SEBASTIAN SCHÄFER

COMPUTER-ASSISTED MOTION COMPENSATION AND
ANALYSIS OF PERFUSION ULTRASOUND DATA

COMPUTER-ASSISTED MOTION COMPENSATION AND ANALYSIS OF PERFUSION ULTRASOUND DATA

Dissertation
zur Erlangung des akademischen Grades
Doktoringenieur (Dr.-Ing.)

angenommen durch die Fakultät für Informatik
der Otto-von-Guericke-Universität Magdeburg

von DIPL.-ING. SEBASTIAN SCHÄFER
geb. am 18.08.1982 in Kassel

Gutachter
Prof. Dr. Klaus D. Tönnies
Prof. Dr. Thomas M. Deserno
Prof. Dr. Johan G. Bosch

Magdeburg, den 29.01.2014

ABSTRACT

Contrast-enhanced ultrasound (CEUS) is a medical imaging technique used to reveal tissue perfusion processes in the human body. The properties of perfusion such as the amount of blood within a tissue region or the blood flow velocity provide important information about various diseases and can be used for diagnosis and treatment planning. CEUS is radiation-free and inexpensive compared to other perfusion imaging techniques. However, the CEUS image acquisition suffers from several quality restrictions. Among these are the low signal-to-noise ratio of the measurements and motion influences caused by patient and probe movements. Both influence the validity of derived perfusion observations.

In this thesis, a systematic approach for perfusion analysis is presented. It contains different preprocessing steps before the analysis is performed. As the most important aspect of preprocessing, motion compensation is applied to a sequence of CEUS frames. It uses Markov random field based image registration taking the CEUS specific restrictions into account. Automatically generated spatial regions are used to derive averaged measurements of perfusion. The approach ensures that only areas of homogeneous perfusion are combined to preserve important characteristics. The system establishes a standardized workflow reducing the amount of time needed by the physician to prepare the data. Furthermore, the motion compensation does not require user interaction.

Multiple experiments using a total number of 40 clinical CEUS datasets of the bowel and the pancreas have been conducted to assess the performance of the system. The proposed approach was able to reduce the overall processing time by 41.0 % in comparison to manual processing. The quality of the motion compensation result was measured using a pharmacokinetic model as physiological reference for observed perfusion. On average, the presented system was able to improve the perfusion measurements in terms of the similarity to the pharmacokinetic model by 36.8 % for pancreas datasets and by 7.0 % for bowel datasets. Additionally, it achieved similar results for 3 out of 4 different patient cohorts compared to manual processing.

ZUSAMMENFASSUNG

Ultraschall mit Kontrastverstärkung (CEUS) ist ein bildgebendes Verfahren in der Medizin, um Gewebeperfusion im menschlichen Körper sichtbar zu machen. Die Eigenschaften von Perfusion, wie beispielsweise die Menge an Blut innerhalb einer bestimmten Region oder die Geschwindigkeit des Blutflusses geben Aufschluss über Krankheiten und deren Verlauf und helfen, die Behandlung abzustimmen. CEUS ist strahlungsfrei und verhältnismäßig günstig im Vergleich zu anderen Bildgebungsverfahren für Perfusion. Allerdings unterliegt die CEUS-Bildgebung unterschiedlichen Störeinflüssen, insbesondere einem niedrigen Signal-Rausch-Verhältnis und Bewegungseinflüssen ausgelöst durch den Patienten oder Lageveränderungen des Ultraschallkopfes. Diese Einflüsse verfälschen die Messungen der Perfusion in den Daten.

In dieser Dissertation wird ein systematischer Ansatz zur Analyse von Perfusion in CEUS-Daten vorgestellt. Als zentraler Aspekt der Vorverarbeitung werden die CEUS-Daten einer Bewegungskorrektur unterzogen. Dazu wird ein Bildregistrierungsverfahren verwendet, welches auf einem Markov Random Field basiert und die speziellen Restriktionen bezüglich der Bewegung in CEUS-Daten berücksichtigt. Automatisch generierte Regionen werden verwendet, um darin enthaltene Werte zur Perfusionsanalyse zu mitteln. Dabei stellt das Verfahren sicher, dass ausschließlich Bereiche homogener Perfusion zusammengefasst werden, damit wichtige Charakteristiken in den Daten erhalten bleiben. Das System legt einen standardisierten Arbeitsablauf fest, der die Bearbeitungszeit der Daten durch den Arzt verkürzt. Zudem arbeitet die Bewegungskorrektur ohne Benutzerinteraktion.

Zur Beurteilung der Leistung des Systems wurden unterschiedliche Experimente durchgeführt. Dazu standen 40 klinische CEUS-Datensätze der Darmwand sowie der Bauchspeicheldrüse zur Verfügung. Die Bearbeitungszeit im Vergleich zur manuellen Bearbeitung der Datensätze wurde um 41 % reduziert. Die Qualität des Verfahrens zur Bewegungskorrektur wurde mithilfe eines pharmakokinetischen Modells gemessen. Dieses Modell dient als physiologische Referenz für die in den Daten gemessene Perfusion. Die automatische Bewegungskorrektur verbesserte die Perfusionsmessungen in Bezug auf die Übereinstimmung mit dem Modell im Durchschnitt um 36.8 % für die Daten der Bauchspeicheldrüse und um 7.0 % für die Daten der Darmwand. Außerdem erzielte das Sys-

tem vergleichbare Resultate im Vergleich zu der manuellen Bearbeitung durch einen Experten für 3 von 4 Patientenkohorten.

PUBLICATIONS

Some ideas and figures have appeared previously in the following publications:

Schäfer, S.; Preim, U.; Glaßer S.; Preim, B. and Tönnies, K.: Local similarity measures for lesion registration in DCE-MRI of the breast. *Annals of the BMVA*, vol. 2011, no. 3, pp. 1–13, 2011.

Schäfer, S.; Angelelli, P.; Nylund, K.; Gilja, O. H. and Tönnies, K.: Registration of ultrasonography sequences based on temporal regions. In *7th Intl. Symp. on Image and Signal Processing and Analysis*, pages 749–759, Dubrovnik, 2011.

Schäfer, S.; Nylund, K.; Gilja, O. H. and Tönnies, K.: Motion compensation of ultrasonic perfusion images. In *Proceedings of SPIE Medical Imaging: Ultrasonic Imaging, Tomography, and Therapy*, volume 8320, San Diego, 2012.

Schäfer, S. and Tönnies, K.: Registration of temporal ultrasonic image sequences using Markov random fields. In *Vision Modelling and Visualization*, pages 151–158, Magdeburg, 2012.

Schäfer, S.: Motion compensation of ultrasonic perfusion images using MRFs and coupled segmentation. In *1. Doktorandentagung Magdeburger-Informatik-Tage*, pages 59–65, Magdeburg, 2012.

DANKSAGUNG

Die Energie und die Motivation, die nötig waren, um diese Arbeit zu erstellen, verdanke ich vielen Menschen. Dabei habe ich Unterstützung auf unterschiedlichste Weise erfahren:

Ich bedanke mich bei Klaus Tönnies für die sehr gute Betreuung, insbesondere für die vielen kreativen Inputs, die mich immer wieder neu motiviert haben. Auch der AG-BV möchte ich dafür danken, dass sie immer eine Idee für mich parat hatte, wenn ich eine gebraucht habe.

Ein Dank geht an die Kooperationspartner: Kim Nylund, Odd-Helge Gilja, Fredrik Sævik, Trond Engjom, Radovan Jiřík, Martin Mézl, Georg Dimcevski, Paolo Angelelli, Uta Preim, Bernhard Preim, Sylvia Glaßer und Steffen Serowy für die Bereitstellung von Daten, Software und die vielen fruchtbaren Diskussionen. An dieser Stelle möchte ich Kim besonders erwähnen, der unglaublich viel Zeit und Kraft in die medizinischen Experimente gesteckt hat und mir auch die Besuche im Uni-Klinikum in Bergen ermöglicht hat.

Dann möchte ich den Korrekturleserinnen und -lesern herzlich für ihr kritisches und ehrliches Feedback danken. Das sind: Clemens, Victoria, Stephen, Charlotte, André, Katja, Antje, Ursula und Steffen. Clemens hat nicht nur bis zur letzten Minute die Dissertation Korrektur gelesen, sondern auch alle meine Veröffentlichungen, die in den letzten Jahren so entstanden sind. Victoria hat die gesamte Arbeit gelesen und tapfer mit konstanter Akribie Fehler gefunden und elegante Formulierungen erarbeitet. Stephen hat jede Formel dieser Arbeit intensiv studiert und Verbesserungsvorschläge gemacht. Außerdem hat er mich im Vorfeld beraten, wie die Sachverhalte am besten zu formalisieren sind. Ich betrachte den Umfang eurer Hilfe nicht als selbstverständlich. Ich möchte euch drei besonderen Dank aussprechen.

Für die letzten 5 Jahre war das ISG quasi eines meiner Zuhause. Ich habe mich hier sehr wohl gefühlt, viele nette Leute kennengelernt und bei jedem Problem, dass ich hatte, sofort Hilfe bekommen. Für die Unterstützung möchte ich mich besonders beim Sekretariat und bei der ISG-Technik bedanken. Ob ich ein solches Arbeitsumfeld noch einmal irgendwo vorfinden werde, wage ich zu bezweifeln.

Meine lieben Freunde, vielen Dank für die schöne Zeit in Magdeburg. Ihr wart immer für mich da, auch wenn ich mal Hilfe oder

CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

AIF	Arterial input function
B-mode	Brightness modulation
CA	Contrast agent
CEUS	Contrast-enhanced ultrasound
CT	Computed tomography
FN	False negative
FP	False positive
FPS	Frames per second
ICM	Iterated conditional modes
KNN	K-nearest neighbors classification
MAD	Mean average difference
MI	Mutual information
MR	Magnetic resonance
MRF	Markov random field
MRI	Magnetic resonance imaging
MSD	Mean squared distance
MTT	Mean transit time
NCC	Normalized cross correlation
NMI	Nuclear medicine imaging
PCA	Principal component analysis
PET	Positron emission tomography
RMS	Root mean squared
ROI	Region of interest
SD	Standard deviation
SNR	Signal-to-noise ratio
SSD	Sum of squared distances
SVM	Support vector machine
TIC	Time intensity curve
TN	True negative
TP	True positive
US	Ultrasound

LIST OF SYMBOLS

Δ_x	The translation in x-direction
Δ_y	The translation in y-direction
η	A parameter to weight the prior energy relative to the data energy in the MRF
γ	A parameter to weight the temporal edge energies relative to the spatial edge energies in the MRF
$\mathcal{C}$	A set containing all edges which are selected in the graph cut
$\mathcal{E}$	The set of edges in an undirected graph $\mathcal{G}$
$\mathcal{E}_s$	The set of all edges in a graph for graph cut processing except the terminal edges
$\mathcal{E}_t$	The set of all terminal edges in a graph for graph cut processing
$\mathcal{G}$	An undirected graph
$\mathcal{L}$	The set of labels to be assigned to vertices in $\mathcal{V}$
$\mathcal{N}$	The neighborhood system of the MRF derived from $\mathcal{E}$ such that $\mathcal{N}_p$ represents all neighbors of p
$\mathcal{R}$	A set of locations from $\{1, ..., x_{max}\} \times \{1, ..., y_{max}\}$ within the image representing a region of interest
$\mathcal{S}$	Segmentation result containing different components C with each component representing a set of frames
$\mathcal{V}$	The set of vertices in an undirected graph $\mathcal{G}$
$\mathcal{W}$	The codomain of the image function I representing the set of possible image intensities
$\mathscr{C}$	The set of cliques in the undirected graph of the MRF whereas cliques of size one ($\mathscr{C}_1$) and of size two ($\mathscr{C}_2$) are used
Ω	The domain of the image function I containing three dimensions: X, Y, T
C	A component consisting of a subset of frames in a dataset

D_p	The disagreement measure at a vertex p in the MRF (for registration, mostly dissimilarity is used)
$E(l_1, l_2)$	The prior energy for two labels l_1 and l_2
E_{data}	The singleton energy of all vertices in an MRF
E_{global}	The overall energy in an MRF
E_{prior}	The doubleton energy of all vertices in an MRF
F	A random function mapping vertices from $\mathcal{V}$ to labels from $\mathcal{L}$
f	A specific label configuration such that f_p yields the current label of p
$f^{\mathcal{C}}$	The label configuration corresponding to a graph cut $\mathcal{C}$
I	The image function mapping values from $\Omega \to \mathcal{W}$
k1, k2	Parameters for semi-automatic frame selection steps 1 and 2 to control the size of the components
l	A label from $\mathcal{L}$
l_p	The label currently assigned to vertex $p \in \mathcal{V}$
M_{sim}	The similarity matrix representing all possible co-similarities of frames in a sequence
p	A vertex in $\mathcal{V}$
P(f)	The likelihood of a label configuration f, i. e., the global energy in the MRF for f
T	The temporal image dimension
t	A value in temporal image dimension T
t_{max}	The number of temporal frames in an image I
t_{ref}	The reference time point (frame) for a frame set
V_{c1}	The energy potential of a clique $c1 \in \mathscr{C}_1$ which is equivalent to the E_{data} term
V_{c2}	The energy potential of a clique $c2 \in \mathscr{C}_2$ which is equivalent to the E_{prior} term
$w(e)$	The weight assigned to an edge e in an undirected graph
X	The horizontal spatial image dimension

x	A value in spatial image dimension X (horizontal direction)
x_{max}	The size of an image I in horizontal spatial image dimension X
Y	The vertical spatial image dimension
y	A value in spatial image dimension Y (vertical direction)
y_{max}	The size of an image I in vertical spatial image dimension Y

INTRODUCTION

For over a century, medical images have gradually gained importance in health care [LeVine, 2010]. In medical research and clinical practice today, many different techniques are used to acquire image data of the human body [Hendee and Ritenour, 2003]. The images are analyzed by physicians and medical experts to obtain an understanding of the structure and the function of observed organs or tissue for disease diagnosis and treatment planning. Medical image processing is an important part of it as it helps with the computational interpretation of medical images by process automation, standardization or the derivation of new measurements [Dougherty, 2011].

The amount of data which is currently produced in clinical routine, science and technology requires the automatic and computer-assisted management and also the assistance in the analysis of this kind of data [Bankman, 2000]. Another objective of computer assistance is to save time as the data need to be prepared for diagnosis in a way that they comprise new information which is quantitative, objective and reproducible [Sonka and Fitzpatrick, 2000]. In this way physicians can spend most of the time for the diagnosis instead of addressing preprocessing and measuring tasks which may be time consuming and prone to intra- and inter-operator variability.

Medical image processing is challenging as the data acquired by medical imaging techniques may be subject to acquisition related deficiencies. Most frequent disturbances are noise, attenuation and motion [Birkfellner, 2010]. Image quality has to be improved in order to extract meaningful parameters, such as the size of a pathology or the amount of contrast uptake. Physicians have been trained for the analysis of medical image data and they also have gained experience over their time of practice. They are able to cope with poor image quality and deficiencies in the data.

The computer has to be able to efficiently and reliably interpret the data for the correct calculation of measures and for the derivation of characteristics in order to assist physicians in clinical routine. Quality improvement of data is inevitable for subsequent image analysis tasks. Also, standards have to be developed and represented by the system to make measurements comparable over different observations.

Contrast-enhanced ultrasound (CEUS) imaging allows for the observation of perfusion such as the blood flow over time. It is becoming increasingly popular because it is inexpensive, radiation free and

highly portable [Bridal and Correas, 2003; Merritt, 1989]. However, there are few approaches dealing with an automated and standardized preprocessing and analysis of CEUS data.

In this thesis, a system to process CEUS image sequences is presented. It includes the compensation of motion artifacts to establish valid correspondences of measurements over time and a method for the analysis of perfusion in the data. The compensation of motion artifacts which is performed by image registration takes the specific demands of CEUS imaging into account and introduces suitable motion constraints. The proposed analysis method exploits the characteristics of perfusion to group areas and establish analysis regions which are suggested for diagnosis. The system and its different parts are tested on 40 clinical CEUS datasets from the bowel and the pancreas.

Medical image processing comprises different methodological areas [Deserno, 2011]. This starts with the formation and management of medical image data and continues with the enhancement, analysis and visualization of acquired data (see Figure 1.1). The thesis makes use of different areas of medical image processing and covers the domain of image enhancement and image analysis whereas the first is used as preprocessing of the latter. The workflow in the image processing pipeline is marked in Figure 1.1.

1.1 MEDICAL IMAGE ACQUISITION

Medical image acquisition is used to give detailed information about structure, function and pathology of the human body [Bankman, 2000]. Depending on the exact purpose of the task different modalities are applied based on physical techniques such as X-ray computed tomography (CT), ultrasound (US), magnetic resonance (MR) and radioactivity in nuclear medicine imaging (NMI) [Iniewski, 2009].

In dynamic medical imaging multiple images are acquired at different times [Vannier, 2000]. This can be done to observe biological processes such as perfusion for functional analysis [Detre et al., 1992]. To enhance and visualize the diffusion of blood, a pharmaceutical is injected intravenously into the patient and is distributed through the blood flow. The contrast-enhancement technique is also able to reveal diseases in early stages as the analysis of perfusion characteristics helps to distinguish pathologies with similar morphology [Preim et al., 2009].

CEUS image sequences are produced using gas-filled micro bubbles as contrast agent which have a high degree of echogenicity by increasing the US backscatter [Leighton, 1997]. In a low dose the bubbles are well-tolerated and are degraded by the liver over time. CEUS imaging has become popular with the first US contrast agents (CAs) ratified in 1999 [Harvey et al., 2002]. It has been improved since and
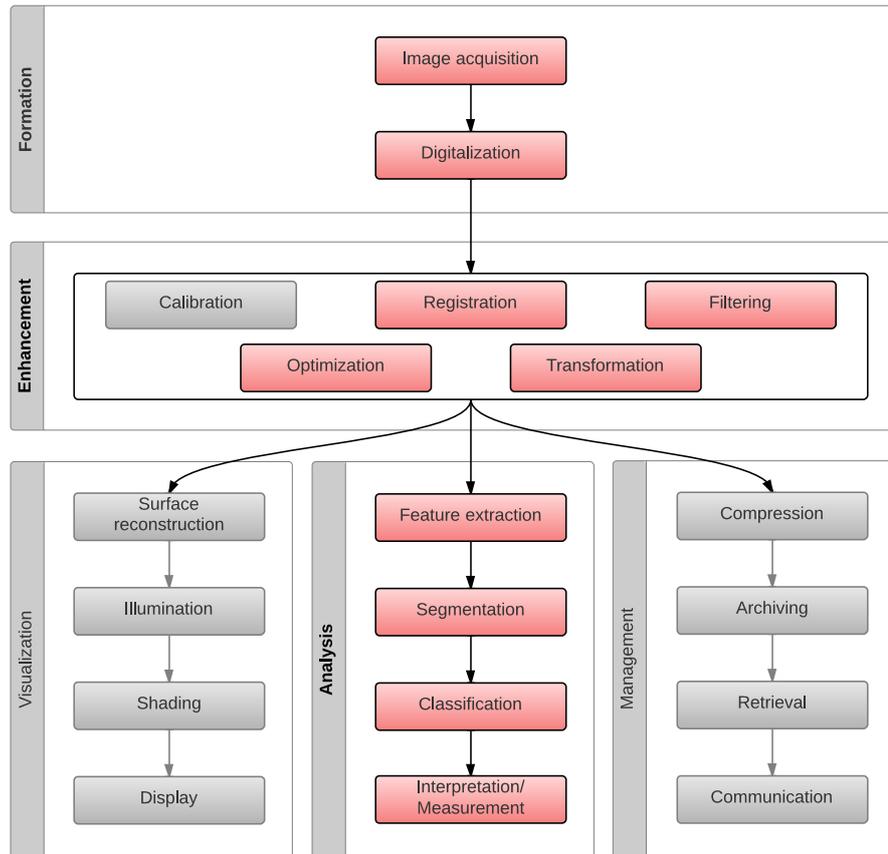
Figure 1.1: The image processing pipeline for medical applications adapted from Deserno [2011] is shown. Modules which are covered in this thesis are depicted in red.

is used in research projects and medical applications [Lindner, 2004; Claudon et al., 2012]. Currently, there is a lot of research targeted on the problems related to ultrasound, the lower resolution, the acquisition artifacts caused by noise and the restrictions when generating 3D volume data [Sanches et al., 2012].

One of the most important restrictions of perfusion imaging is that the recorded scene constantly changes over time. This is caused by motion influence due to patient motion mainly through breathing and organ motion induced by pulsation of blood vessels as well as digestion. In the special case of ultrasound, motion can also be induced by moving or tilting of the transducer. Current medical imaging technologies allow for high quality imaging results and also for quantitatively accurate measurements for diagnosis [Beutel et al., 2000]. However, influences represented in the data also influence the measurements. The deficiencies in the data should be compensated as best as possible and they should be taken into account when measurements are used for diagnosis.

1.2   MEDICAL IMAGE PROCESSING

Medical image enhancement is a preprocessing step to improve the quality of the data for further use. The requirements and suitability of enhancement methods depend on the imaging modality and the application [Dhawan, 2011]. To lower the signal-to-noise ratio (SNR), image filtering or restoration is used. To account for attenuation, specific camera profiles can be applied to recalibrate image intensities [Jan, 2010]. For the compensation of motion, registration techniques can be used to enhance the visual quality and account for acquisition related artifacts [Maintz and Viergever, 1998]. The enhanced images can be used for diagnosis or treatment planning by physicians and by automated methods of image analysis.

Medical image analysis extracts relevant information from medical images. This can be low level information such as structures of a predefined form, edges in general, texture patterns or motion characteristics between frames of an image acquired at different times [Nixon and Aguado, 2008]. These features can be used for segmentation to provide information about the location and size of specific objects or pathologies. The features can be classified according to a-priori known feature distributions and classes for diagnostic purposes. For more specific problems, the results can be enhanced by using high level information about the structures of interest such as structural models representing different shape characteristics and their dependencies [Toennies, 2012].

1.3   ORGANIZATION OF THIS THESIS

This thesis is divided in six chapters. Chapter 2 addresses the medical and technical background of ultrasound and CEUS acquisition. Chapter 3 targets the state-of-the-art image processing techniques for motion analysis and compensation with a special focus on the requirements of CEUS. The analysis system and its components are described in detail in Chapter 4 and the evaluation and the results are described and discussed in Chapter 5. Chapter 6 summarizes the thesis and gives an outlook on future work.

A list of all symbols and abbreviations can be found after the table of contents.

CONTRAST-ENHANCED ULTRASOUND FOR PERFUSION IMAGING

To understand the motivation and application for the analysis system presented in this work, this chapter will give details about the acquisition of perfusion images in general as well as specific details of CEUS image acquisition. The focus is laid on restrictions and problems related to the acquisition of US image sequences with CA, as this is important for constructing and planning the analysis system. Finally, methods for the visual analysis of perfusion imaging are presented.

2.1 PERFUSION IMAGING

Perfusion imaging aims at observing and quantifying fluids passing through an area of an organ or tissue. Primarily, blood is observed being transported to and back from capillaries. It is crucial for an organism to have continuous blood perfusion over time to supply cells with nutrients and oxygen to keep them alive [Hardaway, 1981]. It may be lifesaving to detect any malfunctions leading to shortages in the supply of cells. As an illustrative example for perfusion imaging, Figure 2.1 shows cerebral perfusion measured by an X-Ray-based digital subtraction angiography.

In order to detect and diagnose a disease or to discriminate different pathologies, kinetic properties of the perfusion have to be extracted and quantified [Preim et al., 2009]. Perfusion kinetics describe the amount and the velocity of perfusion to pass a particular barrier [Detre et al., 1992]. Additionally, measuring the characteristics of perfusion in tumor areas gives indication about the condition or permeability of the vessels [Heywang-Köbrunner et al., 1997]. The permeability is a good index measurement of the aggressiveness of cancer (benignancy/malignancy), as it corrodes the vessel wall [McDonald and Baluk, 2002]. Hence, the CA leakage can be used as indicator and is approximated by measuring the CA leakage to extracellular space [Tofts and Kermode, 1991]. In oncology, perfusion is particularly important, as angiogenesis leads to the formation of new vessels within and around tumors [Kuhl, 2007]. This enables the detection of small lesions for an early therapy.

To calculate and derive such perfusion kinetics from image data, perfusion imaging must fulfill two important requirements. First, the imaging system must deliver the perfusion state at different time
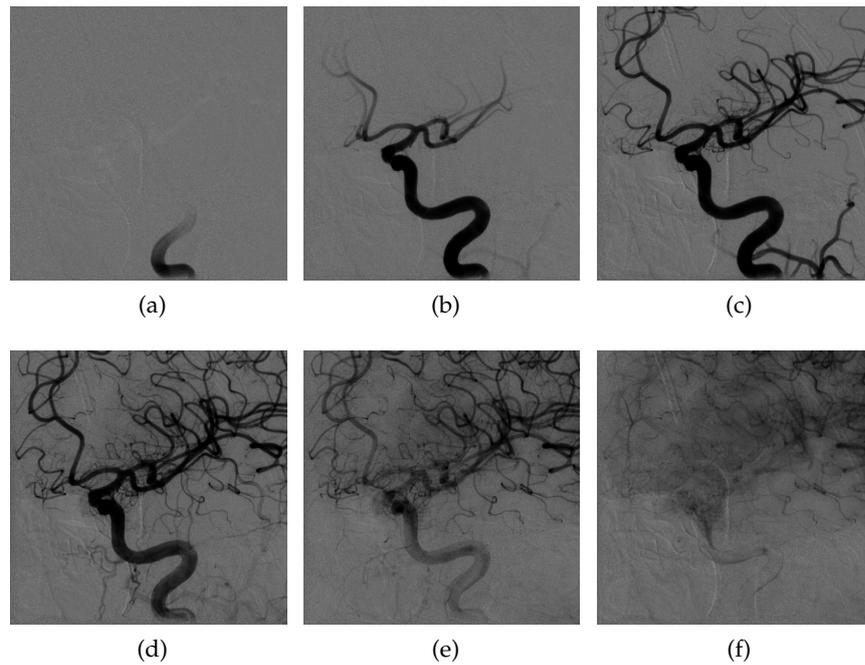
Figure 2.1: Images of a 2D Cerebral angiography with CA injection in the left internal carotid artery. Images in the top row are taken with 0.5 seconds time difference; the images in the bottom row are taken with 1 second time difference. Courtesy of University Hospital Magdeburg, Department of Neuroradiology.

points in order to observe the progression and diffusion of the CA [Axel, 1980]. This also implies that the modality has to be able to guarantee sampling rate which fulfills the requirements of the application. Second, the fluid in the circulation system must be enhanced to contrast from surrounding tissue. To this end, a CA is injected into the organism which can be traced on different time points in a region of interest (ROI) to get information about perfusion dynamics [Detre et al., 1992].

The type of CA has to be chosen according to the modality. For example, in magnetic resonance imaging (MRI) substances such as Gadolinum (GTPA) leading to shorter relaxation times alter the resulting signal and create a contrast between perfused and unperfused tissue. In X-Ray-based imaging iodine or barium are mostly used resulting in an increased absorbability of X-Rays compared to surrounding non-perfused soft tissue [Böhm and Schild, 2006]. In NMI compounds which carry very small doses of radioactive material (radiopharmaceuticals) are used to produce imaging material. The signal is detected by gamma detectors which measure the amount of radioactive decay at different locations [Mettler and Guiberteau, 2012]. CAs may also be constructed such that it binds specific receptors, i. e.cell types or proteins [Morawski et al., 2005]. This enables

another quality of imaging, as diagnosis can be targeted at particular problems or behavior.

In the remainder of this chapter the acquisition of CEUS images with the help of gas filled microbubbles serving as CA is explained as this forms the data basis for the proposed analysis system.

2.2 US AND CEUS IMAGING

US is a widespread, inexpensive and comparatively safe imaging technique [Merritt, 1989]. US acquires 2D brightness modulation (B-mode) images depicting a cutting plane through the object of interest. 3D volume images can be generated from multiple acquisitions of different adjacent planes.

The system presented in this work aims at the analysis of time dependent 2D imaging. Therefore, the focus lies on the features, problems and limitations of 2D US. A 2D US observation is performed in real time, i. e., the examiner gets immediate feedback on a screen about the current image. This enables to correct the US probe position by shifting or tilting to make sure the desired ROI is visible in the imaging plane. In modern US acquisition systems, the observation is recorded as a video stream for later viewing and analysis. Commonly, the image sequence can be exported as video file or DICOM file.

A problem in standard US is that blood and tissue have a similar echogenicity what makes a differentiation difficult. Several techniques address this problem in US, such as Doppler measurements which are used to detect insonated blood cells to determine the flow speed [Loizou et al., 2008]. However, this technique does not produce reliable flow information for all diagnostic applications. Especially, this applies to abdomen imaging [Averkiou et al., 2003]. The blood flow is too slow in the parenchyma with its microvasculature to be measured by the Doppler effect [Lindner, 2004].

To overcome this drawback, CAs for US have been developed for the first time in 1968 to enhance the contrast of several structures [Gramiak and Shah, 1968; Kremkau et al., 1969]. Since then, the quality and usability of both, the CAs and appropriate US acquisition protocols for CEUS, have been improved through continuous research [Leighton, 1997; Averkiou et al., 2003]. In the last decade, the technique has evolved from an experimental tool to a routine diagnostic method in many fields of application [Lindner, 2004]. Its advantages lie in the cost-effectiveness and fewer side-effects compared to, e. g., CT [Correas et al., 2001; Rickes et al., 2007].

Microbubble acoustic stimulation   Microbubble received signal
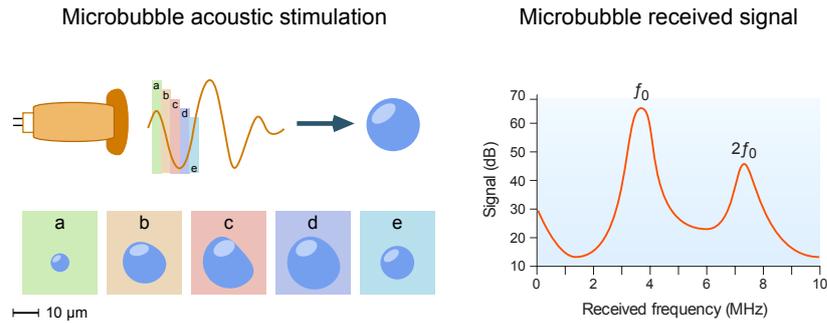
Figure 2.2: The volumetric oscillation of a microbubble under high- and low-pressure phases is shown. The received signal features two signal peaks at the fundamental (f_0) and second harmonic ($2f_0$) frequencies. Adapted from Lindner [2004].

2.2.1 *Microbubbles as Contrast Agent*

To enhance the contrast between blood filled vasculature and tissue, the US signal reflection has to be altered in perfused areas. This can be achieved by the administration of gas-filled microbubbles. These bubbles consist of a compound of air and gases surrounded by a lipid-based shell. By inducing an oscillation frequency pulse the bubbles are alternatingly compressed and expanded. This process changes the density of the compound inside the bubbles producing a specific acoustic reflection. This reflection signal can be well differentiated from normal tissue backscatter (Figure 2.2).

The size of microbubbles varies between 1-4 µm in diameter and is smaller than the size of red blood cells (6-8 µm) so they can easily be transported at microvascular level [Lindner, 2004]. However, they cannot leave the vessel lumen [Postema and Gilja, 2007]. Any echo received from a microbubble gives evidence of the presence of a vessel.

If a high acoustic pulse is applied the shell of a microbubble bursts and the content of the bubble is released into the blood stream [Chomas et al., 2001]. The microbubbles can no longer be observed by US. This phenomenon is used for two purposes. First, bursting of the contrast bubbles within the imaging plane results in a substantial decrease of CA concentration and leads to a replenishment with microbubbles coming from surrounding areas (outside the imaging plane). This effect is used for diagnostic purposes, as it reflects the flow velocity and is called burst-replenishment [Wei et al., 1998]. Second, the compound inside the bubble is equipped with a drug that can specifically be released within a target region [Postema and Gilja, 2011].
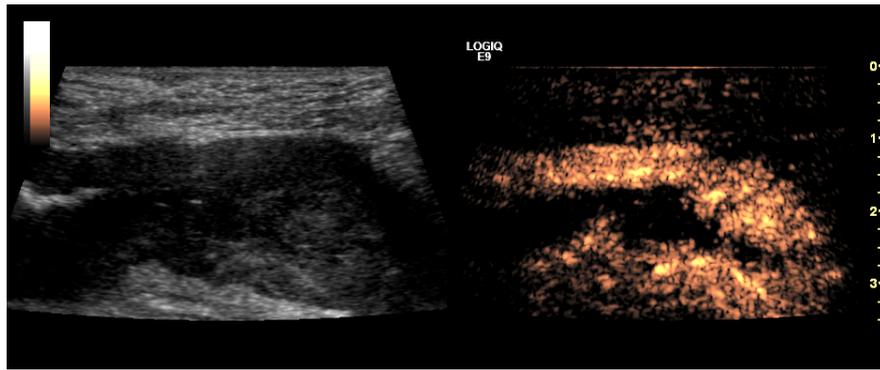
Figure 2.3: CEUS of the small bowel. Left: a single frame of a temporal sequence in B-mode. Right: the corresponding frame in contrast mode. Courtesy of University of Bergen, Norway, Institute of Medicine.

2.2.2 *Imaging Protocol*

CEUS examinations are usually performed freehand by a medical expert. The examiner will try to keep the US transducer as still as possible once the ROI for examination is displayed in the imaging plane. If the ROI leaves the imaging plane, the examiner will try to adjust the view through tilting the US probe to correct for the changes. To standardize the process for different examinations, specific views and transducer positions are intended for different CEUS studies [Nylund et al., 2013].

The ideal resonant frequency of a CA depends on the properties, e. g., the shell stability, size or gas compound [Lindner, 2004]. Given the ideal frequency, filtering can be used to enhance the contrast between background and microbubbles as they generate a specific harmonic pattern (Figure 2.2). The imaging protocol provided by GE Healthcare with US allows for a hybrid mode. I. e., the US device generates frames of simple B-mode and frames with frequency filtered CA detection. Hence, morphological information as well as perfusion information (from contrast enhanced microbubbles) of the same ROI are available and can be displayed side by side (Figure 2.3) or as an overlay. A temporal resolution in both modes of up to 15 frames per second is achieved with current US machines (reference: GE Healthcare LOGIQ® E9). Disruption of the microbubbles is induced application-dependent by the physician at a particular time point during the examination resulting in an abrupt signal increase. Frames affected by the burst impulse should be excluded in later analysis, as comparability is not given due to the large intensity values evoked by the bubble burst.

2.2.3 *Application and current research of CEUS*

In echocardiography, the contrast enhancement of the blood vessels increases the visibility of the endocardium and is used to assess the ventricular size and systolic functionality. It is therefore an important tool for diagnosis of coronary artery disease [Kono et al., 2004]. Besides, CEUS is also used for depicting perfusion at microvascular level, e.g., in abdominal organs such as liver, pancreas, prostate, kidney or the bowel in order to study the pathologic condition in vivo [Albrecht et al., 2005; Claudon et al., 2012]. In addition, the detection of microvasculature by CA enables the finding and diagnosis of primary tumors and metastasis to assist in cancer treatment [Solbiati et al., 2001; Piscaglia et al., 2012].

For the analysis of the bowel, CEUS is used to differentiate between inflammation and fibrosis by quantification of contrast-enhancement [Nylund et al., 2013; Quaia, 2013]. Pancreatic diseases are assessed with CEUS to differentiate between necrosis, cysts, abscesses and normal or inflamed tissue by enhancement characteristics [Erchinger et al., 2011].

A lot of effort is currently put into the development of microbubble CAs [Bloch et al., 2004; Cosgrove and Eckersley, 2006]. For targeted CEUS, ligands are attached to the microbubbles so that they bind to specific substances or cell molecules, e.g., to detect inflammatory processes within the body. For drug and gene delivery the material is included in the compound of the bubble and the burst technique is used for deposition [Postema and Gilja, 2011]. To ensure the drug is delivered at the right location, either targeted bubbles can be used or the target region is imaged for the correct timing of the burst.

2.3 PROBLEMS AND LIMITATIONS IN CEUS

Despite the chances of US there are also drawbacks. Most of these drawbacks are related to the visual quality of the produced image data, such as resolution, noise and motion [Toennies, 2012]. As a consequence, challenges for medical image processing arise, dealing with the correction, compensation or reduction of the influence of quality issues. The most prominent challenges are discussed in this section as they have an influence on the design of the analysis system described in this work.

2.3.1 *Resolution*

Image resolution, i.e., the minimal distance at which two objects can still be separately viewed, depends on the frequency emitted by the transducer in US imaging. To obtain more detailed results, higher

frequencies must be used to produce shorter wavelengths to sample tissue in axial direction (direction of US propagation). Higher frequencies of up to 20 MHz lead to a smaller width of the US beam which in turn results in a higher lateral resolution (right angle to axial direction). However, higher frequencies imply that the signal is not able to reach deeper tissue regions because the penetration of the US beam is proportional to the wave length. Thus, the frequency has to be chosen as a compromise between depth penetration and resolution [Chan and Perlas, 2011]. For abdominal imaging it normally lies between 1 and 6 MHz to ensure that the acquired images show sufficient depth information. A less curve-shaped transducer array is used to induce more parallel beams and to increase lateral resolution especially in the far field.

The temporal resolution describes the time between the acquisition of frames. In general, it depends on the recording area as larger areas lead to longer recording times for a single frame. However, the recorded number of frames per second mostly lies between 10 and 20. This is sufficient to cover contrast uptake characteristics and a higher resolution compared to most other modalities acquiring data over time.

2.3.2 *Noise artifacts*

Noise affects the quality of US imaging. Random noise introduced through distance signal transmission and transformation is present similar to other imaging modalities. Additionally, US images induce a very characteristic noise pattern known as speckle noise. The pattern is composed of relatively small and light blobs superimposing the signal (see Figure 2.3). Strictly speaking, it cannot be referred to as noise, as it is not random. Patterns could be reproduced under the assumption that equal conditions apply to the acquisition each time [Loizou et al., 2008]. Also, the pattern does only slightly change between different temporal acquisitions if the scene does not change.

Random noise increases with depth in the US field, while the speckle effect decreases with the distance to the transducer. Additionally, depth signals have to be amplified because of attenuation, also enhancing the noise signal in those parts of the image [Gibbs et al., 2009].

2.3.3 *Motion artifacts*

Motion is a feature to be observed in time-dependent imaging data, although not all motion effects are desired to be present. In perfusion imaging, the effect of motion induced by CA flow will be further used for analysis. Other motion influences distort the extraction of

perfusion information as inter-pixel correspondences between temporal instances of a sequence may be corrupted.

In US acquisition, different types of motion can be distinguished [Schäfer et al., 2011a]:

- Intrinsic motion induced from inside the organism of the patient

- Extrinsic motion induced from outside, e. g., through probe movement

There are several reasons for the occurrence of intrinsic motion: tissue is shifted, compressed or expanded by breathing. Organs move due to blood flow or digestive activity and the patient involuntarily moves during an examination.

Extrinsic motion is stemming from tilting or shifting of the US probe which additionally occurs to patient movements. The examiner intentionally changes the probe orientation to compensate for motion influences or to adjust the view to the desired ROI. As mentioned earlier, this is necessary as 2D imaging targets a plane and not a volume. To make sure a particular tissue or organ region can be examined correctly over time, it has to stay within the imaging plane.

In the majority of US datasets, motion effects are caused by different sources which overlap and form a specific motion field.

The motion artifacts as well as the other limitations in CEUS are important to be considered when the data are used for diagnostic purposes in perfusion analysis. Therefore, basic concepts of perfusion analysis are explained in the following.

2.4   PERFUSION ANALYSIS

In perfusion imaging, the amount of data to be reviewed by medical experts is larger compared to static imaging and dependencies are also more complex [Preim et al., 2009]. Perfusion imaging introduces a new dimension, time, which has to be taken into account. The task of perfusion analysis provided by the computer is to reduce the complexity of the data by preserving the diagnostic value. This can be done by establishing perfusion parameters which represent the characteristics of perfusion.

2.4.1   *Perfusion Parameters*

Perfusion parameters are derived from the time intensity curve (TIC) representing the enhancement through CA at a specific location as
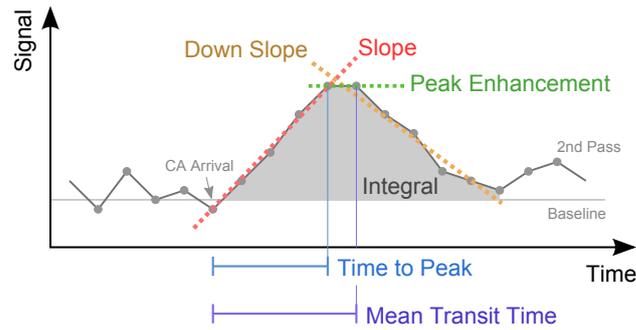
Figure 2.4: Different perfusion parameters derived from the TIC of a pixel/voxel location in the perfusion image or a local ROI (adapted from Oeltze et al. [2007])

a function of time. A perfusion parameter is a scalar value representing a particular aspect of characteristic of perfusion dynamics. Descriptive parameters derived from a TIC mostly used in contrast-enhanced MRI are shown in Figure 2.4: peak enhancement, time to peak, integral, mean transit time, slope and down slope [Oeltze et al., 2007]. Some of them are also used in CEUS, e. g., to assess treatment response [Lassau et al., 2010]. Angelelli et al. [2011] suggest to use perfusion parameters in CEUS for interactive visual exploration. Common practice is to generate parameter maps assigning each location in the image a certain value which is encoded by a color or a graphical primitive. The attention of the examiner is drawn to particular locations and noticeable characteristics are revealed. Another technique, called brushing and linking, is employed to highlight specific areas in the perfusion data by selecting samples with target perfusion parameter values using scatter plots or parallel coordinates [Coto et al., 2005; Oeltze et al., 2007].

Quantitative perfusion parameters are deduced from the TIC using knowledge about the arterial input of CA. These parameters address the absolute quantification of blood going through an organ or an area of tissue per unit time. Usually, this is measured as ml per minute per 100 ml of tissue giving important information about the blood flow [Wei et al., 1998; Nylund et al., 2013]. Pharmacokinetic modeling has been developed to explain and simulate the physiological behavior of the blood flow in the human body.

Tofts and Kermode [1991]; Larsson et al. [1990]; Brix et al. [1991] have almost simultaneously published the first pharmacokinetic models to quantitatively analyze perfusion images. All models are targeting contrast-enhanced MRI and link the physics of the MRI signal acquisition to a physiology model that simulates CA kinetics. The CA concentration is calculated depending on physical properties, acquisition-related parameters and the physiological character of tissue. In general, the latter is unknown. Thus, it is determined by
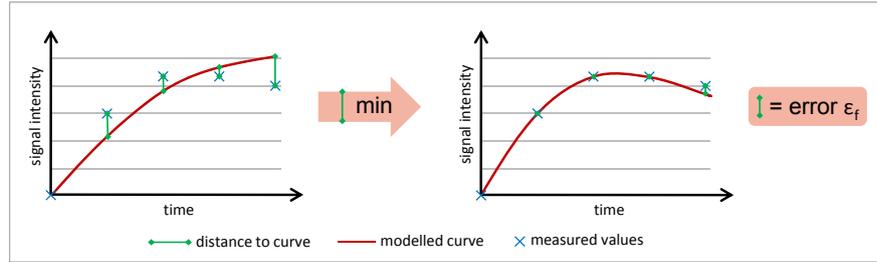
Figure 2.5: The distance of the pharmacokinetic model function to the measured data samples is minimized by variation of the free model parameters.

fitting the pharmacokinetic model function to the concentration measured in the acquired data, leaving physiological values as free parameters. Regression analysis is used to find physiological values for the model to represent the measured perfusion data as best as possible (Figure 2.5). Tofts et al. [1999] establish standardized quantities and symbols for all proposed models which define the physiology of the tissue and permeability properties of vessels.

Models for perfusion analysis in CEUS have also been developed covering different applications such as myocardial blood flow [Vogel et al., 2005] or different techniques such as bolus administration of CA [Strouthos et al., 2010; Gauthier et al., 2012] and the bolus and burst technique [Jiřík et al., 2012]. Jiřík et al. [2013] propose a method to calculate absolute values for mean transit time (MTT), blood flow and blood volume. The model covers bolus administration of CA and the burst-replenishment technique. It approximates the TIC by convolution of the arterial input function (AIF) and the tissue concentration. The approximation is formulated as blind deconvolution.

To account for the afore mentioned limitations in medical imaging data (cf. Section 2.3), both groups of perfusion parameters should be generated from a TIC of a ROI to use average values. This minimizes the influence of noise and motion. Two approaches exist to establish ROI:

- drawing of manual regions by the user and

- automatic aggregation of pixels/voxels with similar perfusion characteristics.

The drawing of manual regions of pixels in 2D images or voxels in 3D images can be established to average perfusion. However, in these regions different perfusion characteristics might be combined and important details might diminish. Additionally, the manual process is tedious. Instead of using manual regions to combine multiple measurements of perfusion, Glaßer et al. [2009] propose an automatic

(a)   (b)
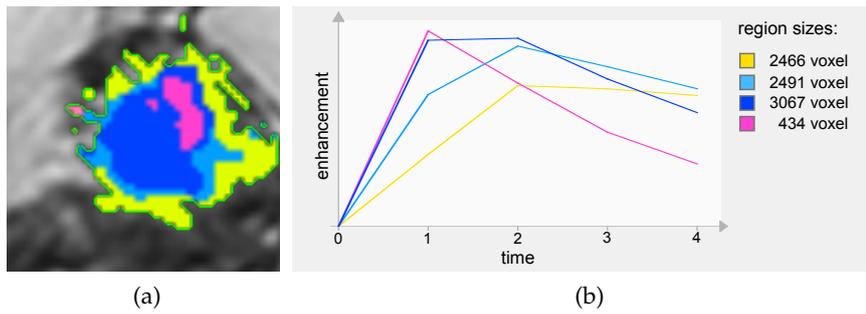
Figure 2.6: (a) shows perfusion regions obtained trough segmentation us-
ing perfusion characteristics as similarity criterion for a tumor
acquired with Perfusion MRI of the breast. (b) depicts the TICs
for the five acquired time points of the different regions from (a)
[Schäfer et al., 2011b].

segmentation approach using perfusion parameters as similarity cri-
terion from contrast-enhanced MRI to automatically generate analysis
regions. Hence, the average is calculated from tissue with compara-
ble perfusion kinetics (cf. Figure 2.6). The procedure involves a re-
gion merging approach with each pixel/voxel being a start region.
Neighboring elements are merged until they exhibit a certain sim-
ilarity, e. g., in terms of sum of squared distances D or normalized
cross correlation (NCC). In this way, tedious work (especially in 3D)
is avoided and regions are guaranteed to contain similar perfusion
and do not mix different characteristics. As an example, Figure 2.6
shows a segmented tumor in an MRI image where each region has
a different TIC. The curves are used to distinguish between malig-
nant and benign tissue. The technique is adapted by Angelelli et al.
[2011] for the analysis of perfusion in CEUS. They use the result of the
segmentation procedure to derive perfusion parameters in order to
enable different visualizations for efficient analysis and comparison.

2.4.2 *Classification Analysis*

For automated analysis, the derived perfusion parameters can be
used in a machine learning approach. Learning combinations of pa-
rameters can be helpful to classify tissue of different types. This has
been applied to breast MRI by using state-of-the-art classification to
investigate the malignancy of breast tumors [Lucht et al., 2001; Nat-
tkemper et al., 2005; Spilker et al., 2005; Levman et al., 2008]. Chen
et al. [2006] present an automatic clustering approach to distinguish
between tumor and background to enhance the analysis of tumors
through their dynamic parameters. Glaßer et al. [2013] classify tu-
mors in MRI data according to their enhancement pattern in benign

and malignant. The enhancement patterns are derived from a clustering analysis of the TICs of the voxels.

2.5 SUMMARY

This chapter introduces perfusion imaging, possible applications and technical information about US and the protocol to acquire CEUS. Despite all advantages given by 2D US acquisition approaches and particularly CEUS, there are several drawbacks which need to be addressed if quantification tasks have to be applied to CEUS datasets. These are the restrictions in quality, specifically noise and motion and furthermore, the problem of the area of interest moving outside the imaging plane.

In the following chapter, strategies and methods are discussed to approach the problems of motion analysis and compensation.

MOTION ANALYSIS IN MEDICAL IMAGING

3

This chapter summarizes related work of preprocessing, analysis and compensation of motion in medical images in general and the application to perfusion imaging and CEUS data in particular.

The quality and validity of medical image analysis is often reduced by motion influence stemming from different sources. This is not only a problem in CEUS as described in Section 2.3, but a general problem in dynamic medical imaging including perfusion imaging [Preim et al., 2009].

Motion analysis in medical images is important because of two reasons. The first one is to compensate for motion influence to allow for a valid analysis of time dependent features and parameters. The second reason is to find out about the amount of motion and its characteristics giving important diagnostic information such as the motion of the heart. In literature the task of motion compensation or correction is commonly known as medical image registration.

On the other hand, pure motion analysis typically focuses on objects or organs where motion plays a key role, e.g., the heart or the lungs. Motion analysis is mostly solved using motion tracking for an overall analysis or object tracking if particular organ characteristics or behavior can be anticipated.

Both, registration and tracking, are techniques which are closely related. Registration covers a much broader scope of problems and can be used for object tracking in some cases as well. Object tracking usually solves very specific application problems. But motion characteristics revealed by those approaches can be used to compensate the motion influence.

In the following sections, important and relevant work in this context is presented to evaluate the suitability for the use in CEUS image motion compensation. First, motion effects and their analysis and the closely related optical flow topic are discussed. Second, image registration is reviewed and basic concepts are described. Third, motion and object tracking approaches based on optical flow are presented.

3.1 MOTION AND OPTICAL FLOW

Image motion is defined as the displacement of objects relative to the observer which is in this case the imaging sensor [Beauchemin and Barron, 1995]. In order to observe motion in recorded data, observations at different times are required. These observations are denoted

image sequence or video sequence. An important parameter for the interpretation of motion is the density of temporal samples in a sequence of images which is the time between acquisitions of frames. This is also valid for human observation or problems in computer vision.

Optical flow, a vector field containing motion direction and velocity of an image in a sequence, can be used to formally describe motion [Gibson, 1977]. However, not all motion effects are covered by optical flow measures. The optical flow is based on the assumption that objects need to exhibit characteristic textures to be traceable over time and brightness constancy of corresponding areas [Horn and Schunck, 1981]. It is calculated using the partial derivatives of the image function in spatial and temporal dimensions. However, such methods need further preconditions to be solved because the problem itself is ill-posed (aperture problem) [Uras et al., 1988]. Horn and Schunck [1981] introduce a global approach restricting the optical flow field to be smooth in space. In contrast, Lucas and Kanade [1981] choose a local strategy by assuming constant velocity within a window area of defined size. Improvements and specializations of global and local approaches based on derivatives have been published [Bruhn et al., 2003].

Glocker et al. [2008] determine the optical flow using discrete optimization. The search space is quantized and each pixel is assigned a displacement label. The optimal configuration of these displacements is determined by energy minimization of a Markov random field (MRF). The same approach has also been used for image matching or registration tasks [Shekhovtsov et al., 2008; Glocker et al., 2008; Mahapatra and Sun, 2012].

The computation of optical flow is closely related to image registration as well as motion and object tracking and thus, the analysis of motion in images. In some cases optical flow serves as basis or feature space for tracking purposes [Cremers and Schnörr, 2003]. In the following sections, the different objectives pursued by registration and tracking are discussed.

3.2   REGISTRATION

In general, image registration is regarded as the process of transforming multiple instances of images or image volumes from different sensors, times or viewpoints into a common coordinate system or common alignment. To establish valid registration, it is important that the same parts of organs or tissue within an image instance are represented in all other instances. This is particularly important in 2D image registration, where a single imaging plane is used which can change between different image acquisitions.
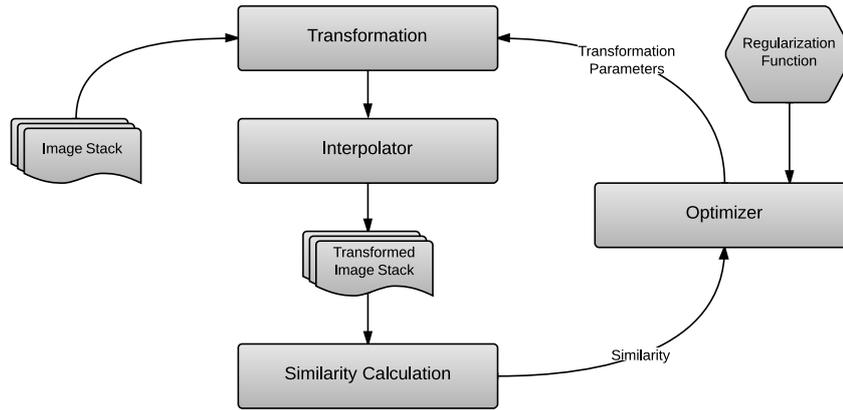
Figure 3.1: The different registration modules and their dependencies carrying out the subtasks. The transformation is applied with initial parameters to the image stack using an interpolation function. Similarity of the images in the stack is calculated and delivered to the optimizer. The optimizer finds new transformation parameters maximizing the similarity under the terms of the regularization function.

Comprehensive surveys and books on registration have been published by Hill et al. [2001]; Zitova and Flusser [2003]; Goshtasby [2005]. In the following, a review covering the most important aspects is given.

3.2.1 *Basics*

Commonly, automatic registration is split into different modules representing the core functionality. These are transformation, similarity calculation, optimization, regularization and interpolation (see Figure 3.1). The transformation defines the function to deform images according to a set of parameters. The similarity defines a scalar value evaluating the match of two images when applying a particular transformation. The optimizer searches for the best set of parameters for a transformation in terms of the similarity. The optimizer also guarantees compliance to the conditions of regularization. Regularization constrains the parameters of a transformation to fulfill any preconditions. The interpolation function is used to calculate the transformed images and especially covers calculation of sub-pixel displacements. The modular design illustrates exchangeability of the registration technique in subtasks which have to be chosen according to the application.

Transformation functions are divided into rigid, affine and non-affine (or mostly referred to as non-rigid or elastic) transformations (see Figure 3.2). Rigid body transformation has a maximum of 6 degrees of freedom (translation and rotation in 3D) and ensures con-

(a)   (b)   (c)
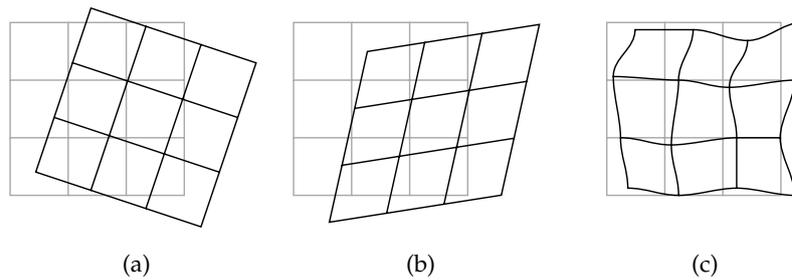
Figure 3.2: The different types of transformations are illustrated using a grid structure: rigid (a), affine (b) and non-rigid or elastic (c) transformation.

stant distances. Affine transformations additionally include scaling and shearing and can still be expressed by matrix operations preserving parallelism [Hill et al., 2001]. In 3D they have a maximum of 12 degrees of freedom[1]. Non-affine or elastic transformations, in contrast, need particular functions such as thin-plate splines or B-splines to map locations in one image to corresponding ones in another image [Bookstein, 1989; Rueckert et al., 1999]. The restrictions, e. g., the size or the smoothness of a transformation are included in the regularization module. In most cases the regularization is directly implemented in the objective function for optimization to penalize parameter configurations which are undesirable [Zitova and Flusser, 2003]. Fluid mechanics models have been adapted to serve for the transformation of images [Christensen et al., 1996]. They offer a high degree of flexibility and are able to cover large deformations, but on the other hand it is more difficult to formulate constraints for fluid mechanics models [Rohr, 2000].

Another important aspect of the objective function for optimization is given by the similarity determination of the transformed images. The similarity function has to be chosen according to the image data to be registered. In medical imaging, area-based measures of the image function are often used to determine similarity [Zitova and Flusser, 2003]. In general, the mean squared distance (MSD) yields stable results for single-modal images with constant intensities for corresponding tissue. It is able to cope with noise and small differences in intensities due to attenuation. More elaborate measures to cover specific problems such as multi-modal image registration include NCC and mutual information (MI). Apart from area-based measures, feature detection methods such as edge detection or combined characteristic features can be employed and matched with similar features in other images. Most prominent representatives of feature

---

1 three from translation and nine from a 3×3 matrix combining rotation, scaling and shearing

detectors are SIFT and SURF features [Lowe, 1999; Bay et al., 2008]. They are invariant with regard to the scale and the orientation of an object or a part of it.

Finally, the optimizer determines the best parameters for the transformation by evaluating the objective function. The choice for the optimizer depends on the following circumstances: the number of parameters to be optimized, the derivability of the objective function and the size of the search space.

Commonly, a registration problem is formulated for only two images. However, multiple images can be registered as well, either by consecutively performing multiple pairwise registration tasks or by solving the problem as a combined objective function. The latter is computationally more complex but enables to formulate dependencies between the registration of individual images. The formulation of dependencies is essential when dealing with poor data quality because the poor quality can be compensated with knowledge about the transformation characteristic derived from images in the neighborhood [Ledesma-Carbayo et al., 2005].

In the field of medical imaging, registration plays an important role especially when collections of the same organ or the same tissue have to be compared. In this case, images are acquired at different time points or with different modalities.

Multi modality registration merges information from different scanning devices which usually cover different aspects of the human body [Rohr, 2000]. However, there are devices which acquire images with different techniques at the same time (e. g., CT and positron emission tomography (PET) scanner) which makes registration obsolete. In practice however, the acquisition of images is often performed in two steps with different setups due to technical reasons or cost efficiency. In this case, a transformation of the images into the same space or coordinate system is required for accurate analysis.

Images of the same patient may also be acquired at different time points using the same imaging device and parameters. A follow-up study is required to take images, e. g., to observe effectiveness of a drug treatment, with even longer intervals. Registration is needed to align the relevant structures and, e. g., compare tissue growth or shrinkage [Sowell et al., 2004]. To capture motion or contrast perfusion in acquired datasets the interval between acquisitions must be in the range of minutes or even seconds. As a consequence, the interval between acquisitions can vary between milliseconds and hours. The actual time interval depends on the processing technique of the scanning device. In general, registration is easier for short time intervals because equal conditions apply as the patient does not change the position and does not leave the scanner. In addition, assumptions

about the continuity can be made and taken into account by the registration framework [Frangi et al., 2003; Ledesma-Carbayo et al., 2005].

3.2.2 *Applications in 2D Ultrasound*

A number of applications of motion registration to US target cardiac analysis [Shekhar et al., 2004; Ledesma-Carbayo et al., 2005]. Registration is needed to assess myocardial elasticity and contractility and to find motion abnormalities in both stress and rest scenario. Ledesma-Carbayo et al. [2005] introduce a B-spline-driven motion model to minimize the average of the SSD considering spatial and temporal continuity of the motion field. Due to the high temporal resolution in the range of milliseconds between acquisitions in US the continuity assumption holds.

Block matching is used to find motion disparities in subsequent frames for thyroid cancer diagnosis by Basarab et al. [2008] and for the carotid artery wall by Golemati et al. [2007]. In both methods constant intensities are assumed and that the speckle pattern movement corresponds to real motion.

Frangi et al. [2003] apply a two step registration strategy to quantify dilation of the brachial artery in US data. The first step corrects for global motion in the image sequence using rigid transformation. The second step estimates the dilation of the brachial artery by applying an affine transformation. The parameters of both registration steps are constrained by a Kalman filter to ensure temporal continuity.

When registering 2D US datasets, special restrictions apply (cf. Section 2.3). The first restriction is the data quality in terms of a low SNR. NCC has been widely used for registration of US data [Rohling et al., 1998; Golemati et al., 2007; Luo and Konofagou, 2010]. A fast calculation method for NCC is proposed by Luo and Konofagou [2010] addressing motion estimation in time-dependent Doppler and elastography US imaging. Cohen and Dinstein [2002] present a new motion estimation measure, CD_2, using a maximum likelihood method and assuming Rayleigh distributed noise in the images. Woo et al. [2008] introduce local phase information of the Fourier transformed signal as registration feature combined with normal intensity information to register US data with non-rigid transformation.

Another restriction is that US devices are limited to the acquisition of 2D cutting planes of 3D objects. During acquisition these cutting planes can change due to motion influence (cf. Section 2.3.3). This out-of-plane motion changes the appearance of organs or structures and causes the decorrelation of speckle structures as well [Yeung et al., 1998]. As a result, this means that valid registration or motion analysis cannot be obtained when out-of-plane motion is present.

Chen et al. [1997] suggest to perform a speckle decorrelation analysis to assess different motion influences and account for these in potential motion correction steps. Learning of different speckle characteristics from synthetic datasets with known ground truth can be used to differentiate between motion influence characteristics and particularly identify out-of-plane motion [Laporte and Arbel, 2011].

Only little attention is paid to this circumstance by most applications in literature. Renault et al. [2005] select frames of similar time points in the respiration cycle through factor analysis of the independent components. However, this approach assumes that the liver always returns to the same location after a respiration cycle. Frouin et al. [2006] apply motion compensation to muscle perfusion studies. If the calculated motion of two image pairs exceeds a certain threshold, the result of the registration is rejected before any analysis is performed.

If spatial and especially temporal continuity is assumed for transformation parameters, image frames which are out-of-plane will have a negative impact also on the registration of valid image frames. Therefore, the detection of out-of-plane motion and the exclusion of frames exhibiting such motion is a key point addressed in this thesis.

3.3 OBJECT AND MOTION TRACKING

Registration methods target the removal of differences between images, e. g., the removal of motion. The registration transformation is just the description of motion in the scene. In contrast, object and motion tracking primarily focuses on defining location, velocity and motion trajectories of one or more objects in the scene [Challa, 2011]. Tracking methods used for motion analysis are able to cope with several (additional) problems:

- tracking of objects over large distances with complex motion,

- discontinuities in the motion field and

- partial or full object occlusion.

However, most of the tracking approaches need a decent number of temporal observations with only small changes in location and appearance to successfully track objects or points of interest over large distances [Comaniciu et al., 2003]. The following section describes the basics of object or motion tracking approaches to assess possibilities to transfer the techniques to the problem of motion compensation of US data.

3.3.1   *Basics*

In addition to registration, a-priori knowledge can be integrated in an object model specifying, e. g., the shape of the object to be tracked. Object representation can range from using simple characteristic points or basic geometric shapes to complex shape and statistical appearance models [Yilmaz et al., 2006]. According to the object or motion representation, features are extracted to identify the objects or motion characteristics. Common features are intensity, edges, optical flow and texture [Yilmaz et al., 2006].

The tracking process covers the determination of a trajectory of the object of interest by tracking characteristic features in subsequent instances of the imaged sequence using an object model and a motion model. The basic problem, which is already known from registration, is the same: the correspondence problem to find out where parts of an image are located in subsequent images. According to Yilmaz et al. [2006], three different strategies can be distinguished to solve the correspondence problem and keep track of the object features:

- Point tracking

- Kernel tracking

- Silhouette tracking

POINT TRACKING   solves the correspondence problem of multiple instances of feature points in subsequent frames of a sequence (see Figure 3.3a). This is a non-trivial problem, as the occlusion, appearance and disappearance of objects, and therefore of feature points has to be considered. That means that not all feature points will have correspondences in subsequent images.

KERNEL TRACKING   uses the region of an object or a specified area around a feature point. This region can be used as a template to find corresponding locations in consecutive frames (see Figure 3.3b). This method is analogue compared to similarity calculation in registration. Derived features such as histograms, texture or mixture models can also be used to determine the location.

SILHOUETTE TRACKING   requires a contour or silhouette based representation or model of the object (see Figure 3.3c). The object contour location of the object can be found by minimizing the distance between model and data, e. g., by using the Hausdorff distance.

In all of these cases, prior knowledge about the motion (a motion model) is required, in order to reduce the optimization space to a
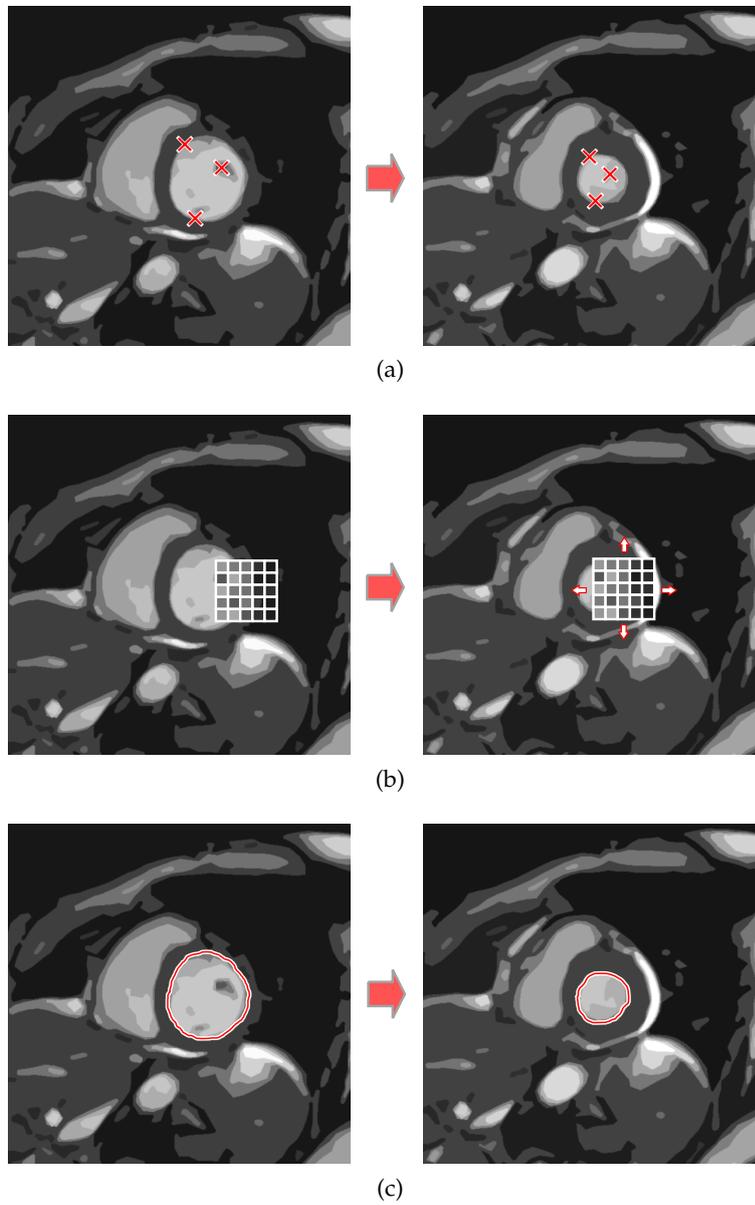
Figure 3.3: The different tracking strategies are demonstrated using different time points of a beating heart (schematic illustration): (a) point tracking, (b) kernel tracking and (c) silhouette tracking.

feasible size. This can be low-level prior knowledge such as inertia or smoothness, as well as high-level knowledge about characteristic movements of certain objects in the scene [Veenman et al., 2001].

A commonly used strategy to include low-level knowledge is to estimate direction and velocity of feature point trajectories from the previous states to enforce continuity. This is achieved using Bayesian filtering [Chen, 2003]. In a prediction step, knowledge about state changes relative to the previous state is used. In an update step, the likelihood function of the current observation is employed to calculate posterior probability.

If the likelihood function is assumed to be Gaussian, the system is denoted as Kalman filter [Kalman, 1960; Challa, 2011]. If discrete states and a finite number is assumed the problem can also be described as a hidden Markov model. The states are random samples with attached likelihood functions to represent the motion model in relation to the previous state. A new configuration is calculated by optimizing the likelihood functions [Comaniciu et al., 2003].

3.3.2 *Applications in 2D Ultrasound*

Many applications performing tracking in US make use of the speckle noise effect. Speckle is a characteristic pattern caused by sound reflection and attenuation (cf. Section 2.3.2). In general, it can be assumed as stable feature for motion correction if basic conditions remain the same, i.e., viewing angle of the US probe, position of organs reflecting US (cf. Section 3.2.2).

Yeung et al. [1998] propose a deformable mesh and use kernel-based tracking. The mesh is adapted to the temporal speckle pattern. Nodes are placed at positions where features are temporally stable and the mesh is refined in regions with high feature energy. The deformation of the mesh is regularized using Finite Elements to overcome deficiencies in the speckle signal, e.g., due to out-of-plane motion. In Revell et al. [2005] motion is estimated in US elastography datasets by measuring either NCC or CD_2 [Cohen and Dinstein, 2002] depending on the SNR of the local image signal. Coherent local image trajectories are enforced to obtain smooth motion and take account for outliers. Korstanje et al. [2010] also use kernel-based tracking with variant kernel shape and size to quantify tendon displacement by NCC in US with sub-pixel accuracy. Their goal is to drop the use of manual anatomic landmarks for exact motion determination.

Other work addresses the motion analysis of particular structures or organs. Bosch et al. [2002] introduce an active appearance model which allows for robust tracking and segmentation of the left ventricle in echocardiographic datasets. They integrate the characteristic shape of the organ and its variation, local image features for

edges learned from expert segmentations and continuity assumptions which hold for the adjacent frames acquired over time. Li et al. [2005] track the contour of the tongue using a Snake-based energy minimization. Instead of using homogeneity as criterion along edges, intensity profiles are used to account for US imaging specifics. Guerrero et al. [2007] incorporate an elliptical model to track vessels in US using Kalman filtering. Nascimento and Marques [2008] address the problem of heart motion tracking where the data are characterized by motion with abrupt changes, occluded features and weak boundaries. They suggest a multiple dynamic model to represent systolic (contracting contour) and diastolic (expanding contour) motion characteristics. The tracking problem is solved using different Kalman filtering steps for each model to find the parameters for a mixture of Gaussian.

3.4 TRACKING VS. REGISTRATION

The main difference between tracking and registration is the motivation behind the problem to be solved. Tracking is used to follow objects in a scene, often involving large changes in the appearance or context in which the object appears. Registration is employed to transform images to a common basis or to reduce dissimilarities between instances of images. However, both are closely related and when applied to medical applications differentiating between tracking and registration is difficult. This especially applies to kernel tracking and registration to assess motion.

The motion compensation of the proposed system will be referred to as a registration approach with an area-based similarity determination. The key aspect of this approach is to remove motion or compensate for motion influence while the amount and characteristic of the motion is of secondary concern. Additionally, elastic registration is well suited to be used for a physically-based transformation. An elastic transformation model is able to stabilize the motion compensation approach and reduce the computational costs, as correspondence does not have to be calculated for each image pixel but can be interpolated by the transformation function.

An important advantage of tracking is the integration of knowledge about shape or motion of objects in the scene. The system proposed in this work targets different applications of CEUS. Although the creation of a-priori models for different application areas (e.g., heart, kidney, bowel, pancreas) would be feasible, the acquisition of different US frames does not meet the same preconditions regarding the viewing angle, image resolution according to the depth of field and the location of objects. Especially different viewing angles pose a problem here because the same object may look very different from

different angles. As a consequence, a new model may be required for each CEUS acquisition.

Feature point approaches require that the scene can be represented by characteristic points. Motion influenced US data which are characterized by noise influence and remaining out-of-plane motion make it difficult to establish such characteristic feature points which remain stable over the sequence of frames.

The silhouette tracking approach is unfeasible for the application in most US data, because edge features are weak due to speckle, noise and acoustic shadowing [Toennies, 2012; Frangi et al., 2003].

Kernel-based tracking is very similar to image registration based on the minimization of an intensity difference. The new position of a local intensity profile of a location determined by tracking can also be used as transformation parameter to compensate motion. However, for tracking the locations or areas to be tracked have to be determined somehow. The approach used for the motion compensation in this thesis is therefore determined on the basis of image registration.

3.5   MOTION COMPENSATION IN PERFUSION IMAGING

In the previous section, motion compensation has primarily been presented for general medical applications and US data. However, motion influences also pose a problem in perfusion imaging. Perfusion kinetics monitored by perfusion imaging play an inherent role in diagnosis [Detre et al., 1992; Pandharipande et al., 2005]. In order to be able to extract valid quantitative perfusion parameters (cf. Section 2.4) from perfusion data by the help of computer assisted analysis, it is required that the same object or tissue is regarded in each instance of a time dependent recording.

As a consequence, motion should be removed or at least compensated as best as possible [Metz et al., 2010]. When tracking approaches or registration is applied to perfusion images, special attention has to be paid to the feature correspondence and the similarity determination [Melbourne et al., 2007].

A specific problem in perfusion imaging is that different structures might become visible at different times. Some modalities depict perfusion information solely, like the contrast sequence of CEUS data or NMI imaging. Other modalities, e. g., contrast-enhanced MRI, include morphological information. Morphological information is important for feature correspondence and similarity to get robust measures over the time sequence.

Feature correspondence and similarity calculation for perfusion imaging is handled differently, but two strategies can be differentiated [Metz et al., 2010]:

- approaches trying to minimize influence from perfusion effects and

- approaches introducing new knowledge to interpret the perfusion effects.

The first strategy assumes that perfusion information is not present in all areas and that sufficient morphological information for registration remains when ignoring areas of perfusion [Schäfer et al., 2011b]. This, however, requires that the modality measuring the perfusion also contains morphological information, e. g., CT or MRI.

Additional knowledge for the interpretation of perfusion is incorporated in the second strategy, e. g., through segmentation which is created based on pharmacokinetic model assumptions about perfusion [Hayton et al., 1997; Xiaohua et al., 2004]. Knowledge derived from such a perfusion model can also be directly used as part of the objective function for registration [Adluru et al., 2006]. This assumes that for well registered image data the perfusion time curves are smooth and thus can be well fitted to the perfusion model.

Both strategies are not feasible to be applied to contrast sequences of CEUS. The image frames in the contrast sequence do not show morphological information which can be seen when analyzing pre-contrast image frames. Additionally, there is almost no intensity information in the pre-contrast phase. Knowledge about perfusion is available, e. g., in terms of a pharmacokinetic perfusion model [Jiřík et al., 2012]. However, the model describes perfusion only for a specific area of tissue or organ. This is why differently parameterized model functions have to be used for different parts of the image. Therefore, the image sequence has to be divided in different compartments to establish model-based perfusion. This is unfeasible, especially when motion influence is still present.

As a consequence, the information to compensate for motion in CEUS is extracted from the B-mode sequence where constant intensity for corresponding tissue can be assumed over time.

3.6 SYSTEMS TO PROCESS CEUS IMAGE SEQUENCES

The main developers of US scanning devices also offer workstation software to acquire, display, organize and also to analyze US data afterwards. The LOGIQ® Series (GE Healthcare, Milwaukee, WI, USA) is equipped with software supporting the analysis of CEUS image sequences. BRACCO VueBox^{TM} Quantification Toolbox[2] is an analysis software supporting data from any manufacturer. It is specialized on the analysis of CEUS image sequences especially the extraction of per-

---

2 http://vuebox.bracco.ch/en/index.php, 01.07.2013

fusion quantification parameters. Workstation software from other manufacturers was not available for testing.

We tested both applications, the GE software on a LOGIQ® E9 machine and BRACCO's VueBox™ with a special focus on the analysis and compensation of motion. The GE software does offer the possibility to motion compensate acquired CEUS data by repositioning a ROI in each frame manually with mouse input. It is also possible to declare single frames or groups of consecutive frames as out-of-plane. These frames are not included when a TIC for a location over time is extracted. A standardized workflow is not provided and the manual motion correction is laborious and time consuming.

VueBox™ offers a rigid motion compensation tool. The user has to define a polygonal ROI for which the transformation is calculated. However, the tool does not offer parameters to control the process. When the motion compensation fails, the ROI can be adjusted in the hope that the changed ROI improves the result. A manual exclusion of out-of-plane frames for the analysis is possible. In addition, the toolbox provides numerous modeling techniques to approximate perfusion parameters, such as blood flow and mean transit time (cf. Section 2.4.1).

Both applications enable freehand drawing of ROI to extract TICs for specific regions. However, there is no assistance on generating those regions in terms of information about the underlying perfusion such as perfusion homogeneity or relative enhancement in percent. Only the B-mode and contrast data are shown.

The system proposed in this thesis aims at simplifying the steps for analysis and making them more robust towards inter- and intra-observer variability. This is possible by reducing the steps where excessive manual input is required, e. g., to align frames or search for out-of-plane motion in a sequence of CEUS image sequences. A standardized workflow is proposed.

3.7 SUMMARY

In this chapter, related work to the components of the CEUS analysis system has been presented. Two important strategies for the analysis of motion in time-dependent image data were described: registration and tracking. Representative work applying these strategies to problems in US were also presented. Subsequently, the strategies have been compared and reasons for the choice of registration to serve as tool for processing motion compensation of CEUS data have been listed. The next chapter describes the different parts of the analysis system for CEUS image sequences.

4

ANALYSIS OF CEUS IMAGE SEQUENCES

Difficulties arising when data from CEUS are analyzed have been discussed in Chapter 2 and Chapter 3. As the central part of thes thesis, this chapter proposes a prototype system. It consists of different subtasks which address the specific problems of CEUS in order to establish valid analysis results from acquired CEUS patient data. Evaluation of each subtask is described in Chapter 5.

At the beginning, the workflow and interactions of the different subtasks are outlined, requirements for the data that are processed by the system are stated and guidelines are established to define the application field. Below, the subtasks are explained in detail.

4.1 SYSTEM OVERVIEW

The system is divided into three principal parts, sequentially processing CEUS sequences of frames (see Figure 4.1):

- **Temporal Frame Selection**: The sequence of images is analyzed for different kind of motion. Frame sets are established forming subsets of frames which show the same area of interest, most notably not containing out-of-plane motion (cf. Section 2.3).

- **Motion Compensation**: Frame sets from the previous step are motion compensated in a two-step strategy addressing rigid and non-rigid motion, respectively.

- **Region Analysis**: Analysis regions are formed to derive perfusion characteristics from motion compensated data. The regions are robust against noise influence from US acquisition and residual motion artifacts.

4.1.1 *Data Acquisition Requirements*

CEUS acquisition is required to fulfill some preconditions so the system presented in this work is suitable to process the data. These requirements have been established in cooperation with a clinical partner in Bergen, Norway. At the time, US equipment from GE (LOGIQ® 9 and LOGIQ® E9) was used in clinical studies and clinical practice.

A key feature which has to be provided by the US device is the dual view technique in the CEUS acquisition mode. This means that
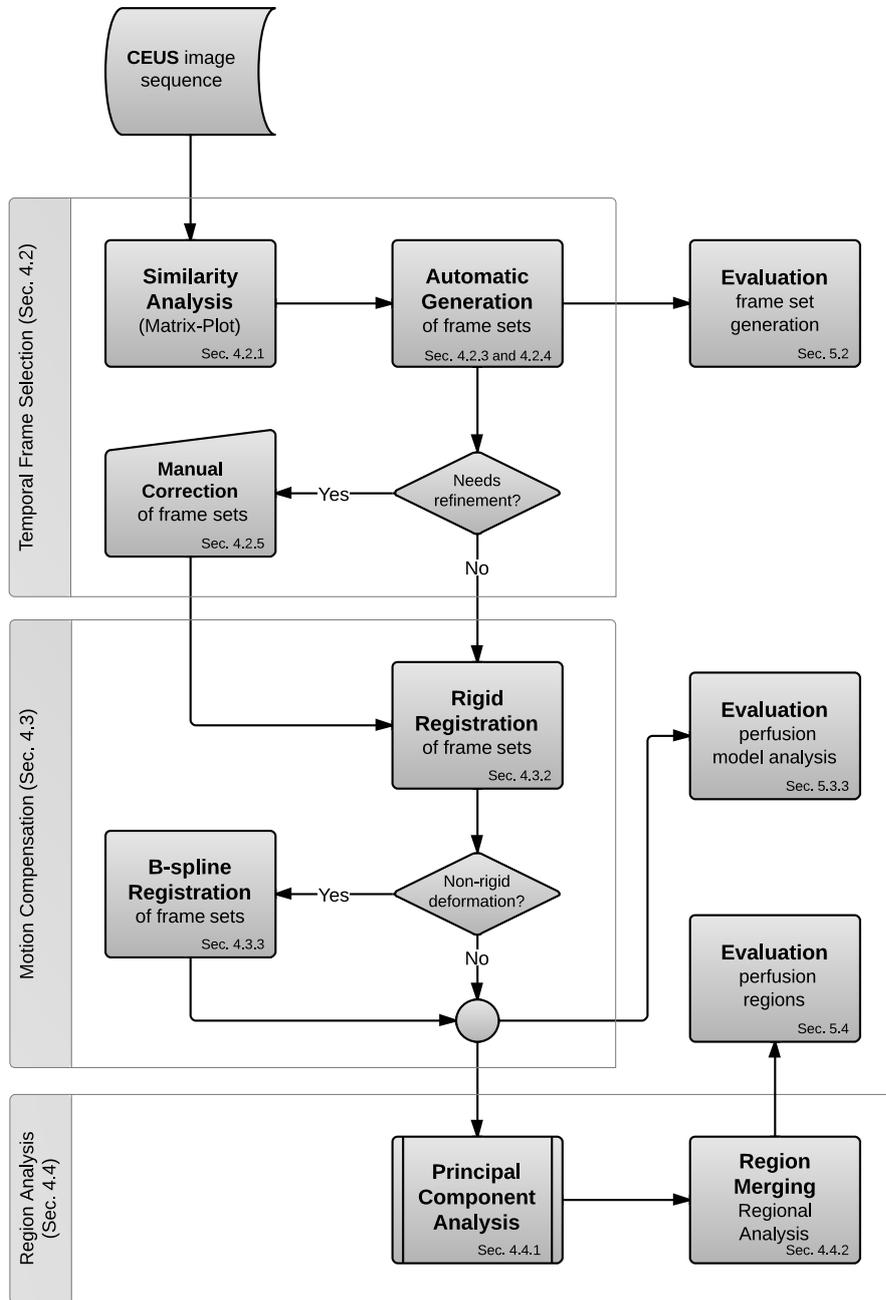
Figure 4.1: The workflow chart depicts the different tasks of the system and their relation grouped in three principal parts. Evaluation of each part will be addressed in Chapter 5.

both 2D B-mode and contrast mode images are recorded at the same time and that they are depicted side-by-side. The B-mode sequence is an important feature used in the temporal frame selection and the motion compensation process.

In order to transfer datasets and to load them into the system, the workstation software needs to provide an export function which includes both sequences, B-mode and contrast. Uncompressed formats, e. g., packed into DICOM exchange formats are preferred to maintain the original image quality.

In contrast to many other medical imaging technologies, CEUS data are acquired hand-held and thus, the quality strongly depends on the knowledge and expertise of the sonographer. The European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) has established different levels of training combined with recommendation guidelines to ensure standardized and correct diagnostic output [Piscaglia et al., 2012]. As a prerequisite for the proposed system, it is important that the sonographer has experience in the domain of application. This is necessary to choose an image plane where out-of-plane motion is less likely to occur. If the area of interest moves out-of-plane, the sonographer should be able to take countermeasures so the area of interest is visible again. Only this way, valid analysis results can be produced with the recorded sequence.

Additionally, parameters of the system are adjusted in order to be suitable for a frame rate of about 2-4 frames per second (FPS). This is the frame rate which has been chosen to be sufficient for diagnosis in the datasets of the studies presented later in this thesis. Processing of datasets with higher FPS is possible, however, a few parameters have to be adapted.

4.2 TEMPORAL FRAME SELECTION

In Section 2.3.3 and Section 3.2.2, the effect and problem of out-of-plane motion has been discussed. Due to properties and limitations in the acquisition of CEUS images, out-of-plane motion artifacts can hardly be avoided. Moreover, the acquisition cannot be arbitrarily often repeated because it is cost-intensive and implies additional strain for the patient. It is therefore desirable that analysis accounts for out-of-plane motion and extracts as much valid information as possible. Currently, common practice is to examine the sequence frame-by-frame and define frames exhibiting out-of-plane motion manually.

The temporal frame selection proposes to group frames according to the area they show (area of interest) or specific projection which is covered by the imaging plane of the US probe. These groups of frames are denoted as frame sets. It should be noted that such sets do not necessarily contain consecutive frames but that they also contain

frames from different areas of the sequence which are split by sets representing another area of interest. Resulting sets which combine less than three frames are discarded. They do not have any diagnostic relevance because perfusion effects cannot be evaluated over such a short time interval.

In order to set up frame sets, the type of motion (primarily the direction) and the amount of relative motion between frames of the sequence have to be evaluated. Section 3.3 presents different approaches to determine a motion field for continuously sampled data acquisition. These approaches require that features in the image are traceable. When out-of-plane motion is present, this assumption does not hold because structures, objects or parts of objects appear and disappear with respect to the imaging plane. Instead, the sequence is analyzed by similarity comparisons to find out if motion has taken place at all between all possible combinations of two frames of the sequence. Local characteristics about the motion are deduced from a matrix view.

4.2.1   *Similarity Matrix View*

A domain Ω and a set of intensities $\mathcal{W}$ is considered. The function domain Ω is integer-based and finite with the two spatial dimensions X, Y and a temporal dimension T representing the different frames in a sequence, in our case the CEUS sequence:

$$\Omega = \{1, ..., x_{max}\} \times \{1, ..., y_{max}\} \times \{1, ..., t_{max}\}. \tag{4.1}$$

Accordingly, x_{max}, y_{max} and t_{max} are the sizes of the respective dimensions of the image sequence. The set of normalized image intensities is defined as:

$$\mathcal{W} = [0, 1] \subset \mathbb{R}. \tag{4.2}$$

A B-mode image of a CEUS image sequence is now denoted as a function $I : \Omega \rightarrow \mathcal{W}$.

In order to build the similarity matrix view, inter-frame similarities have to be calculated. For n frames in a sequence, this will result in $n \times n$ similarity calculations. Results are stored in a matrix with n rows and n columns. The item at row i and column j holds the similarity of frame i and frame j. As the similarity is commutative, the matrix is mirrored along the diagonal.

As measure for similarity, NCC in the B-mode sequences is used. The B-mode data provide morphological information which, in general,

does not change over time. NCC yields the following definition of the similarity matrix:

$$M_{sim}(t, t') = \frac{\sum_{r \in \mathcal{R}} (I(r, t) \cdot I(r, t'))}{\sqrt{\sum_{r \in \mathcal{R}} I(r, t)^2 \cdot \sum_{r \in \mathcal{R}} I(r, t')^2}}, \qquad (4.3)$$

with $\mathcal{R}$ being the set of pixel coordinates $r = (x, y) \in \mathcal{R}$ of a spatial ROI focusing on parts with diagnostic relevance. x and y are coordinate points in dimension X and Y. t and t' are time points in dimension T. The ROI can be roughly defined in a representative frame and is then used in the whole sequence. This is also important to omit image information from the epithelium as there is no or almost no motion present because it lies directly under the US probe. When curved array transducers (cf. Section 2.3.1) are used, the resulting image area is not rectangular but fan-shaped. In this case, it is important to establish a ROI that masks out areas with no image information.

The advantage of NCC is that similarities are scaled between 0 and 1 and thus, are comparable across datasets. This is important for the users to learn to understand how much dissimilarity is present between frames in a dataset. The values of the similarity matrix can be depicted in an image if the similarities are assigned to brightness values. Values in the vicinity of the matrix diagonal represent local similarities within the neighborhood. Observation of this area is important to reveal abrupt changes in the similarity which can be a possible indication for motion. Areas in the matrix which are not directly located at the diagonal are important to find out if two separated frame sets depict the same area of interest, or in other words, if the area of interest has moved back into the image plane after motion disturbances.

The advantages of the similarity matrix view are manifold. First, the user can navigate to areas in the sequence where higher dissimilarities are reported to observe motion effects. Second, homogeneous areas (meaning areas with high similarity) in the matrix can be combined into a frame set and they do not have to be observed in detail by the user. Third, patterns in the similarity matrix reveal the type of motion which is present in the sequence.

4.2.2 *Types of Motion*

As stated earlier, motion in 2D US sequences is a result of different overlapping motion influences (cf. Section 2.3.3). However, in most cases there is a predominant influence depending on the area of investigation. The type of motion does not give information if the motion moves the area of interest out-of-plane but it helps a) to automatically find different categories of frames (or groups of frames)
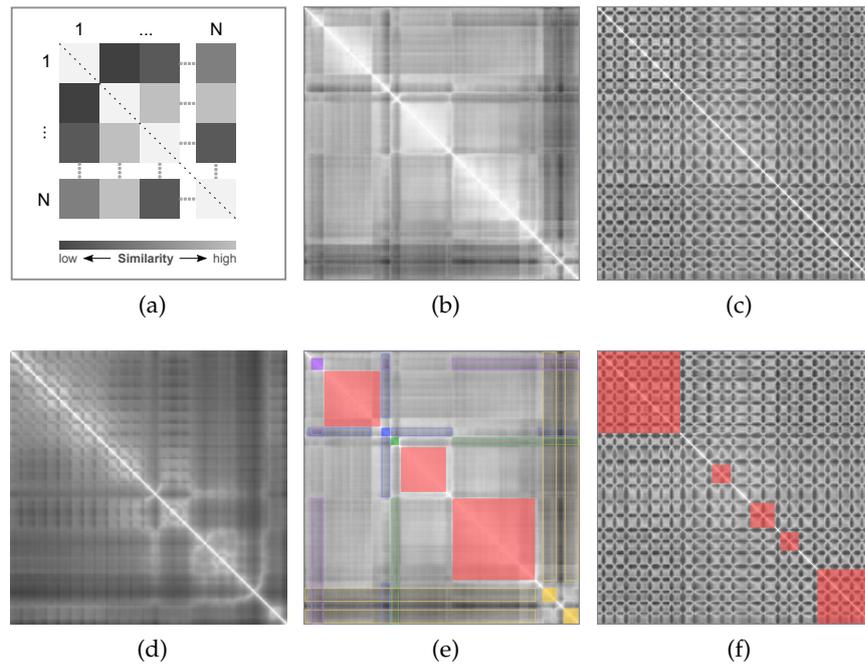
Figure 4.2: (a) schematic description of the generation of similarity matrices. (b) Similarity matrix of a CEUS sequence of the kidney. (c) Similarity matrix of a CEUS sequence of the pancreas. (d) Similarity matrix of another CEUS sequence of the pancreas. (e) Matrix from (b) with possible frame sets (filled) and in-between set-similarities (unfilled areas) indicated. (f) depicts the matrix from (c) with possible sets indicated.

and b) the user to get a quick and precise overview of the whole sequence. During the investigation of 2D US sequences of kidney, pancreas, small bowel and colon, two major types of motion can be distinguished:

- individually occurring motion shifts (Figure 4.2b)

- periodically reoccurring regular motion (Figure 4.2c)

For the first type, frame sets can be defined in the squared homogeneous areas between the areas of higher dissimilarity (Figure 4.2e filled regions). For each of the areas with higher dissimilarity the user should check if motion shifts are out-of-plane. If this is not the case, the frames can manually be assigned to one of the frame sets. If frame sets are defined, it should be checked which sets represent the same area of interest. This can be done by looking at between-set-similarities (indicated in Figure 4.2e by framed regions). For the second type, the pattern of the dissimilarities gives indication about the periodic motion. For an area with a regular occurring pattern, it must be decided if the periodic motion is in-plane or out-of-plane. If the motion occurs in-plane, the frame set can be defined over the whole

pattern area (Figure 4.2f). If not, the area can either be discarded or small sets with high similarity between frames can be generated.

There are similarity characteristics in datasets that cannot be assigned to one of the groups unambiguously. In this case, user investigation and interaction is required to define frame sets in the similarity matrix view. Figure 4.2d shows a mixture of both types of motion influence.

The definition of the frame sets is a tedious process, as a CEUS sequence has to be examined in detail by an expert. The similarity matrix view is able to identify areas where the scene changes. Mostly these changes are caused by motion. Thus, this simplifies manual definition of regions and also ensures that important information is not overlooked. The possibility of a semi-automatic and an automatic method to generate frame sets using the similarity matrix is investigated in this work. Such methods would standardize the process and save valuable time for medical experts. If a perfect result cannot be reached by the methods, it is interesting if at least a good initialization is provided which can be completed by the user.

Below, two methods are introduced to assist in the generation of frame sets. The semi-automatic approach uses a graph-based segmentation to find homogeneous regions in the similarity view. The fully automatic approach learns texture features to distinguish between in-plane and out-of-plane motion frames. Comparative results are shown in Figure 4.4.

4.2.3 *Semi-Automatic Frame Selection*

The semi-automatic frame selection uses the similarity correlation matrix as described above. In order to establish frame sets, a two step strategy is pursued:

1. Formation of local regions by grouping neighboring frames according to their similarity

2. Merging the groups of frames established in 1. if inter-regional similarity is high

Both steps make use of a simple but elegant technique originating from the image segmentation approach by Felzenszwalb and Huttenlocher [2004]. An important property of this approach is that it is able to find weak boundaries in low variability regions in image data. Additionally, it ignores them in high variability regions. Elements to be grouped into sets are interpreted as an undirected graph structure $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with vertices $v \in \mathcal{V}$ and edges $\langle v_i, v_j \rangle \in \mathcal{E}$. The elements are represented by vertices in $\mathcal{V}$ and edges in $\mathcal{E}$ express neighbor relations between elements. Each edge is assigned a weight

$w(e) \,\widehat{=}\, w(\langle v_i, v_j \rangle)$ which expresses the dissimilarity of the connected elements. Now, a segmentation $\mathcal{S}$ is being sought that divides $\mathcal{V}$ into regions or components C:

$$\mathcal{S} = \{C_1, C_2, ..., C_n\}, \quad C \subseteq \mathcal{V} \text{ with } C_i \cap C_j = \{\emptyset\} \text{ for } i \neq j. \quad (4.4)$$

Therefore, a predicate D is formulated to evaluate if there is a boundary between two components $C_1, C_2 \in \mathcal{S}$ or not:

$$D : \wp(\mathcal{V}) \times \wp(\mathcal{V}) \rightarrow \{\text{true}, \text{false}\}, \quad (4.5)$$

with $\wp(\mathcal{V})$ being the power set containing all possible subsets of $\mathcal{V}$. Two important pieces of information are required to calculate the predicate: the internal difference of a component and the difference between two components. The internal difference is defined as the maximum weight in the minimum spanning tree $MST(C, \mathcal{E})$ of a component:

$$Int(C) = \max_{e \in MST(C, \mathcal{E})} w(e). \quad (4.6)$$

The difference between two components C_1 and C_2 is defined as the minimum edge weight between the components:

$$Dif(C_1, C_2) = \min_{v_i \in C_1, v_j \in C_2, (v_i, v_j) \in \mathcal{E}} w(v_i, v_j). \quad (4.7)$$

The predicate to decide if there is a boundary between components checks whether the between-components-difference is larger than the difference within one of the components:

$$D(C_1, C_2) = \begin{cases} \text{true} & \text{if } Dif(C_1, C_2) > MInt(C_1, C_2) \\ \text{false} & \text{otherwise.} \end{cases} \quad (4.8)$$

MInt is the minimal internal difference of two components:

$$MInt(C_1, C_2) = \min(Int(C_1) + \tau(C_1), Int(C_2) + \tau(C_2)), \quad (4.9)$$

with a threshold function τ controlling the sensitivity of a boundary:

$$\tau(C) = \frac{k}{|C|}. \quad (4.10)$$

$|C|$ denotes the size of the component and k is the control parameter. For larger k, bigger components or segmentation regions are obtained in the end. For details about this segmentation method, please see Felzenszwalb and Huttenlocher [2004].

For step 1, the elements to be grouped are the frames of the sequence. So for n frames in an image sequence, there will be n vertices $\{v_1, v_2, ..., v_n\} = \mathcal{V}$ connected by $n - 1$ edges as each frame is connected to its successor:

$$\mathcal{E} = \{(v_i, v_{i+1}) \mid v_i \in \mathcal{V}, i = 1, ..., n - 1\}. \quad (4.11)$$

The edge weights w are taken from the similarity matrix M_{sim} (cf. Section 4.2.1):

$$w(v_i, v_j) = M_{sim}(i, j), \quad i, j = 1, ..., t_{max}. \tag{4.12}$$

The parameter k_1 (parameter k for step one) is user controlled and the segmentation result is computed in real time. So, the user gets immediate feedback in form of graphically displayed regions. The result $\mathcal{S}$ in form of the found components is used to initialize the second step.

For step two, the elements (vertices) to be grouped are the regions found in the first step:

$$\mathcal{V} = \mathcal{S}. \tag{4.13}$$

A new graph $\mathcal{G}(\mathcal{V}, \mathcal{E})$ is induced which is complete, i. e., fully connected:

$$\mathcal{E} = \left\{ (u, u') \in \mathcal{V} \times \mathcal{V} \mid u \neq u' \right\}. \tag{4.14}$$

The edge weights for step two will be calculated by averaging the dissimilarity measure between contained frame elements:

$$w(C_1, C_2) = \frac{1}{|C_1| \cdot |C_2|} \sum_{g \in C_1} \sum_{h \in C_2} M_{sim}(g, h), \tag{4.15}$$

where g and h are frames in the corresponding regions C_1 and C_2. The parameter k_2 (parameter k for step two) controls how many regions are merged together. In the second step, adjacent regions as well as remote regions can be merged. As in step one, the user gets immediate feedback when choosing the parameter k_2 and moreover, different combinations of parameters for the first and the second step can be tested at once. This enables the tool to be a fast method to initialize frame sets for motion compensation (cf. Figure 4.4b).

Note that this approach may generate frame sets representing different areas of interest depending on the parameter k_2. In clinical practice, the expert normally excludes all frames not belonging to the desired area of interest. The information that frame sets may represent different areas of interest has to be understood as recommendation for the user. One can either choose an area of interest and manually delete sets representing different areas or one can merge the sets by increasing the parameter k_2 or manually defining two sets to be merged in a subsequent correction step (cf. Section 4.2.5).

4.2.4 *Automatic Frame Selection*

As alternative to the user-controlled process that assists in finding frame sets in an US sequence, a method which does not need user

interaction is included in the system. The suggested method is based on texture feature analysis which is used to distinguish between in-plane and out-of-plane frames. A classifier algorithm is trained with ground truth data and is then able to decide whether a frame is declared in-plane or out-of-plane. In-plane frames are characterized by homogeneous areas in the similarity matrix or a regular pattern representing periodic motion. For the automatic method, it is assumed that an area of a regular pattern (see Figure 4.2) is regarded as a whole region. I. e., the periodic motion shall be removed by motion compensation. This assumption is important when training a classifier either to declare periodic motion as in-plane or out-of-plane class.

Two different texture features are considered to analyze the local neighborhood in the similarity matrix of a frame: Haralick texture features and a 2D Gabor filter. Haralick features have been chosen as representative of local features in the spatial domain. The 2D Gabor filter is representative for feature computation in the frequency domain.

Haralick features are computed using a co-occurrence matrix of pixel intensities [Haralick et al., 1973; Haralick, 1979]. Therefore, the local neighborhood around the frame in the similarity matrix is compared to a shifted version of the local neighborhood in both directions X and Y. Each co-occurring intensity combination increments the respective element entry of the co-occurrence matrix. Haralick et al. [1973]; Stachowiak et al. [2005] suggest different features which are calculated from the co-occurrence matrix. In preliminary tests, the different features were tested on similarity matrices of CEUS image data. The cluster shade from Equation (4.16) and cluster prominence from Equation (4.17) suggested by Arivazhagan and Ganesan [2003] showed most promising results to distinguish between in-plane and out-of-plane frames:

$$CS = \sum_{i,j=1}^{N} (i - M_x + j - M_y)^3 \cdot COM(i,j), \tag{4.16}$$

$$CP = \sum_{i,j=1}^{N} (i - M_x + j - M_y)^4 \cdot COM(i,j), \tag{4.17}$$

where COM is the co-occurrence matrix of size N, $M_x = \sum_{i,j=1}^{N} i \cdot COM(i,j)$ and $M_y = \sum_{i,j=1}^{N} j \cdot COM(i,j)$. The offset of the shifted version is varied between 4 and 17 pixels with two intermediate steps. According to the requirement of 2-4 FPS (Section 4.1.1), this covers a time range between 1 and 8 seconds. This is well suited to address the detection of periodic motion induced by the heart beat or breath-

ing. This setup leads to four different feature values for CS and four different values for CP.

The Gabor filter is a linear filter applied to an image [Jaini and Farrokhnia, 1991; Unser, 1995]. The filter kernel (wavelet) is constructed using a sine function with given direction, amplitude, frequency and phase shift multiplied with a Gaussian function according to Equation (4.18).

$$g(v, w, \lambda, \theta, \psi, \sigma) = \exp\left(-\frac{h1(v, w, \theta)^2 + h2(v, w, \theta)^2}{2\sigma^2}\right) \cdot \cos\left(2\pi \frac{h1(v, w, \theta)}{\lambda} + \psi\right) \tag{4.18}$$

$$h1(a, b, \theta) = a\frac{2}{M-1}\cos(\theta) + b\frac{2}{M-1}\sin(\theta) \tag{4.19}$$

$$h2(a, b, \theta) = -a\frac{2}{M-1}\sin(\theta) + b\frac{2}{M-1}\cos(\theta) \tag{4.20}$$

v and w are the coordinates of the Gabor kernel, λ is the wave length, θ represents the orientation of the wave, ψ the phase offset and σ is the standard deviation of the Gaussian function.

By the help of these parameters, specific characteristics can be extracted from images by applying a Gabor filter kernel and checking the mean and standard deviation (SD) of the filter response. Four gabor kernels are created by varying the wave length λ between 0.05 and 0.5, four other kernels are created by varying the SD of the Gaussian function σ between 5 and 15 and four kernels are created by varying the phase shift ψ between 0 and 180. This leads to 12 Gabor features overall Figure 4.3.

A local neighborhood of 30 frames is regarded when calculating texture features on the similarity matrix. The 20 overall features (12 Gabor kernels and 8 Haralick features) are used to train models for multiple classifiers using the WEKA data mining software [Hall et al., 2009]. For training, frame sets manually defined by a domain expert using the similarity matrices are used from 40 datasets. Representatives of the most prominent groups of classifiers found in literature are chosen and standard parameterizations from WEKA are used [Wu et al., 2007]:

- **Basic:** k-nearest neighbors classification (KNN), Naive Bayes classification,

- **Boosting:** Logit boost [Friedman et al., 2000], AdaBoost [Bauer and Kohavi, 1999],

- **Tree-based:** Alternating decision tree [Freund and Mason, 1999], C4,5 [Quinlan, 1993],

- **Support vector machine (SVM)** [Burges, 1998].
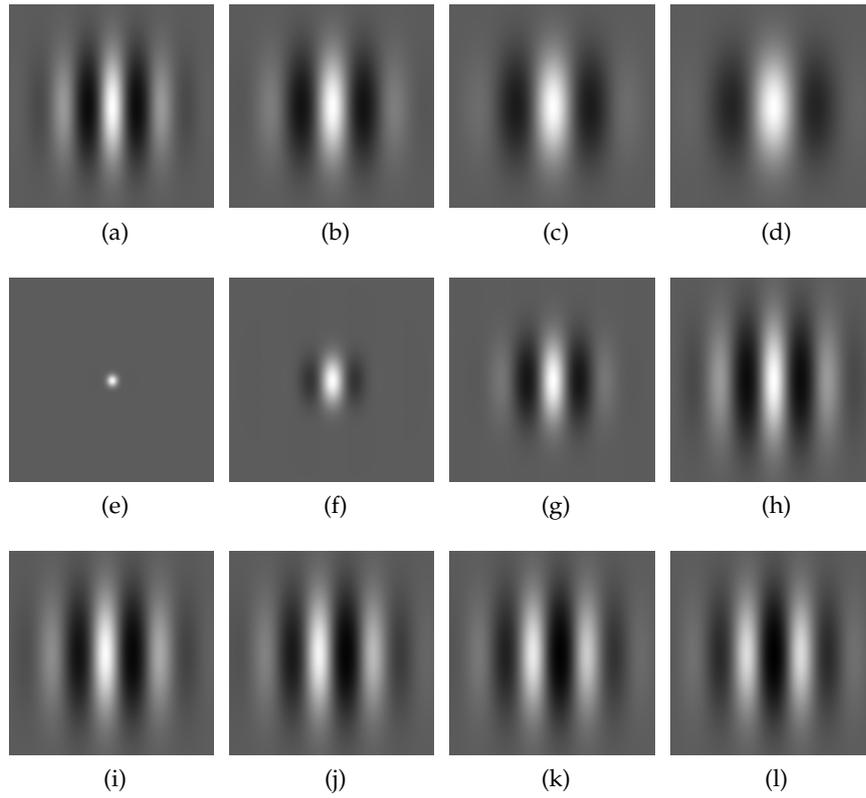
Figure 4.3: The applied 12 Gabor filter kernels. In the first row, λ is set between 0.05 and 0.5 from left to right. In the second row σ is set between 5 and 50 from left to right and in the third row ψ is set between 45 and 180 degrees from left to right. θ is set to 0.

For classification of a new dataset for automatic frame selection, a voting approach is used incorporating the results of all classifiers listed above [Kittler et al., 1998]. For the voting, the majority vote rule is applied, i. e., a sample is assigned to the class with the majority of the votes from the different classifiers. In this way, the risk of misclassification due to a single classifier performance and its parameterization is minimized. This is important, because the data which are originating from the similarity matrix are subject to change with newly acquired datasets.

Another problem is the imbalance of the two classes *out-of-plane* and *in-plane*. The training data contain much more samples (6 to 8 times) for the *in-plane* class. To overcome this problem, three different techniques are worth considering: subsampling, a combination of super- and subsampling and a cost-weighting approach [He and Garcia, 2009]. The first approach just subsamples the majority class. In ther super- and subsampling combination, the minority class is supersampled by generating random samples between existing samples and the majority class is subsampled so that both classes are

| (a) | (b) | (c) |
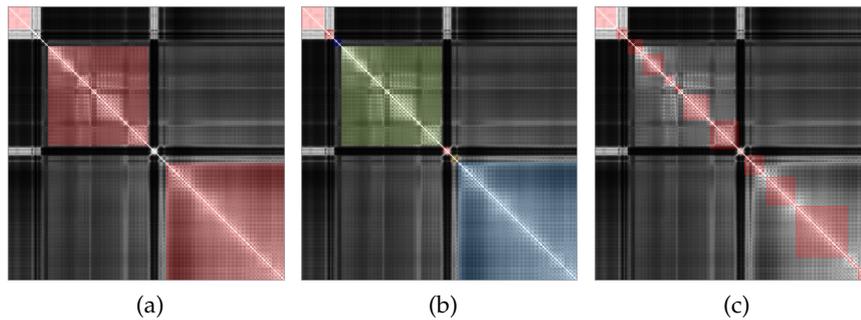
Figure 4.4: The ground truth frame set for a CEUS acquisition of the pancreas defined by a medical expert is depicted in (a). The result generated by the semi-automatic approach is shown in (b) and the result of the automatic approach is shown in (c).

represented by an equal number of representatives. Cost-weighting introduces different misclassification costs for the classes to eliminate a bias.

After performing preliminary tests, we found that the combination of super- and subsampling leads to the most promising results. Therefore, this technique is used to compensate for the imbalance. An exemplary result is shown in Figure 4.4c.

4.2.5 *Manual Selection and Correction*

After using the semi-automatic or the automatic method to generate frame sets the user has the possibility to correct the suggested sets. An interactive version of the similarity matrix view allows to quickly navigate to a frame for visual inspection. Moreover, existing sets can be resized, moved, merged or deleted. New sets can be interactively drawn onto the similarity matrix. In this way, a correct selection of temporal frames for the subsequent motion compensation is guaranteed by user supervision.

4.3 MOTION COMPENSATION

After the frame sets have been selected, motion compensation is applied to the CEUS sequences (cf. Figure 4.1). The motion compensation is composed of two steps conducted consecutively:

- Translation motion compensation: globally occurring motion shifts in X- and Y-direction are compensated

- Non-rigid motion compensation: non-linear motion effects are compensated using B-spline functions

Both strategies are applied independently from each other. The rigid step addresses elementary motion effects which are equally present in the whole ROI to be motion compensated and should be applied first. The non-rigid step addresses local motion effects and deformation, particularly those caused by intrinsic patient motion, such as heart beat, aorta pulsation or digestive activity. As a basis for motion compensation, the frame sets defined in the previous step are used. For each frame set, a fixed frame is declared as reference for the registration procedures. The fixed frame is defined using the similarity matrix data from Section 4.2.1 by searching the highest similarity of a region member to all other frame members of that region. This attempts to keep the extent of the overall transformations minimal.

For both steps, a priori assumptions about the motion are established. These assumptions concern the dependency of the parameters organized within a motion field. This is the temporal stability of motion within a local neighborhood of frames. For the second step targeting non-rigid motion, this is also spatial smoothness of motion within frames. One way to model these assumptions is to use an MRF. An MRF is a system of random variables which are conditionally independent except for those parameters connected with each other [Durrett, 2010]. Additionally, the size of displacement being accounted for by motion compensation should be bounded. The MRF is used for optimization of discrete values so the assumptions about the solution space can be integrated easily.

The concept of an MRF is described below. It allows to optimize an objective function and intuitively integrating the constraints described above. Afterwards, the application of MRF for the motion compensation steps for rigid and non-rigid motion effects is presented in detail.

4.3.1 *Optimization using MRF*

Numerous methods in computer vision use energy minimization strategies (cf. Section 3.1 and Section 3.3) to solve a vision problem such as segmentation, image restoration or registration [Veksler, 1999]. Energy in this case can be viewed as a quality measure for different possible solutions of a problem. The quality of a solution can be determined in different ways. First, a disagreement between the observed data and the solution-generated data can be assessed. Let us denote this as the data energy, E_{data}. Second, the composition and state of the solution can be subject to special rules or assumptions which are known a priori. Let us denote this as the prior energy, E_{prior}. The incorporation of prior model knowledge helps to overcome uncertainties stemming from the data measurement (cf. Section 2.3). The overall objective function E_{global} can be described

as a weighted sum of E_{data} and E_{prior} with a weighting parameter η:

$$E_{global} = E_{data} + \eta \cdot E_{prior}. \tag{4.21}$$

In case of segmentation or image restoration, E_{data} is measured by defining an accordance of pixel intensity to a segment definition or the proximity to a label [Perez et al., 1998]. For registration, E_{data} is determined by measuring the similarity between the transformed image using the current solution and a reference image [Glocker et al., 2011].

There are many possibilities to establish prior knowledge about the composition of the solution. Spatially piecewise smooth intensities or segment labels are typically assumed for segmentation and restoration. For image registration, the solution can be regularized to achieve a piecewise smooth motion vector field in spatial and temporal direction.

Let us consider the generation of a solution for Equation (4.21) as a labeling problem [Li, 1994]. Then, each variable of the system is represented as a vertex $p \in \mathcal{V}$ in an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with edges $e \in \mathcal{E}$ connecting exactly two vertices $e = \langle p, q \rangle, \quad p, q \in \mathcal{V}$. A label $l \in \mathcal{L}$ is assigned to each vertex from $\mathcal{V}$ randomly. The configuration is represented by the random function F, mapping the vertices to the assigned labels. F takes values from the function space $\mathcal{L}^{\mathcal{V}}$. As a consequence, $F(p)$ is the random variable describing the label of vertex $p \in \mathcal{V}$. Let us denote $F(p)$ as F_p for reasons of brevity. The domain of F can be restricted to a new domain of vertices $\mathcal{X} \subseteq \mathcal{V}$:

$$F_{\mathcal{X}} = F|_{\mathcal{X}} \ \forall \mathcal{X} \subseteq \mathcal{V}. \tag{4.22}$$

$f \in \mathcal{L}^{\mathcal{V}}$ defines a specific labeling, called a configuration, with $f_p = f(p)$ giving the current label at node p. f can also be restricted, so that $f_{\mathcal{X}} = f|_{\mathcal{X}} \ \forall \mathcal{X} \subseteq \mathcal{V}$ (cp. Equation (4.22)).

For F to be an MRF, the following two conditions have to be met:

$$P(f) > 0 \quad \forall f \in \mathcal{L}^{\mathcal{V}} \tag{4.23}$$

$$P\left(f_p \mid f_{\mathcal{V}-\{p\}}\right) = P\left(f_p | f_{\mathcal{N}_p}\right) \tag{4.24}$$

$P(F = f)$ is abbreviated as $P(f)$. The neighborhood system $\mathcal{N}$ states the neighbors of each vertex $p \in \mathcal{V}$ such that

$$\mathcal{N}_p = \{q \in \mathcal{V} \mid \langle p, q \rangle \in \mathcal{E} \vee \langle q, p \rangle \in \mathcal{E}\} \quad \forall p \in \mathcal{V}. \tag{4.25}$$

The first condition (Equation 4.23) exists to assure positivity also for the joint probabilities [Besag, 1974]. The second condition (Equation 4.24) states that the random variable f_p of a vertex p only depends on its neighbor vertices in $\mathcal{N}_p$.

The Hammersley-Clifford theorem proves that the probability distribution of an MRF can be expressed by a Gibbs distribution [Besag, 1974; Clifford, 1990]. Consequently, F is an MRF if $P(f)$ follows a Gibbs distribution:

$$P(f) = \frac{1}{Z} \cdot exp\left(-\sum_{c \in \mathscr{C}} V_c(f)\right).$$ (4.26)

$\mathscr{C}$ is the set of cliques in the graph $\mathcal{G}$, Z is a normalizing constant and $V_c(f)$ calculates the energy for the configuration f. In computer vision tasks, mostly cliques of size one and two are used, denoted as $\mathscr{C}_1$ and $\mathscr{C}_2$. This describes a neighborhood system of pairwise interaction.

For the application to be an energy minimization problem, the clique potentials for the cliques $\mathscr{C} = \mathscr{C}_1 \cup \mathscr{C}_2$ need to be established. The energy potentials of cliques of size 1 are defined by the E_{data} term which is associated with the disagreement measure D_p between data and labels:

$$E_{data}(f) = \sum_{c_1 \in \mathscr{C}_1} V_{c_1}(f) = \sum_{p \in \mathcal{V}} D_p(f_p).$$ (4.27)

The energy potentials of cliques of size 2 for pairwise interaction are defined by the E_{prior} term. E_{prior} is representing the prior knowledge (energy) $E(\cdot, \cdot)$ only associating dependencies of two vertices in the neighborhood system, which is represented in $\mathcal{E}$:

$$E_{prior}(f) = \sum_{c_2 \in \mathscr{C}_2} V_{c_2}(f) = \sum_{\langle p,q \rangle \in \mathcal{E}} E(f_p, f_q).$$ (4.28)

$\mathcal{L}$ contains a discretized set of possible values. This helps to bound the solution space to represent only plausible values. In order to get maximum probability for a configuration f, the energy term E_{global} needs to be minimized.

For that, different approaches have been proposed. Iterated conditional modes (ICM) locally optimize the energy at a vertex p by finding the best label l according to the best energy decrease [Besag, 1974]. The process is repeated iteratively until the global energy converges. This is a greedy strategy resulting in a local minimum and thus, the solution is highly depending on initialization [Boykov et al., 2001]. In contrast, simulated annealing, originating from an analogy in metallurgy, excepts an increase of the energy with a certain probability [Geman and Geman, 1984]. This probability is controlled by a "temperature" parameter, whose value decreases with time preventing the global energy to increase again. This technique allows the optimization to leave the area of a local minimum. However, the result is highly depending on the temperature. Consequently, the risk

of getting stuck in another local minimum is high and it is compara-
tively expensive to be computed.

In recent years, energy minimization with graph cuts has become
very efficient in computation and accuracy terms [Kolmogorov and
Zabih, 2004; Kohli and Torr, 2005; Szeliski et al., 2008].

Graph Cuts for MRF optimization

For graph cut-based energy minimization, two extra vertices are in-
serted into $\mathcal{V}$ as terminal vertices α and β. For now, the graph cut
problem is considered as binary decision or two label problem. The
terminal vertices represent two different labels $l_\alpha, l_\beta \in \mathcal{L}$. All ver-
tices in the graph except the terminal vertices are denoted as $\mathcal{V}_s$. The
edges connecting the vertices in $\mathcal{V}_s$ to form the neighborhood sys-
tem according to the application problem are denoted $e_s \in \mathcal{E}_s$. Each
edge $e_s = \langle p, q \rangle$ is assigned a weight value through a function $w(e_s)$
which calculates the prior energy $E(f_p, f_q)$. The terminal vertices α
and β are connected to each vertex in $\mathcal{V}_s$ through the terminal edges
$\langle \alpha, p \rangle$ and $\langle \beta, p \rangle$ with $p \in \mathcal{V}_s$. The set of terminal edges is called
$\mathcal{E}_t$. The edge weights $w(e_t)$ for terminal edges $e_t \in \mathcal{E}_t$ are defined
by the disagreement measure D_p, whereas edge weights $w(e_s)$ for
two connected vertices by an edge $e_s \in \mathcal{E}_s$ are defined by the prior
knowledge E.

A graph cut is solved by finding a cut $\mathcal{C}$ (of edges) in the graph
which is minimal with regard to the assigned edge weights or cost
values. Furthermore, a cut $\mathcal{C}$ has to contain exactly one edge to a
terminal vertex for each vertex in $\mathcal{V}_s$. Considering two vertices p and
q connected by an edge, the following possibilities exist:

$$\mathcal{C} = \bigcup_{e_s = \langle p, q \rangle \in \mathcal{E}_s} \begin{cases} \{\langle \alpha, p \rangle, \langle \alpha, q \rangle\} \\ \{\langle \beta, p \rangle, \langle \beta, q \rangle\} \\ \{\langle \beta, p \rangle, \langle \alpha, q \rangle, e_s\} \\ \{\langle \alpha, p \rangle, \langle \beta, q \rangle, e_s\} \end{cases} . \qquad (4.29)$$

If $\langle \alpha, p \rangle \in \mathcal{C}$ then p is assigned the label l_α. The same applies to label
l_β and the connected terminal edges. If the minimum cut includes
terminal edges of different labels for two vertices p and q both in
$\mathcal{V}_s$ and $e_s = \langle p, q \rangle \in \mathcal{E}_s$, the connecting edge e_s has to be contained
in the cut as well and its weight is contained in the energy of the
minimum cut (see Figure 4.5). p and q are not assigned the same
labels.

Boykov et al. [2001] propose an efficient way to solve MRFs via
graph cut techniques using an arbitrary number of labels. They in-
troduce the concept of α-β *swap* moves and *α-expansion* moves. A
move is a change of the label configuration for which a new energy
evaluation is conducted. To be able to solve multi-label problems,
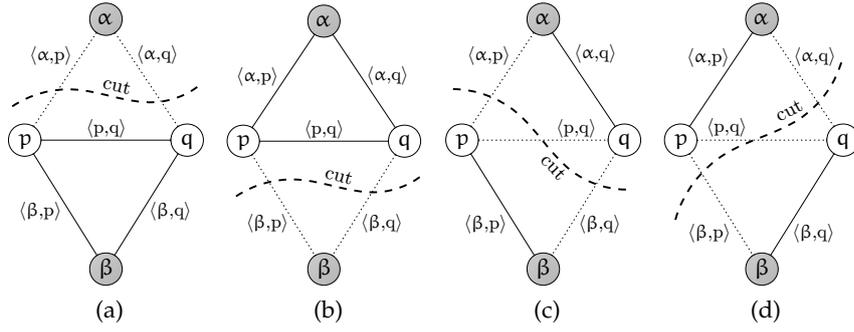
Figure 4.5: Possibilities for a cut $\mathcal{C}$ on a graph $\mathcal{G}$ for two connected vertices $\langle p, q \rangle \in \mathcal{E}_s$. The dashed edges are contained in $\mathcal{C}$ and express membership of the vertices represented by p and q to the label of the corresponding terminal vertex (adapted from [Boykov et al., 2001]).

subsequent move cycles are performed for each label pair (α-β *swap*) and each label (α-*expansion*), respectively. Methods to achieve an optimal move in terms of energy minimization of the current cycle are presented by Boykov et al. [2001]; Boykov and Kolmogorov [2004]. The approach uses either α-β *swap* moves or α-*expansion* moves to solve the MRF problem both using different intermediate graphs.

α-β *swap* moves change labels of a labeling f of vertices with label l_α to l_β and vice versa yielding labeling f'. Vertices with a label different from l_α and l_β in the current labeling f are not included in the graph for the current cycle and keep their current label:

$$\left\{ p \mid f_p \neq l_\alpha, l_\beta \right\} = \left\{ p \mid f'_p \neq l_\alpha, l_\beta \right\}. \tag{4.30}$$

For each α-β *swap* cycle, a new graph cut problem with a new graph is constructed for the two specific labels l_α and l_β according to Figure 4.5. The weight $w(\cdot)$ for all $e_t \in \mathcal{E}_t$ and the edges $e_s \in \mathcal{E}_s$ is defined to account for interaction between all labels (Table 4.1). The minimum cut for a move is determined by using a min-cut/max-flow algorithm [Boykov and Kolmogorov, 2004]. Move cycles are performed for each label pair until the energy of the labeling defined by the minimum cut $f^{\mathcal{C}}$ cannot be improved anymore.

α-*expansion* moves change a labeling f to a new labeling f' by assigning the label l_α to a subset of vertices, i. e.:

$$\begin{align} \{p \mid f_p = l_\alpha\} &\subseteq \{p \mid f'_p = l_\alpha\} \quad \text{and} \tag{4.31} \\ \{p \mid f_p \neq l_\alpha\} &\supseteq \{p \mid f'_p \neq l_\alpha\}. \tag{4.32} \end{align}$$

For the α-*expansion* moves, a different graph structure has to be established [Boykov et al., 2001]. The expansion move technique can only be used if the prior energy term E_{prior} from Equation (4.28) is a metric. For each cycle, a new graph $\mathcal{G}_\alpha = \langle \mathcal{V}_\alpha, \mathcal{E}_\alpha \rangle$ is constructed

Table 4.1: Each edge in the graph $\mathcal{G}$ is assigned a weight so that a minimum cut can be determined. $\mathcal{N}_p$ contains the neighboring vertices of p according to Equation (4.25) (adapted from [Boykov et al., 2001]).

edge	weight	condition
$\langle \alpha, p \rangle$	$D_p(l_\alpha) + \sum\limits_{\substack{q \in \mathcal{N}_p \\ f_q \neq l_\alpha, l_\beta}} E(l_\alpha, f_q)$	$p \in \mathcal{V}_s, f_p = l_\alpha \vee f_p = l_\beta$
$\langle \beta, p \rangle$	$D_p(l_\beta) + \sum\limits_{\substack{q \in \mathcal{N}_p \\ f_q \neq l_\alpha, l_\beta}} E(l_\beta, f_q)$	$p \in \mathcal{V}_s, f_p = \alpha \vee f_p = l_\beta$
$\langle p, q \rangle$	$E(l_\alpha, l_\beta)$	$\langle p, q \rangle \in \mathcal{E}_s$ $f_p = l_\alpha \vee f_p = l_\beta$ $f_q = l_\alpha \vee f_q = l_\beta$

and the current labeling f is used as input (see Figure 4.6). The set of vertices for the graph is defined as

$$\mathcal{V}_\alpha = \{\alpha, \bar{\alpha}\} \quad \cup \quad \mathcal{V}_s \quad \cup \quad \bigcup_{\substack{\langle p,q \rangle \in \mathcal{E}_s \\ f_p \neq f_q}} a_{p,q}. \tag{4.33}$$

In addition to the terminal vertices α and $\bar{\alpha}$ and the vertices of the MRF for each pair of connected vertices p and q in the MRF having different labels ($f_p \neq f_q$), an auxiliary vertex $a_{p,q}$ is inserted. The vertices in $\mathcal{V}_\alpha$ are connected through edges in

$$\mathcal{E}_\alpha = \bigcup_{p \in \mathcal{V}} \{\langle \alpha, p \rangle, \langle \bar{\alpha}, p \rangle\} \quad \cup \quad \bigcup_{\substack{\langle p,q \rangle \in \mathcal{E}_s \\ f_p \neq f_q}} \mathcal{E}_{p,q}$$

$$\cup \quad \{\langle p, q \rangle \in \mathcal{E}_s \mid f_p = f_q\}, \tag{4.34}$$

with a triplet of edges

$$\mathcal{E}_{p,q} = \{\langle p, a_{p,q} \rangle, \langle a_{p,q}, q \rangle, \langle \bar{\alpha}, a_{p,q} \rangle\}. \tag{4.35}$$

$\mathcal{E}_\alpha$ comprises the terminal edges $\langle \alpha, p \rangle$ and $\langle \bar{\alpha}, p \rangle$ from all $p \in \mathcal{V}_s$ to both terminal vertices α and $\bar{\alpha}$ (see Figure 4.6). For vertices p and q having different labels ($f_p \neq f_q$) it contains the edges $\langle p, a_{p,q} \rangle$ and $\langle a_{p,q}, q \rangle$ to the corresponding auxiliary vertex $a_{p,q}$ and edges from the auxiliary vertex to the terminal vertex $\langle \bar{\alpha}, a_{p,q} \rangle$. It also contains the edges $\langle p, q \rangle$ between all vertices having the same label ($f_p = f_q$).

Note that the label $\bar{\alpha}$ represents the initial label of a vertex and can be read as 'not α'. All edges in the graph are assigned edge weights $w(\cdot)$ which are listed in Table 4.2.
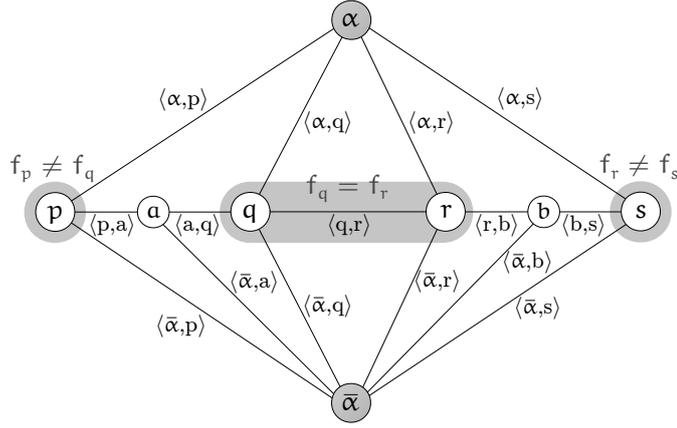
Figure 4.6: Exemplary, a graph $\mathcal{G}_\alpha$ is shown to solve a problem for four variables p, q, r and s in a 1D image. q and r currently share the same label, whereas the label from p is different from q and the label from r is different from the label of s. Auxiliary vertices $a \hat{=} a_{p,q}$ and $b \hat{=} a_{r,s}$ are appropriately inserted into the graph between the vertices with different label (adapted from [Boykov et al., 2001]).

Again, a cut $\mathcal{C}$ on the graph $\mathcal{G}_\alpha$ must exactly contain one terminal edge for each MRF vertex $p \in \mathcal{V}$. The new labeling $f_p^{\mathcal{C}}$ derived from the cut $\mathcal{C}$ is defined as:

$$f_p^{\mathcal{C}} = \begin{cases} l_\alpha & \text{if} \quad \langle \alpha, p \rangle \in \mathcal{C} \\ f_p & \text{if} \quad \langle \bar{\alpha}, p \rangle \in \mathcal{C} \end{cases} \quad \forall p \in \mathcal{V} . \tag{4.36}$$

For each pair of adjacent vertices, edges to be contained in the cut $\mathcal{C}$ have to be determined. For two vertices p and q sharing the same label ($f_p = f_q$), this can be solved analogous to the solution in Figures 4.5a to 4.5d, except that label l_β is substituted by label $\bar{\alpha}$. For two vertices p and q having different input labels ($f_p \neq f_q$), different possibilities to include edges in the cut $\mathcal{C}$ exist (Figure 4.6 and Figure 4.7):

$$\mathcal{C} = \bigcup_{\substack{\{p,q\} \in \mathcal{E}_s \\ f_p \neq f_q}} \begin{cases} \{\langle \alpha, p \rangle, \langle \alpha, q \rangle\} \\ \{\langle \bar{\alpha}, p \rangle, \langle \bar{\alpha}, q \rangle, \langle \bar{\alpha}, a_{p,q} \rangle\} \\ \{\langle \bar{\alpha}, p \rangle, \langle p, a_{p,q} \rangle, \langle \alpha, q \rangle\} \\ \{\langle \alpha, p \rangle, \langle a_{p,q}, q \rangle, \langle \bar{\alpha}, q \rangle\} \end{cases}$$

$$\cup \bigcup_{\substack{\{q,r\} \in \mathcal{E}_s \\ f_q = f_r}} \begin{cases} \{\langle \alpha, q \rangle, \langle \alpha, r \rangle\} \\ \{\langle \bar{\alpha}, q \rangle, \langle \bar{\alpha}, r \rangle\} \\ \{\langle \bar{\alpha}, q \rangle, \langle \alpha, r \rangle, \langle q, r \rangle\} \\ \{\langle \alpha, q \rangle, \langle \bar{\alpha}, r \rangle, \langle q, r \rangle\} \end{cases} , \tag{4.37}$$

leading to a different local energy for the current cycle. The minimum cut for a move on the graph $\mathcal{G}_\alpha$ is also determined by using a

Table 4.2: Each of the edges in the graph $\mathcal{G}_\alpha$ is assigned a weight so a minimum cut can be determined (adapted from [Boykov et al., 2001]).

edge	weight	condition
$\langle \bar{\alpha}, p \rangle$	∞	$f_p = l_\alpha$
$\langle \bar{\alpha}, p \rangle$	$D_p(f_p)$	$f_p \neq l_\alpha$
$\langle \alpha, p \rangle$	$D_p(l_\alpha)$	$p \in \mathcal{V}_s$
$\langle p, a_{p,q} \rangle$	$E(f_p, l_\alpha)$	$\langle p, q \rangle \in \mathcal{E}_s, f_p \neq f_q$
$\langle \bar{\alpha}, a_{p,q} \rangle$	$E(f_p, f_q)$	
$\langle p, q \rangle$	$E(f_p, l_\alpha)$	$\langle p, q \rangle \in \mathcal{E}_s, f_p = f_q$

(a)   (b)   (c)   (d)
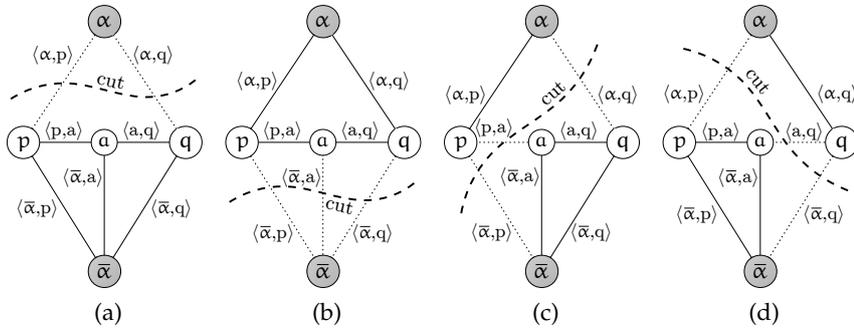
Figure 4.7: Possibilities for a cut $\mathcal{C}$ on a graph $\mathcal{G}_\alpha$ for *α-expansion* for two connected vertices p, q. The dashed edges are contained in $\mathcal{C}$ and express membership of the vertices p and q, $\langle p, q \rangle \in \mathcal{E}_s$ in the label of the corresponding terminal vertex (adapted from [Boykov et al., 2001]).

min-cut/max-flow algorithm [Boykov and Kolmogorov, 2004]. Move cycles for each label $l \in \mathcal{L}$ are performed until the global energy of a labeling defined by the minimum cut $f^{\mathcal{C}}$ cannot be improved any further.

The proposed algorithms for both move types are optimal with respect to a single move (on a local vertex neighborhood). I. e., the final solution is within bounds to the global optimum. For *α-expansion* moves the solution is proven to be within the factor of two of the global optimum [Boykov et al., 2001].

In comparison, ICM uses standard moves which only affect a single vertex in the MRF before a new energy evaluation. This is slow, the result relies on initialization and is often far away from the global optimum [Szeliski et al., 2008]. Simulated annealing is able to compute the global optimum with a suitable annealing schedule. This however, requires very long computation times and is therefore infeasi-
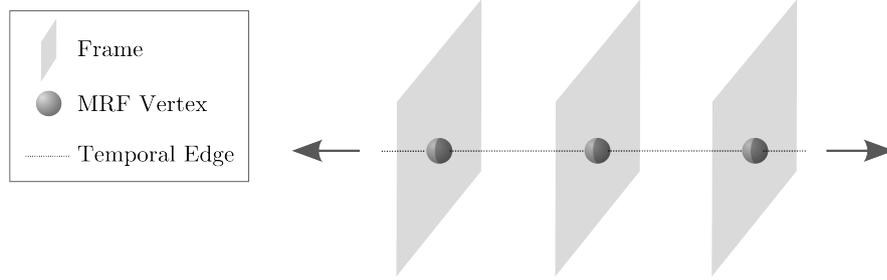
Figure 4.8: Each translation parameter for a frame is represented by a vertex. The vertices are connected by edges to achieve temporal smoothness of the parameters.

ble. For simulated annealing to terminate within reasonable time, it is only able to produce approximations to the optimum [Granville et al., 1994].

From Temporal Frame Selection to MRF formulation

For each group of connected frames defined in the temporal frame selection, a single MRF is formulated. In this way, temporal independence is given between regions separated by frames which are left out in motion compensation. A reference frame t_{ref} is assigned to each MRF. This frame may be included in the current frame set or a remote frame set representing the same area of interest.

4.3.2   *Translation motion*

The first step of the motion compensation aims to reduce the influence of global motion, i.e., motion artifacts which can be compensated by translational shifts of the frames. In other words, two parameters have to be determined for each image frame in I, the translation in spatial dimension in x and y direction, Δ_x and Δ_y. This results in a single vertex p representing the transformation of a frame and being assigned a label $l \in \mathcal{L}$, with

$$\mathcal{L} = \{(\Delta_x, \Delta_y) \in \mathbb{Z} \times \mathbb{Z} \mid -\omega \leqslant \Delta_x \leqslant \omega, -\omega \leqslant \Delta_y \leqslant \omega\}. \quad (4.38)$$

The parameters are discretized to take integer values for pixel translations between a specified bound between $-\omega$ and ω in each spatial dimension.

The data energy E_{data} is calculated by similarity analysis to the predefined reference frame at t_{ref}. The SSD measure is chosen as it is fast in computation and delivers robust results for B-mode data. Only

pixels within the user-defined ROI $\mathcal{R} \subseteq \{1, ..., x_{max}\} \times \{1, ..., y_{max}\}$ are taken into account:

$$D_p(l) = \frac{1}{|\mathcal{R}|} \sum_{r \in \mathcal{R}} (I(r_x + l_x, r_y + l_y, p) - I(r_x, r_y, t_{ref}))^2 \quad (4.39)$$

A simple neighborhood system is used in the translation motion case. It connects adjacent vertices (frames) to express the local temporal dependency of the parameters. This represents a special case of an MRF called a Markov chain (Figure 4.8). The weight $w(\cdot)$ assigned to an edge $e_s = \langle p, q \rangle \in \mathcal{E}_s$ connecting vertices p and q is calculated through the Euclidean distance of the currently assigned labels f_p and f_q:

$$w(e_s) = \|l_p - l_q\|_2 \quad (4.40)$$

The sum over all edge weights is equivalent to the prior energy $E_{prior}(\cdot)$ in Equation (4.28). To balance between similarity information and a priori knowledge, the weight η from Equation (4.21) has to be set appropriately. Standard values have been empirically found for different emphasis, either to be more flexible in terms of parameter changes over time or more towards stable transformation parameters. The different standard values have to be chosen by the user and can also be adapted if the result is not satisfactory.

4.3.3 Non-rigid motion

In the second step of motion compensation, non-rigid local motion is addressed. To compute the transformation for each location in the image, B-spline functions are used [Rueckert et al., 1999]. To control the B-spline based transformation of an image, equally distributed control points forming a grid are used. The transformation is determined by the position of these control points. The number of control points defines the ability to adapt to local motion effects. It has to be chosen according to the application and the present motion.

In order to solve the non-rigid motion compensation, a number of vertices v per frame determined by the number of control points is regarded. A label $l \in \mathcal{L}$ has to be assigned to each vertex, representing a control point of the B-spline transform. $\mathcal{L}$ is defined similarly to the translation motion case, however, different bounds have to be used to preserve the transformation in order to be diffeomorphic [Rueckert et al., 2006]:

$$\mathcal{L} = \{(\Delta_x, \Delta_y) \in \mathbb{Z} \times \mathbb{Z} \mid -\delta \leqslant \Delta_x \leqslant \delta, -\delta \leqslant \Delta_y \leqslant \delta\}. \quad (4.41)$$

The parameters are again discretized to take full pixel translations for each control point. The translation of the points is bounded by δ:

$$\delta = \frac{\min(x_{max}, y_{max}) \cdot 2}{n}, \quad (4.42)$$
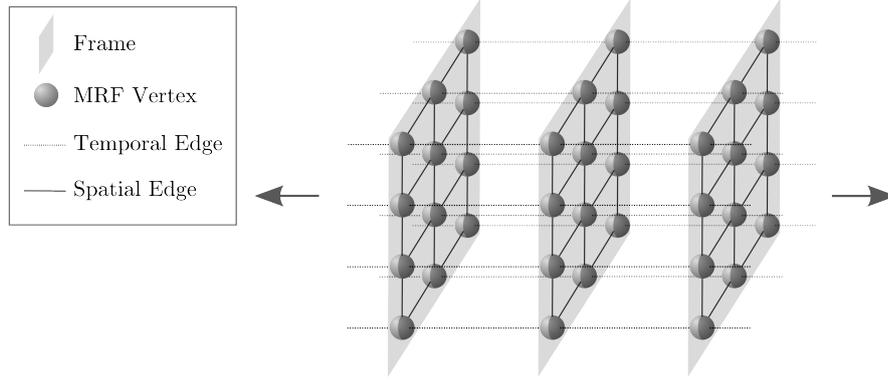
Figure 4.9: Each transformation parameter (B-spline control point) is represented by a vertex. The vertices are connected by spatial edges to achieve spatial continuity and by temporal edges to achieve temporal smoothness of the parameters.

with n being the number of control points in the grid per row or column, respectively. Control points cannot change their relative positions in the grid and the transformation is invertible for all locations.

The neighborhood system connects vertices in spatial and temporal dimensions (Figure 4.9). In the spatial dimensions, the von-Neumann neighborhood is used to connect adjacent vertices. In the temporal dimension, corresponding vertices (at the same grid position) in adjacent frames are connected. This way, the local dependency of transformation parameters is taken into account as smoothness prior. The cost (weight) assigned to the edges is defined in Equation (4.40). However, spatial and temporal edge costs have to be weighted differently. In MRF terms, the prior energy is divided into spatial and temporal prior energy

$$E_{prior}(f) = \sum_{\langle p,q \rangle \in N_s} E(f_p, f_q) + \gamma \cdot \sum_{\langle p,q \rangle \in N_t} E(f_p, f_q), \qquad (4.43)$$

where $N_s \subseteq \mathcal{E}$ contains all spatial edges with vertices representing the same frame and $N_t \subseteq \mathcal{E}$ contains all temporal edges with vertices representing different frames. γ is the weighting parameter, set according to the control point spacing and a user parameter $\kappa \in \mathbb{R}^+$:

$$\gamma = \frac{100}{0.5 \left(\frac{x_{max}}{n-1} + \frac{y_{max}}{n-1} \right)} \cdot \kappa. \qquad (4.44)$$

$w(\langle p, q \rangle) = E(f_p, f_q)$ is the prior energy of two vertices p and q and n is the number of control points in x and y dimension, respectively.

The data energy E_{data} or disagreement measure is again calculated using the SSD measure. This time a 7×7 neighborhood K_p of pixels around the B-spline control points is used to find local motion patterns. According to empirical tests with artificial data with

known ground truth, a 7×7 neighborhood was found to produce robust information about motion shifts [Schäfer and Toennies, 2012]:

$$D_p(l) = \frac{1}{|K_p|} \sum_{r \in K_p} (I(r_x + l_x, r_y + l_y, p) - I(r_x, r_y, t_{ref}))^2, \quad (4.45)$$

with r being the pixels in the local neighborhood K_p of a vertex p and I_{ref} the reference frame image in the current MRF.

In both motion compensation steps, energy minimization is achieved using the *α-expansion* technique of the graph cuts implementation *GCoptimization* [Boykov et al., 2001; Kolmogorov and Zabih, 2004; Boykov and Kolmogorov, 2004]. The weighting parameter η from Equation (4.21) should be set to a value between 1 and 40 for translation motion registration. For the non-rigid registration η should be set to a value between 4 and 20. A higher value will achieve higher smoothness of the transformation parameters. Lower values will allow local adaptions per frame. The weighting parameter κ from Equation (4.44) should be set to a value between 1 and 10. A higher value for κ will produce more stability of the B-spline control points at the same locations over time. A smaller value for κ will allow the B-spline control points to be more flexible in spatial direction.

4.3.4 *Generation of Motion Compensated Contrast Sequence*

According to the acquisition procedure (cf. Section 2.2) B-mode and contrast data are recorded simultaneously. After the transformation has been calculated from B-mode data, it can be applied to the contrast frames of the sequence. As a result, the contrast sequence is motion compensated and valid spatial correspondences of contrast measurements over time are established. The derivation of contrast-uptake measurements — the TIC — for a location or a ROI is possible (cf. Figure 4.10). Note that there is no transformation available for frames which are not contained in a frame set defined in Section 4.2. Hence, excluded frames should not be included in a subsequent analysis of the motion compensated sequence.

The following section describes an approach for the analysis of the contrast information of CEUS datasets through the formation of perfusion coherent regions.

4.4 REGION ANALYSIS WITH NOISE REDUCTION

As already stated in Section 2.4, a common way in medical image analysis to obtain reliable measurements is to manually draw regions
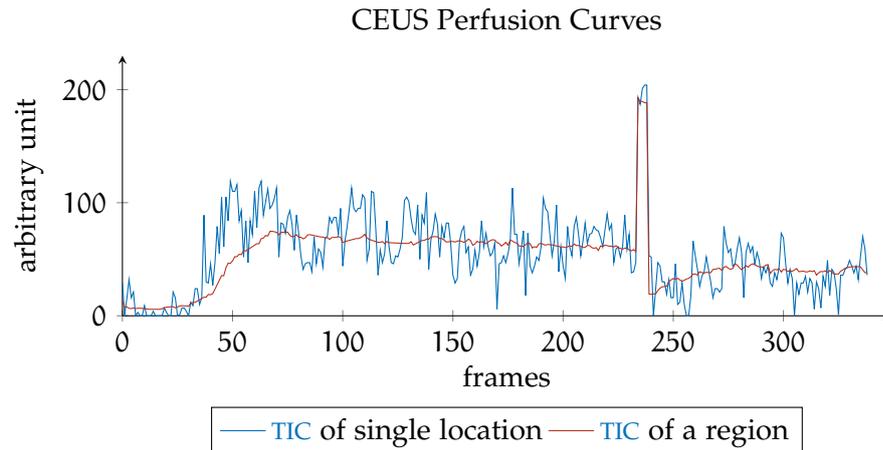
Figure 4.10: The blue curve represents measured intensity of arbitrary unit at a single location (pixel) over time. The red curve represents averaged measured intensities of multiple observations of a circular region around the location of the blue curve over time. The high peak is caused by the burst of the microbubbles which is followed by a replenishment of CA in the region.

and to calculate the average of the contained samples [Quaia, 2013]. This is done for two reasons. First, the examiner wants to assure that only the target area is observed without including irrelevant parts. Second, evaluating measurements from single data points (pixels) is subject to noise influence. Calculating the average within a defined region levels out most of the noise (cf. Figure 4.10).

The manual establishment of regions bears the risk of joining measurements with different perfusion information. As a consequence, not only noise is eliminated but also important details of perfusion may be removed. The idea behind the region analysis approach is to automate the generation of such analysis regions. These regions are not designed to represent specific objects. They state that contained data points are characterized by similar perfusion [Glaßer et al., 2009]. Whether they carry diagnostic information or not has to be decided by a medical expert. They can also be used as template to manually create other regions. The advantage is that the examiner is informed about underlying characteristics.

When such regions are created the measurements over time of different locations have to be compared. This is again impeded by noise. The following sections describe a method to reduce the influence of noise when assessing perfusion parameters. It uses a classic segmentation approach to derive regions of similar perfusion characteristics from the data.
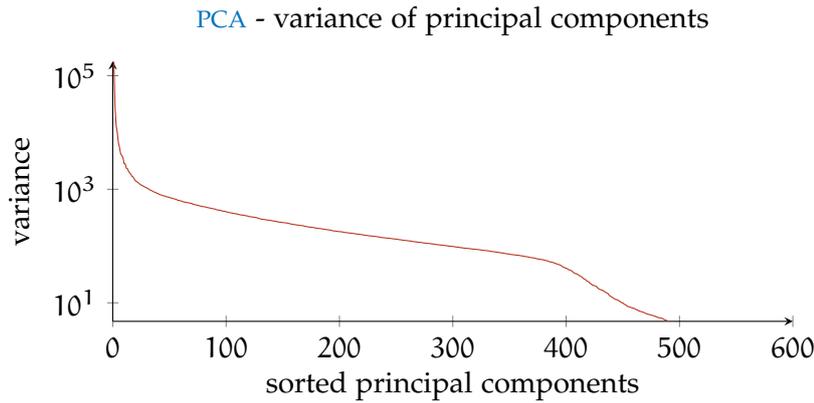
Figure 4.11: The variance in terms of the eigenvalues is showed contained in the principal components acquired through PCA of a CEUS contrast sequence. Note that the variance is displayed using a logarithmic scale.

4.4.1 *PCA for noise canceling*

Principal component analysis (PCA) is a statistical technique used to extract most important information from a set of observed dependent variables [Hotelling, 1933; Vaseghi, 2008]. It uses a set of orthogonal (linearly uncorrelated) vectors — the principal components — to represent the observed data:

$$X = PQ^T. \tag{4.46}$$

X is the matrix of observed data with I observations represented as rows and J variables represented as columns. P is the matrix of principal components with size $I \times \min(I, J)$ onto which the data have been projected. Q contains the coefficients for each variable which reconstruct the data together with the principal components. The size of Q is $J \times \min(I, J)$. The components in P are calculated by singular value decomposition of the covariance matrix $X^T X$ of the data [Abdi and Williams, 2010]. They are ordered according to the contained variance (see Figure 4.11).

In case all components are retained, the original data are reconstructed. If the L components with highest variance are retained only the most important information from the data is reconstructed:

$$Y = P_L Q_L^T. \tag{4.47}$$

The sizes of P_L and Q_L are reduced to $I \times L$ and $J \times L$, respectively. Y is of original size $I \times J$ and contains the reconstructed data. Figure 4.12 shows the TIC of a location in contrast CEUS as observed
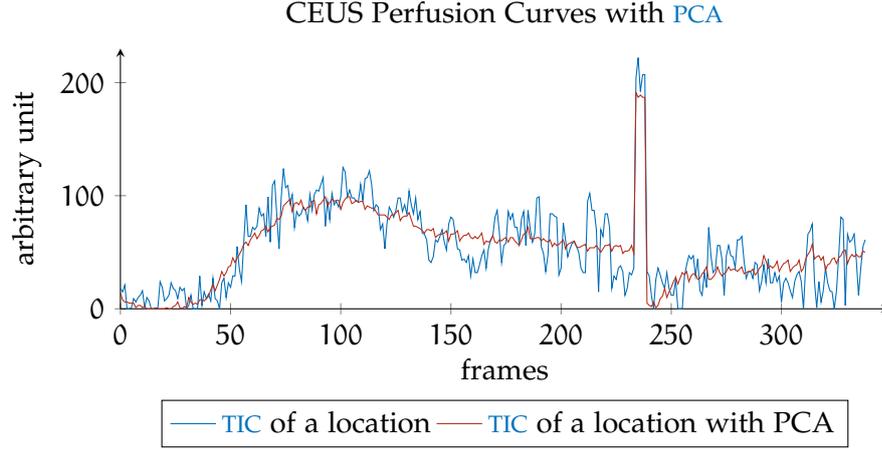
Figure 4.12: The blue curve represents measured intensity of arbitrary unit at a single location (pixel) over time. The red curve represents the PCA reconstructed signal (the largest five components retained) of the same location used for the blue curve.

and reconstructed with a PCA by retaining the five most important components. An example frame from this sequence is shown in Figure 4.13.

Under the assumption that noise is mostly represented by components with smaller variance, it will be removed by this operation. However, it should be noted that some details of perfusion might be removed as well. Therefore, PCA reconstructed data are suitable to be used for region generating but should not be directly used to deduce diagnostic conclusions.

4.4.2 *Region Merging using Perfusion Characteristics*

Region Merging is a global image segmentation approach using region homogeneity as criterion to establish segments [Jayaraman et al., 2011]. A similar procedure is used in our method to generate 2D regions representing the same perfusion characteristics.

An undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is used for representation. For each pixel in the CEUS sequence, a vertex $v \in \mathcal{V}$ is introduced being a start pixel for a region. So, each vertex represents a region of pixels and gets assigned the temporal vector of intensities from the PCA reconstructed contrast sequence within the regions, the TIC. If the region contains multiple pixels, the TIC is averaged. All vertex pairs whose assigned start pixels are adjacent are connected through an edge $e = \langle p, q \rangle \in \mathcal{E}$. Each edge e is assigned an edge weight:

$$w(\langle p, q \rangle) = \frac{1}{T} \sum_{i=0}^{T-1} |R_p(i) - R_q(i)|, \tag{4.48}$$
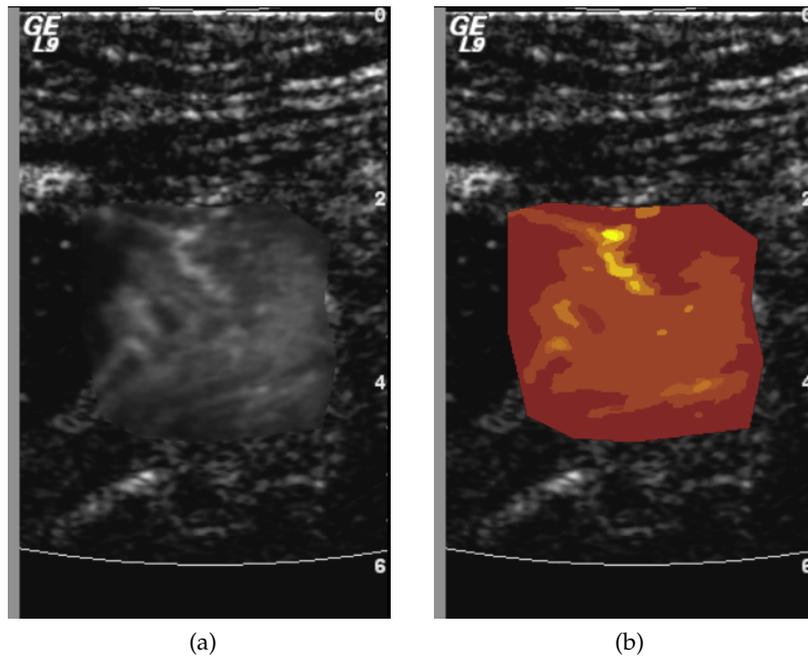
Figure 4.13: (a) A single CEUS contrast frame from a motion compensated dataset of the bowel is shown with a ROI (red overlay) reconstructed by PCA and the five components with highest variance. In (b), the PCA reconstructed image has been used to establish perfusion regions with the region merging procedure. Corresponding TICs to the regions are displayed in Figure 4.14. Images courtesy of the University of Bergen, Norway.
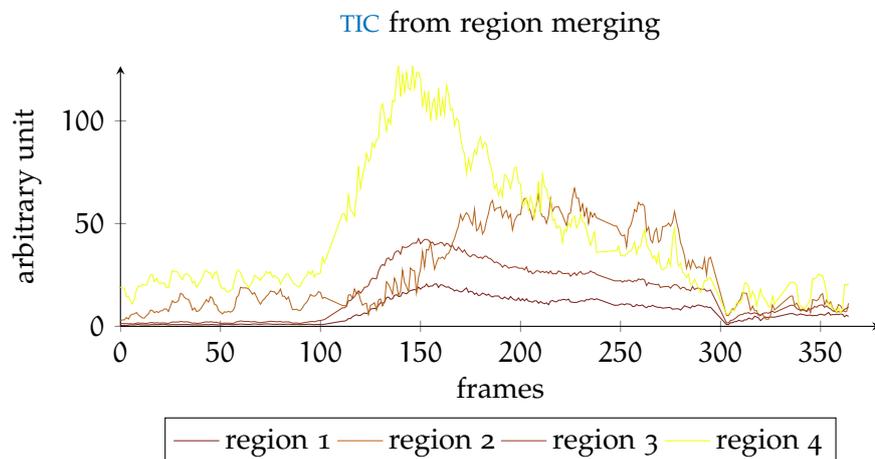
Figure 4.14: The TICs of regions from Figure 4.13 are shown using the same color coding. Note that in areas where frames have been declared out-of-plane, the TIC is interpolated between the last and the next valid perfusion measurement.

with R_v being the TIC of a vertex v. All edges in $\mathcal{E}$ are sorted according to their edge weight w. Two connected vertices are merged

until a predefined threshold of dissimilarity has been reached, starting with edges with the lowest weights. Merging two vertices p and q representing two regions implies:

- add the contained pixels of the region associated with q to the region of p,

- change all edges $e = \langle q, s \rangle \in \mathcal{E}$, $s \in \mathcal{V}$ and $e = \langle s, q \rangle \in \mathcal{E}$ to $e = \langle p, s \rangle$ and $e = \langle s, p \rangle$, respectively,

- remove duplicates in $\mathcal{E}$ and

- delete vertex q.

This approach has two important advantages. It always combines those two regions which have the highest similarity and it only depends on one parameter — the threshold until merging is applied. This parameter is user-controlled.

After the segmentation has been calculated from PCA-reconstructed data, the average perfusion TIC is calculated for each region using the normal contrast sequence information. This ensures that all important details are represented in the averaged region curves.

4.5 SUMMARY

This chapter describes a prototype system for the analysis of perfusion in CEUS image data. It introduces specialized methods addressing problems related to the acquisition of CEUS. In particular, these are

- the effect of out-of-plane motion due to the acquisition of a 2D image plane,

- the effect of motion from different sources and

- the low image quality (SNR) in CEUS.

The system comprises three main components which are consecutively applied:

- For the temporal frame selection, out-of-plane frames can be excluded by two different approaches. A segmentation-based approach using similarity data requires user interaction to establish the regions and to identify the regions representing out-of-plane motion. A classification approach uses texture features to distinguish between in-plane and out-of-plane regions. The second approach works fully automatically whereas the first approach is semi-automatic.

- Motion compensation is performed using an MRF-based optimization framework and the defined frame sets. Translation as well as elastic transformations are subsequently applied. The high temporal resolution of the CEUS sequences is used to constrain the transformations to be smooth. The constraints stabilize the approach in the temporal dimension against influences of noise.

- Finally, the motion compensated contrast data are segmented to find areas of similar perfusion. PCA is used to extract principal components of the perfusion. The regions serve as suggestion for analysis and assist the physician to define additional analysis regions.

RESULTS AND EVALUATION

In this chapter, the proposed system for CEUS data analysis is evaluated. Each of the three parts (frame selection, motion compensation and perfusion region generation) is separately tested to reveal their contribution. Additionally, experiments are conducted to assess the suitability and the improvement of the whole system.

First, the datasets used for evaluation are described in detail. Then, the following experiments are described and their results are presented:

- Evaluation of temporal frame selection (cf. Section 4.2)
  - Experiment 1: Semi-automatic method: comparison of the results with manually generated regions (cf. Section 4.2.3)
  - Experiment 2: Automatic method: cross-validation of the classification results using manually generated regions (cf. Section 4.2.4)

- Evaluation of motion compensation (cf. Section 4.3)
  - Experiment 3: Proof-of-concept: analysis of intensity profiles in B-mode and contrast data
  - Experiment 4: Robustness of weighting parameters
  - Experiment 5: Pharmacokinetic model: similarity of measured perfusion with the model before and after registration

- Experiment 6: Evaluation of regional analysis (cf. Section 4.4): dissimilarity analysis of generated regions

- Experiment 7: Evaluation of the system workflow: time consumption of the proposed approach in comparison to manual processing

The results of each experiment are discussed at the end of each section. At the end of the chapter, all results are discussed in an overall context. Selected measures are shown in tables and graphs in this chapter. Complete results for the experiments can be found in the Appendix A.

For evaluation purposes, two different databases have been used. Database I was used for studies conducted by the medical cooperation partner in Norway. Database II was used for experiments conducted at the University of Magdeburg. Database I comprises 40 datasets of which 20 depict the pancreas and 20 depict sections of the bowel and the bowel wall. These two applications in the abdominal area have been chosen because they represent two different challenges. The pancreas datasets are characterized by strong and different motion influences. The bowel datasets are characterized by a low contrast. It is also difficult to image the bowel wall, which only measures a few millimeters.

The datasets are divided into groups of equal size, representing different diseases as well as recordings from healthy volunteers as control group. The pancreatic datasets include images of 10 patients with cystic fibrosis and known exocrine pancreatic failure. The datasets of the bowel include recordings from 10 patients suffering from an acute exacerbation of Crohn's disease. All bowel datasets of the patients show the area of inflammation, while the 10 datasets of the control group show the ascending colon.

Database II is a subgroup of database I, randomly representing 12 datasets, 8 datasets showing the pancreas and 4 showing a section of the bowel.

All datasets were acquired with a LOGIQ® 9 or LOGIQ® E9 US scanner (GE Healthcare, Milwaukee, WI, USA). Observations of the pancreas were acquired for about 90 seconds at a frame rate of 9 FPS using a curvilinear probe (C1-5). Datasets of the bowel were acquired for about 90 seconds at a frame rate of 11 FPS using a linear probe (L9). A bolus and burst regime was applied during the acquisition [Nylund et al., 2013]. A bolus of SonoVue® (BRACCO, Milan, Italy) was given intravenously for two seconds followed by an intravenous bolus of saline for four seconds. The burst of the bubbles within the imaging plane was induced 45 seconds and 60 seconds after CA injection for pancreas and bowel data, respectively. This leads to a significant change of intensities in the contrast sequence and a subsequent replenishment of CA (cf. Section 2.2.1). For the studies included in this thesis, the frame rate in the datasets has been reduced to about 2-4 FPS. This is enough to assess relevant perfusion characteristics.

The sonographer tried to observe the same imaging plane throughout the whole examination. The bowel datasets were imaged in the longitudinal direction. For the pancreatic datasets a section exhibiting as few motion as possible, using the left liver lobe as an acoustic window, was chosen. If the area of interest moved out-of-plane dur-

ing acquisition, the sonographer tried to readjust the probe to get the area back into focus.

A comprehensive study using database I was performed by medical experts of the cooperation partner from the University of Bergen, Norway, by using the approaches presented in this thesis. Additionally, software by a collaboration partner from the University of Technology, Brno, Czech Republic was used for pharmacokinetic evaluation [Jiřík et al., 2013].

5.2 EVALUATION OF TEMPORAL FRAME SELECTION

The semi-automatic and automatic approach of the temporal frame selection use knowledge extracted from the similarity matrix view. They aim at assisting the physician in the *out-of-plane* frame exclusion process. To evaluate the accuracy of both approaches the generated frame sets are compared to ground truth data. The ground truth is derived by manually defined sets of frames by medical experts. These frame sets have been established using the similarity matrix view (cf. Section 4.2.1) and a detailed assessment of the CEUS sequences. To evaluate the classification approach, frames included in a ground truth region are assigned to the group *in-plane*; frames not included in a set are assigned to the group *out-of-plane*. Database I is taken as data basis.

5.2.1 *Experiment 1: Semi-automatic frame selection accuracy*

The semi-automatic frame selection maintains different regions of frames indicated by a different color if the similarity between the regions gives evidence that they do not represent the same area (cf. Section 4.2.2 and Figure 4.2e). This happens for example when motion occurs which does not move the area of interest back into the image plane. The expert declares all frames *out-of-plane* which do not represent the area of interest (ground truth regions). The semi-automatic approach suggests regions which show different areas of interest leaving the decision to the user if and which regions have to be discarded.

For evaluation it was therefore tested, if and how precise a frame set from the ground truth was found by the semi-automatic approach. Regions which were additionally found by the approach were only penalized if they interfere with the ground truth regions (see Figure 5.1b). In order to assess the quality of the results from the semi-automatic approach the overlap ratio is defined as:

$$r_o = \sum_{i=0}^{|\mathcal{T}|} \frac{\left|\mathcal{T}_i \bigcap \mathcal{D}_j\right|}{|\mathcal{T}_i| + \mathrm{OLP}(\mathcal{T}_i, \mathcal{D}_j)}, \tag{5.1}$$
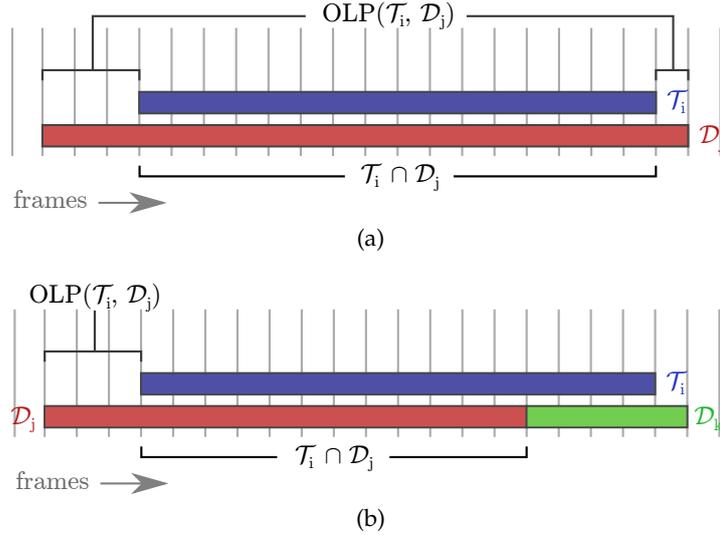
Figure 5.1: Two different cases of the determination of $\mathrm{OLP}(\mathcal{T}_i, \mathcal{D}_j)$ for the overlap percentage to evaluate the semi-automatic frame selection approach. In (a) the protrusion of $\mathcal{D}_j$ is on both sides. In (b) the overlap is calculated between $\mathcal{T}_i$ and $\mathcal{D}_j$ as they intersect most. $\mathcal{D}_k$ interferes with the ground truth region $\mathcal{T}_i$.

with $\mathcal{T}$ being the ground truth frame set and $\mathcal{D}$ being the set of frames found by the semi-automatic approach. The index j was chosen so that $\left|\mathcal{T}_i \bigcap \mathcal{D}_j\right|$ was maximal. $\mathrm{OLP}(\mathcal{T}_i, \mathcal{D}_j)$ calculated the protrusion of $\mathcal{D}_j$ in relation to $\mathcal{T}_i$ (see Figure 5.1).

RESULTS   The semi-automatic frame selection accuracy was computed using database II. For each dataset the best two parameters k_1 and k_2 from the semi-automatic approach (cf. Section 4.2.3) were determined in terms of the overlapping percentage $O_\mathcal{T}$. The mean of the overlapping percentage of the 40 datasets was 75.8 % with an SD of 14.5 percentage points. Eight datasets out of 40 had an overlapping percentage higher than 95 % while only 3 datasets had an overlapping percentage below 50 %. The comprehensive results for the datasets can be found in Tables A.1 to A.2 in the Appendix A.

DISCUSSION   For 37 of 40 datasets the results only require little manual refinement. In 20 % of the cases, the overlapping percentage is even sufficient that the user only needs to manually delete some regions, but it is not necessary to adapt the size or the position of regions found by the semi-automatic approach. In all cases, the semi-automatic approach is able to deliver a good initialization as basis for the frame sets.

In discussions with medical experts we found that the similarity matrix view is able to give a good overview about possible motion in a dataset. This also simplifies the manual selection of out-of-plane

frames. The experts stated that navigating to specific frames using the similarity matrix enables fast comparison of frames and areas of interest. The time advantage of the similarity matrix approach for manual frame selection is assessed in Section 5.5.

5.2.2 *Experiment 2: Automatic frame selection accuracy*

The goal of Experiment 2 is to find out, how the voting approach used for the automatic frame selection performs in comparison to individually using the classifiers. To quantify the accuracy of a classification result, the *g-mean* (g_m) measure is calculated that has been proposed to measure the quality of classifiers of imbalanced data [Kubat et al., 1998]. It takes account of the sensitivity and the specificity. A high specificity is essential because it is more important to avoid out-of-plane frames to be classified as in-plane frame than vice versa. However, the specificity would take a value of 1 if all samples are classified as in-plane frame. The *g-mean* is defined as:

$$g_\mathrm{m} = \sqrt{se \cdot sp} = \sqrt{\frac{TP}{TP + FN} \cdot \frac{TN}{TN + FP}}. \tag{5.2}$$

A true positive (TP) is a sample that is correctly classified as out-of-plane frame, a true negative (TN) is a sample that is correctly classified as in-plane frame, an false positive (FP) is a sample erroneously classified as out-of-plane frame and an false negative (FN) is a sample erroneously classified as in-plane frame.

To establish results the WEKA data mining software [Hall et al., 2009] was used. The list of used classifiers includes representatives of the groups from Section 4.2.4. If no specific parameterization is given, the suggested standard parameters from WEKA were used. The classifiers were also used in the voting approach:

- Naive Bayes

- KNN (k = 10)

- LogitBoost [Friedman et al., 2000]

- AdaBoost [Bauer and Kohavi, 1999]

- Alternating decision tree (ADTree) (Iterations: 100) [Freund and Mason, 1999]

- C4,5 [Quinlan, 1993]

- SVM (c = 1, linear kernel) [Burges, 1998]

Table 5.1: G-mean of classification results from ten-fold cross valida-
tion using all 20 features and the 8 Haralick features only as
well as for super- and subsampling and cost-weighting as ap-
proach to account for imbalanced data. The best results are
printed in bold.

| | all features | | 8 Haralick features | |
Classifier	sampled	cost-weight	sampled	cost-weight
Naive Bayes	0.60	0.60	0.55	0.55
KNN (k=10)	**0.74**	**0.75**	**0.77**	**0.77**
LogitBoost	0.65	0.64	0.64	0.64
ADTree	0.71	0.72	0.71	0.70
SVM (linear)	0.53	0.52	0.50	0.55
C4,5	0.73	0.70	0.73	0.71
AdaBoost	0.63	0.63	0.00	0.65
Voting approach	0.70	0.71	0.72	0.72

Ten-fold cross validation was performed with training on the su-
per- and subsampled dataset and cost-weighting to compensate the
class imbalance [Witten et al., 2011]. In order to find out if there is a
texture feature contributing more than others to successful classifica-
tion, all measurements were performed using the 20 Gabor filter and
Haralick features and using the 8 Haralick features solely.

RESULTS   Results in terms of the *g-mean* for the four different se-
tups are shown in Table 5.1. KNN produced the best overall results in
terms of g_m yielding 0.77 using the eight Haralick features only. Sim-
ilar results were produced by ADTree and C4,5 ($g_m > 0.7$). The voting
approach combining several classifiers was close to the best results
(between 0.69 and 0.73). The voting approach performed slightly bet-
ter when the eight Haralick features are used only. In general, little
differences were found between the results using all 20 features com-
pared to the results using the eight Haralick features only. Detailed
results for all classifiers are shown in the Appendix A in Tables A.3
to A.6.

DISCUSSION   The voting approach is able to achieve results which
are comparable to the best classifier results (KNN). The advantage
of the voting approach is that the result does not depend on the
performance of a single classifier. However, the feature distribution
can be separated best by KNN classification. According to the results

achieved with the 40 datasets of database I, KNN should be used in the future for classification.

The Haralick features used solely are able to achieve the same results compared to a combination of Haralick features and Gabor filter kernels. This leads to the conclusion that the use of the different Haralick features would be sufficient.

Overall, the automatic frame selection approach using classification can only be regarded as initialization for frame sets. Mostly, manual corrections are required. The semi-automatic approach (cf. Section 5.2.2) has proven to be capable of finding the desired regions. The possibility to influence the results through interaction represents another advantage of the semi-automatic approach.

5.3 EVALUATION OF MOTION COMPENSATION

The evaluation of motion compensation tasks is difficult, mostly because a ground truth about motion is not available. Despite that no true motion information is available, different approaches have been proposed that directly assess the quality of the transformation for motion compensation: fiducial markers and simulated data with known transformation [Hill et al., 2001; Crum et al., 2004]. Both strategies have disadvantages, particularly if motion compensation is applied to a huge amount of data like CEUS sequences. Fiducial markers have to be explicitly set in between 400 and 1000 consecutive frames. This is a tedious process and subject to intra- and inter-observer variability. Additionally, the quality of the transformation is only measured at the marker positions. There are approaches to simulate US data including noise and speckle [Jensen, 1996]. However, the question is if a realistic environment can be established with reasonable computational effort. In particular, the B-mode sequence from a CEUS acquisition also shows effects of CA administration which cannot yet be regarded in simulation.

To be able to qualitatively evaluate the results of motion compensation and the different parameterizations of the method the approach was indirectly evaluated. Therefore, the plausibility of the results was examined as a proof-of-concept evaluation in the third experiment. The remaining intensity dissimilarity in B-mode data was measured as well as the smoothness TIC in contrast data [Schäfer and Toennies, 2012]. Experiment 4 used pharmacokinetic modeling of the perfusion TICs to check if perfusion can be described more precisely by the model after motion compensation [Schäfer et al., 2011b].

Results of SD Measurements in B-mode
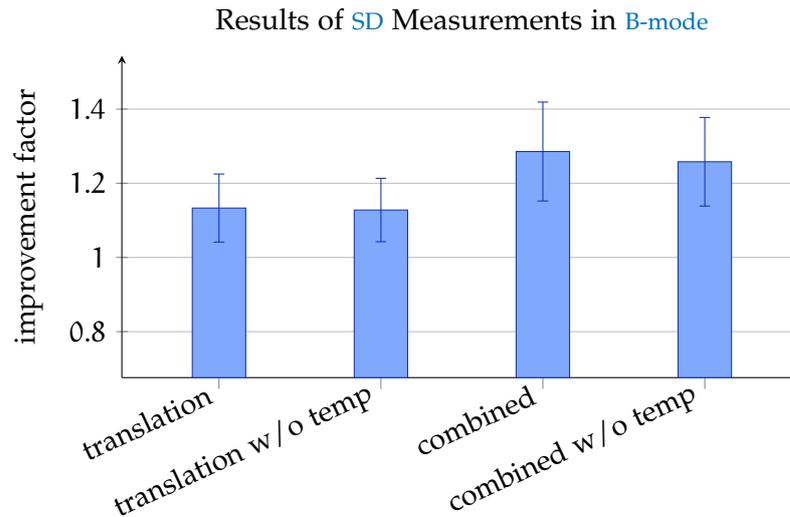
Figure 5.2: The improvement of the average SD in diagnostic regions of the 12 datasets of database II is shown for different configurations of the algorithm in relation to the SD before motion compensation: translation with and without temporal constraints and combined translation and B-spline based transformation with and without temporal constraints. The error bars indicate the SD of the results.

5.3.1   *Experiment 3: Proof-of-concept for Motion Compensation*

This experiment is intended to proof the concept of motion compensation with special focus on the temporal continuity assumptions. The aim of motion compensation is to maximize similarity in B-mode data. As a result of the motion compensation, the deviation of intensity at the same location in different time frames should decrease over time. Therefore, the average SD was measured in diagnostic regions within the B-mode data [Schäfer et al., 2012].

Another assumption about motion compensated CEUS data are that the perfusion signal is improved. Improvement in this context means that intensity variation introduced by motion is reduced. The concentration change of CA in the vascular system over time is a continuous process and therefore, the TIC course is expected to be smooth over time. The second part of this experiment examined the smoothness of perfusion in the diagnostic regions. The smoothness was measured by the mean average difference (MAD) of adjacent measurements and is expected to decrease as a result of motion compensation.

To assess the contribution of the temporal constraints, measurements were performed after translation motion compensation with and without temporal constraints as well as after non-rigid motion compensation with and without temporal constraints. The measure-

Results of MAD Measurements of TIC
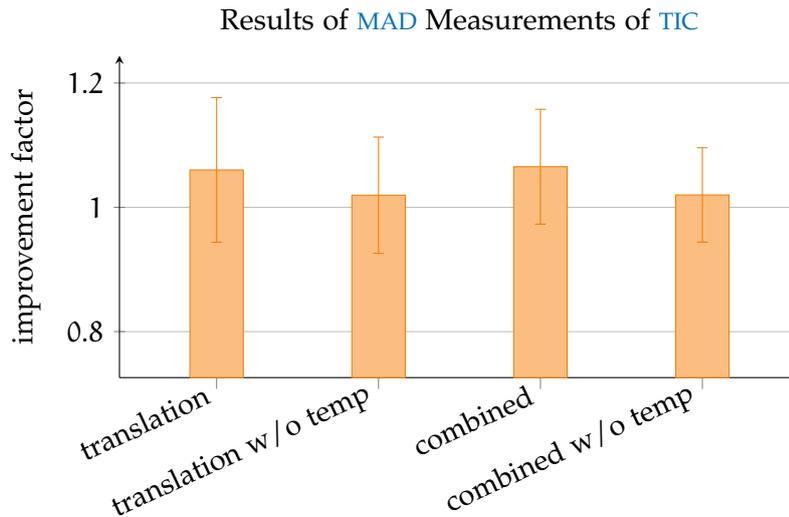
Figure 5.3: The improvement of the MAD in contrast data in diagnostic regions measured in 12 datasets is shown for different configurations of the algorithm in relation to the MAD of the TIC before motion compensation: translation transformation with and without temporal constraints and combined translation and B-spline based transformation with and without temporal constraints.

ments were compared to a measurement taken before any motion compensation. Evaluation results are indicated as improvement factor, meaning the improvement of a measure compared to the value before motion compensation. For this experiment the 12 CEUS datasets from database II were assessed showing examinations of the bowel and the pancreas.

RESULTS   In both tests of Experiment 3, consecutive translation transformation and B-spline based motion compensation with temporal constraints was able to achieve the best results (Figure 5.2). In terms of the average SD for regions in B-mode the improvement was 28.6 %. This doubled the improvement compared to using translation transformation only (13.3 %). The results of the same experiments without using temporal constraints were slightly worse. Full results for this experiment can be found in Table A.7 in the Appendix A.

In terms of the perfusion curve smoothness, the MAD of the TIC improved by 6.5 %. The result was only slightly worse (6.0 %) when using translation transformation only (Figure 5.3). Results computed without temporal constraints were not able to perform equally; only 1.9 % and 2.0 % were achieved with translation transformation and combined translation and B-spline based transformation, respectively. Full results for this experiment can be found in Table A.8 in the Appendix A.
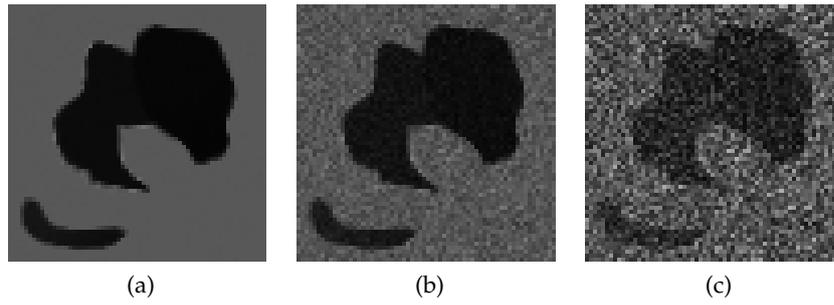
Figure 5.4: The first frame of an artificial image sequence with no noise (a) and noise added leading to an SNR of 2 (b) and 3 (c).

DISCUSSION   The SD measurements in B-mode data show that the variation of intensity over the whole sequence is reduced. This indicates that the temporal constraints help to find a good global solution with reduced variability in terms of intensity. This gives proof of the concept of temporal stability of transformation parameters.

The perfusion curve smoothness gives an indication about the plausibility of a physical process. Noise influence, which also influences the smoothness, is reduced by measuring the averaged smoothness in relevant regions. Remaining abrupt changes suggest that there are errors in the data due to motion. The smoothness of perfusion data has been improved by motion compensation. Surprisingly, translation as well as combined transformations improve the smoothness in a similar way. However, the temporal constraints used for both transformations have significant influence on the perfusion curve smoothness (Figure 5.3).

5.3.2   *Experiment 4: Robustness of weighting parameters*

The automatic motion compensation approach based on an MRF depends on two weighting parameters: first, η from Equation (4.21) weighting between the data term energy and the prior term energy. Second, the parameter κ from Equation (4.44) used for B-spline based registration to weight between the prior energy of spatial and temporal edges.

Artificially transformed data were used, in order to find out how the outcome of the motion compensation was altered by changing the weighting parameters. The ground truth of the transformation is known and can therefore be used to qualitatively assess the result (MSD error) of the motion compensation.

The robustness of the weighting parameters was tested for three different cases:

- η for translation motion compensation,

Robustness of η for translation motion compensation
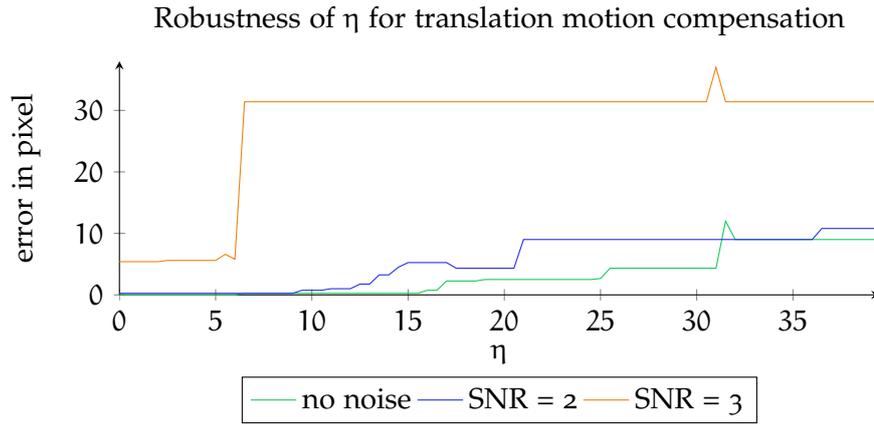
Figure 5.5: The MSD error in terms of average pixel distances from the ground truth after translation motion compensation is plotted against the value of η for artificial images with no noise, an SNR of two and an SNR of three.

Robustness of η for non-rigid motion compensation
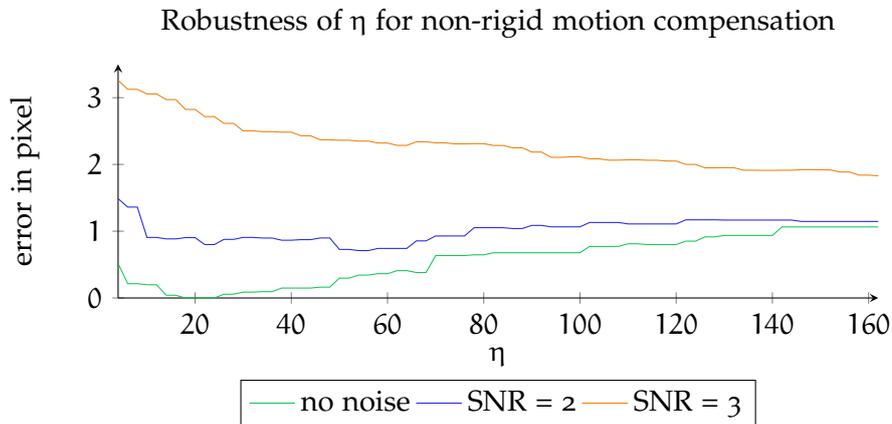
Figure 5.6: The MSD error in terms of average pixel distances from the ground truth after B-spline based motion compensation is plotted against the value of η for artificial images with no noise, an SNR of two and an SNR of three.

- η for non-rigid motion compensation and

- κ for non-rigid motion compensation.

A dataset of five frames with identical artificial data was generated. The frames were shifted with a known translation to simulate rigid motion. For non-rigid motion, the artificial sequence was also transformed using known translations of B-spline control points. Additionally, in both artificial sequences Gaussian noise was added leading to an SNR of two and an SNR of three (see Figure 5.4). The parameters ranged from zero to a value where no further changes of the error were observed. The parameters have been sampled so that 80 measures were obtained for each evaluation.

Robustness of κ for non-rigid motion compensation
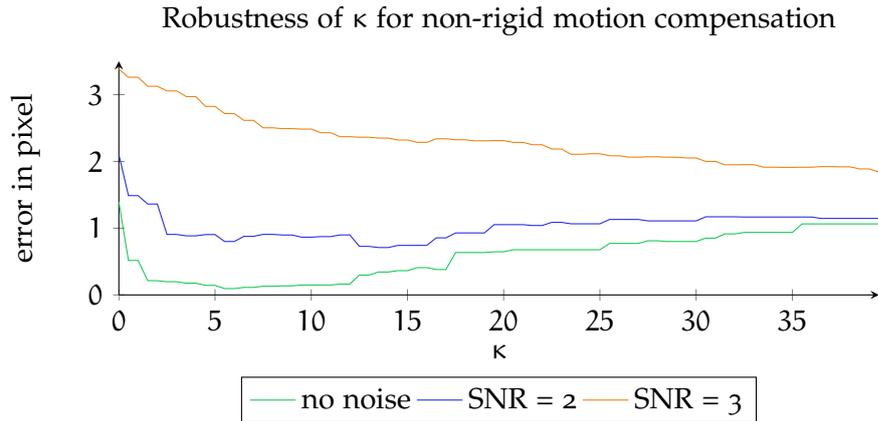
Figure 5.7: The MSD error in terms of average pixel distances from the ground truth after B-spline based motion compensation is plotted against the value of κ for artificial images with no noise, an SNR of two and an SNR of three.

RESULTS   The results of the robustness analysis for the weighting parameters η and κ are shown in Figures 5.5 to 5.7. All results show reasonable small registration errors within large ranges of the weighting parameters. For translation motion compensation an η value larger than 15 led to higher errors in terms of the MSD to the ground truth. For the artificial image with added noise of an SNR of 3 the error rapidly increased at a value of six for η. The weighting parameters η and κ for non-rigid motion compensation showed very similar characteristics. The range for minimal errors for the artificial images with no noise and with an SNR of two lied in between 10 and 65 for η and between 3 and 16 for κ. For the artificial image with an SNR of 3 optimal values were much higher, for η larger than 140 and for κ larger than 34.

DISCUSSION   For all three cases, parameters are robust against changes. For the artificial image sequence with an SNR of three a higher η in translation motion compensation leads to underweighting of the data term energy. The strong prior energy term produces equal translation parameters for all frames. The non-rigid motion compensation even seems to work best for the image sequence with an SNR of three with very high η or κ. This is caused by very weak data term energies as the similarity at the control points is only evaluated in a local neighborhood. The high prior energy prevents the control points from adapting to wrong locations.

The observations in Figure 5.6 and Figure 5.7 are very similar except for κ = 0 and η = 0. This can be explained by the fact that η also controls the weighting of the temporal edge weights. The important aspect is that spatial or temporal edge weights increase in relation

to the data term energy. The similarity of the results for η and κ for the non-rigid transformation gives evidence that it is important to have the correct weighting between data term and prior term but not between spatial and temporal prior terms. As the results have only been tested on artificial datasets it is not guaranteed that this holds true for clinical datasets.

The main interest of diagnosis by the help of CEUS imaging is the perfusion analysis. That is why in the next experiment an additional analysis of the influence of motion compensation on perfusion data quality was performed. Pharmacokinetic modeling was used to assess perfusion characteristics.

5.3.3 *Experiment 5: Pharmacokinetic Modeling of contrast data*

The distribution, absorption and depletion of substances in an organism depend on physiological conditions and the consistency of the liquid. There are models simulating this behavior, with special focus on the simulation of the distribution of pharmacokinetics such as CAs known as pharmacokinetic models [Brix et al., 1991; Tofts, 2005].

Such a model is controlled by a number of fixed parameters, defined by the acquisition technique, protocol and the administration technique of the pharmaceutical or CA. It also depends on a number of free physiological parameters of an organism which are initially unknown. These parameters give important diagnostic information, such as the capillary permeability which is essential for tumor analysis or the MTT of the pharmaceutical in a specified area [Radjenovic et al., 2008; Kolář et al., 2010]. In order to extract these results, the physiological parameters are sought, which yield a model perfusion curve representing best the observed perfusion measurements (cf. Section 2.4.1 and Figure 2.5).

For the evaluation of the motion compensation results, pharmacokinetic modeling was used to obtain reliability and plausibility of the perfusion measurements. Strouthos et al. [2010] assessed different functions to approximate perfusion characteristic in US perfusion data and to extract parameters such as mean transit time. However, these models are not able to represent the characteristics of burst-replenishment. Jiřík et al. [2012] presented a model to assess perfusion parameters (mean transit time, blood volume, blood flow) in CEUS imaging which is able to simulate bolus administration of CA and the replenishment phase after a micro bubble burst has been triggered. Both phases are modeled using a combination of an arterial input function and a tissue residue function. As the datasets of the study are recorded using burst-replenishment, a perfusion model which is able to represent this characteristic TIC is required.

The datasets of database I were analyzed using the model of Jiřík et al. [2013]. The best accuracy of fit of the model to measured perfusion data within a ROI was calculated in terms of the MAD error (ϵ) through regression analysis. These error values were acquired for three different working steps:

- the original CEUS sequence without motion compensation applied (ϵ_{orig}),

- manual processing (*manual approach*) of the datasets (ϵ_{man}) and

- processing by the proposed approach (*automatic approach*) of the CEUS sequences (ϵ_{auto}).

Measurements from the unprocessed sequence were obtained as reference for an improvement in the *manual* and the *automatic approach*. The frame selection of the *manual approach* was done by a frame-by frame analysis. Motion compensation was performed by defining a translation through clicking a reference point in each of the frames. For the *automatic approach* the frame selection was done using the similarity matrix view for navigation in the sequence and manually establishing frame sets (cf. Section 4.2.1). Motion compensation of the frame sets was applied using translation transformations. If residual motion was present, non-rigid transformation was subsequently applied. All manual steps were performed by medical experts. The improvement of the accuracy of fit for the *manual approach* was defined as:

$$v_{man} = \frac{\epsilon_{orig}}{\epsilon_{man}} \cdot 100, \tag{5.3}$$

and for the *automatic approach*:

$$v_{auto} = \frac{\epsilon_{orig}}{\epsilon_{auto}} \cdot 100. \tag{5.4}$$

In addition, the performance was evaluated for different groups of pancreatic and bowel datasets:

- pathological data and

- healthy control group.

A Wilcoxon signed rank test was performed to find significant differences between *manual* and *automatic* processing for each result group.

This experiment has been carried out by domain experts at Haukeland University Hospital in Bergen, Norway. For calculation of the perfusion model, software from the University of Technology in Brno, Czech Republic, was used [Jiřík et al., 2013].

Table 5.2: The improvements in % v_{man} for the *manual approach* and v_{auto} for the *automatic approach* of the averaged MAD error of the best accuracy achieved with the perfusion model are listed dividing pancreas and bowel datasets as well as pathological datasets and datasets from healthy volunteers.

group	manual approach	automatic approach
	v_{man}	v_{auto}
pancreas datasets		
sick	44.9	46.5
healthy	29.6	27.1
both	37.3	36.8
bowel datasets		
sick	-6.6	-1.7
healthy	14.2	15.7
both	3.8	7.0

RESULTS   The accuracy of fit improved using the *automatic approach* for pancreas and bowel datasets in terms of the MAD error (see Table 5.2). The *manual approach* was also able to improve the accuracy but in terms of the average MAD values it did not perform as well as the *automatic approach*. The statistical plots of the data reveal equal median values for both approaches for bowel datasets and inferior median values of the *automatic approach* for pancreas datasets. The results for all datasets are listed in Tables A.9 to A.16 in the Appendix A.

To be able to analyze distribution and variance of the results, box plot graphs (see Figure 5.8) are shown for both pancreas (Figure 5.9) datasets and bowel dataset (Figure 5.10). They show the average, the quantiles at 25 %, 50 % (median) and 75 % as well as the last measurements still within 1.5 of the inter-quartile range in positive and negative direction. Measurements outside 1.5 of the inter-quartile range are interpreted as outliers.

The most noticeable difference was observed between the performance for pancreas datasets and for bowel datasets. This holds true for the *automatic* and the *manual approach*. The *automatic approach* yielded 36.8 % of improved quality of fit (MAD) to the pharmacokinetic model for the pancreas datasets compared to 7.0 % for the
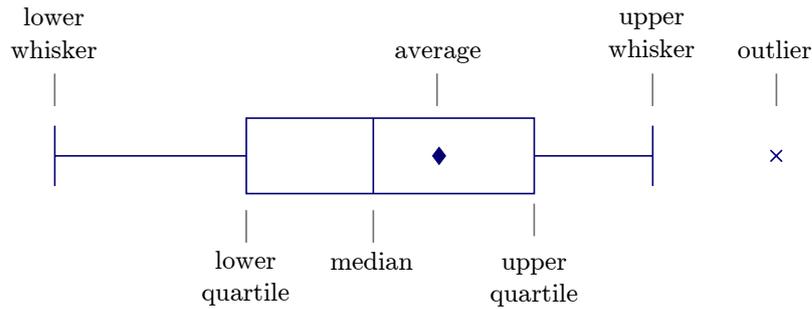
Figure 5.8: Schematic representation of the box plots: the boxes show measurements within the lower and upper quartile and are split at the median. The whiskers are showing measurements still within 1.5 of the inter-quartile range. If there are measurements outside, they are shown as outliers (x mark). The average is shown as diamond mark.

bowel datasets. The *manual approach* reached 37.3 % for the pancreas datasets compared to 3.8 % improvement of the MAD measure for the bowel datasets.

For pancreas datasets, the accuracy of fit to the model was improved for both groups and by both approaches. 46.5 % and 44.9 % was computed for the pathological group by *automatic* and *manual* processing, respectively. For the healthy control group 27.1 % and 29.6 % was reached for the *automatic* and the *manual approach*, respectively.

The box plots in Figure 5.9 show that the variance for the pancreas results of the *automatic approach* was larger for both groups. Median values for the pathological group indicated that the *manual approach* produced superior results to the *automatic* one. The Wilcoxon signed rank test revealed that the differences were significant. For the healthy control group the median indicated superior results in favor for the *automatic approach*. The results for both approaches did not significantly differ for the healthy control group.

For the bowel datasets different results have been computed. Both approaches were able to improve the accuracy of fit to the model for the control group datasets. For the pathological datasets the accuracy of fit was not improved (-1.7 %). For the *manual approach* the results were slightly worse (-6.6 %). Thus, the *automatic approach* almost achieved equal results compared to the datasets without motion compensation on average. Differences for both groups between *manual* and *automatic approach* were not significant.

The box plots in Figure 5.10 show that the variance for the bowel results was larger for the control group of healthy volunteers compared to the pathological group. The *automatic approach* was able to produce slightly better results in terms of the average values in all cases compared to the *manual approach*. The median values were the same for

Results motion compensation for pancreas datasets - MAD measures
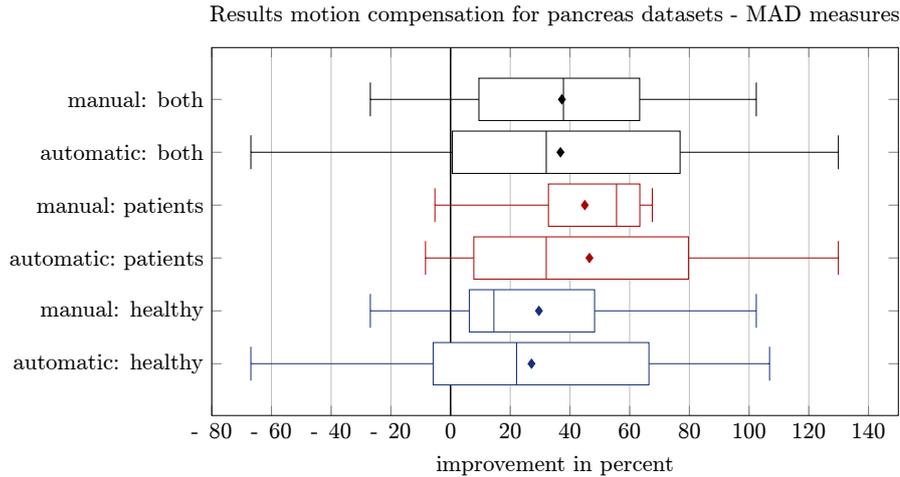
Figure 5.9: Boxplot evaluation with pharmacokinetic model for pancreatic datasets motion compensated by the *manual* and the *automatic approach*. Improvements of the MAD measures are shown in percent for the pathological group of patients, the group of healthy volunteers as well as for both together. See Figure 5.8 for explanation on box plots.

Results motion compensation for bowel datasets - MAD measures
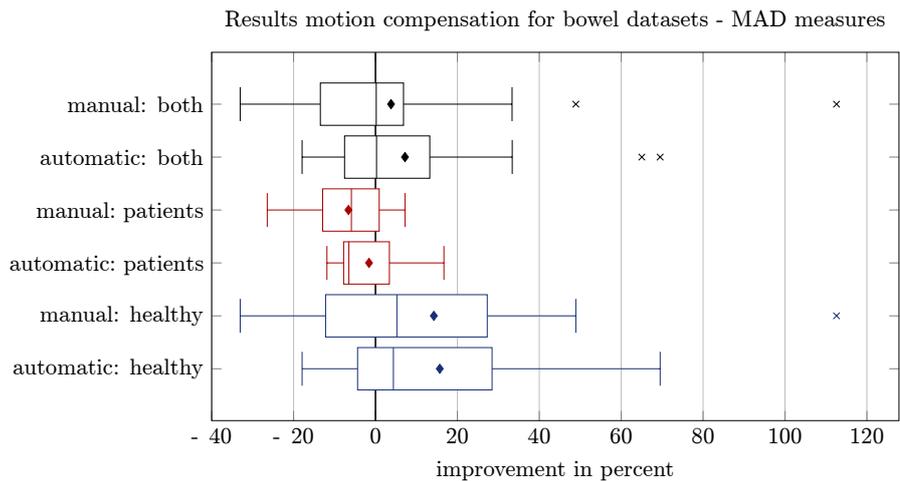
Figure 5.10: Boxplot of evaluation with pharmacokinetic model for motion compensated bowel datasets by the *manual* and the *automatic approach*. Improvements of the MAD measures are shown in percent for the pathological group of patients, the group of healthy volunteers as well as for both together. See Figure 5.8 for explanation on box plots.

the *manual* and *automatic approach*. However, for the group of pathological datasets the MAD measurements were worse than those from the uncompensated datasets.

Figure 5.11 shows an exemplary result where the quality of the pharmacokinetic model fitting was improved in comparison to the unprocessed dataset. Figure 5.12 shows an exemplary result where
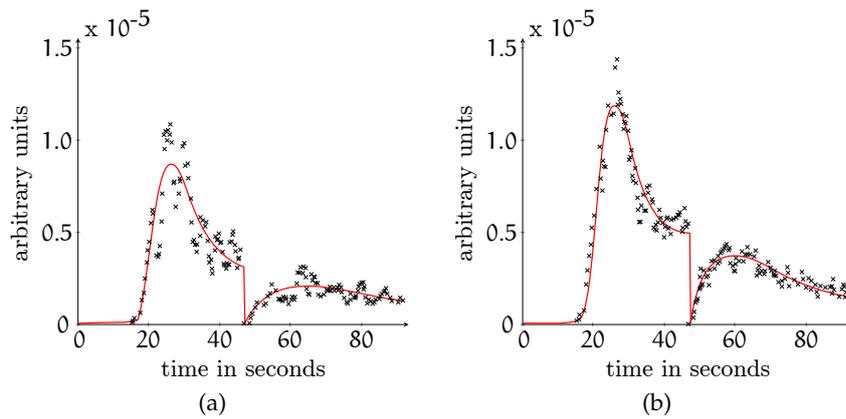
Figure 5.11: The pharmacokinetic model with best fitting parameters is shown (red curve) with the data points of the TIC (crosses). (a) shows a representative region from uncompensated pancreas data. (b) shows the same region in a dataset processed with automatic motion compensation. The data have been log-converted and are therefore declared as arbitrary unit.
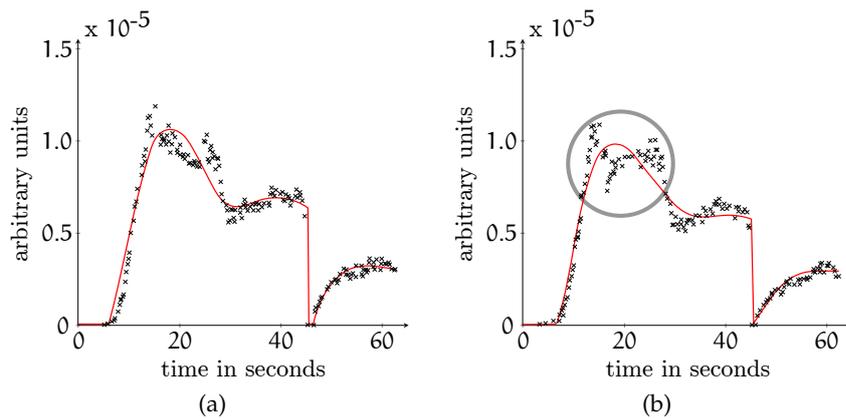
Figure 5.12: An example is shown where the fitting quality to the pharmacokinetic model of the automatic motion compensation is worse compared to manual compensation. In (a) the data points of the TIC (crosses) are shown from a region of a manually compensated dataset with the resulting pharmacokinetic curve. In (b), the data points from a region of an automatically compensated dataset with the resulting pharmacokinetic model curve is depicted. The gray circle points at differences between the measurements. The data have been log-converted and are therefore declared as arbitrary unit.

the *automatic approach* delivered inferior results to the *manual approach*. The deterioration was caused by residual motion in the automatic compensation result. The approach was not able to completely align the frame set. This results in two adjacent peaks in the TIC which can-

not be represented by the pharmacokinetic model (cf. Figure 5.12b). The fitting quality was lower compared to the manual result where no residual motion is present.

DISCUSSION   The evaluation of contrast data using pharmacokinetic modeling shows that the plausibility of the data towards perfusion physiologies is increased by the proposed approach. The measured improvement was much higher for pancreas datasets than for bowel datasets (36.8 % improvement compared to 7.0 % on average for MAD measure). This is caused by the amount of motion, which is much higher in pancreatic datasets due to the proximity to a pulsating aorta and the diaphragm. The large vessels surrounding the pancreas also exhibit contrast enhancement. If motion is present, the resulting TIC for a location might be mixed by inclusion of different perfusion characteristics. The increased amount of motion also leads to more frames declared out-of-plane (19.7 % compared to only 6.9 % for bowel datasets).

The *manual* and the *automatic approach* are able to improve the accuracy of fit to the perfusion model. The average improvement for both approaches are approximately at the same level. However, there are large differences between pancreas and bowel datasets.

The box plots in Figure 5.9 for pancreas datasets reveal that the *automatic approach* produces more variability. In case of the healthy control group, the results of both approaches are comparable. There is no significant difference, the average values are almost equal and the median is slightly better for *automatic* processing. In case of the pathological group, the *manual* compensation produces significantly better results with a median at 55.5 % compared to 32.1 % for *automatic* compensation. The non-rigid transformation has difficulties in some of the cases leading to inferior results compared to *manual* motion compensation. The motion in pancreatic acquisitions sometimes is very large and makes it difficult to establish non-rigid deformations. Contrast enhancement in the pancreas with cystic fibrosis is higher compared to observations of the healthy pancreas. This also reveals a coarser structure of the tissue. This might be a reason why the datasets can be well aligned by the manual reference point method by medical experts leading to superior results.

The results for bowel datasets have different characteristics. Comparing the median values, both approaches show an equal performance (see Figure 5.10). This can be explained by the fact that the bowel datasets mostly exhibit linear motion effects. These can be compensated equally well by both approaches. A difference is observed between the results for the pathological group and the healthy control group. The accuracy of fit for the control group improves by 15.7 %. For the pathological group it deteriorates slightly by 1.7 %.

This can be explained by differences in the datasets. In the control group with healthy bowel the intestinal wall measures just a few millimeters. As a consequence, the ROI where the contrast uptake is deducted from is also small and thin. Motion will have a larger effect, giving the motion compensation the chance to improve the signal in terms of the perfusion model. In the pathological group, the ROI of contrast uptake is larger and thus, the motion compensation has smaller influence. Compared to the pancreatic datasets the amount of motion influence is smaller and there are no other objects in the surroundings which enhance. Accuracy of fit improvement to the perfusion model is less likely. The variance of the results for the *automatic approach* is always larger in positive direction (above median). The upper quartile range is larger compared to the results of the *manual approach* and the lower quartile range is smaller. The *automatic approach* produces less variability of the accuracy of fit with respect to the deterioration of the results.

Overall, the *automatic approach* produces results of at least the quality of a manual expert driven approach in terms of an evaluation with a pharmacokinetic model for the two bowel dataset groups and the healthy control group of pancreas datasets. For the pathological group of the pancreas datasets, the *automatic approach* is able to improve the quality of fit to the pharmacokinetic model. It is, however, not capable of achieving comparable results to manual motion compensation.

5.4   EXPERIMENT 6: EVALUATION OF PERFUSION REGIONS

As subsequent step to the motion compensation, perfusion regions are established in the proposed CEUS analysis system through a region merging procedure (cf. Section 4.4.2). The regions are intended to segment the perfusion data for analysis of the TIC. The goal is to combine locations of similar TICs into one region in order to decrease the influence of noise by averaging and to avoid mixing locations with different TIC characteristics. The regions can either be used directly or the user can create new analysis regions by employing the perfusion information of the generated regions.

The evaluation of the perfusion regions aims at finding out if there is reason for the established regions. If the procedure defines a border between groups of perfusion measurements, there should be a difference between the samples.

Therefore, statistical difference analysis of the regions with respect to the TIC was performed. Given a set of perfusion regions two tests were conducted to measure the differences:

- inter-regional difference: for each two neighbored regions it was determined if they significantly differ
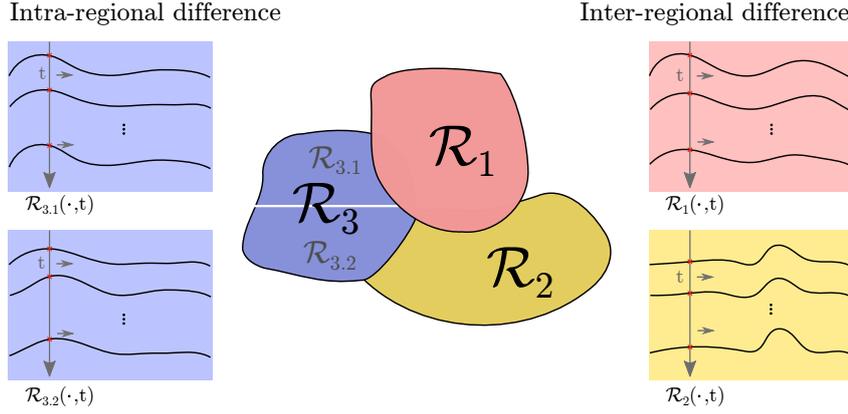
Figure 5.13: The inter-regional difference of two regions and the intra-regional difference of a region are schematically drawn.

- intra-regional difference: each region was randomly divided into two regions and it was tested whether they significantly differ

The difference test was performed as follows. For two regions $\mathcal{R}_1$ and $\mathcal{R}_2$ the contained TIC values for each point in time t were analyzed by two-sampled independent Student t-tests [Howell, 2011]. The null hypothesis was that $\mathcal{R}_1$ and $\mathcal{R}_2$ are similar assuming that the TIC values have not been generated by different distributions. The average of the p-values of the tests was determined:

$$\overline{p} = \frac{1}{t_{max} - |\Upsilon_f|} \sum_{t=1}^{t_{max}} \begin{cases} 0 & \text{if } t \in \Upsilon_f \\ STT(\mathcal{R}_1(\cdot, t), \mathcal{R}_2(\cdot, t)) & \text{else} \end{cases} . \quad (5.5)$$

t_{max} is the number of frames in the sequence. Υ_f contains all indices of excluded frames. $STT(\cdot, \cdot)$ calculates the p-value of a two-sampled independent t-test of two data vectors and $\mathcal{R}(\cdot, t))$ represents the tth values of all TICs in a region $\mathcal{R}$ (see Figure 5.13). If $\overline{p}$ was below a significance level of 0.05 the null hypothesis was rejected and both regions differ significantly.

For regions generated by the region merging procedure the ratio r_{br} of the number of significant and non-significant inter-regional differences was calculated as well as the ratio r_{wr} of the number of significant and non-significant intra-regional differences:

$$r_{br} = \frac{\text{\# significant inter-regional differences}}{\text{\# of non-significant inter-regional differences}} \quad (5.6)$$

and

$$r_{wr} = \frac{\text{\# significant intra-regional differences}}{\text{\# of non-significant intra-regional differences}}. \quad (5.7)$$

It is expected that r_{br} is considerably larger than r_{wr} to express the inhomogeneity between regions and the homogeneity inside regions.

Table 5.3: The average inter-regional ratio r_{br} and intra-regional ratio r_{wr} of the 36 (3 threshold parameter for each of the 12 datasets) runs measured in the original contrast data and the PCA-reconstructed contrast data.

	average r_{br}	average r_{wr}
Original data	174.1	8.6
PCA-reconstructed	7.2	0.9

The statistical difference analysis was applied to the original TIC from contrast data and the PCA-reconstructed signals from contrast data which has also been used in the region merging procedure (cf. Section 4.4.1). This is done to assess differences caused by PCA information reduction. Additionally, three different parameters for the similarity threshold were used for the region merging procedure so that approximately 200, 150 and 100 regions were generated, respectively. This was done to assess if the results were depending on the region merging parameter. The 12 datasets from database II were motion compensated using the proposed approach from Section 4.3 and were used for the analysis region evaluation.

RESULTS   The results are listed in Table 5.3. The average inter-regional ratio in original data was 174.1 and in the PCA-reconstructed data it was 7.2. Using the original data, the average intra-regional ratio was 7.2 and using the PCA-reconstructed data the average intra-regional ratio was 0.9. For original data and PCA-reconstructed data, the average inter-regional ratios are much higher than the average intra-regional ratios.

The variation of all ratios was very high (up to the fifth of the mean). However, the inter-regional ratio was always higher than the intra-regional ratio except for two runs using the original data and for one run using the PCA-reconstructed data. Results for the single datasets can be found in Table A.17 for PCA-reconstructed contrast data and in Table A.18 for original contrast data.

DISCUSSION   The evaluation states that there is a difference when comparing TICs inside a region or TICs of a region with neighbored regions. The results demonstrate plausibility for the regions generated by the region merging procedure. There are approximately 20 times more significant differences between region TICs compared to intra-regional differences. The PCA reconstruction used for region generation is suitable to create the analysis regions as shown by measurements in the original data. However, the results decrease when

the significance analysis is performed using the PCA-reconstructed image sequences. There are only seven times more significant differences between region TICs than intra-regional differences. Thus, it cannot be ruled out that the PCA eliminates information other than noise. Further, it is not guaranteed that the best possible regions in terms of our test method are found.

5.5   EXPERIMENT 7: EVALUATION OF SYSTEM WORKFLOW

This experiment addresses the time consumption of the whole system. To compare the components of the proposed approach, the time consumption of a manual processing was assessed as well. All steps have been conducted by a domain expert. The manual approach consisted of two steps:

- manual exclusion of *out-of-plane* frames by a frame-by-frame analysis and

- manual correction of motion shifts by clicking a reference point in each frame.

In comparison to the manual processing, the system workflow for this experiment consisted of the two steps:

- manual generation (drawing) of frame sets to exclude *out-of-plane* motion frames using the similarity matrix view and

- automatic motion compensation using the proposed approach from Section 4.3.

The first step models the worst case scenario. I. e., the assumption is that both automatic and semi-automatic approaches fail and the user has to establish regions manually just by the help the similarity matrix view. The time measurements for the second step include the setting of parameters for the automatic motion compensation.

RESULTS   Measurements have been performed using a quad-core Intel® i7-2720QM™ computer with 8 GB RAM and the datasets of database I. The proposed system was able to save 41 % of time compared to the manual processing (see Figure 5.14). The automatic motion compensation step saved 54 % of the time for the manual compensation. The frame selection using the similarity matrix view saved 21 % compared to the time it takes to manually exclude *out-of-plane* frames. According to a Wilcoxon signed rank test for paired data with independent distribution the improvements of the processing times were significant.

Average time consumption of the workflow
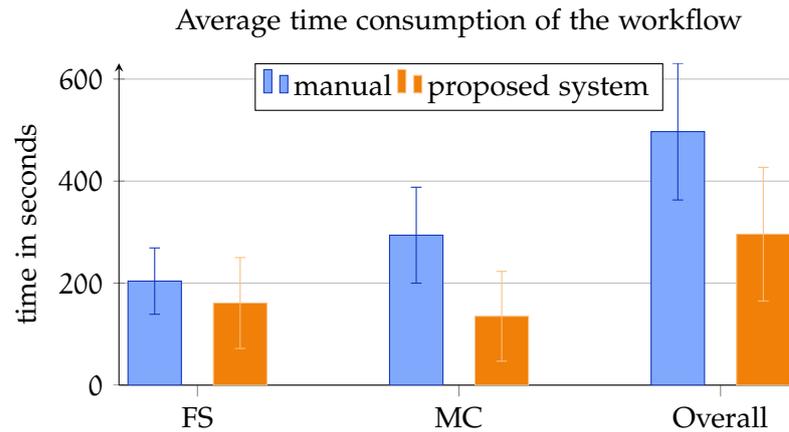
Figure 5.14: Time consumption of the frame selection (FS) and the motion compensation (MC) for manual processing and for the proposed system. The error bars indicate the SD of the measurements.

DISCUSSION   The proposed methods for frame selection and motion compensation are able to save valuable time. The time for manual generation of regions has been assessed as worst case scenario. The use of the suggested methods for automatic and semi-automatic region generation (cf. Section 4.2) is able to give good initializations so that in practice the saving of time should be even larger.

As the MRF-based motion compensation approach makes use of the multi-core architecture of the computer, the computation times for motion compensation can be further improved as well. This can be done by using a computer with more than four cores or a high performance computer cluster allowing the calculation of a high number of parallel processes.

5.6 DISCUSSION

The different experiments and their results previously described prove the suitability of the CEUS analysis system. As central part, the motion compensation is able to improve results compared to uncompensated data for the majority of the datasets. For each group of the datasets, the proposed system yields comparable results to manual motion compensation in terms of the pharmacokinetic model evaluation except for the group of pathological datasets from the pancreas. However, the automatic motion compensation significantly improves the quality of fit to the model compared to uncompensated datasets, it is not able to achieve comparable results to manual motion compensation for this particular group.

The automatic approach runs unsupervised once the parameters have been set. Transformation parameters such as the number of

B-spline control points can be set according to empirical tests. The amount of parameters to be set by the user is minimal. Most of the parameters have shown to be robust against changes.

Multiple datasets can be motion compensated by a batch process defining parameters before-hand. This is a huge advantage compared to manual processing. Another advantage is that the proposed system introduces non-rigid transformation to account for local deformation. This type of transformation can hardly be obtained by user input, at least not without investing a lot of manual work and time. The results of the proof-of-concept evaluation give evidence that the motion compensation result is improved by the non-rigid registration step.

For the frame selection and motion compensation the required processing time is almost halved with respect to a comparable manual process eliminating inter- and intra-observer variability.

The process of the frame selection is simplified by the generation of the similarity matrix view. This is confirmed by discussions with medical experts and the reduced processing time. Experts also stated that it is easier to navigate to areas where motion is present and to go back and forth to establish and place regions in the similarity matrix view. The suggested semi-automatic region generation is able to provide good initializations which are adaptable afterwards. For the semi-automatic approach the user is able to reduce the sensitivity by changing the parameters and getting immediate feedback on the resulting regions. The fully automatic approach tends to generate small regions. This is good to prevent erroneously classifying *out-of-plane* motion frames as *in-plane* frames. However, there are numerous cases where the sensitivity has to be adapted to include more frames.

The system generates analysis regions with similar perfusion characteristics. The evaluation of this component is experimental, as it does not make use of physiological knowledge. However, regions proposed by the region merging procedure are proven to have less significant intra-regional differences than significant differences between adjacent regions. Considering the fact that regions are used to extract averaged TICs for analysis, the proposed regions are superior to manually drawn regions. The user can redefine own regions if the proposed regions are not suitable by being fully aware about the underlying perfusion characteristics. It is therefore possible to draw manual regions using the underlying knowledge of the region merging result.

SUMMARY AND FUTURE WORK

This thesis presents a system for the analysis of CEUS sequences. First, the medical and technical background of CEUS acquisition is described. Second, methods addressing the deficiencies such as motion are discussed for the scenario of CEUS sequence analysis. Subsequently, a prototypical analysis system for CEUS sequences for perfusion imaging is described. It is the first system to establish an automated and standardized work flow addressing the pre-processing and analysis of CEUS after acquisition. The most important aims are to support medical experts and researchers by efficient automatic or semi-automatic methods in order to minimize intra- and inter-observer variability and to save time.

The presented system comprises techniques for the selection of sequence frames to exclude frames with out-of-plane motion. Motion compensation is subsequently applied to the selected frames using translation and non-rigid transformations. The optimal transformation parameters are found by discrete optimization using an MRF. The MRF incorporates spatial and temporal restrictions to adjacent transformation parameters to overcome weak information from noisy data. In the last step of the system, the motion compensated datasets are segmented according to their perfusion characteristics. These automatic regions can be used to derive quantitative parameters for diagnosis.

The quality of the different parts of the system for automatic and semi-automatic processing was evaluated in different experiments. For evaluation either 40 patient datasets of the pancreas and the bowel or a subset of 12 random datasets have been used.

A matrix view of inter-frame similarities was proposed for the frame selection. A semi-automatic and an automatic approach were used to identify sets of connected frames. The semi-automatic approach proved to be superior in comparison to the full-automatic approach using two parameters which can be easily found as the feedback is given in real-time. On average, the semi-automatic approach was able to find 75.8 % of the ground truth frame sets defined by a domain expert.

The temporal restrictions in the motion compensation approach increased the quality of the results compared to unconstrained processing of the datasets. The parameters for motion compensation were robust.

A combination of manual frame selection using the proposed similarity matrix view and automatic motion compensation was tested against manual processing of 40 clinical datasets. The automatic approach was able to save 41 % of time overall. As this approach runs unsupervised, the user is not required during the motion compensation process in contrast to the manual approach where extensive user input is required.

Motion compensated perfusion data have been evaluated using a pharmacokinetic model as an indicator of improved validity. For bowel datasets, the automatic approach was able to achieve equal results compared to manual expert driven motion compensation. On average, the automatic method improved the quality of fit to the model by 7.0 % compared to uncompensated CEUS data.

For the pancreas datasets which are more challenging due to the bigger motion impact, the automatic motion compensation only produces equal results for the group of healthy pancreas, 27.1 % on average. For the pathological group, it is not able to achieve equal results compared to manual compensation. However, the quality of fit to the perfusion model was still improved by 46.5 % compared to the uncompensated datasets.

Overall, the presented system showed very good and promising results for two different exemplary applications, bowel and pancreatic datasets. The analysis of CEUS data is a very challenging task due to the specific restrictions of the acquisition. In particular this is the low SNR in US, the motion influence and the problem of out-of-plane motion in 2D imaging. In clinical practice, the analysis of CEUS sequences requires a lot of training and expertise to be used in diagnosis.

The results of the automatic motion compensation for bowel datasets are remarkable, as the approach is able to keep up with the quality of manual processing where the experience of a skilled physician is used to interpret and process the data. For pancreas datasets, good results have been achieved. Although, it is not able to reach similar results compared to manual processing for both groups, the proposed approach has the ability to cover specific non-linear deformation fields. In combination with the time saving aspect, the automatic regions combining perfusion characteristics establish the basis for automated and efficient analysis and studies of CEUS data.

The main contribution of this pioneer work is the introduction of a compact and prototypic system for the analysis of perfusion in CEUS. It allows the standardized processing of these datasets for clinical research. With the integrated workflow it is possible to conduct different studies with a large number of datasets by easy automation of the most time consuming processes. Furthermore, it is the first approach proposing an automatic method to select frames for diagnosis

which have not been affected by out-of-plane motion. In the absence of ground truth for the evaluation of the motion compensation, extensive experiments to prove the proposed concept and to assess its performance were conducted for two different applications.

6.1 FUTURE WORK

To improve the motion compensation quality for datasets with large motion influence, the introduction of a more complex motion model is required [Weruaga et al., 2003; Vedam et al., 2004]. A motion model simulating respiratory motion effects and periodically induced pulsation motion can help to overcome the restricted image quality and the noise artifacts in US data. However, in most publications this requires the segmentation of the data in order to identify organs and to assign a predefined model. This is not suitable for CEUS images because the results of the segmentation will depend on the deficiencies in the data as well.

For the application to densely perfused CEUS datasets, the use of the contrast sequence can help to improve the registration quality further. This requires a model to interpret the differences in intensity between different frames such as a pharmacokinetic model.

Regarding the analysis of perfusion regions, physiological and medical experiments have to be conducted to find out if the regions are depicting different types of pathologies. For example, this could be achieved by comparing the physiological parameters of the pharmacokinetic model between the different analysis regions or between different patients and pathologies [Radjenovic et al., 2008]. Histopathologic findings can be used as ground truth for evaluation or training.

Other applications of CEUS need to be separately addressed to evaluate the performance of the proposed system. This is particularly necessary to use the system outside of medical research. Additionally, the study should be extended to contain more datasets in order to confirm the findings.

The use of the proposed approach in other imaging modalities and applications is also possible, for example in angiographic datasets of coronary arteries. Currently the system is designed to process 2D image data acquired over time. Of course, the most interesting aspect is the application to datasets where out-of-plane motion should be considered before motion compensation and analysis.

APPENDIX

This chapter lists all results from the experiments described in Chapter 5 in the following tables. See Sections 5.2 to 5.5 for the description of the respective experiments.

Table A.1: The overlapping percentage of frame sets generated with the semi-automatic approach with ground truth frame sets for two groups of bowel datasets.

healthy control group	overlapping percentage
HB01	0.746
HB07	0.971
HB12	0.874
HB15	0.629
HB17	0.806
HB19	0.847
HB21	0.745
HB27	0.966
HB29	0.942
HB30	0.673
patholocial group	overlapping percentage
SB15	0.878
SB16	0.891
SB17	0.911
SB18	0.861
SB20	0.980
SB22	0.948
SB23	0.727
SB24	0.929
SB25	0.712
SB26	0.652
average	0.795
SD	0.214

Table A.2: The overlapping percentage of frame sets generated with the semi-automatic approach with ground truth frame sets for two groups of pancreatic datasets.

healthy control group	overlapping percentage
HP02	0.447
HP04	0.523
HP06	0.688
HP09	0.603
HP12	0.780
HP16	0.689
HP18	0.744
HP20	0.705
HP23	0.711
HP24	0.727
pathological group	overlapping percentage
SP01	0.794
SP13	0.878
SP15	0.499
SP23	0.735
SP26	0.913
SP27	0.737
SP29	0.531
SP31	0.490
SP33	0.612
SP34	0.825
average	0.649
SD	0.197

Table A.3: Performance of different classifiers for all 40 datasets of database I using 12 Gabor filter kernels and 8 different Haralick features and super/subsampling.

classifier	TP	FN	FP	TN	g-mean
Naive Bayes	731	656	2764	6241	0.60
KNN (k=10)	1030	357	2315	6690	0.74
LogitBoost	1070	317	4119	4886	0.65
ADTree (iter.: 100)	1012	375	2743	6262	0.71
SVM (c 1, radial)	0	1387	0	9005	0.00
SVM (c 1, polyn.)	0	1387	0	9005	0.00
SVM (c 1, linear)	703	684	4023	4982	0.53
J48	880	507	1362	7643	0.73
AdaBoost	945	442	3682	5323	0.63
Voting approach	996	391	2740	6265	0.70

Table A.4: Performance of different classifiers for all 40 datasets of database I using 12 Gabor filter kernels and 8 different Haralick features and cost weighting.

classifier	TP	FN	FP	TN	g-mean
Naive Bayes	726	661	2736	6269	0.60
KNN (k=10)	1090	297	2643	6362	0.75
LogitBoost	1142	245	4475	4530	0.64
ADTree (iter.: 100)	1085	302	3109	5896	0.72
SVM (c 1, radial)	0	1387	0	9005	0.00
SVM (c 1, polyn.)	0	1387	0	9005	0.00
SVM (c 1, linear)	808	579	4878	4127	0.52
J48	785	602	1254	7751	0.70
AdaBoost	947	440	3691	5314	0.63
Voting approach	1089	298	3240	5765	0.71

Table A.5: Performance of different classifiers for all 40 datasets of database I using 8 different Haralick features and super/subsampling.

classifier	TP	FN	FP	TN	g-mean
Naive Bayes	523	864	1879	7126	0.55
KNN (k=10)	1120	267	2364	6641	0.77
LogitBoost	1141	246	4510	4495	0.64
ADTree (iter.: 100)	1041	346	3010	5995	0.71
SVM (c 1, radial)	0	1387	0	9005	0.00
SVM (c 1, polyn.)	0	1387	0	9005	0.00
SVM (c 1, linear)	752	635	4894	4111	0.50
J48	906	481	1663	7342	0.73
AdaBoost	0	1387	0	9005	0.00
Voting approach	1032	355	2756	6249	0.72

Table A.6: Performance of different classifiers for all 40 datasets of database I using 8 different Haralick features and cost weighting.

classifier	TP	FN	FP	TN	g-mean
Naive Bayes	520	867	1859	7146	0.55
KNN (k=10)	1121	266	2440	6565	0.77
LogitBoost	1132	255	4427	4578	0.64
ADTree (iter.: 100)	1061	326	3171	5834	0.70
SVM (c 1, radial)	0	1387	0	9005	0.00
SVM (c 1, polyn.)	139	1248	7204	1801	0.14
SVM (c 1, linear)	841	546	4573	4432	0.55
J48	899	488	1986	7019	0.71
AdaBoost	1010	377	3802	5203	0.65
Voting approach	1120	267	3201	5804	0.72

Table A.7: Evaluation of the intensity profiles in terms of the SD in B-mode data for different motion compensation setups using the 12 datasets of database II. The data in the columns *non-rigid* has been generated using translation transformation followed by B-spline based transformation.

dataset	transl.	transl. w/o t.[a]	non-rigid	non-rigid w/o t.[a]
1	1.300	1.263	1.475	1.246
2	1.072	1.074	1.207	1.214
3	1.026	1.023	1.117	1.121
4	1.178	1.178	1.328	1.325
5	1.218	1.220	1.396	1.396
6	1.153	1.149	1.295	1.260
7	1.083	1.072	1.200	1.189
8	1.056	1.061	1.158	1.162
9	1.241	1.237	1.547	1.536
10	1.066	1.025	1.202	1.134
11	1.185	1.172	1.320	1.310
12	1.019	1.060	1.182	1.202
average	1.133	1.128	1.286	1.258
SD	0.092	0.085	0.133	0.119

[a] without temporal constraints

Table A.8: Evaluation of the smoothness of intensity profiles in contrast data for different motion compensation setups using the 12 datasets of database II. The data in the columns *non-rigid* has been generated using translation transformation followed by B-spline based transformation.

dataset	transl.	transl. w/o t.[a]	non-rigid	non-rigid w/o t.[a]
1	1.073	1.033	1.083	1.131
2	0.954	0.942	0.968	0.954
3	0.983	0.952	1.009	0.945
4	1.013	0.985	1.001	0.941
5	1.011	0.975	1.004	0.970
6	1.036	0.932	1.050	1.030
7	1.409	1.273	1.321	1.168
8	1.050	1.016	1.063	1.004
9	1.062	1.059	1.064	1.057
10	1.003	0.958	1.012	0.943
11	1.088	1.089	1.076	1.075
12	1.039	1.021	1.132	1.019
average	1.060	1.019	1.065	1.020
SD	0.116	0.094	0.092	0.076

[a] without temporal constraints

Table A.9: Results for the *automatic approach* are shown for the bowel datasets of the healthy control group. The quality of fit to the pharmacokinetic model is expressed by the MAD and the RMS error.

Patient	time fe[a] (s)	time mp[b] (s)	MAD imp[c]	RMS imp	MTT	# excl. frames	# total frames
HB01	121	82	1.695	1.772	4.210	11	248
HB07	54	88	0.933	1.038	6.300	3	281
HB12	57	182	1.025	1.068	7.050	3	274
HB15	192	87	1.137	1.049	4.320	47	286
HB17	245	45	0.887	0.953	0.850	101	284
HB19	95	134	1.055	1.055	5.820	3	273
HB21	81	104	0.821	0.873	9.270	16	273
HB27	87	143	1.650	1.128	5.150	3	251
HB29	75	231	1.334	1.414	3.820	3	269
HB30	270	371	1.033	1.252	8.160	81	269
average	127.7	146.7	1.157	1.160	5.495	27.1	270.8
median	91.0	119.0	1.044	1.062	5.485	7.0	273.0

[a] frame exclusion
[b] motion compensation
[c] improvement factor

Table A.10: Results for the *manual approach* are shown for the bowel datasets of the healthy control group. The quality of fit to the pharmacokinetic model is expressed by the MAD and the RMS error.

Patient	time fe[a] (s)	time mp[b] (s)	MAD imp[c]	RMS imp	MTT	# excl. frames	# total frames
HB01	202	264	1.489	1.636	6.15	5	248
HB07	116	261	0.825	0.954	4.21	0	281
HB12	120	290	1.037	1.010	5.73	0	274
HB15	134	232	1.091	1.154	10.3	5	286
HB17	199	268	1.067	1.213	1.14	85	284
HB19	101	276	0.749	0.835	5.08	0	273
HB21	125	257	1.037	1.024	6.21	0	273
HB27	103	301	2.126	1.582	5.66	0	251
HB29	114	296	1.334	1.324	2.96	0	269
HB30	422	561	0.670	0.821	1.64	60	269
average	163.6	300.6	1.142	1.155	4.908	15.5	270.8
median	122.5	272.0	1.052	1.089	5.370	0.0	273.0

[a] frame exclusion
[b] motion compensation
[c] improvement factor

Table A.11: Results for the *automatic approach* are shown for the bowel datasets of the pathological group. The quality of fit to the pharmacokinetic model is expressed by the MAD and the RMS error.

Patient	time fe[a] (s)	time mp[b] (s)	MAD imp[c]	RMS imp	MTT	# excl. frames	# total frames
SB15	63	129	1.167	1.295	3.58	8	206
SB16	59	55	0.824	1.128	5.3	3	200
SB17	33	70	0.926	0.936	4.63	3	192
SB18	34	96	0.963	0.944	2.47	3	220
SB20	43	71	0.937	0.962	3.9	4	275
SB22	58.8	127	0.934	0.963	4.35	3	260
SB23	39	452	0.980	1.028	5.15	3	267
SB24	64	72	1.131	1.031	2.67	4	288
SB25	120	62	0.915	1.061	3.55	32	280
SB26	94	41	1.052	1.008	5.09	23	269
average	60.8	117.5	0.983	1.036	4.069	8.6	245.7
median	58.9	71.5	0.950	1.018	4.125	3.5	263.5

[a] frame exclusion
[b] motion compensation
[c] improvement factor

Table A.12: Results for the *manual approach* are shown for the bowel datasets of the pathological group. The quality of fit to the pharmacokinetic model is expressed by the MAD and the RMS error.

Patient	time fe[a] (s)	time mp[b] (s)	MAD imp[c]	RMS imp	MTT	# excl. frames	# total frames
SB15	127	160	0.975	0.947	3.29	27	206
SB16	88	197	0.736	0.834	4.15	0	200
SB17	67	126	1.016	0.989	4.28	0	192
SB18	91	339	0.868	0.910	1.87	0	220
SB20	113	144	0.908	0.921	3.43	0	275
SB22	109	257	0.881	0.947	4.01	0	260
SB23	115	179	1.072	1.158	5.33	0	267
SB24	126	278	0.859	0.875	2.42	0	288
SB25	178	293	1.039	1.178	1.41	74	280
SB26	161	183	0.987	0.960	4.33	34	269
average	117.5	215.6	0.934	0.972	3.452	13.5	245.7
median	114.0	190.0	0.941	0.947	3.720	0.0	263.5

[a] frame exclusion
[b] motion compensation
[c] improvement factor

Table A.13: Results for the *automatic approach* are shown for the pancreatic datasets of the healthy control group. The quality of fit to the pharmacokinetic model is expressed by the MAD and the RMS error.

Patient	time fe[a] (s)	time mp[b] (s)	MAD imp[c]	RMS imp	MTT	# excl. frames	# total frames
HP02	243	88	0.916	0.861	5.89	86	281
HP04	275	269	1.080	1.228	2.41	44	263
HP06	311	51.6	0.331	0.324	8.31	40	287
HP09	193	115	0.700	0.744	5.67	53	310
HP12	185	67	1.485	1.590	1.6	41	273
HP16	164	149	1.724	1.930	2.09	60	296
HP18	221	335	2.069	1.841	4.81	25	289
HP20	162	80	1.363	1.310	2.67	39	294
HP23	426	158	2.024	2.316	7.33	46	292
HP24	276	96	1.020	1.285	0.96	35	278
average	245.6	140.9	1.271	1.343	4.174	46.9	286.3
median	232.0	105.5	1.221	1.298	3.740	42.5	288.0

[a] frame exclusion
[b] motion compensation
[c] improvement factor

Table A.14: Results for the *manual approach* are shown for the pancreatic datasets of the healthy control group. The quality of fit to the pharmacokinetic model is expressed by the MAD and the RMS error.

Patient	time fe[a] (s)	time mp[b] (s)	MAD imp[c]	RMS imp	MTT	# excl. frames	# total frames
HP02	306	216	1.116	1.022	11.93	39	281
HP04	301	285	1.055	1.175	1.99	60	263
HP06	252	201	0.991	1.106	3.03	31	287
HP09	272	349	0.731	0.767	2.79	112	310
HP12	220	239	1.573	1.443	2.03	85	273
HP16	220	282	2.000	2.178	2.78	69	296
HP18	210	365	1.086	1.026	0.85	48	289
HP20	233	283	1.174	1.218	3.27	43	294
HP23	234	377	2.024	2.330	5.36	50	292
HP24	299	362	1.209	1.213	1.27	64	278
average	254.7	295.9	1.296	1.348	3.530	60.1	286.3
median	243.0	284.0	1.145	1.194	2.785	55.0	288.0

[a] frame exclusion
[b] motion compensation
[c] improvement factor

Table A.15: Results for the *automatic approach* are shown for the pancreatic datasets of the pathological group. The quality of fit to the pharmacokinetic model is expressed by the MAD and the RMS error.

Patient	time fe[a] (s)	time mp[b] (s)	MAD imp[c]	RMS imp	MTT	# excl. frames	# total frames
SP01	112	49	2.299	1.598	5.72	30	181
SP13	54	111	0.964	1.210	5.96	4	340
SP15	338	75	2.226	3.779	5.64	143	351
SP23	236	97	1.479	0.943	8.4	68	367
SP26	223	81	1.348	0.962	5.59	51	211
SP27	268	154	1.182	1.333	1.04	56	286
SP29	380	202	1.903	1.891	36.87	114	298
SP31	236	198	0.916	0.923	36.86	109	286
SP33	254	126	1.043	1.140	6.16	62	327
SP34	143	78	1.293	1.307	7	33	279
average	224.4	117.1	1.465	1.509	11.924	67.0	292.6
median	236.0	104.0	1.321	1.259	6.060	59.0	292.0

[a] frame exclusion
[b] motion compensation
[c] improvement factor

Table A.16: Results for the *manual approach* are shown for the pancreatic datasets of the pathological group. The quality of fit to the pharmacokinetic model is expressed by the MAD and the RMS error.

Patient	time fe[a] (s)	time mp[b] (s)	MAD imp[c]	RMS imp	MTT	# excl. frames	# total frames
SP01	184	197	1.633	1.347	7.2	26	181
SP13	135	380	1.278	1.424	4.17	0	340
SP15	333	345	1.676	1.046	11.2	170	351
SP23	331	362	1.554	1.110	10.84	154	367
SP26	322	374	1.634	1.281	12.2	65	211
SP27	288	304	1.479	1.558	2.93	113	286
SP29	306	296	1.558	1.472	13.97	91	298
SP31	253	378	0.947	0.939	12.17	84	286
SP33	356	334	1.098	1.159	5.57	68	327
SP34	284	654	1.640	1.592	8.63	61	279
average	279.2	362.4	1.450	1.293	8.888	83.2	292.6
median	297.0	353.5	1.556	1.314	9.735	76.0	292.0

[a] frame exclusion
[b] motion compensation
[c] improvement factor

Table A.17: Results for all datasets of database II for the inter-
regional and intra-regional ratios of the segmentation
of perfusion characteristics. The segmentation was cal-
culated using PCA-reconstructed contrast data.

dataset	threshold[a]	# regions	ratio (r_{br})[b]	ratio (r_{wr})[c]
1	3.0	152	0.04	0.03
	4.0	61	0.07	0.03
	5.0	35	0.13	0.03
2	7.5	206	1.59	0.67
	8.5	144	1.67	0.69
	10.5	75	1.85	0.88
3	2.0	429	0.10	0.09
	3.0	116	0.20	0.10
	4.0	57	0.32	0.16
4	7.5	187	1.45	0.56
	8.5	116	1.60	0.55
	9.5	82	2.12	0.82
5	3.0	345	0.82	0.36
	4.0	138	1.73	0.38
	5.0	75	1.01	0.42
6	3.0	245	1.72	0.55
	3.5	147	2.15	0.63
	4.0	59	2.70	0.84
7	2.0	328	0.07	0.10
	3.0	97	0.19	0.09
	4.0	56	0.30	0.19
8	11.0	281	0.81	0.19
	12.0	208	0.95	0.24
	13.0	164	1.13	0.29
9	9.0	131	0.54	0.28
	10.0	93	0.74	0.37
	11.0	73	0.75	0.35

[a] threshold parameter of region merging appraoch

[b] inter-regional ratio of significant and non-significant differences

[c] intra-regional ratio of significant and non-significant differences

Table A.17: Results perfusion segmentation using PCA-reconstructed contrast data. (continued)

dataset	threshold	# regions	ratio (r_{br})	ratio (r_{wr})
10	4.0	143	1.80	0.83
	5.0	91	1.93	0.65
	6.0	53	2.72	0.83
11	4.0	167	70.82	6.59
	5.0	69	53.56	5.90
	6.0	44	93.60	7.80
12	4.0	107	1.93	0.75
	5.0	64	2.61	1.00
	6.0	36	4.25	1.25
average			7.22	0.99

Table A.18: Results for all datasets of database II for the inter-regional and intra-regional ratios of the segmentation of perfusion characteristics. The segmentation was calculated using original contrast data.

dataset	threshold[a]	# regions	ratio (r_{br})[b]	ratio (r_{wr})[c]
1	3	152	2.08	0.57
	4	61	3.96	0.85
	5	35	4.78	1.06
2	7.5	206	10.46	2.61
	8.5	144	9.58	2.69
	10.5	75	8.13	2.95
3	2	429	0.13	0.09
	3	116	0.21	0.10
	4	57	0.39	0.16

[a] threshold parameter of region merging appraoch
[b] inter-regional ratio of significant and non-significant differences
[c] intra-regional ratio of significant and non-significant differences

Table A.18: Results perfusion segmentation using original contrast data. (continued)

dataset	threshold	# regions	ratio (r_{br})	ratio (r_{wr})
4	7.5	187	10.59	1.37
	8.5	116	22.91	1.58
	9.5	82	18.20	3.10
5	3	345	9.90	1.57
	4	138	89.89	3.93
	5	75	27.91	5.25
6	3	245	21.57	1.82
	3.5	147	44.66	2.87
	4	59	189.11	6.38
7	2	328	6.22	1.78
	3	97	464.60	7.08
	4	56	13.13	17.67
8	11	281	1.01	0.22
	12	208	1.13	0.26
	13	164	1.34	0.32
9	9	131	19.77	12.10
	10	93	21.05	12.29
	11	73	30.66	9.43
10	4	143	8.86	4.96
	5	91	16.21	3.14
	6	53	15.60	2.53
11	4	167	3464.25	17.56
	5	69	782.00	68.00
	6	44	236.50	44.00
12	4	107	51.51	25.75
	5	64	30.50	31.00
	6	36	630.00	12.00
average			174.13	8.58

Abdi, H. and Williams, L. J.: Principal component analysis, *Wiley Interdisciplinary Reviews: Computational Statistics*, 2(4): 433–459, 2010. (Cited on page 57.)

Adluru, G.; DiBella, E. V. and Schabel, M. C.: Model-based registration for dynamic cardiac perfusion MRI, *Journal of Magnetic Resonance Imaging*, 24(5): 1062–1070, 2006. (Cited on page 29.)

Albrecht, T.; Thorelius, L.; Solbiati, L.; Cova, L. and Frauscher, F.: *Contrast-enhanced ultrasound in clinical practice: liver, prostate, pancreas, kidney and lymph nodes*, Springer, 2005. (Cited on page 10.)

Angelelli, P.; Nylund, K.; Gilja, O. H. and Hauser, H.: Interactive visual analysis of contrast-enhanced ultrasound data based on small neighborhood statistics, *Computers & Graphics*, 35(2): 218–226, 2011. (Cited on pages 13 and 15.)

Arivazhagan, S. and Ganesan, L.: Texture segmentation using wavelet transform, *Pattern Recognition Letters*, 24(16): 3197–3203, 2003. (Cited on page 40.)

Averkiou, M.; Powers, J.; Skyba, D.; Bruce, M. and Jensen, S.: Ultrasound contrast imaging research, *Ultrasound Quarterly*, 19(1): 27–37, 2003. (Cited on page 7.)

Axel, L.: Cerebral blood flow determination by rapid-sequence computed tomography: theoretical analysis, *Radiology*, 137(3): 679–686, 1980. (Cited on page 6.)

Bankman, I. N.: *Handbook of Medical Imaging: Processing and Analysis*, Academic Press, 2000. (Cited on pages 1 and 2.)

Basarab, A.; Lyshchik, A.; Grava, C.; Buzuloiu, V. and Delachartre, P.: Ultrasound image sequence registration and its application for thyroid nodular disease, *Journal of Signal Processing Systems*, 55(1-3): 127–137, 2008. (Cited on page 22.)

Bauer, E. and Kohavi, R.: An empirical comparison of voting classification algorithms: Bagging, boosting, and variants, *Machine Learning*, 36(1-2): 105–142, 1999. (Cited on pages 41 and 67.)

Bay, H.; Ess, A.; Tuytelaars, T. and Gool, L. V.: Speeded-up robust features (SURF), *Computer Vision and Image Understanding*, 110(3): 346–359, 2008. (Cited on page 21.)

Beauchemin, S. S. and Barron, J. L.: The computation of optical flow, *ACM Computing Surveys*, 27(3): 433–466, 1995. (Cited on page 17.)

Besag, J.: Spatial interaction and the statistical analysis of lattice systems, *Journal of the Royal Statistical Society. Series B (Methodological)*, 36(2): 192–236, 1974. (Cited on pages 45 and 46.)

Beutel, J.; Kundel, H. and Van Metter, R. L.: *Handbook of medical imaging: physics and psychophysics*, number 1 in Handbook of Medical Imaging, SPIE Press, 2000. (Cited on page 3.)

Birkfellner, W.: *Applied medical image processing: a basic course*, Biomedical Engineering, Taylor & Francis, 2010. (Cited on page 1.)

Bloch, S. H.; Dayton, P. A. and Ferrara, K. W.: Targeted imaging using ultrasound contrast agents, *IEEE Engineering in Medicine and Biology*, 23(5): 18–29, 2004. (Cited on page 10.)

Böhm, I. and Schild, H. H.: A practical guide to diagnose lesser-known immediate and delayed contrast media-induced adverse cutaneous reactions, *European Radiology*, 16(7): 1570–9, 2006. (Cited on page 6.)

Bookstein, F. L.: Principal warps: thin-plate splines and the decomposition of deformations, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6): 567–585, 1989. (Cited on page 20.)

Bosch, J. G.; Mitchell, S. C.; Lelieveldt, B. P. F.; Nijland, F.; Kamp, O.; Sonka, M. and Reiber, J. H. C.: Automatic segmentation of echocardiographic sequences by active appearance motion models, *IEEE Transactions on Medical Imaging*, 21(11): 1374–83, 2002. (Cited on page 26.)

Boykov, Y. and Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(9): 1124–37, 2004. (Cited on pages 48, 51, and 55.)

Boykov, Y.; Veksler, O. and Zabih, R.: Fast approximate energy minimization via graph cuts, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11): 1–18, 2001. (Cited on pages 46, 47, 48, 49, 50, 51, and 55.)

Bridal, S. and Correas, J.: Milestones on the road to higher resolution, quantitative, and functional ultrasonic imaging, *Proceedings of the IEEE*, 91(10): 1543–1561, 2003. (Cited on page 2.)

Brix, G.; Semmler, W.; Port, R.; Schad, L. R.; Layer, G. and Lorenz, W. J.: Pharmacokinetic parameters in CNS Gd-DTPA enhanced MR imaging, *Journal of Computer-assisted Tomography*, 15(4): 621–628, 1991. (Cited on pages 13 and 75.)

Bruhn, A.; Weickert, J. and Schnörr, C.: Lucas/Kanade meets Horn/Schunck: combining local and global optic flow methods, *International Journal of Computer Vision*, 61(3): 211–231, 2003. (Cited on page 18.)

Burges, C.: A tutorial on support vector machines for pattern recognition, *Data Mining and Knowledge Discovery*, 2(2): 121–167, 1998. (Cited on pages 41 and 67.)

Challa, S.: *Fundamentals of object tracking*, Cambridge books online, Cambridge University Press, 2011. (Cited on pages 23 and 26.)

Chan, V. and Perlas, A.: *Basics of ultrasound imaging*, Springer, 2011. (Cited on page 11.)

Chen, J.-F.; Fowlkes, J. B.; Carson, P. L. and Rubin, J. M.: Determination of scan-plane motion using speckle decorrelation: theoretical considerations and initial test, *International Journal of Imaging Systems and Technology*, 8(1): 38–44, 1997. (Cited on page 22.)

Chen, W.; Giger, M. and Bick, U.: A Fuzzy C-Means (FCM)-Based Approach for Computerized Segmentation of Breast Lesions in Dynamic Contrast-Enhanced MR Images, *Academic Radiology*, 13(1): 63–72, 2006. (Cited on page 15.)

Chen, Z.: Bayesian filtering: from Kalman filters to particle filters, and beyond, *Statistics*, 182(1): 1–69, 2003. (Cited on page 26.)

Chomas, J. E.; Dayton, P.; Allen, J.; Morgan, K. and Ferrara, K. W.: Mechanisms of contrast agent destruction, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 48(1): 232–48, 2001. (Cited on page 8.)

Christensen, G.; Miller, M.; Vannier, M. and Grenander, U.: Individualizing neuro-anatomical atlases using a massively parallel computer, *Computer*, 29(1): 32–38, 1996. (Cited on page 20.)

Claudon, M.; Dietrich, C. F.; Choi, B. I.; Cosgrove, D. O.; Kudo, M.; Nolsøe, C. P.; Piscaglia, F.; Wilson, S. R.; Barr, R. G.; Chammas, M. C.; Chaubal, N. G.; Chen, M. H.; Clevert, D. A.; Correas, M.; Ding, H.; Forsberg, F.; Fowlkes, J. B.; Gibson, R. N.; Goldberg, B. B.; Lassau, N.; Leen, E. L. S.; Mattrey, R. F.; Moriyasu, F.; Solibiati, L.; Weskott, H. P. and Xu, H. X.: Guidelines and good clinical practice recommendations for contrast enhanced ultrasound (CEUS) in the liver - update 2012: a WFUMB-EFSUMB initiative in cooperation with representatives of AFSUMB, AIUM, ASUM, FLAUS and ICUS, *Ultrasound in Medicine & Biology*, 39(2): 187–210, 2012. (Cited on pages 3 and 10.)

Clifford, P.: Markov random fields in statistics, in *Disorder in Physical Systems*, 19–32, Oxford Univ. Press, 1990. (Cited on page 46.)

Cohen, B. and Dinstein, I.: New maximum likelihood motion estimation schemes for noisy ultrasound images, *Pattern Recognition*, 35(2): 455–463, 2002. (Cited on pages 22 and 26.)

Comaniciu, D.; Ramesh, V. and Meer, P.: Kernel-based object tracking, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(5): 564–577, 2003. (Cited on pages 23 and 26.)

Correas, J.-M.; Bridal, L.; Lesavre, A.; Méjean, A.; Claudon, M. and Hélénon, O.: Ultrasound contrast agents: properties, principles of action, tolerance, and artifacts, *European Radiology*, 11(8): 1316–1328, 2001. (Cited on page 7.)

Cosgrove, D. and Eckersley, R.: Contrast-enhanced ultrasound: basic physics and technology overview, in R. Lencioni, editor, *Enhancing the Role of Ultrasound with Contrast Agents*, 3–14, Springer, 2006. (Cited on page 10.)

Coto, E.; Grimm, S.; Bruckner, S.; Gröller, E.; Kanitsar, A. and Rodriguez, O.: MammoExplorer: An advanced CAD application for breast DCE-MRI, *Proceedings of Vision, Modelling, and Visualization 2005*, 91–98, 2005. (Cited on page 13.)

Cremers, D. and Schnörr, C.: Statistical shape knowledge in variational motion segmentation, *Image and Vision Computing*, 21(1): 77–86, 2003. (Cited on page 18.)

Crum, W.; Hartkens, T. and Hill, D.: Non-rigid image registration: theory and practice, *British Journal of Radiology*, 77(Special Issue 2): S140, 2004. (Cited on page 69.)

Deserno, T. M.: *Biomedical image processing*, Biological and Medical Physics, Biomedical Engineering, Springer, 2011. (Cited on pages 2 and 3.)

Detre, J. A.; Leigh, J. S.; Williams, D. S. and Koretsky, A. P.: Perfusion imaging, *Magnetic Resonance in Medicine*, 23(1): 37–45, 1992. (Cited on pages 2, 5, 6, and 28.)

Dhawan, A. P.: *Medical image analysis*, IEEE Press Series on Biomedical Engineering, Wiley, 2011. (Cited on page 4.)

Dougherty, G.: *Medical image processing: techniques and applications*, Biological and Medical Physics, Biomedical Engineering, Springer, 2011. (Cited on page 1.)

Durrett, R.: *Probability: theory and examples*, number v. 3 in Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2010. (Cited on page 44.)

Erchinger, F.; Dimcevski, G.; Engjom, T. and Gilja, O.: Transabdominal ultrasonography of the pancreas: basic and new aspects, *Imaging in Medicine*, 3(4): 411–422, 2011. (Cited on page 10.)

Felzenszwalb, P. F. and Huttenlocher, D. P.: Efficient graph-based image segmentation, *International Journal of Computer Vision*, 59(2): 167–181, 2004. (Cited on pages 37 and 38.)

Frangi, A. F.; Laclaustra, M. and Lamata, P.: A registration-based approach to quantify flow-mediated dilation (FMD) of the brachial artery in ultrasound image sequences, *IEEE Transactions on Medical Imaging*, 22(11): 1458–69, 2003. (Cited on pages 22 and 28.)

Freund, Y. and Mason, L.: The alternating decision tree learning algorithm, in *Machine Learning International Workshop*, 124–133, 1999. (Cited on pages 41 and 67.)

Friedman, J.; Hastie, T. and Tibshirani, R.: Additive logistic regression: a statistical view of boosting, *The Annals of Statistics*, 28(2): 337–407, 2000. (Cited on pages 41 and 67.)

Frouin, F.; Duteil, S.; Lesage, D.; Carlier, P. G.; Herment, A. and Leroy-Willig, A.: An automated image-processing strategy to analyze dynamic arterial spin labeling perfusion studies. Application to human skeletal muscle under stress., *Magnetic Resonance Imaging*, 24(7): 941–51, 2006. (Cited on page 23.)

Gauthier, M.; Tabarout, F.; Leguerney, I.; Polrot, M.; Pitre, S.; Peronneau, P. and Lassau, N.: Assessment of quantitative perfusion parameters by dynamic contrast-enhanced sonography using a deconvolution method, *Ultrasound in Medicine*, 31(4): 595–608, 2012. (Cited on page 14.)

Geman, S. and Geman, D.: Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6: 721–741, 1984. (Cited on page 46.)

Gibbs, V.; Cole, D. and Sassano, A.: *Ultrasound physics and technology: how, why, and when*, How, Why and When Series, Churchill Livingstone, 2009. (Cited on page 11.)

Gibson, J. J.: On the analysis of change in the optic array, *Scandinavian Journal of Psychology*, 18(1): 161–163, 1977. (Cited on page 18.)

Glaßer, S.; Niemann, U.; Preim, U.; Spiliopoulou, M. and Preim, B.: Classification of benign and malignant DCE-MRI breast tumors by analyzing the most suspect region, in *Bildverarbeitung für die Medizin (BVM)*, 45–50, 2013. (Cited on page 15.)

Glaßer, S.; Schäfer, S.; Oeltze, S.; Preim, U.; Tönnies, K. and Preim, B.: A visual analytics approach to diagnosis of breast DCE-MRI data, in *Vision Modelling and Visualization*, 351–362, Braunschweig, 2009. (Cited on pages 14 and 56.)

Glocker, B.; Paragios, N.; Komodakis, N.; Tziritas, G. and Navab, N.: Optical flow estimation with uncertainties through dynamic MRFs, in *Computer Vision and Pattern Recognition (CVPR)*, 1–8, 2008. (Cited on page 18.)

Glocker, B.; Sotiras, A.; Komodakis, N. and Paragios, N.: Deformable medical image registration: setting the state of the art with discrete methods, *Annual Review of Biomedical Engineering*, 13: 219–44, 2011. (Cited on page 45.)

Golemati, S.; Stoitsis, J. and Nikita, K. S.: Motion analysis of the carotid artery wall and plaque using B-mode ultrasound, *Vascular Disease Prevention*, 4(4): 296–302, 2007. (Cited on page 22.)

Goshtasby, A.: *2-D and 3-D image registration: for medical, remote sensing, and industrial applications*, Wiley-Interscience, 2005. (Cited on page 19.)

Gramiak, R. and Shah, P. M.: Echocardiography of the aortic root, *Investigative Radiology*, 3(5): 356–366, 1968. (Cited on page 7.)

Granville, V.; Krivanek, M. and Rasson, J.-P.: Simulated annealing: a proof of convergence, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(6): 652–656, 1994. (Cited on page 52.)

Guerrero, J.; Salcudean, S. E.; McEwen, J. a.; Masri, B. a. and Nicolaou, S.: Real-time vessel segmentation and tracking for ultrasound imaging applications, *IEEE Transactions on Medical Imaging*, 26(8): 1079–90, 2007. (Cited on page 27.)

Hall, M.; Frank, E.; Holmes, G. and Pfahringer, B.: The WEKA data mining software: an update, *ACM SIGKDD Explorations Newsletter*, 11(1): 10–18, 2009. (Cited on pages 41 and 67.)

Haralick, R. M.: Statistical and structural approaches to texture, *Proceedings of the IEEE*, 67(5), 1979. (Cited on page 40.)

Haralick, R. M.; Shanmugam, K. and Dinstein, I.: Textural features for image classification, *IEEE Transactions on Systems, Man and Cybernetics*, SMC-3(6): 610–621, 1973. (Cited on page 40.)

Hardaway, R. M.: *Capillary perfusion in health and disease*, Futura Pub. Co., 1981. (Cited on page 5.)

Harvey, C. J.; Pilcher, J. M.; Eckersley, R. J.; Blomley, M. J. K. and Cosgrove, D. O.: Advances in ultrasound, *Clinical Radiology*, 57(3): 157–77, 2002. (Cited on page 2.)

Hayton, P.; Brady, M.; Tarassenko, L. and Moore, N.: Analysis of dynamic MR breast images using a model of contrast enhancement, *Medical Image Analysis*, 1(3): 207–224, 1997. (Cited on page 29.)

He, H. and Garcia, E. A.: Learning from imbalanced data, *IEEE Transactions on Knowledge and Data Engineering*, 21(9): 1263–1284, 2009. (Cited on page 42.)

Hendee, W. R. and Ritenour, E. R.: *Medical imaging physics*, Wiley, 2003. (Cited on page 1.)

Heywang-Köbrunner, S. H.; Viehweg, P.; Heinig, A. and Küchler, C.: Contrast-enhanced MRI of the breast: accuracy, value, controversies, solutions, *European Journal of Radiology*, 24(2): 94–108, 1997. (Cited on page 5.)

Hill, D. L. G.; Batchelor, P. G.; Holden, M. and Hawkes, D. J.: Medical image registration, *Physics in Medicine and Biology*, 46(3): R1–45, 2001. (Cited on pages 19, 20, and 69.)

Horn, B. and Schunck, B.: Determining optical flow, *Artificial Intelligence*, 17(1): 185–203, 1981. (Cited on page 18.)

Hotelling, H.: *Analysis of a complex of statistical variables into principal components*, Warwick & York, 1933. (Cited on page 57.)

Howell, D. C.: *Statistical Methods for Psychology*, PSY 613 Qualitative Research and Analysis in Psychology Series, Wadsworth Cengage Learning, 8 edition, 2011. (Cited on page 83.)

Iniewski, K.: *Medical Imaging: Principles, Detectors, and Electronics*, Wiley, 2009. (Cited on page 2.)

Jaini, A. K. and Farrokhnia, F.: Unsupervised texture segmentation using Gabor filters, *Pattern Recognition*, 24(12): 1167–1186, 1991. (Cited on page 41.)

Jan, J.: *Medical image processing, reconstruction and restoration: concepts and methods*, Signal Processing and Communications, Taylor & Francis, 2010. (Cited on page 4.)

Jayaraman, S.; Esakkirajan, S. and Veerakumar, T.: *Digital image processing*, Tata McGraw Hill Education, 2011. (Cited on page 58.)

Jensen, J. A.: FIELD: a program for simulating ultrasound systems, in *10th Nordicbaltic Conference on Biomedical Imaging*, 351–353, 1996. (Cited on page 69.)

Jiřík, R.; Nylund, K.; Gilja, O. H.; Mézl, M.; Harabiš, V.; Kolář, R.; Standara, M. and Taxt, T.: Ultrasound perfusion analysis combining bolus-tracking and burst-replenishment, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 60(2): 310–9, 2013. (Cited on pages 14, 65, and 76.)

Jiřík, R.; Nylund, K.; Taxt, T.; Mézl, M.; Hausken, T.; Harabiš, V.; Kolář, R.; Standara, M. and Gilja, O. H.: Parametric ultrasound perfusion analysis combining bolus tracking and replenishment, in *Proceedings of the IEEE International Ultrasonics Symposium*, 5–8, 2012. (Cited on pages 14, 29, and 75.)

Kalman, R.: A new approach to linear filtering and prediction problems, *Journal of Basic Engineering*, 82(Series D): 35–45, 1960. (Cited on page 26.)

Kittler, J.; Hatef, M.; Duin, R. P. W. and Matas, J.: On combining classifiers, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3): 226–239, 1998. (Cited on page 42.)

Kohli, P. and Torr, P. H. S.: Efficiently solving dynamic markov random fields using graph cuts, in *Computer Vision (ICCV)2*, 922–929, 2005. (Cited on page 47.)

Kolář, R.; Jiřík, R.; Harabiš, V.; Mézl, M. and Bartoš, M.: Advanced methods for perfusion analysis in echocardiography, *Physiological Research / Academia Scientiarum Bohemoslovaca*, 59 Suppl 1: S33–41, 2010. (Cited on page 75.)

Kolmogorov, V. and Zabih, R.: What energy functions can be minimized via graph cuts?, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(2): 147–59, 2004. (Cited on pages 47 and 55.)

Kono, Y.; Pinnell, S. P.; Sirlin, C. B.; Sparks, S. R.; Georgy, B.; Wong, W. and Mattrey, R. F.: Radiology carotid arteries: contrast-enhanced US angiography - preliminary clinical experience, *Radiology*, 230(2): 561–568, 2004. (Cited on page 10.)

Korstanje, J.-W. H.; Selles, R. W.; Stam, H. J.; Hovius, S. E. R. and Bosch, J. G.: Development and validation of ultrasound speckle tracking to quantify tendon displacement, *Journal of Biomechanics*, 43(7): 1373–9, 2010. (Cited on page 26.)

Kremkau, F. W.; Gramiak, R.; Carstensen, E. L.; Shah, P. M. and Kramer, D. H.: Ultrasonic detection of cavitation at catheter tips,

The Journal of the Acoustical Society of America, 45: 340, 1969. (Cited on page 7.)

Kubat, M.; Holte, R. and Matwin, S.: Machine learning for the detection of oil spills in satellite radar images, *Machine Learning*, 30(2-3): 195–215, 1998. (Cited on page 67.)

Kuhl, C. K.: The current status of breast MR imaging, part I, *Radiology*, 244(2): 356–378, 2007. (Cited on page 5.)

Laporte, C. and Arbel, T.: Learning to estimate out-of-plane motion in ultrasound imagery of real tissue, *Medical Image Analysis*, 15(2): 202–13, 2011. (Cited on page 23.)

Larsson, H. B. W.; Stubgaard, M.; Frederiksen, J. L.; Jensen, M.; Henriksen, O. and Paulson, O. B.: Quantitation of blood-brain barrier defect by magnetic resonance imaging and gadolinium-DTPA in patients with multiple sclerosis and brain tumors, *Magnetic Resonance in Medicine*, 16(1): 117–31, 1990. (Cited on page 13.)

Lassau, N.; Koscielny, S.; Albiges, L.; Chami, L.; Benatsou, B.; Chebil, M.; Roche, A. and Escudier, B. J.: Metastatic renal cell carcinoma treated with sunitinib: early evaluation of treatment response using dynamic contrast-enhanced ultrasonography, *Clinical Cancer Research*, 16(4): 1216–25, 2010. (Cited on page 13.)

Ledesma-Carbayo, M. J.; Kybic, J.; Desco, M.; Santos, A.; Sühling, M.; Hunziker, P. and Unser, M.: Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation, *IEEE Transactions on Medical Imaging*, 24(9): 1113–26, 2005. (Cited on pages 21 and 22.)

Leighton, T. G.: *The acoustic bubble*, Academic Press, 1997. (Cited on pages 2 and 7.)

LeVine, H.: *Medical Imaging*, Health and Medical Issues Today, Greenwood Publishing Group Incorporated, 2010. (Cited on page 1.)

Levman, J.; Leung, T.; Causer, P.; Plewes, D. and Martel, A. L.: Classification of dynamic contrast-enhanced magnetic resonance breast lesions by support vector machines, *IEEE Transactions on Medical Imaging*, 27(5): 688–696, 2008. (Cited on page 15.)

Li, M.; Kambhamettu, C. and Stone, M.: Automatic contour tracking in ultrasound images, *Clinical Linguistics & Phonetics*, 19(6-7): 545–554, 2005. (Cited on page 27.)

Li, S.: Markov random field models in computer vision, in *Computer Vision ECCV*, volume 801, 361–370, Springer, 1994. (Cited on page 45.)

Lindner, J. R.: Microbubbles in medical imaging: current applications and future directions, *Nature Reviews. Drug Discovery*, 3(6): 527–32, 2004. (Cited on pages 3, 7, 8, and 9.)

Loizou, C.; Pattichis, C. S. and Pattichis, C.: *Despeckle filtering algorithms and software for ultrasound imaging*, Synthesis lectures on algorithms and software in engineering, Morgan & Claypool, 2008. (Cited on pages 7 and 11.)

Lowe, D.: Object recognition from local scale-invariant features, in *Proceedings of the Seventh IEEE International Conference on Computer Vision*, 1150–1157 vol.2, Ieee, 1999. (Cited on page 21.)

Lucas, B. and Kanade, T.: An iterative image registration technique with an application to stereo vision, in *Proceedings of Image Understanding Workshop*, volume 130, 121–130, 1981. (Cited on page 18.)

Lucht, R. E.; Knopp, M. V. and Brix, G.: Classification of signal-time curves from dynamic MR mammography by neural networks, *Magnetic Resonance Imaging*, 19(1): 51–7, 2001. (Cited on page 15.)

Luo, J. and Konofagou, E.: A fast normalized cross-correlation calculation method for motion estimation, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 57(6): 1347–57, 2010. (Cited on page 22.)

Mahapatra, D. and Sun, Y.: Integrating segmentation information for improved MRF-based elastic image registration, *IEEE Transactions on Image Processing*, 21(1): 170–83, 2012. (Cited on page 18.)

Maintz, J. and Viergever, M. a.: A survey of medical image registration, *Medical Image Analysis*, 2(1): 1–36, 1998. (Cited on page 4.)

McDonald, D. M. and Baluk, P.: Significance of blood vessel leakiness in cancer, *Cancer Research*, 62(18): 5381–5, 2002. (Cited on page 5.)

Melbourne, A.; Atkinson, D.; White, M. J.; Collins, D.; Leach, M. and Hawkes, D.: Registration of dynamic contrast-enhanced MRI using a progressive principal component registration (PPCR), *Physics in Medicine and Biology*, 52(17): 5147–5156, 2007. (Cited on page 28.)

Merritt, C. R. B.: Ultrasound safety: what are the issues?, *Radiology*, 173(2): 304–306, 1989. (Cited on pages 2 and 7.)

Mettler, F. A. and Guiberteau, M. J.: *Essentials of Nuclear Medicine Imaging*, Elsevier/Saunders, 2012. (Cited on page 6.)

Metz, C. T.; Klein, S.; Schaap, M.; van Walsum, T. and Niessen, W. J.: Nonrigid registration of dynamic medical imaging data using nD+t B-splines and a groupwise optimization approach, *Medical Image Analysis*, 2010. (Cited on page 28.)

Morawski, A. M.; Lanza, G. A. and Wickline, S. A.: Targeted contrast agents for magnetic resonance imaging and ultrasound, *Current Opinion in Biotechnology*, 16(1): 89–92, 2005. (Cited on page 6.)

Nascimento, J. and Marques, J.: Robust shape tracking with multiple models in ultrasound images, *IEEE Transactions on Image Processing*, 17(3): 392–406, 2008. (Cited on page 27.)

Nattkemper, T. W.; Arnrich, B.; Lichte, O.; Timm, W.; Degenhard, A.; Pointon, L.; Hayes, C. and Leach, M. O.: Evaluation of radiological features for breast tumour classification in clinical screening with machine learning methods, *Artificial Intelligence in Medicine*, 34(2): 129–39, 2005. (Cited on page 15.)

Nixon, M. and Aguado, A. S.: *Feature extraction & image processing*, Feature Extraction and Image Processing Series, Elsevier Science, 2008. (Cited on page 4.)

Nylund, K.; Jiřík, R.; Mézl, M.; Leh, S.; Hausken, T.; Pfeffer, F.; Ødegaard, S.; Taxt, T. and Gilja, O. H.: Quantitative contrast-enhanced ultrasound comparison between inflammatory and fibrotic lesions in patients with Crohn's disease, *Ultrasound in Medicine & Biology*, 39(7): 1197–1206, 2013. (Cited on pages 9, 10, 13, and 64.)

Oeltze, S.; Doleisch, H.; Hauser, H.; Muigg, P. and Preim, B.: Interactive visual analysis of perfusion data, *IEEE Transactions on Visualization and Computer Graphics*, 13(6): 1392–9, 2007. (Cited on page 13.)

Pandharipande, P. V.; Krinsky, G. a.; Rusinek, H. and Lee, V. S.: Perfusion imaging of the liver: current challenges and future goals, *Radiology*, 234(3): 661–73, 2005. (Cited on page 28.)

Perez, P.; Inria, I. and Beaulieu, C. D.: Markov random fields and images, *CWI Quarterly*, 11(4): 413–437, 1998. (Cited on page 45.)

Piscaglia, F.; Nolsøe, C.; Dietrich, C. F.; Cosgrove, D. O.; Gilja, O. H.; Bachmann Nielsen, M.; Albrecht, T.; Barozzi, L.; Bertolotto, M.; Catalano, O.; Claudon, M.; Clevert, D. A.; Correas, J. M.; D'Onofrio, M.; Drudi, F. M.; Eyding, J.; Giovannini, M.; Hocke, M.; Ignee, A.; Jung, E. M.; Klauser, A. S.; Lassau, N.; Leen, E.; Mathis, G.; Saftoiu, A.; Seidel, G.; Sidhu, P. S.; ter Haar, G.; Timmerman, D. and Weskott, H. P.: The EFSUMB guidelines and recommendations on the clinical practice of contrast enhanced ultrasound (CEUS): update 2011 on non-hepatic applications, *Ultraschall in der Medizin*, 33(1): 33–59, 2012. (Cited on pages 10 and 33.)

Postema, M. and Gilja, O.: Contrast-enhanced and targeted ultrasound, *World Journal of Gastroenterology*, 17(1): 28–41, 2011. (Cited on pages 8 and 10.)

Postema, M. and Gilja, O. H.: Ultrasound-directed drug delivery, *Current Pharmaceutical Biotechnology*, 8(6): 355–361, 2007. (Cited on page 8.)

Preim, B.; Oeltze, S.; Mlejnek, M.; Gröeller, E.; Hennemuth, A. and Behrens, S.: Survey of the visual exploration and analysis of perfusion data, *IEEE Transactions on Visualization and Computer Graphics*, 15(2): 205–20, 2009. (Cited on pages 2, 5, 12, and 17.)

Quaia, E.: Contrast-enhanced ultrasound of the small bowel in Crohn's disease, *Abdominal Imaging*, 2013. (Cited on pages 10 and 56.)

Quinlan, J. R.: *C4.5: programs for machine learning*, volume 1, Morgan Kaufmann, 1993. (Cited on pages 41 and 67.)

Radjenovic, A.; Dall, B. J.; Ridgway, J. P. and Smith, M. A.: Measurement of pharmacokinetic parameters in histologically graded invasive breast tumours using dynamic contrast-enhanced MRI, *British Journal of Radiology*, 81(962): 120–128, 2008. (Cited on pages 75 and 91.)

Renault, G.; Tranquart, F.; Perlbarg, V.; Bleuzen, A.; Herment, A. and Frouin, F.: A posteriori respiratory gating in contrast ultrasound for assessment of hepatic perfusion, *Physics in Medicine and Biology*, 50(19): 4465–80, 2005. (Cited on page 23.)

Revell, J.; Mirmehdi, M. and McNally, D.: Computer vision elastography: speckle adaptive motion estimation for elastography using ultrasound sequences, *IEEE Transactions on Medical Imaging*, 24(6): 755–66, 2005. (Cited on page 26.)

Rickes, S.; Mönkemüller, K. and Malfertheiner, P.: Acute severe pancreatitis: contrast-enhanced sonography, *Abdominal Imaging*, 32(3): 362–364, 2007. (Cited on page 7.)

Rohling, R.; Gee, A. and Berman, L.: Automatic registration of 3-D ultrasound images, *Ultrasound in Medicine & Biology*, 24(6): 841–854, 1998. (Cited on page 22.)

Rohr, K.: Elastic registration of multimodal medical images: a survey, *Künstliche Intelligenz*, 14(3): 11–17, 2000. (Cited on pages 20 and 21.)

Rueckert, D.; Aljabar, P.; Heckemann, R. a.; Hajnal, J. V. and Hammers, A.: Diffeomorphic registration using b-splines, in *Medical Image Computing and Computer-assisted Intervention (MICCAI)*, 702–9, 2006. (Cited on page 53.)

Rueckert, D.; Sonoda, L.; Hayes, C.; Hill, D.; Leach, M. and Hawkes, D.: Nonrigid registration using free-form deformations: application to breast MR images, *IEEE Transactions on Medical Imaging*, 18(8): 712–721, 1999. (Cited on pages 20 and 53.)

Sanches, J. M.; Laine, A. F. and Suri, J. S.: *Ultrasound imaging: advances and applications*, Springer New York, 2012. (Cited on page 3.)

Schäfer, S.; Angelelli, P.; Nylund, K.; Gilja, O. H. and Tönnies, K.: Registration of ultrasonography sequences based on temporal regions, in *7th Intl. Symp. on Image and Signal Processing and Analysis*, 749–759, 2011a. (Cited on page 12.)

Schäfer, S.; Nylund, K.; Gilja, O. H. and Tönnies, K.: Motion compensation of ultrasonic perfusion images, in *Proceedings of SPIE Medical Imaging: Ultrasonic Imaging, Tomography, and Therapy*, 2012. (Cited on page 70.)

Schäfer, S.; Preim, U.; Glaßer, S.; Preim, B. and Tönnies, K.: Local similarity measures for lesion registration in DCE-MRI of the breast, *Annals of the BMVA*, 2011(3): 1–13, 2011b. (Cited on pages 15, 29, and 69.)

Schäfer, S. and Toennies, K.: Registration of temporal ultrasonic image sequences using Markov random fields, in *Vision Modelling and Visualization*, 151–158, 2012. (Cited on pages 55 and 69.)

Shekhar, R.; Zagrodsky, V.; Garcia, M. J. and Thomas, J. D.: Registration of real-time 3-D ultrasound images of the heart for novel 3-D stress echocardiography, *IEEE Transactions on Medical Imaging*, 23(9): 1141–9, 2004. (Cited on page 22.)

Shekhovtsov, A.; Kovtun, I. and Hlaváč, V.: Efficient MRF deformation model for non-rigid image matching, *Computer Vision and Image Understanding*, 112(1): 91–99, 2008. (Cited on page 18.)

Solbiati, L.; Tonolini, M.; Cova, L. and Goldberg, S. N.: The role of contrast-enhanced ultrasound in the detection of focal liver lesions, *European Radiology*, 11(3 Suppl.): E15–E25, 2001. (Cited on page 10.)

Sonka, M. and Fitzpatrick, J. M., editors: *Handbook of medical imaging: medical image processing and analysis*, volume 2, International Society for Optical Engineering (SPIE), 2000. (Cited on page 1.)

Sowell, E. R.; Thompson, P. M.; Leonard, C. M.; Welcome, S. E.; Kan, E. and Toga, A. W.: Longitudinal mapping of cortical thickness and brain growth in normal children, *The Journal of Neuroscience*, 24(38): 8223–31, 2004. (Cited on page 21.)

Spilker, M. E.; Seng, K.-Y.; Yao, A. A.; Daldrup-Link, H. E.; Shames, D. M.; Brasch, R. C. and Vicini, P.: Mixture model approach to tumor classification based on pharmacokinetic measures of tumor permeability, *Journal of Magnetic Resonance Imaging*, 22(4): 549–58, 2005. (Cited on page 15.)

Stachowiak, G. P.; Podsiadlo, P. and Stachowiak, G. W.: A comparison of texture feature extraction methods for machine condition monitoring and failure analysis, *Tribology Letters*, 20(2): 133–147, 2005. (Cited on page 40.)

Strouthos, C.; Lampaskis, M.; Sboros, V.; McNeilly, A. and Averkiou, M.: Indicator dilution models for the quantification of microvascular blood flow with bolus administration of ultrasound contrast agents, *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 57(6): 1296–310, 2010. (Cited on pages 14 and 75.)

Szeliski, R.; Zabih, R.; Scharstein, D.; Veksler, O.; Kolmogorov, V.; Agarwala, A.; Tappen, M. and Rother, C.: A comparative study of energy minimization methods for Markov random fields with smoothness-based priors, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(6): 1068–80, 2008. (Cited on pages 47 and 51.)

Toennies, K. D.: *Guide to medical image analysis: methods and algorithms*, Advances in Computer Vision and Pattern Recognition, Springer, 2012. (Cited on pages 4, 10, and 28.)

Tofts, P. S.: Modeling tracer kinetics in dynamic Gd-DTPA MR imaging, *Journal of Magnetic Resonance Imaging*, 7(1): 91–101, 2005. (Cited on page 75.)

Tofts, P. S.; Brix, G.; Buckley, D. L.; Evelhoch, J. L.; Henderson, E.; Knopp, M. V.; Larsson, H. B. W.; Lee, T.-Y.; Mayr, N. A.; Parker, G. J. M.; Port, R. E.; Taylor, J. and Weisskoff, R. M.: Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI of a diffusable tracer: standardized quantities and symbols, *Journal of Magnetic Resonance Imaging*, 232: 223–232, 1999. (Cited on page 14.)

Tofts, P. S. and Kermode, a. G.: Measurement of the blood-brain barrier permeability and leakage space using dynamic MR imaging. 1. Fundamental concepts, *Magnetic Resonance in Medicine*, 17(2): 357–67, 1991. (Cited on pages 5 and 13.)

Unser, M.: Texture classification and segmentation using wavelet frames, *IEEE Transactions on Image Processing*, 4(11), 1995. (Cited on page 41.)

Uras, S.; Girosi, F.; Verri, A. and Torre, V.: A computational approach to motion perception, *Biological Cybernetics*, 60(2): 79–87, 1988. (Cited on page 18.)

Vannier, M. W.: Tumor imaging, analysis, and treatment planning, in M. Sonka and J. M. Fitzpatrick, editors, *Handbook of Medical Imaging: medical image processing and analysis*, chapter 18, 1131–1152, SPIE Press, 2000. (Cited on page 2.)

Vaseghi, S. V.: *Advanced digital signal processing and noise reduction*, Wiley, 2008. (Cited on page 57.)

Vedam, S. S.; Keall, P. J.; Docef, A.; Todor, D. a.; Kini, V. R. and Mohan, R.: Predicting respiratory motion for four-dimensional radiotherapy, *Medical Physics*, 31(8): 2274–2283, 2004. (Cited on page 91.)

Veenman, C.; Reinders, M. and Backer, E.: Resolving motion correspondence for densely moving points, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(1): 54–72, 2001. (Cited on page 26.)

Veksler, O.: *Efficient graph-based energy minimization methods in computer vision*, Ph.D. thesis, Cornell University, 1999. (Cited on page 44.)

Vogel, R.; Indermühle, A.; Reinhardt, J.; Meier, P.; Siegrist, P. T.; Namdar, M.; Kaufmann, P. a. and Seiler, C.: The quantification of absolute myocardial perfusion in humans by contrast echocardiography, *Journal of the American College of Cardiology*, 45(5): 754–62, 2005. (Cited on page 14.)

Wei, K.; Jayaweera, a. R.; Firoozan, S.; Linka, A.; Skyba, D. M. and Kaul, S.: Quantification of myocardial blood flow with ultrasound-induced destruction of microbubbles administered as a constant venous infusion, *Circulation*, 97(5): 473–83, 1998. (Cited on pages 8 and 13.)

Weruaga, L.; Morales, J.; Núñez, L. and Verdú, R.: Estimating volumetric motion in human thorax with parametric matching constraints, *IEEE Transactions on Medical Imaging*, 22(6): 766–772, 2003. (Cited on page 91.)

Witten, I. H.; Frank, E. and Hall, M. A.: *Data mining: practical machine learning tools and techniques*, The Morgan Kaufmann Series in Data Management Systems, Elsevier Science, 2011. (Cited on page 68.)

Woo, J.; Hong, B.-W.; Hu, C.-H.; Shung, K. K.; Kuo, C.-C. J. and Slomka, P. J.: Non-rigid ultrasound image registration based on intensity and local phase information, *Journal of Signal Processing Systems*, 54(1-3): 33–43, 2008. (Cited on page 22.)

Wu, X.; Kumar, V.; Ross Quinlan, J.; Ghosh, J.; Yang, Q.; Motoda, H.; McLachlan, G. J.; Ng, A.; Liu, B.; Yu, P. S.; Zhou, Z.-H.; Steinbach, M.; Hand, D. J. and Steinberg, D.: Top 10 algorithms in data mining, *Knowledge and Information Systems*, 14(1): 1–37, 2007. (Cited on page 41.)

Xiaohua, C.; Brady, M. and Rueckert, D.: Simultaneous segmentation and registration for medical image, in *Medical Image Computing and Computer-assisted Intervention (MICCAI)*, 663–670, Springer, 2004. (Cited on page 29.)

Yeung, F.; Levinson, S. F.; Fu, D. and Parker, K. J.: Feature-adaptive motion tracking of ultrasound image sequences using a deformable mesh, *IEEE Transactions on Medical Imaging*, 17(6): 945–56, 1998. (Cited on pages 22 and 26.)

Yilmaz, A.; Javed, O. and Shah, M.: Object tracking: a survey, *ACM Computing Surveys*, 38(4), 2006. (Cited on page 24.)

Zitova, B. and Flusser, J.: Image registration methods: a survey, *Image and Vision Computing*, 21(11): 977–1000, 2003. (Cited on pages 19 and 20.)

EHRENERKLÄRUNG

Ich versichere hiermit, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; verwendete fremde und eigene Quellen sind als solche kenntlich gemacht. Insbesondere habe ich nicht die Hilfe eines kommerziellen Promotionsberaters in Anspruch genommen. Dritte haben von mir weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Ich habe insbesondere nicht wissentlich:

- Ergebnisse erfunden oder widersprüchliche Ergebnisse verschwiegen,

- statistische Verfahren absichtlich missbraucht, um Daten in ungerechtfertigter Weise zu interpretieren,

- fremde Ergebnisse oder Veröffentlichungen plagiiert,

- fremde Forschungsergebnisse verzerrt wiedergegeben.

Mir ist bekannt, dass Verstöße gegen das Urheberrecht Unterlassungs- und Schadensersatzansprüche des Urhebers sowie eine strafrechtliche Ahndung durch die Strafverfolgungsbehörden begründen kann. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form als Dissertation eingereicht und ist als Ganzes auch noch nicht veröffentlicht.

Magdeburg, den 29.01.2014

_____

Sebastian Schäfer