Kreuzkorrelation zwischen der Reflexionscharakteristik der Referenz und den Reflexionscharakteristiken aller Voxel durchgeführt. Als Ergebnis wird für jeden Voxel eine zweidimensionale Matrix zurückgeliefert. Sie

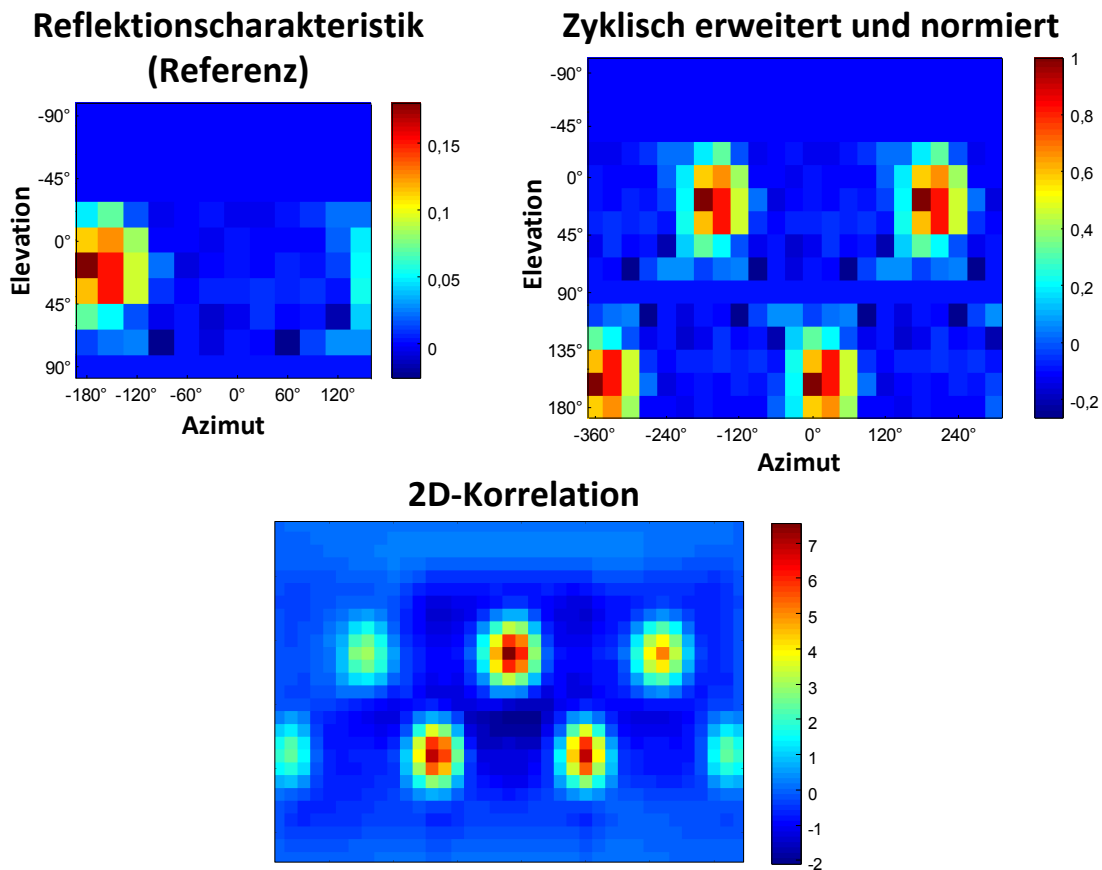


Abbildung 9.18: Reflexionscharakteristik des Referenz-Voxels (oben links), zyklisch erweitert und normiert (oben rechts) und Ergebnis der 2D-Korrelation von beiden (unten).

beinhaltet für jede mögliche Rotation in Elevation- und Azimut-Richtung einen Korrelationswert. Beispielhaft ist in Abbildung 9.18 unten das Ergebnis der Korrelation mit der Referenz-Reflexionscharakteristik dargestellt.

7. In dieser Ergebnis-Matrix wird das Maximum der Kreuzkorrelationsmatrix gesucht.
8. Das Maximum stellt die maximale Ähnlichkeit zum Referenzvoxel für alle Rotationen in Elevation- und Azimut-Richtung dar und wird daher als Ergebnis für das jeweilige Voxel in das Ähnlichkeitsvolumen übernommen.

Der Vergleich für ein XZ-Schnittbild ist in Abbildung 9.19 für zwei verschiedene Referenzvoxel dargestellt. Zum Vergleich ist in Abbildung 9.19 oben die Standard-SAFT-Rekonstruktion dargestellt. Für die erste Referenz wurde ein Voxel im Drüsengewebe gewählt (siehe Markierung x_1 Abbildung 9.19 oben). Das Ergebnis zeigt eine hohe Ähnlichkeit zu reflektierenden Strukturen des Drüsengewebes und der Brustoberfläche. Auch tiefere Strukturen werden hervorgehoben. Der Hintergrund ist deutlich abgegrenzt. Für die zweite Referenz wurde ein Voxel in einem Bereich aus dem Hintergrund gewählt, der keine stark streuende Gewebearten beinhaltet (siehe Markierung x_2 Abbildung 9.19 oben). Auch hier zeigt sich eine deutliche Ähnlichkeit des Hintergrundes, selbst auf der rechten Seite innerhalb der Brust. Und eine gute Abgrenzung zu reflektierenden Strukturen wie dem Drüsengewebe oder der Brustoberfläche.

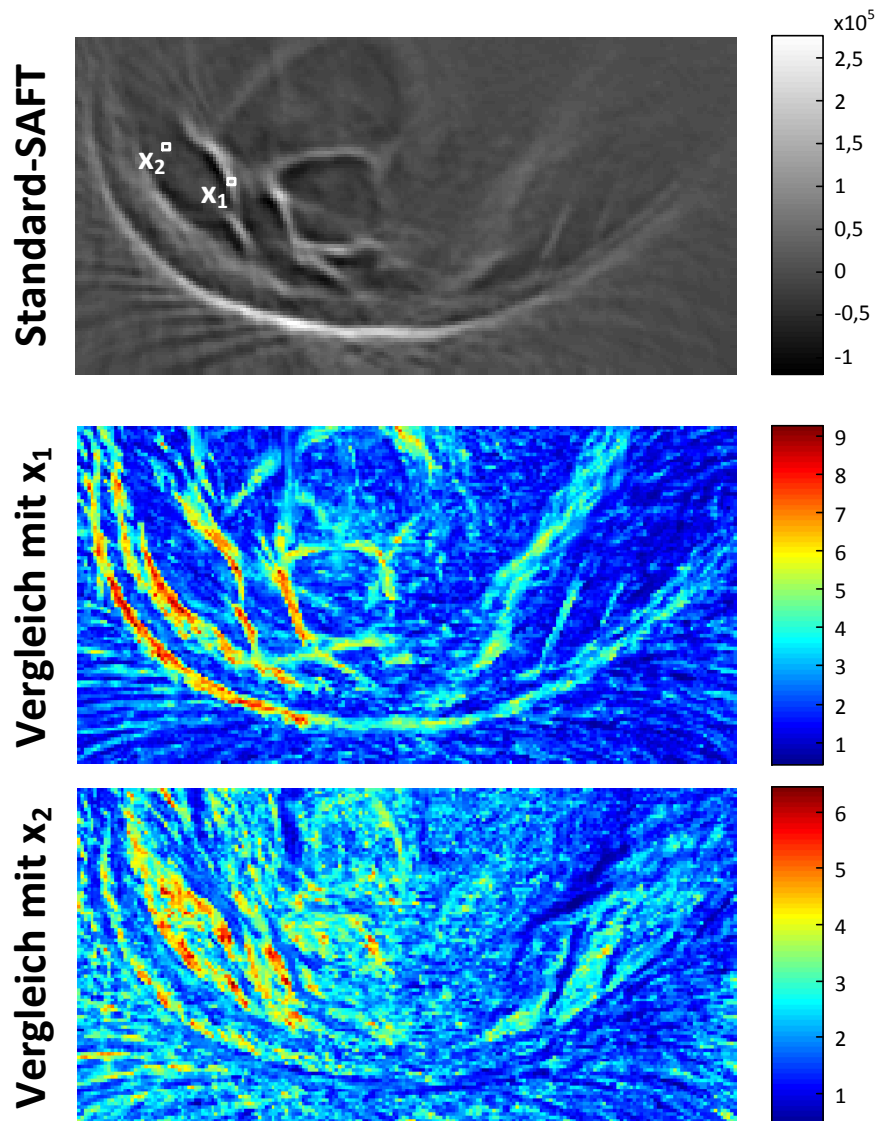


Abbildung 9.19: xz -Schnittbilder einer realen Brust mit der Standard-SAFT (oben) und die Ähnlichkeit der Reflexionscharakteristiken pro Voxel mit den Reflexionscharakteristiken zweier Referenzvoxel. Referenzvoxel aus dem Drüsengewebe (mittig) und aus Fettgewebe (unten). Nach [14].

9.3.3 Performanz

Die maximale Rechengeschwindigkeit für die Berechnung der Reflexionscharakteristiken ist davon abhängig, welche Geometrie und welcher Richtungsvektor genutzt werden (siehe Tabelle 9.3). Die Komplexität ergibt sich im Allgemeinen zu $O(V \cdot A \cdot F)$. Dabei bezeichnet V die Anzahl der zu rekonstruierenden Voxel im Reflexionsvolumen, A die Anzahl der genutzten A-Scans und F die Anzahl der Flächen der genutzten Geometrie.

Dabei erreicht die Zuordnung mit dem Richtungsvektor \vec{S} , bei einer diffusen Annahme, jeweils die höchste Rechengeschwindigkeit. Der Grund liegt in der Reihenfolge, in der die A-Scans der SEP abgearbeitet werden. Diese liegen so geordnet vor, dass für einen Sender erst alle Empfänger abgearbeitet werden, bevor der nächste Sender vorkommt. Dadurch muss erst jedem neuen Sender der Richtungsvektor neu

Geometrie	Richtungsvektor $\vec{V} \in (\vec{S}, \vec{H}, \vec{E})$	Flächen	Performanz	
			GVA/s	GVAF/s
Dodekaeder	diffus (\vec{S})	12	4,0	47,5
	spiegelnd (\vec{H})		2,7	31,9
	Voxel-Empfänger (\vec{E})		3,2	38,7
Ikosaeder	diffus (\vec{S})	20	2,2	43,1
	spiegelnd (\vec{H})		0,8	16,5
	Voxel-Empfänger (\vec{E})		0,9	17,1
Gradnetz (30°, 18°)	diffus (\vec{S})	110	0,4	47,3
	spiegelnd (\vec{H})		0,2	21,0
	Voxel-Empfänger (\vec{E})		0,2	21,4

Tabelle 9.3: Überblick der Rechengeschwindigkeiten bei verschiedenen Geometrien und Richtungsvektoren auf einer GTX TITAN GPU. Die Daten wurden in [102] bei einem Volumen von 238x238x119 Voxel und 30.000 A-Scans ermittelt.

berechnet und die Zuordnung zu einer Fläche bestimmt werden. Im Idealfall gilt die Zuordnung für 1413 SEP für einen Sender und 1413 Empfänger. Die Zuordnung für den Richtungsvektor \vec{E} erreicht die zweithöchste Rechengeschwindigkeit. Durch die Koordinaten des Empfängers kann die Zuordnung zu einer Fläche direkt bestimmt werden. Durch die Reihenfolge der A-Scans muss diese jedoch für jedes SEP neu berechnet werden. Für die Zuordnung mit dem Richtungsvektor \vec{H} , bei einer spiegelnden Annahme, muss bei jedem SEP zusätzlich die Winkelhalbierende \vec{H} neu bestimmt werden. Das führt zu der geringsten Rechengeschwindigkeit.

Die Berechnungen für die Geometrie des Dodekaeders können aufgrund der geringen Flächenanzahl von zwölf Flächen am schnellsten durchgeführt werden. Beim Ikosaeder mit 20 Flächen sinkt die Rechengeschwindigkeit, da pro A-Scan mehr Richtungsvektoren verglichen werden müssen. Da die Rechenzeit für den Vergleich mit den Kugelabschnittsvektoren \vec{R}_j überwiegt, ist der Unterschied zwischen den Richtungsvektoren \vec{E} und \vec{H} nur noch gering. Beim Gradnetz kann die Zuordnung der Sendevektoren direkt mit einem Mapping durchgeführt werden. Durch die größere Datenmengen mit 110 bzw. 2522 Flächen wird die Rechengeschwindigkeit vom Datentransfer begrenzt. Allein das Cachingverhalten für den Vektor zum Sender führt hier noch zu einem Unterschied von Faktor 2.

Wird die Fläche in die Rechengeschwindigkeit mit einbezogen, so erhält man die Metrik (Giga·Voxel·A-Scans·Flächen)/Sekunde (GVAF/s). Es zeigt sich, dass die Rechengeschwindigkeit für die diffuse Annahme nahezu gleich bleibt. Für eine hohe Anzahl an Flächen lohnt es sich aus der Sicht der Rechengeschwindigkeit auf ein Gradnetz überzugehen, da dort die Zuordnung deutlich schneller durchgeführt werden kann. Dieser Performanzgewinn wird jedoch damit erkauft, dass die Kugelabschnittsflächen nicht mehr gleich groß sind.

10. Zusammenfassung, Diskussion und Ausblick

Das Ziel dieser Arbeit war es hochauflösende SAFT-Bilder für heterogene Medien wie der Brust zu ermöglichen. Dafür wurden in dieser Arbeit erfolgreich vielfältige Methoden entwickelt und analysiert, die die Rekonstruktion mit der Standard-SAFT erweitern. Ein besonderes Augenmerk wurde bei allen Methoden auf die Beschleunigung mittels GPUs gelegt, da die Rekonstruktion sehr rechenintensiv ist und für eine klinische Anwendbarkeit deutlich beschleunigt werden musste. Durch die Ausnutzung der Richtungsinformation von Reflexionen konnten in der 3D-USCT gänzlich neuartige Daten gewonnen werden, die weitere Informationen über das vermessene Medium preisgeben. Diese Reflexionscharakteristiken können mit einer Vielzahl an Methoden untersucht und weiterverwendet werden.

Erweiterung der 3D-SAFT-Rekonstruktion für heterogene Medien

Der Standardversion der SAFT liegen vereinfachende Annahmen zugrunde, die für Untersuchungen der Brust nicht geeignet sind, da die Brust ein komplexes und heterogenes Medium ist. Der Einfluss dieser Annahmen auf die Bildqualität wurde anhand einer eigens entwickelten Simulation sowie an experimentellen Messdaten unterschiedlicher Komplexität detailliert untersucht. Die dafür entwickelte Simulation basiert vereinfacht auf der Strahlverfolgung. Weitere physikalische Effekte wie Beugung oder Brechung wurden vernachlässigt, da diese für die Bewertung der Schallgeschwindigkeits- und Dämpfungskorrektur, die in dieser Arbeit durchgeführt wurde, nicht im Fokus standen. Es wurde gezeigt, dass die Annahme einer konstanten Schallgeschwindigkeit die Bildqualität signifikant verschlechtert. Dabei verursacht beispielsweise eine Abweichung der Schallgeschwindigkeit von 40 m/s bereits im Mittel eine Verschiebung von 5 mm , eine Kontrastminderung auf $4,5\%$ und eine Verschlechterung des Auflösungsvermögens auf $0,5\text{ mm}$. Es kann zu Verformungen einer streuenden Struktur und zu Mehrfachabbildungen insbesondere von kleineren Strukturen wie Punktstreuern kommen. Aufgrund der lokalen Abhängigkeit der Bildqualität wurde die Bildqualität dabei für mehrere Punkte im Raum untersucht. Für eine kohärente Bildgebung wurde der maximal erlaubte Fehler für die mittlere Schallgeschwindigkeit analytisch hergeleitet. Dieser liegt bei der aktuellen Geometrie bei $0,39\text{ m/s}$. Ebenso wurde der Einfluss der Annahme einer vernachlässigbaren und konstanten Dämpfung auf die Bildqualität analysiert. Bedingt durch die Dämpfung innerhalb der Brust kommt es zu einer deutlichen Abschattung sowie zu einer Kontrastreduzierung der inneren Bereiche der Brust. Die Analyse der Bildqualität macht deutlich, dass für den Anwendungsfall der Brustuntersuchung eine Korrektur der Annahmen von SAFT essentiell wichtig ist.

Daher wurde die SAFT um eine Schallgeschwindigkeits- und Dämpfungskorrektur erweitert, wodurch die Begrenzungen durch die vereinfachenden Annahmen der SAFT erfolgreich überwunden werden konnten. Dazu wurde der Vorteil der 3D-USCT ausgenutzt, die eine Rekonstruktion von Schallgeschwindigkeits- und Dämpfungsvolumen mit der Transmissionstomographie ermöglicht. In dieser Arbeit konnte an

komplexeren Phantomen gezeigt werden, dass die durchgeführte Schallgeschwindigkeitskorrektur die Bildqualität signifikant verbessern kann. Die mittlere Verschiebung konnte so auf 0,06 mm reduziert, der Kontrast auf 93,6% angehoben und das räumliche Auflösungsvermögen auf 0,28 mm verbessert werden. Damit kann, was den Rekonstruktionsalgorithmus betrifft, in der 3D-USCT erstmals bei heterogenen Medien eine räumliche Auflösung von 0,28 mm erreicht werden. Durch die Dämpfungskorrektur ist es nun möglich die gedämpften Bereiche anzuheben und räumlich homogen darzustellen. Da die Verstärkung der Signale für gedämpfte Bereiche ebenfalls das Rauschen verstärkt, kann es bedingt durch das Signal-zu-Rausch-Verhältnis in den Messsignalen zu Artefakten im Bild kommen, die wiederum den Kontrast reduzieren. Daher werden die Messsignale vor der Rekonstruktion analysiert und die Dämpfungskorrektur entsprechend limitiert. Somit kann die Dämpfungskorrektur abhängig vom Objekt optimiert durchgeführt werden.

Trotz der durchgeführten Korrekturen konnte bei einer realen Brust bisher die maximale Auflösung von 0,24 mm noch nicht erreicht werden. Ein großer Einflussfaktor stellt hierbei die Genauigkeit und Auflösung der vorhandenen Schallgeschwindigkeitsbilder dar. So liegt die Genauigkeit laut [92] mit ca. 3 m/s bereits um einen Faktor 7,7 oberhalb der ermittelten Obergrenze von 0,39 m/s. Ebenso liegt die Auflösung der Transmissionsbilder aktuell mit ca. 1,2 cm um einen Faktor 50 höher als die maximale Auflösung der Reflexionstomographie.

Daher wird aktuell in der Forschungsgruppe an neuen Algorithmen für die Transmissionstomographie gearbeitet. Sobald verbesserte Bilder zur Verfügung stehen, wird aus den simulierten Daten erwartet, dass die Auflösung der verbesserten SAFT-Rekonstruktion weiter ansteigt.

Reduzierung der Rekonstruktionsartefakte in der 3D-USCT

Es wurde eine neue Methode entwickelt und analysiert, um die SAFT-typischen Rekonstruktionsartefakte in den Reflexionsbildern zu reduzieren. Unter der Annahme, dass die Reflexion einer Struktur über mehrere Empfänger hinweg sichtbar sein sollte, können Fehlerbeiträge von anderen Streuern mithilfe eines laufenden Medianfilters herausgefiltert werden. Diese Erweiterung lässt sich bei einem nur geringen Speicherbedarf direkt in den Programmablauf der SAFT-Rekonstruktion integrieren. Es wurde zudem gezeigt, dass sich durch eine einfache Umsortierung der Messsignale vor der Rekonstruktion das Ergebnis des Medianfilters verbessern lässt. Die Reduzierung der Rekonstruktionsartefakte führt zu einem um rund Faktor zwei erhöhten Kontrast und trägt somit dazu bei, auch kleinere streuende Strukturen zu erkennen.

Beschleunigung der SAFT-Algorithmen

Um den Nutzen aus der GPU-beschleunigten SAFT-Rekonstruktion in der Praxis ziehen zu können, wurde eine effiziente Schnittstelle zwischen der GPU und der aktuell eingesetzten Entwicklungsumgebung in MATLAB konzipiert und implementiert. Zudem wurde der Rekonstruktionsserver der 3D-USCT erstmals über eine externe GPU-Box mit acht GPUs erweitert, womit ein weiterer Beschleunigungsfaktor von 7,8 erreicht werden konnte.

Durch die Schallgeschwindigkeits- und Dämpfungskorrektur steigt die Rechenkomplexität der SAFT weiter an. Daher wurden mehrere Rechenstrategien bezüglich

ihrer Bildqualität und Rechengeschwindigkeit hin analysiert, um einen optimierten Kompromiss zu finden. Es konnte gezeigt werden, dass durch eine geeignete Vorberechnung der Teilpfade die nötigen Korrekturen für heterogene Medien sehr effizient berechnet werden können. Die Vorberechnung benötigt dabei nur 1,5% der Rekonstruktionszeit bei einer Rechengeschwindigkeit von 7,3 GVA/s auf einer GTX 590 GPU. Mit zusätzlicher Dämpfungskorrektur sinkt die maximale Rechengeschwindigkeit auf 4,3 GVA/s. Hierbei wurde gezeigt, dass insbesondere die Textureinheiten der GPU effizient für die 3D-USCT genutzt werden können, über die eine Interpolation direkt in Hardware durchgeführt werden kann. Diese GPU-beschleunigten Schallgeschwindigkeits- und Dämpfungskorrekturen ermöglichen es erstmals für die 3D-USCT eine korrigierte Rekonstruktion mit einer Auflösung von $(0,28 \text{ mm})^3$ innerhalb einer praktikablen Zeit von 2 h durchzuführen.

Für eine klinische Anwendbarkeit der 3D-USCT, bei der eine Rekonstruktion innerhalb von 15 min möglich sein muss, wird bei acht GPUs eine Rechengeschwindigkeit von mindestens 32,3 GVA/s pro GPU benötigt. Daher wurde der Einsatz von neueren GPUs wie der GTX TITAN untersucht, die jedoch nur eine Rechengeschwindigkeit von 7,8 GVA/s mit der korrigierten SAFT erreicht. Da somit eine weitere Beschleunigung nötig ist, wurde der TOFI-Algorithmus als neuartige SAFT-Rekonstruktionsvariante für heterogene Medien entwickelt und bezüglich ihrer Bildqualität und Performanz analysiert. Sie reduziert durch eine geeignete Approximation die Rechenkomplexität der Rekonstruktion und erreicht einen signifikanten Performanzgewinn um den Faktor 7,1 auf 55,3 GVA/s. Die Bildqualität bleibt dabei nahezu identisch. Dadurch ist es erstmals möglich, die korrigierte Rekonstruktion für heterogene Medien auf einem Rekonstruktionsserver mit acht GPUs innerhalb von 9 min durchzuführen.

Die in Kapitel 2.6 beschriebenen USCT-Systeme von Delphinus und Techniscan nutzen ebenfalls eine Schallgeschwindigkeits- und Dämpfungskorrektur. Allerdings werden diese nur zweidimensional durchgeführt und besitzen somit eine geringere Rechenkomplexität. Es werden zwar Gesamtrekonstruktionszeiten genannt, die jedoch keinen direkten Vergleich der Rechengeschwindigkeiten zulassen, da nicht angegeben ist, wie viele Messsignale zur Rekonstruktion der Reflexionsbilder genutzt werden.

Der ermittelte Performanzwert von 32,3 GVA/s, für eine klinisch relevante Rekonstruktion innerhalb von 15 min, bezieht sich auf den reinen Rechenaufwand für die Rekonstruktion. Dabei wird davon ausgegangen, dass die Messdaten bereits vollständig im Speicher vorliegen und die Rekonstruktion bei einer hinreichend großen Problemgröße ausgeführt wird, so dass die GPUs in der Sättigung betrieben werden. Der Datentransfer von der DAQ-Hardware zum Rekonstruktionsserver ist aktuell durch eine 100 MBit-Netzwerkverbindung beschränkt. Da die reine Auslesezeit aktuell 47 Minuten benötigt und bereits über den geforderten 15 min liegt, konnte diese Auslesezeit in dieser Arbeit nicht mitbetrachtet werden. Hier besteht für den Aufbau der Hardware Optimierungsbedarf, um für einen klinischen Einsatz einen höheren Patientendurchsatz zu erreichen. Dies wird aktuell durch die Entwicklung eines neuen DAQ-Systems in Angriff genommen, die pro Karte eine 40 GBit-Ethernet-Netzwerkverbindung mit einschließt. Damit wäre bei zwölf geplanten Karten eine theoretische Nettodatenrate von 451,2 GBit/s möglich. Der gesamte Speicher des DAQ-Systems mit 86 GB, was den Rohdaten für 47 AP entspricht, könnte so un-

ter 2 Sekunden übertragen werden. Jedoch müssten diese auch schnell genug vom Rekonstruktionsserver entgegengenommen werden können.

Berechnung der Reflexionscharakteristik

Darüber hinaus wurde eine neuartige Idee für die SAFT-Rekonstruktion getestet und auf ihr Potential hin in der 3D-USCT evaluiert. Dabei werden die bisher vernachlässigten Richtungsinformationen, die durch die Sender- und Empfänger-Positionen gegeben sind, ausgenutzt, um pro Voxel eine Richtcharakteristik zu gewinnen. Diese Richtcharakteristik gibt Aufschluss darüber, wie stark eine Struktur innerhalb des Voxels in einzelne Richtungen streut. Durch diese neuartigen Informationen ergeben sich vielfältige Möglichkeiten der Nutzbarkeit und Auswertung. So konnte z.B. die Richtung der Oberflächennormalen von reflektierenden Strukturen bestimmt werden. Durch eine einfache Maximumdetektion konnten die reflektierenden Strukturen hervorgehoben und mit einem höheren Kontrast von bis zu 53% dargestellt werden. Weiterhin zeigen die Daten das Potential, eine Klassifizierung einzelner Voxel anhand ihrer Reflexionscharakteristiken durchzuführen.

Für die Evaluierung der Reflexionscharakteristik pro Voxel wurden aufgrund der vielfältigen Möglichkeiten bisher nur erste Untersuchungen durchgeführt. Dabei zeigte sich bereits ein großes Potential, obwohl die Reflexionscharakteristiken mit der Standard-SAFT berechnet wurden. Eine Einbettung der Berechnung der Reflexionscharakteristiken in die Rekonstruktion mit Schallgeschwindigkeits- und Dämpfungskorrektur lässt eine deutliche Verbesserung der Ergebnisse erwarten. Die korrigierte Schallgeschwindigkeit wird benötigt, um die Reflexionscharakteristik eines Voxels korrekt berechnen zu können, da die Messsignale in die korrekten Voxel zurückprojiziert werden müssen. Zum anderen kann durch die Dämpfungskorrektur die am Voxel auftretende Größe der Reflexion exakter ermittelt werden. Somit ebnet diese Methoden den Weg für die Berechnung quantitativer Reflektivitätswerte. Da die Berechnung der Daten trotz GPU-Beschleunigung zeitaufwändig ist, stellt sie sich als eine Möglichkeit dar, um weitergehende Untersuchungen nach der Rekonstruktion der Reflexionsbilder in einem zweiten Schritt für einen als interessant markierten Bereich durchzuführen. Die Reflexionscharakteristik erweitert die 3D-USCT neben der Schallgeschwindigkeit, Dämpfung und Reflektivität um eine gänzlich neue Modalität und nutzt dafür die weltweit einzigartige Stärke der dreidimensionalen USCT voll aus. Die Möglichkeiten, diese Daten zu verarbeiten, sind sehr vielfältig und eröffnen neue Forschungsgebiete. Aufgrund der Richtcharakteristiken pro Voxel ist eine Unterscheidung des Mediums möglich und wird für eine Klassifizierung des Mediums in verschiedene Gewebetypen von großer Bedeutung sein.

Ausblick

Verbesserung der Algorithmik

Durch die erzielte Beschleunigung der SAFT-Rekonstruktion rückt die Erforschung von komplexeren und aufwändigeren Algorithmen in den Vordergrund, die die Bildqualität weiter verbessern können.

Weitere Untersuchungsmethoden des Mediums ergeben sich insbesondere durch die neuartig gewonnenen Reflexionscharakteristiken, für die in dieser Arbeit bereits ein hohes Potential für weitere vielfältige Forschungsgebiete aufgezeigt werden konnte. Ein interessantes Forschungsgebiet stellt insbesondere die Klassifizierung und Segmentierung des Mediums dar. Dabei besteht die Möglichkeit, dass Tumorgewebe an seiner Reflexionscharakteristik erkannt werden kann. Dazu werden jedoch mehr Messungen benötigt, die in weiteren Studien durchgeführt werden können. Weitere Untersuchungen ergeben sich in einer Optimierung der genutzten Geometrie. Mit einer höheren Anzahl an Flächen steigt die Auflösung, jedoch sinkt dabei die Performanz. Neben Geometrien mit einer gleichmäßigen Verteilung können auch Geometrien mit einer ungleichmäßigen Verteilung wie beim Gradnetz eingesetzt werden. Ebenso ist eine Rekonstruktionsartefaktreduzierung anhand der Reflexionscharakteristik vorstellbar. Durch die Reflexionscharakteristik könnten die Informationen, im Vergleich zum 1D-Medianfilter, aus allen Richtungen genutzt werden, um eine verbesserte Artefaktreduzierung zu erreichen.

In dieser Arbeit wurde die Bildqualität der Rekonstruktion mit SAFT für heterogene Medien untersucht. Die Bildqualität ist jedoch in großem Maß von der Qualität der Messsignale abhängig. Eine verbesserte Signaldetektion und Trennbarkeit der einzelnen Reflexionssignale bei limitierter Bandbreite ist ein aktuelles Forschungsgebiet. In [89] wurde bereits gezeigt, dass durch den Einsatz von Huffman Codes beim Senden Verbesserungen erreicht werden können. Eine Verbesserung der Signaldetektion ließe sich des Weiteren durch eine Wavelet-basierte Signalverarbeitung erreichen, mit der Rauscheinflüsse deutlich verringert und die Zeitauflösung der Reflexionssignale verbessert werden können [109].

Beschleunigung des Rekonstruktionsablaufs

Eine Möglichkeit zur Beschleunigung des gesamten Rekonstruktionsablaufs ergibt sich während der Rekonstruktion auf den GPUs. Während die Rekonstruktion auf den GPUs durchgeführt wird, ist die CPU frei für weitere Aufgaben und könnte parallel zur SAFT-Rekonstruktion die Signalvorverarbeitung für den nächsten A-Scandatenblock durchführen. Diese wird aktuell von mehreren CPU-Kernen durchgeführt, wobei die Dauer der Signalvorverarbeitung in etwa der Rekonstruktionsdauer entspricht.

Aktuell wird am IPE ein neues DAQ-System auf Basis des MTCA.4-Systemstandards entwickelt und soll eine 40 GBit-Ethernet-Netzwerkverbindung enthalten. Damit ergibt sich die Möglichkeit, die Messdaten mittels eines DMA-Transfers direkt in den Speicher der GPUs zu transferieren, ohne dabei die CPU zu belasten. Dieses Verfahren wird von NVIDIA durch die Technologie „GPUDirect“ unterstützt. Diese Unterstützung gilt jedoch aktuell nur für die teureren Tesla-GPUs von NVIDIA. Weiterhin sollen in dem neuen DAQ-System zwölf Kintex-7-410 FPGAs verbaut

werden. Auf diesen dürften, von dem Ressourcenverbrauch des SAFT-Algorithmus für den Virtex-7-FPGA ausgehend, jeweils 55 SAFT-Units Platz finden [110], womit in etwa eine Rechengeschwindigkeit von 201 GVA/s für den Standard-SAFT erreicht werden kann. Damit liegt sie knapp unterhalb der Rechenleistung, von acht GTX TITAN GPUs, mit denen für den Standard-SAFT eine maximale Rechenleistung von 210 GVA/s erreicht werden kann. Da sich die Daten nach der Messung auf dem DAQ-System befinden, ist eine Rekonfiguration der FPGAs für eine anschließende Rekonstruktion eine attraktive Möglichkeit, um die Bilder direkt darstellen zu können. Zudem ist ein Virtex-7-FPGA laut [9] um den Faktor 2,2 energiesparender als vergleichbare GPUs, wodurch sich für ein laufendes System Kosten einsparen ließen. Die Rechenleistung, die sich für die korrigierten SAFT-Algorithmen auf FPGAs erreichen lässt, muss jedoch erst noch untersucht werden. Dabei lassen sich, wie für die GPU bereits gezeigt wurde, die korrigierten SAFT-Algorithmen gut in mehreren Schritten verarbeiten. Damit besteht die Möglichkeit, die Berechnungen in einer anwendungsspezifischen Pipeline für FPGAs effizient durchzuführen, um eine hohe Hardwareauslastung und einen hohen Durchsatz zu erreichen.

Literaturverzeichnis

- [1] F. Bray, J. S. Ren, E. Masuyer und J. Ferlay: Global estimates of cancer prevalence for 27 sites in the adult population in 2008. *International journal of cancer. Journal international du cancer*, 132(5):1133–45, 3/2013.
- [2] J. Michaelson, E. Halpern, A. Garland, D.B. Kopans, R. Moore, G. Weber und S. Satija: Estimates of the Sizes at Which Breast Cancers Become Detectable on Mammographic and Clinical Grounds. *Journal of Women's Imaging*, 5(1):11–19, 2003.
- [3] H. Gemmeke und N.V. Ruiter: 3D ultrasound computer tomography for medical imaging. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 580(2):1057–1065, 10/2007.
- [4] T. Hopp: *Multimodal Registration of X-Ray Mammograms with 3D Volume Datasets*. Dissertation, Karlsruher Institut für Technologie, 2012.
- [5] C. Calderon, D. Vilkomerson, R. Mezrich, K.F. Etzold, B. Kingsley und M. Haskin: Differences in the attenuation of ultrasound by normal, benign, and malignant breast tissue. *Journal of Clinical Ultrasound*, 4(4):249–254, 8/1976.
- [6] S.A. Johnson, T. Abbott, R. Bell, M. Berggren, D. Borup, D. Robinson, J. Wiskin, S. Olsen und B. Hanover: Non-Invasive Breast Tissue Characterization Using Ultrasound Speed and Attenuation. *Acoustical Imaging*, Seiten 147–154, 2007.
- [7] N.V. Ruiter, M. Zapf, R. Dapp, T. Hopp, W.A. Kaiser und H. Gemmeke: *First results of a clinical study with 3D ultrasound computer tomography*. In: *IEEE International Ultrasonics Symposium (IUS)*, Seiten 651–654, 7/2013.
- [8] S.R. Doctor, T.E. Hall und L.D. Reid: SAFT – the evolution of a signal processing technology for ultrasonic testing. *NDT International*, 19(3):163–167, 6/1986.
- [9] M. Birk: *Effiziente Datenverarbeitung auf heterogenen Rechnerarchitekturen für die 3D-Ultraschall-Computertomographie*. Dissertation, Karlsruher Institut für Technologie, 2014.
- [10] E. Kretzek, M. Zapf, M. Birk, H. Gemmeke und N.V. Ruiter: *GPU based acceleration of 3D USCT image reconstruction with efficient integration into MATLAB*. In: *Proceedings of SPIE*, Seite 86750O, 3/2013.
- [11] E. Kretzek, R. Dapp, M. Zapf, M. Birk und N.V. Ruiter: *Evaluation of phase aberration correction for a 3D USCT using a ray trace based simulation*. In: *IEEE International Ultrasonics Symposium (IUS)*, Seiten 378–381, 7/2013.

- [12] E. Kretzek und N.V. Ruiter: *GPU based 3D SAFT reconstruction including phase aberration*. In: *Proceedings of SPIE*, Band 9040, Seiten 90400W–90400W–11, 3/2014.
- [13] E. Kretzek, T. Hopp und N.V. Ruiter: *GPU-based 3D SAFT reconstruction including attenuation correction*. In: *SPIE Medical Imaging*, Seite 94190E, 3/2015.
- [14] E. Kretzek, P. Hucker, M. Zapf und N.V. Ruiter: *Evaluation of directional reflectivity characteristics as new modality for 3D Ultrasound Computer Tomography*. In: *Ultrasonics Symposium (IUS), 2015 IEEE International*, Seiten 1–4, Oct 2015.
- [15] B. Povh: *Anschauliche Physik: für Naturwissenschaftler*. Springer, Berlin, Heidelberg, 2011.
- [16] U. S. Albert *et al.*: Kurzfassung der aktualisierten Stufe-3-Leitlinie Brustkrebs-Früherkennung in Deutschland 2008. *Geburtshilfe und Frauenheilkunde*, 68(3):251–261, 3/2008.
- [17] *Brustkrebs: Behandlungsverfahren - eine Übersicht*. <https://www.krebsinformationsdienst.de/tumorarten/brustkrebs/behandlung-uebersicht.php>, besucht: 2015-06-10.
- [18] J. Ferlay, I. Soerjomataram, M. Ervik, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D.M. Parkin, D. Forman und F. Bray: *GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11*, 2013. <http://globocan.iarc.fr>, besucht: 2015-06-10.
- [19] M. Malvezzi, P. Bertuccio, T. Rosso, M. Rota, F. Levi, C. La Vecchia und E. Negri: European cancer mortality predictions for the year 2015: does lung cancer have the highest death rate in EU women? *Annals of Oncology*, 2015.
- [20] R. Sivaramakrishna und R. Gordon: Detection of breast cancer at a smaller size can reduce the likelihood of metastatic spread: a quantitative analysis. *Academic radiology*, 4(1):8–12, 1/1997.
- [21] H. Friedmann: *Einführung in die Kernphysik*. Wiley, 2015.
- [22] D. von Fournier, H. W. Anton, H. Junkermann und G. Bastert: *Breast Cancer Screening*. In: *Cancer Diagnosis*, Seiten 78–87. Springer Berlin Heidelberg, 1992.
- [23] *Häufige Fragen zur Magnetresonanztomographie (MRT)*. <https://www.krebsinformationsdienst.de/untersuchung/kernspintomographie-faq.php>, besucht: 2015-06-08.
- [24] K. Fujii, T. Aoyama, C. Yamauchi-Kawaura, S. Koyama, M. Yamauchi, S. Ko, K. Akahane und K. Nishizawa: Radiation dose evaluation in 64-slice CT examinations with adult and paediatric anthropomorphic phantoms. *The British Journal of Radiology*, 82(984):1010–8, 12/2009.
- [25] N. Fliege: *Systemtheorie*. Teubner, Stuttgart, 1991.

- [26] P. Suetens: *Fundamentals of Medical Imaging*. Cambridge medicine. Cambridge University Press, 2. Auflage, 2009.
- [27] OpenStax College: *Limits of Resolution: The Rayleigh Criterion*. http://cnx.org/contents/f403618a-c8e1-4b20-8737-6164df697aea@5/Limits_of_Resolution:_The_Rayl, besucht: 2015-03-18.
- [28] Wikipedia: *Halbwertsbreite*. <http://de.wikipedia.org/wiki/Halbwertsbreite>, besucht: 2015-03-18.
- [29] E. Hecht: *Optik*. Oldenbourg, München, 5. verbesserte Auflage, 2009.
- [30] O. Dössel: *Bildgebende Verfahren in der Medizin: von der Technik zur medizinischen Anwendung*. Springer, Berlin, 2000.
- [31] Y. Erez, Y. Schechner und D. Adam: *Ultrasound Image Denoising by Spatially Varying Frequency Compounding*. In: *Pattern Recognition*, Band 4174 der Reihe *Lecture Notes in Computer Science*, Seiten 1–10. Springer Berlin Heidelberg, 2006.
- [32] R.R. Entekin, B.A. Porter, H.H. Sillesen, A.D. Wong, P.L. Cooperberg und C.H. Fix: Real-time spatial compound imaging: application to breast, vascular, and musculoskeletal ultrasound. *Seminars in ultrasound, CT, and MR*, 22(1):50–64, 2/2001.
- [33] Asbjørn Støylen: *Basic ultrasound, echocardiography and Doppler ultrasound*, 2015. <http://folk.ntnu.no/stoylen/strainrate/Ultrasound/>, besucht: 2015-03-04.
- [34] S. Schmidt, N. Duric, C. Li, O. Roy und Z. F. Huang: Modification of Kirchhoff migration with variable sound speed and attenuation for acoustic imaging of media and application to tomographic imaging of the breast. *Medical Physics*, 38(2):998–1007, 2011.
- [35] N. Webster: Abdominal Ultrasonography: Practice makes perfect sense. *FOCUS Ultrasonography*, Seiten 14–18, 2008.
- [36] F. Simonetti, L. Huang und N. Duric: On the spatial sampling of wave fields with circular ring apertures. *Journal of Applied Physics*, 101(8):083103, 2007.
- [37] N.V. Ruiter, R. Dapp, M. Zapf und H. Gemmeke: *A new method for grating lobe reduction for 3D synthetic aperture imaging with ultrasound computer tomography*. In: *IEEE International Ultrasonics Symposium (IUS)*, Seiten 2372–2375, 2010.
- [38] G.F. Schwarzenberg: *Untersuchung der Abbildungseigenschaften eines 3D-Ultraschall-Computertomographen zur Berechnung der 3D-Abbildungsfunktion und Herleitung einer optimierten Sensorgeometrie*. Dissertation, Karlsruher Institut für Technologie, 2009.
- [39] NVIDIA Corporation: *CUDA C Programming Guide v7.0*, 2015. http://docs.nvidia.com/cuda/pdf/CUDA_C_Programming_Guide.pdf, besucht: 2015-05-06.

- [40] NVIDIA Corporation: *NVIDIAs Next Generation CUDA Compute Architecture: Kepler GK110*. Technischer Bericht, 2012. <http://www.nvidia.de/content/PDF/kepler/NVIDIA-Kepler-GK110-Architecture-Whitepaper.pdf>.
- [41] M.J. Flynn: Some Computer Organizations and Their Effectiveness. *IEEE Transactions on Computers*, C-21(9):948–960, 9/1972.
- [42] *TUNING CUDA APPLICATIONS FOR KEPLER*. Technischer Bericht, NVIDIA Corporation, 2015. http://docs.nvidia.com/cuda/pdf/Kepler_Tuning_Guide.pdf.
- [43] *NVIDIAs Next Generation CUDA Compute Architecture Fermi*. Technischer Bericht, NVIDIA Corporation, 2009.
- [44] F. Simonetti und L. Huang: Synthetic aperture diffraction tomography for three-dimensional imaging. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 465(2109):2877–2895, 7/2009.
- [45] C. Hansen *et al.*: *Ultrasound breast imaging using Full Angle Spatial Compounding: In-vivo results*. In: *IEEE International Ultrasonics Symposium (IUS)*, Seiten 54–57. IEEE, 11/2008.
- [46] *Embryo1*. <http://www.gynpraxis-wagner.de>, besucht: 13.06.2015.
- [47] *Embryo2*. <http://www.praenatal-central.de>, besucht: 13.06.2015.
- [48] M. Pharr und G. Humphreys: *Physically Based Rendering: From Theory to Implementation*. Morgan Kaufmann, 2010.
- [49] Delphinus: *SOFTVUE SYSTEM*, 2015. <http://www.delphinusmt.com/our-technology/softvue-system>, besucht: 14.06.2015.
- [50] N. Duric, P. Littrup, S. Schmidt, C. Li, O. Roy, L. Bey-Knight, R. Janer, D. Kunz, X. Chen, J. Goll, A. Wallen, F. Zafar, V. Allada, E. West, I. Jovanovic, K. Li und W. Greenway: *Breast imaging with the SoftVue imaging system: first results*. In: *Proceedings of SPIE*, Band 8675, Seiten 86750K–86750K–8, 3/2013.
- [51] O. Roy, S. Schmidt, C. Li, V. Allada, E. West, D. Kunz und N. Duric: *Breast imaging using ultrasound tomography: From clinical requirements to system design*. In: *IEEE International Ultrasonics Symposium (IUS)*, Seiten 1174–1177. IEEE, 7/2013.
- [52] P. Pellegretti, M. Vicari, M. Zani, M. Weigel, D. Borup, J. Wiskin, U. Saueressig, E. Kotter und M. Langer: *A clinical experience of a prototype automated breast ultrasound system combining transmission and reflection 3D imaging*. In: *IEEE International Ultrasonics Symposium (IUS)*, Seiten 1407–1410. IEEE, 10/2011.
- [53] M. André, J. Wiskin, D. Borup, S. Johnson, H. Ojeda-Fournier und L. Olson: Quantitative volumetric breast imaging with 3D inverse scatter computed tomography. *Conference proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, 2012:1110–3, 1/2012.

- [54] *IEEE Standard Letter Designations for Radar-Frequency Bands*, 2003.
- [55] D.C. Giancoli: *Physik: Lehr- und Übungsbuch*. Pearson Deutschland GmbH, München, 3. erweiterte Auflage, 2010.
- [56] D. McGowen und R. Morris: Choosing Side Scan Sonar Frequencies. *Sea Technology*, 54(4):44, 2013.
- [57] M.P. Hayes und P.T. Gough: Synthetic Aperture Sonar: A Review of Current Status. *IEEE Journal of Oceanic Engineering*, 34(3):207–224, 7/2009.
- [58] F. Baralli, M. Couillard, J. Ortiz und D.G. Caldwell: *GPU-based real-time synthetic aperture sonar processing on-board autonomous underwater vehicles*. In: *MTS/IEEE OCEANS - Bergen*, Seiten 1–8. IEEE, 6/2013.
- [59] S. Norton und M. Linzer: Ultrasonic reflectivity imaging in three dimensions: Reconstruction with spherical transducer arrays. *Ultrasonic Imaging*, 1(3):210–231, 7/1979.
- [60] N.V. Ruiter, G. Göbel, L. Berger, M. Zapf und H. Gemmeke: *Realization of an optimized 3D USCT*. In: *Proceedings of SPIE*, Band 7968, Seiten 796805–796805–8, 3/2011.
- [61] H. Gemmeke, L. Berger, M. Birk, G. Gobel, A. Menshikov, D. Tcherniakhovski, M. Zapf und N.V. Ruiter: *Hardware setup for the next generation of 3D Ultrasound Computer Tomography*. In: *IEEE Nuclear Science Symposium & Medical Imaging Conference*, Seiten 2449–2454, 10/2010.
- [62] P.F. León und H. Jäkel: *Signale und Systeme*. De Gruyter Oldenbourg, Berlin, 6. überarbeitete Auflage, 2015.
- [63] N.V. Ruiter, G.F. Schwarzenberg, M. Zapf und H. Gemmeke: *Improvement of 3D ultrasound computer tomography images by signal pre-processing*. In: *IEEE International Ultrasonics Symposium (IUS)*, Seiten 852–855, 11/2008.
- [64] D.G. Manolakis, V.K. Ingle und S.M. Kogon: *Statistical and adaptive signal processing: spectral estimation, signal modeling, adaptive filtering, and array processing*, Band 46. Artech House Norwood, 2005.
- [65] S.J. Norton und M. Linzer: Ultrasonic reflectivity tomography: Reconstruction with circular transducer arrays. *Ultrasonic Imaging*, 1(2):154–184, 4/1979.
- [66] S.Norton und M. Linzer: Ultrasonic reflectivity imaging in three dimensions: Reconstruction with spherical transducer arrays. *Ultrasonic Imaging*, 1(3):210–231, 7/1979.
- [67] O. Dössel *et al.*: *Biomedizinische Technik - Medizinische Bildgebung*. De Gruyter, Berlin, Biomedizin Auflage, 2014.
- [68] N. Gebhardt: *Einige BRDF Modelle*. <http://www.irrlicht3d.org/papers/BrdfModelle.pdf>.
- [69] J.H. Lambert und E. Anding: *Ostwalds Klassiker der exakten Wissenschaften: Lamberts Photometried (1760)*. Engelmann, W., Leipzig, 1892.

- [70] H.A. Stuart und G. Klages: *Kurzes Lehrbuch der Physik*. Springer, Berlin, 18. aktualisierte Auflage, 2006.
- [71] M. Hardt: *Distributed Simulations for 3D Ultrasound Computer Tomography*. Dissertation, Karlsruher Institut für Technologie, 2012.
- [72] S.J. Norton und M. Linzer: Ultrasonic reflectivity tomography: Reconstruction with circular transducer arrays. *Ultrasonic Imaging*, 1(2):154–184, 4/1979.
- [73] N.V. Ruiter, M. Zapf, T. Hopp und H. Gemmeke: *Experimental evaluation of noise generated by grating lobes for a sparse 3D ultrasound computer tomography system*. In: *Proceedings of SPIE*, Seiten 86750N–86750N–8, 3/2013.
- [74] S.A.K. Ubaidullah: Accelerating MATLAB slow loop execution with CUDA. *7th International Conference on Engineering and Technology (ICET)*, Seiten 1–4, 2011.
- [75] J. Huang, W. Yuqiang und X. Xishen: A Research of Mixed Programming between VC++ and Matlab. *Computer Development and Applications*, 2008(10):38–41, 2008.
- [76] M. Fatica: Accelerating MATLAB with CUDA. *Proceedings of the Eleventh Annual High Performance Embedded Computing Workshop Lexington Massachusetts*, Seiten 2–3, 2007.
- [77] V. Simek und R.R. Asn: *GPU Acceleration of 2D-DWT Image Compression in MATLAB with CUDA*. In: *Second UKSIM European Symposium on Computer Modeling and Simulation*, Seiten 274–277. IEEE, 9/2008.
- [78] MathWorks: *Product Documentation - Introducing MEX-Files*, 2015. http://de.mathworks.com/help/matlab/matlab_external/introducing-mex-files.html, besucht: 2015-02-20.
- [79] One Stop Systems: *2U Expansion System*. http://www.onestopsystems.com/documents/OSS-PCIE-2U-ENCL-EXP-x-x_001.pdf.
- [80] N.V. Ruiter, M. Zapf, T. Hopp, R. Dapp und Hartmut H. Gemmeke: *Phantom image results of an optimized full 3D USCT*. In: *Proceedings of SPIE*, Band 6, Seiten 832005–832005–6, 2/2012.
- [81] F. Anderson und F. Morgan: Active Imaging Analysis via Ellipsoidal Projections. *Acoustical Imaging*, 21:171–182, 1995.
- [82] F.L. Anderson: *3D ellipsoidal backprojection images from large arrays (II)*. In: *IEEE International Ultrasonics Symposium (IUS)*, Seiten 1223–1226, 1992.
- [83] F.L. Anderson: *Three-dimensional real-time ultrasonic imaging using ellipsoidal backprojection*. In: *Medical Imaging*, Seiten 62–80, 7/1991.
- [84] G.F. Schwarzenberg, R. Stotzka, M. Zapf, H. Gemmeke und N.V. Ruiter: *Resolution assessment of a 3D ultrasound computer tomograph using ellipsoidal backprojection*. In: *IEEE International Ultrasonics Symposium (IUS)*, Band 1, Seiten 1979–1982, 2006.

- [85] Q. Huynh-Thu und M. Ghanbari: Scope of validity of PSNR in image/video quality assessment. *Electronics Letters*, 44(13):800, 2008.
- [86] T. Stathaki: *Image fusion: Algorithms and applications*. Elsevier, Amsterdam, 1. Auflage, 2008.
- [87] Z.L. Zhang, S.H. Sun und F.C. Zheng: Image fusion based on median filters and SOFM neural networks:. *Signal Processing*, 81(6):1325–1330, 6/2001.
- [88] G.R. Sinha und B.C. Patel: *Medical Image Processing*. PHI Learning, 2014.
- [89] S. Gupta, M. Zapf, H. Krauss und N.V. Ruitter: *Design of huffman sequences with limited bandwidth*. In: *IEEE International Ultrasonics Symposium (IUS)*, Seiten 1089–1092. IEEE, 9/2014.
- [90] R. Hofmann, A. Hegele und A. Honacker: *Ultraschall in der Urologie*, Band 1. 2012.
- [91] N.V. Ruitter, R. Schnell, M. Zapf und H. Gemmeke: *P3D-2 Phase Aberration Correction for 3D Ultrasound Computer Tomography Images*. In: *IEEE International Ultrasonics Symposium (IUS)*, Seiten 1808–1811, 10/2007.
- [92] R. Dapp: *Abbildungsmethoden für die Brust mit einem 3D-Ultraschall-Computertomographen*. Dissertation, Karlsruher Institut für Technologie, 2013.
- [93] J.E. Bresenham: Algorithm for computer control of a digital plotter. *IBM Systems Journal*, 4(1):25–30, 1965.
- [94] B.A.J. Angelsen: *Ultrasound imaging: waves, signals, and signal processing. 2. Propagation and scattering in heterogeneous, nonlinear tissue with contrast agent imaging and Doppler measurements*. Ultrasound Imaging: Waves, Signals, and Signal Processing. Emantec, 2000.
- [95] F.A. Duck: *Physical properties of tissue: a comprehensive reference book*. Academic Pr., London, 1990.
- [96] J.T. Bushberg, J.A. Seibert, E.M. Leidholdt, J.M. Boone und M. Mahesh: *The Essential Physics of Medical Imaging, Third Edition.*, Band 40. American Association of Physicists in Medicine, 6/2013.
- [97] R. Dapp, H. Gemmeke und N.V. Ruitter: *Attenuation Reconstruction for 3D Ultrasound Computer Tomography*. In: *19th International Conference on Systems, Signals and Image Processing (IWSSIP)*, Seiten 484–487, 4/2012.
- [98] M. Pharr und R. Fernando: *GPU gems 2: programming techniques for high-performance graphics and general-purpose computation*. Addison-Wesley, 1. publ. Auflage, 2005.
- [99] J. Shen, H. Wang, C. Cain und E.S. Ebbini: A post-beamforming processing technique for enhancing conventional pulse-ultrasound imaging contrast resolution. *1995 IEEE Ultrasonics Symposium. Proceedings. An International Symposium*, 2:1319–1322, 1995.

- [100] C. Jülg: *Evaluation einer Migration von GPU-beschleunigten USCT Rekonstruktionsalgorithmen nach OpenCL anhand einer Grating-Lobe-Reduktion*. Studienarbeit, Karlsruher Institut für Technologie (KIT), 2015.
- [101] M. Kachelriess: *Branchless vectorized median filtering*. In: *IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*, Seiten 4099–4105, 10/2009.
- [102] P. Hucker: *Evaluation einer neuen Modalität für die 3D-Ultraschall-Computertomographie mit Hilfe eines Rückstreumodells*. Masterthesis, Karlsruher Institut für Technologie (KIT), 2014.
- [103] B.T. Phong: *Illumination for computer generated pictures*, 6/1975.
- [104] M. Oleg: *Licht und Schatten*. Technischer Bericht, Institut für Visualisierung und Interaktive Systeme, Stuttgart, 2010.
- [105] R. Ziegler: *Platonische Körper: Verwandtschaften, Metamorphosen, Umstülpungen*. Verlag am Goetheanum, 4. Auflage, 2008.
- [106] *Archimedische Körper*. <http://www.math.uni-augsburg.de/~bernt/Archimedes/vorschau.html>, besucht: 16.06.2015.
- [107] E.W. Weisstein: *Sphere Point Picking*. <http://mathworld.wolfram.com/SpherePointPicking.html>, besucht: 2015-06-16.
- [108] G.A.F. Seber und C.J. Wild: *Nonlinear Regression*. Wiley Series in Probability and Statistics. Wiley, 2003.
- [109] L. Rong: *Data Compression in Ultrasound Computed Tomography*. Dissertation, Karlsruher Institut für Technologie, 2011.
- [110] M. Birk, E. Kretzek, P. Figuli, M. Weber, J. Becker und N. Ruiter: High-Speed Medical Imaging in 3D Ultrasound Computer Tomography. *IEEE Transactions on Parallel and Distributed Systems*, PP(99):1–1, 2015.