UNIVERSITÀ CAMPUS BIO-MEDICO DI ROMA

FACOLTÀ DI INGEGNERIA

Laboratorio di Automatica

Dottorato di Ricerca in " Bioingegneria e Bioscienze"

Ciclo XXIX

Techniques
for
Deformable Organ Registration
in DCE Framework

Advisor

Prof. Roberto Setola

Ph.D Candidate

Leda Maria Montoni

Abstract

In the last decades, the development of non-invasive medical imaging modalities enabled remarkable progress in clinical diagnostics. Many diseases are nowadays detected by medical imaging techniques, which allows for earlier diagnosis and therapeutic intervention with a higher efficiency.

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is a clinical technique devoted to the study of lesions in organs. It is based on the analysis of perfusion dynamics of a contrast agent (CA) in the examined tissues by acquiring an image series over time, before, during and after the arrival of the CA. The acquired signal is analyzed for a selected region of interest (ROI) or individual pixels in each 2D slice of the series to generate an enhancement curve (EC), which reflects the tissue's response to the arrival of the CA by time-dependent enhancement values of the detected intensity. These curves provide a set of parameters, commonly visualized by 2D color maps, that yield valuable information for diagnosis. However, for a correct diagnosis, it is crucial to determine the EC with a high accuracy, which is often compromised by artifacts within the acquired time series due to patient motion. DCE algorithms assume that the analyzed ROI does not modify its shape and position over the entire time series, which is usually not a valid assumption, since the examined non-rigid structures are affected by patient motion and can change in position, shape and brightness. The compensation for such artifacts is complex, since the detected signal intensities intrinsically vary over time due to the diffusion of the CA .

In this thesis, we present a novel technique to compensate artifacts in DCE induced by motion/deformation of the ROI. The algorithm uses deformable active contours (AC) ,also called snakes, to define the deformation of the ROI in each image. We optimized the standard AC method by adding an additional term to the energy functional that is minimized during the procedure. Specifically, we included a distance map calculated with the Chamfer transform into the external forces that are responsible for the attraction of the snake to the edges in an image. This results in a considerably more accurate detection of the contour of the ROI by the snake and thus in a better segmentation. Subsequently, the center of mass and the principal axes for each frame are determined and used for a rototranslation compensation of misalignment between the frames. The algorithm was applied to different organs of the abdominal part and the results clearly show an improved accuracy of the corrected EC and the resulting parameters, when compared to the uncorrected EC. In addition to inter-frame misalignments, motion artifacts also cause displacements in the transverse direction relative to the image plane, thereby generating inconsistencies in the shape of the organ profiles in the same temporal series. Therefore, we extended our algorithm to four dimensions by using a hybrid approach based on ACs and a template matching algorithm based on the Chamfer distance. The frame-by-frame method segments the ROI of a particular frame via the AC model, subsequently, the final contour of the time frame is compared to the images of the same and adjacent planes of the successive frame and the best match is assigned to the previous image.

Finally, we introduce a fast and user-friendly tool to analyze 4D DCE-MRI data. The results can be visualized in different 3D representations and provide quantitative volumetric information in order to improve the understanding of the contrast agent perfusion in pathological tissues and the resulting diagnosis.

Contents

Chapter 1

Introduction

Rapid advancements in medical imaging technology have led to the introduction of a broad range of imaging modalities being used in clinical practice. The most common modalities are radiography, X-ray computed tomography (CT), magnetic resonance imaging (MRI) and ultra sound imaging (US). Medical imaging technique are used for several purposes in the medical field, such as image-guided diagnosis, treatment planning, or image-guided surgery.

Since the development of MRI by Lauterbur [1] and Mansfield [2] in the early 70s, it has become one of the widely used imaging modalities in modern medicine. It offers the possibility for non-invasive, high-resolution three-dimensional imaging without using ionizing radiation. The versatile possibilities of designing MRI experiments led to the development of an unparalleled variety of applications. Besides anatomical imaging modes with different contrast modalities, namely, *T1*, *T2* and proton density weighted imaging, a variety of functional imaging modalities can be performed, such as diffusion MRI [3] , functional brain imaging [4, 5]or ventilation imaging on[6] the lung to only name a few applications.

Since Raymond Damadian demonstrated that NMR can be used to distinguish healthy tissues from tumor tissue in 1971 [7], the diagnosis, grading and classification of tumors has benefited considerably from MRI, which is now essential for the adequate diagnosis and treatment of many tumor types. The ability of MRI to demonstrate tumor morphology and the relationships between malignant lesions and neighboring structures provides essential clinical information for both clinical management and surgical planning. The ability to acquire multiplanar images or even volume acquisitions with different contrast modalities is extremely valuable and provides the clinician with a true three-dimensional visualization of tumor and tissue morphology.

The development of cancer in the human body depends on a several factors, such as strongly increased cell replication in the cancerous tissue and the ability to evade programmed cell death [8]. A crucial requirement for the conversion from small malignant lesions ($< 200\mu$m) to a larger rapidly growing invasive neoplasm is the tumors ability to stimulate neighboring vasculature to initiate formation of new blood vessels that can grow into the tumor and supply it with oxygen and nutrients. The development of neovascularization within a tumor is known as angiogenesis[9] and the abnormal high existence of blood vessels in cancerous tissue is called hypervascularity. Furthermore, the morphology of the neovascular network in tumors usually differs significantly

from that seen in normal tissue. Tumor vasculature is often highly heterogeneous and the capillaries are irregular in shape and structure. The capillary walls have numerous openings, widened inter-endothelial junctions and a discontinuous or absent basement membrane. Consequently, tumor capillaries are extremely leaky compared to healthy vessels [10, 9], which results in altered blood perfusion dynamics.

In this context, dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) has become a very powerful and popular technique in oncology, as it allows to analyze blood perfusion dynamics in vessels by using a specific contrast agent (CA). More specifically, DCE-MRI relies on the acquisition of the MRI signal from the tissue of interest over a predefined period of time (time series or time sequence). First, baseline images without contrast enhancement are recorded, subsequently the contrast agent is injected and a series of images are acquired during and after the arrival of the CA in the tissue of interest. From the recorded signal it is possible to generate a time intensity curve, also referred to as enhancement curve (EC). It reflects the response of the tissue to the arrival of the CA in enhancement values, i.e., each data point of the enhancement curve depicts the intensity increase of the detected signal relative to the baseline signal in a single frame of the time series. These curves can be created from the signal of individual pixels or a region of interest in the images. By analyzing ECs, a set of parameters can be extracted, namely, the area under the curve (AUC), the local peak intensity (LPT) and the corresponding time to peak (TTP), and the average rising of the curve (SLOPE). These parameters are usually visualized by 2D color maps and provide important information for clinical and oncological diagnosis, in particular for the characterization of vessel morphology and the discrimination between healthy and malignant tissue.

A major problem regarding the correct determination of enhancement curves is the fact that standard DCE-MRI algorithms assume that the analyzed ROI does not modify its shape and position over the course of the entire time series. Unfortunately, this is not true in most cases, since DCE-MRI measurements are highly susceptible to errors induced by motion artifacts. The organs and regions of interest move and deform over the duration of the measurement due to breathing, heart contraction, organ peristalsis or shifting of the patient. This motion artifacts can cause significant errors in the estimation of an enhancement curve and the extracted parameters, leading to wrong or inaccurate diagnosis. However, motion correction in DCE-MRI is a challenging task, since intensity changes from one frame to the next can be caused by a superposition of changes in contrast agent concentration and motion artifacts. Therefore, common intensity-based registration algorithms are not a reliable approach for motion correction in DCE-MRI.

To overcome these problems, we present a novel technique to compensate for motion artifacts in DCE-MRI and ensure a more accurate diagnosis. In a first step, the proposed algorithm determines the optimal contour of the ROI via active deformable contours (AC) for each image in the sequence. To increase the reliability and precision of the contour extraction we modify the energy functional that is minimized during the execution of the AC algorithm. After the segmentation step, the center of mass and the principal axes for each ROI are calculated and subsequently a rototranslational transformation is applied to the ROIs in order to correct for misalignment between the individual frames. Applying the developed algorithm to clinical data of pathological abdominal organs,

namely, liver, spleen, kidney and bowel wall, clearly shows a clear improvement of the accuracy of the ECs and the characteristic parameters compared to the uncorrected results.

However, the undesired motions during acquisition do not only result in misalignments between the frames of a time sequences, but also cause displacements in the transverse direction relative to the image plane, so- called out-of-plane motion. This induces inconsistencies with regard to the shape of the organ profiles within a temporal sequence that can also compromise a correct data analysis. To account for this misalignments, we extend the algorithm to four dimension, i.e. we include the slices from the adjacent spatial planes into our analysis in order to reassign images that belong to the same temporal sequence. We use a particular temporal sequence that is initially selected manually by the radiologist and apply a hybrid algorithm based on active contour models and an template matching algorithm using the Chamfer distance transform. The active contours are used to automatically segment organs of interest in the respective slides. The final contour of an individual frame, which is a template image containing the shape contour of the organs of interest, is superimposed on the images of the same and the adjacent planes from the successive frame. This is done frame-by frame in order to create time sequences without inconsistencies in the organ profiles.

The described perfusion analysis of DCE-MRI data is usually performed on a temporal series of 2D images, hence, the results are also given in a two-dimensional representation, such as the enhancement curves or 2D color maps representing the associated parameters. To illustrate physiopathological information related to the dynamic properties of the observed process in three dimensions, we developed a tool that includes the volume data of the MRI exam into the DCE analysis, i.e. we use a 4D volume data set of spatially and temporarily ordered images and analyze and reconstruct the data. This allows to generate 3D renderings of the volume of interest (VOI) either in the original intensity representation or in the form of 3D color maps that represent the characteristic parameters of CA perfusion analysis. Furthermore, an enhancement surface can be generated in analogy to the enhancement curve in 3D. For a better visualization of the organ/lesion inner structure, we also performed suitable cutting that shows the core of the ROI. The tool enables fast generation of 3D results and features a graphical user interface for a simple and intuitive user interaction. In this way, it is possible for a medical doctor to retrieve novel information from a DCE-MRI examination, for example, by estimating the maximum intensity in space and locating the section with the highest vascularity.

In Chapter 2, we introduce the technical concepts necessary for understanding image formation in DCE-MRI. After introducing the theoretical foundation of nuclear magnetic resonance (NMR), we explain how NMR can be extended to a three-dimensional imaging modality, namely, MRI. Subsequently, we describe the concepts of DCE-MRI measurements and discuss the common problems of this technique.

An introduction to image processing techniques in medical imaging is presented in Chapter 3. We discuss the basic concepts of segmentation, registration and motion tracking, since these processing steps are essential to comprehend the presented algorithm. Finally, we describe several operations that are particularly important for the processing and visualization of three-dimensional data.

In Chapter 4, we explain the implementation of the developed motion tracking algorithm. First, a detailed mathematical description of the applied active contour model is given, followed by a delineation of the individual methodological steps of our technique. Finally, an approach for out-of-plane motion compensation is presented.

A detailed description of the 3D visualization and analysis tool is given in Chapter 5. We explain the underlying processing operations and show exemplary results of 3D visualizations of different organs. Furthermore, we present the developed graphical user interface that facilitates the user interaction with the program.

In Chapter 6, we present the results from the application of the motion tracking algorithm to clinical DCE-MRI data of abdominal organs, namely, bowel wall, kidney, liver and spleen. In order to quantify the improvement of the results when applying our method, we show a comparison of perfusion analysis results retrieved from data corrected with our algorithm and from data without correction. Moreover, we present results comparing our segmentation data to reference data in order to objectively quantify the performance of our active contour segmentation method.

Finally, in Chapter 7, we sum up the content of this work, discuss the presented methods and results, and propose approaches for further improving the analysis and visualization of DCE-MRI data.

Chapter 2

Theoretical and technical background of NMR, MRI and DCE-MRI

In this chapter the basic concepts of nuclear magnetic resonance (NMR), magnetic resonance imaging (MRI) and Dynamic contrast-enhanced MRI (DCE-MRI) are briefly introduced. For more detailed information the reader is referred to literature. A more extensive description of NMR physics can be found in the textbooks [11, 12]. A comprehensive overview about MRI concepts and sequences is given in [13, 14, 15]. For a more extensive description of DCE-MRI the reader is referred to [16][17].

2.1 Nuclear magnetic resonance

Nuclear magnetic resonance (NMR) was first described and measured by Isidor Rabi [18] in 1938. For his experiments on NMR in molecular beams Rabi was awarded with the Nobel Prize in Physics in 1944. In 1944, Rabi's experiments were carried out on liquids and solids by Felix Bloch [19] and Edward Purcell [20], which was again awarded with the Nobel Prize in Physics in 1952. In 1971, Raymond Damadian demonstrated with his experiments on laboratory animals that NMR can be used to distinguish healthy tissues from tumor tissue, thus encouraging researchers to consider NMR as a tool for disease detection [7]. He showed that $T1$ and $T2$ relaxation times of malignant tumors are significantly longer than for the corresponding normal tissues, thereby demonstrating the great potential of NMR for cancer diagnosis.

2.1.1 Nuclear spin and magnetic moment

In 1922, Otto Stern and Walther Gerlach performed an experiment [21], which is today often used to illustrate basic principles of quantum mechanics. In the Stern-Gerlach experiment, a beam of silver particles is sent through an inhomogeneous magnetic field and subsequently detected on a photographic plate in order to measure the particles' deflection. While classical mechanics predicts an isotropic distribution of angular momenta among the particles, resulting in a homogeneous pat-

tern on the photographic plate, the experiment actually showed two distinct deflection distribution directions. This result constitutes the first evidence of a quantized intrinsic angular momentum, termed spin. Nuclear magnetic resonance (NMR) relies on the magnetic properties of atomic particles such as electrons, protons, and neutrons, which form the building blocks of atoms and molecules. A nucleus of an atom consists of nucleons (neutrons and protons) and can produce an intrinsic angular momentum called the nuclear spin $\boldsymbol{I}$. In case of atoms with an odd number of protons and/or neutrons, such as ^1H, ^{13}C, and ^{19}F, the nuclear spin and the magnetic moment $\boldsymbol{\mu}$ is non-zero (Figure 2.1). The mathematical relation between the nuclear spin of an nucleus and its magnetic moment are given by:

$$\boldsymbol{\mu} = \gamma \boldsymbol{I} \tag{2.1}$$

The gyromagnetic ratio γ is a characteristic constant for every nucleus. For^1H the numerical value of the gyromagnetic ratio is$\gamma/2\pi = 42.58$ MHz/T, which is the highest known value for a stable nuclide. The hydrogen nucleus is the most used nucleus for medical applications of NMR, due to its magnetic spin properties and the fact that it has the highest abundance in biological structures (70%-80% of the human body consists of water).

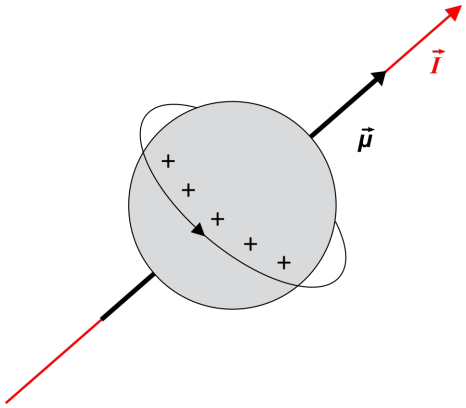

Figure 2.1: Spin and magnetic moment of a charged sphere. An atomic nucleus can be imagined as a rotating, positively charged sphere. The rotation of a charged particle, described by its angular momentum $\boldsymbol{I}$, results in an electric current, which induces a magnetic dipolar field. Direction and magnitude of this field are described by the magnetic moment $\boldsymbol{\mu}$ (figure from [13]).

2.1.2 External magnetic fields and macroscopic magnetization

NMR relies on the interaction of magnetic moments with external magnetic fields. Considering an arbitrary volume V containing N hydrogen atoms (protons), which have individual magnetic moments $\boldsymbol{\mu_i}$. The macroscopic magnetization $\vec{M}$ can be calculated as follows:

$$\boldsymbol{M} = \frac{1}{V} \sum_{i}^{N} \boldsymbol{\mu_i} \tag{2.2}$$

In the absence of an external magnetic field, nuclei do not have any preferred direction of alignment and every orientation is equally probable (left panel of Figure 2.2) . Consequently, the macroscopic magnetization is zero. In case of a stationary magnetic field $\boldsymbol{B}_0 = (0, 0, B_0)$ applied in z-direction, the nuclei precess around the direction of the field with a characteristic frequency (right panel of Figure 2.2). The so-called Larmor frequency ω_0 is directly proportional to the magnitude of the applied field and is given by:

$$\omega_0 = \gamma B_0 \tag{2.3}$$

For a nucleus with the spin quantum number I, $2I + 1$ quantum mechanical eigenstates exist; however, these states are degenerated without an external field. This degeneration vanishes if the nucleus is located inside an external magnetic field; this principle is called the Zeeman effect [22]. For the case[1] H and the external field in z-direction two eigenstates of the magnetic momentum exist:

$$\mu_z = \pm \frac{h}{4\pi}\gamma \tag{2.4}$$

where h is the Planck's constant ($h = 6.627 \cdot 10^{-34}$Js). This results in the two energy levels:

$$E_\pm = \mu_z B_0 \tag{2.5}$$

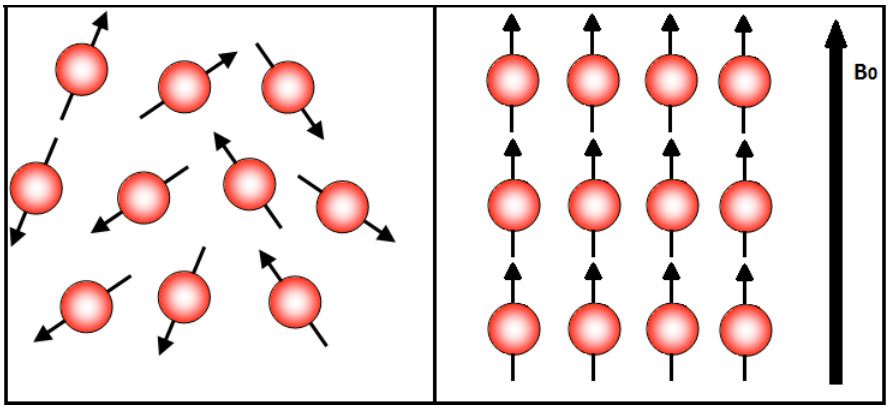

Figure 2.2: Spin orientation. When no external magnetic field is applied and the spins are oriented randomly (left side) . When an external magnetic field $\boldsymbol{B_0}$ is applied the spins precess around the axis of the field (right side).

Transitions between these states are possible, fulfilling the spin and the energy conservation law, which says that a particle with a nuclear spin equal to one (a photon) and with the energy $\triangle E$ and the Larmor frequency ω_0 has to be emitted or absorbed (Figure 2.3). The energy difference ΔE between the states is given by:

$$\Delta E = \frac{h\gamma}{2\pi} B_0 = \frac{h}{2\pi}\omega_0 \tag{2.6}$$

In thermal equilibrium, the ratio of the number of protons in the parallel state N^+(spin-up) and

in the anti-parallel state N^-(spin-down) relative to the magnetic field is given by the Boltzmann distribution:

$$\frac{N^-}{N^+} = \exp\left(-\frac{\Delta E}{kT}\right) \tag{2.7}$$

where k is the Boltzmann constant ($k = 1.38 \cdot 10^{-23}$J/K), T is the temperature in Kelvin and ΔE is the energy difference between the states. As the parallel state N^+ has a lower energy level, it has a slightly higher population, resulting in a net magnetization in z-direction. However, at room temperature (T=293 K) and a typical field strength of clinical MRI scanners ($B_0 = 3$ T) the relative population difference is only a few parts per million (ppm). This leads to a very small net magnetization and explains the inherently low sensitivity of NMR. The macroscopic magnetization $\boldsymbol{M}$, can be decomposed into two components (Figure 2.4):

- A longitudinal component M_z, i.e. parallel to $\boldsymbol{B_0}$
- A transverse component M_{xy}, orthogonal to $\boldsymbol{B_0}$

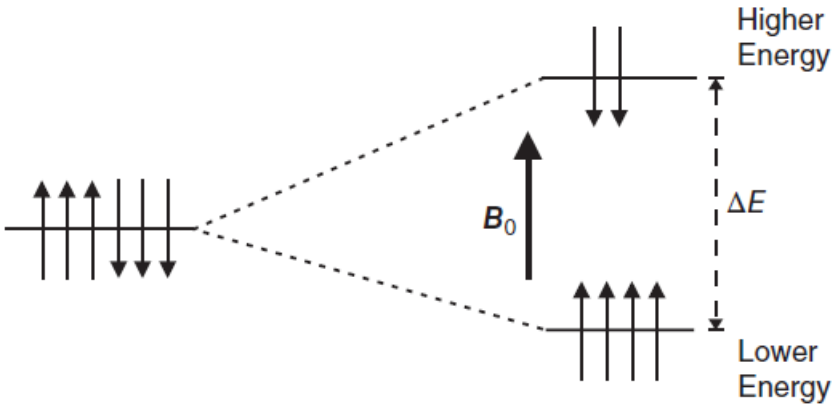

Figure 2.3: Zeeman diagram. In the absence of a magnetic field (left side), a collection of spins will have no preferential alignment between the spin-up and spin-down orientations. In the presence of a magnetic field (right side), the spin-up orientation (parallel to $\boldsymbol{B_0}$) is of lower energy and its configuration contains more spins than the higher-energy spin-down configuration. The difference in energy ΔE between the two levels is proportional to the magnitude B_0 of the static magnetic field.

2.1.3 Excitation phase

In an external static field at equilibrium, the transverse component M_{xy} disappears, since all the individual spins are indeed precessing, but they are all out of phase with each other. To generate a non-zero M_{xy}, the magnetization has to be tipped away from the z-direction. Therefore, the thermal equilibrium has to be disturbed by an additional external electromagnetic field. Applying a radiofrequency (RF) field $\boldsymbol{B_1} = B_1(cos(\omega_1), sin(\omega_1), 0)$ with the Larmor frequency of the nuclei ($\omega_1 = \omega_0$) and polarized orthogonally to $\boldsymbol{B_0}$, causes the net magnetization to flip away from the longitudinal z-direction toward the transverse xy-plane (Figure 2.4). In order to simplify

the following considerations, it is convenient to introduce a transformation from the stationary laboratory system (x, y, z) in a coordinate system rotating with the the Larmor frequency. In the rotating coordinate system, the static magnetic field $\boldsymbol{B_0}$ can be neglected, as it is already accounted for by the frame selection. An external field $\boldsymbol{B_1}$ rotating with the Larmor frequency in the laboratory frame corresponds to a static field in the rotating frame. Here, the trajectory of the magnetization vector can be interpreted as precession around with the Larmor frequency referring to the RF field $\omega_1 = \gamma B_1$. The irradiation with an RF field of duration t_1, rotates the magnetization vector $\boldsymbol{M}$ from the z-axis towards the xy-plane by the flip angle:

$$\alpha = \gamma B_1 t_1 \tag{2.8}$$

In the quantum mechanic picture, this means that the some nuclei are changing from the parallel state $N^+$to the anti-parallel state N^-. The duration of the RF pulse is very short, in the order of microseconds or milliseconds. Starting with a magnetization M_0 in the z-direction, the transverse and longitudinal magnetization can be calculated by:

$$M_{xy} = M_0 sin(\alpha) \tag{2.9}$$

$$M_z = M_0 cos(\alpha) \tag{2.10}$$

The transverse component rotating in the xy-plane can be detected by a RF coil and creates the actual NMR signal.

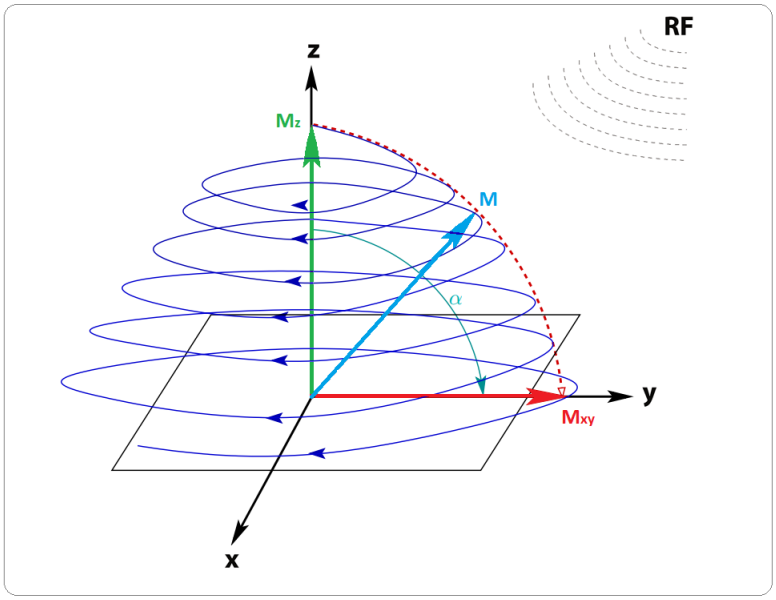

Figure 2.4: Macroscopic magnetization and excitation phase in NMR. The macroscopic magnetization $\boldsymbol{M}$ (blue) is composed of two components: The longitudinal component M_z (green) and the transverse component M_{xy} (red). The The energy of an RF pulse flips the macroscopic magnetization vector $\boldsymbol{M}$ by an angle α.

2.1.4 Relaxation phase

While the considered situation so far describes an ideal system, in reality, due to molecular and chemical exchange processes, nuclei and their respective magnetic moment are in constant thermal motion. The statistical nature of this motion causes an additional, locally fluctuating magnetic field which in turn induces changes in the macroscopic magnetization known as relaxation. Consequently, the magnetization vector in the xy-plane decreases and longitudinal magnetization along the z-direction returns to its initial value. The intensity time patterns of these vectors follow exponential laws, with the characteristic times depending on the nature of the examined tissues and are essential in the production of diagnostic images.

2.1.4.1 Spin-lattice relaxation

The spin-lattice relaxation denotes the effect of the longitudinal magnetization recovery M_z after the application of an RF pulse, bringing the system back to the initial thermodynamic equilibrium condition. The physical reason for the spin–lattice relaxation is energy exchange between the spin and neighboring atoms and molecules (the lattice), or in other words, the spin can transfer energy to other (mechanical) degrees of freedom of the system. This happens by dipole–dipole interactions, chemical shift anisotropies, or electric quadrupole moments. Thus, the spins experience fluctuations in the local field, which result in a non-zero probability of transitions between the energy states and therefore in relaxation of the ensemble. The effect on the macroscopic magnetic moment $\boldsymbol{M}$ is the progressive reduction of the flip angle α towards zero and the recovery to the initial value of M_z before the excitation. This process is modeled by an exponential function characterized by the time constant $T1$, the period for the longitudinal magnetization to recover 63% of its equilibrium value (Figure 2.5). For a 90-degree excitation pulse, the longitudinal magnetization is given by:

$$M_z(t) = M_0(1 - \exp(-\frac{t}{T1})) \tag{2.11}$$

The numerical values of $T1$ of biological tissues depend strongly on the strength of the static magnetic field and the temperature. The value is also a function of the microviscosity of the medium and depends on the mass and the size of the molecules that constitute the tissue. In pure water rotational movements or translation of the molecules are very fast and thus there is only a small energy exchange with the excited hydrogen nuclei. Consequently, the longitudinal recovery time is relatively long. If the examined medium contains medium sized molecules such as lipids, the molecular movements are slower and allow a higher energy exchange, which results in shorter $T1$ relaxation times of the protons.

2.1.4.2 Spin-spin relaxation

The term spin-spin relaxation denotes the decay of the transversal magnetization M_{xy} over time. Spins do not only interact with the lattice but also with each other. Due to this dipole-dipole interactions, each spin rotates with a slightly different precession frequency causing an irreversible loss in phase coherence (dephasing) and as a consequence the loss of macroscopic magnetization in the transversal plane (Figure 2.6). The transversal magnetization decays exponentially and

is characterized by the parameter $T2$, which corresponds to a decrease of 63% of its initial value directly after the RF pulse.

$$M_{xy}(t) = M_0 \exp(-\frac{t}{T2}) \tag{2.12}$$

In biological tissues the numerical values of $T2$ are generally lower than those of the $T1$, e.g., the $T1$ value of blood at 1.5 T is approximately 1200 ms, whereas the corresponding $T2$ value is approximately 300 ms [23].

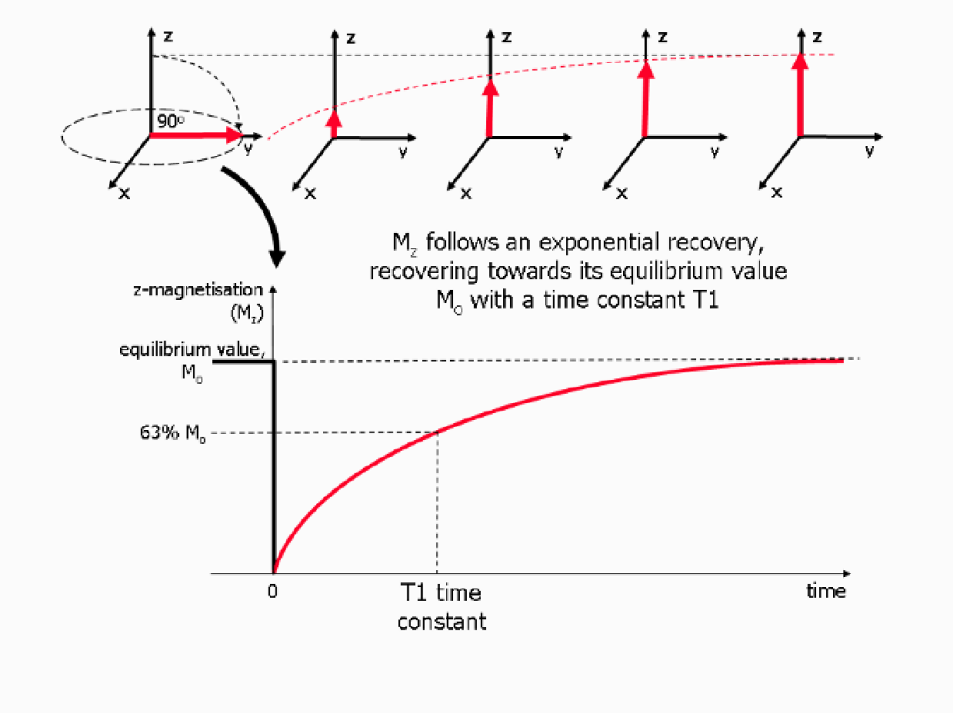

Figure 2.5: $T1$ relaxation process. Diagram showing the process of $T1$ relaxation after a 90° RF pulse that is applied in the thermal equilibrium state. Directly after the pulse the z-component of the macroscopic magnetization M_z is reduced to zero, but then recovers gradually back to its equilibrium value if no further RF pulses are applied. The recovery of M_z is an exponential process with a time constant $T1$. This is the time at which the magnetization has recovered to 63% of its equilibrium value (figure from[24]).

2.1.4.3 Free induction decay

In addition to the effect of the introduced fluctuating fields, in real systems imperfections in the homogeneity of $\boldsymbol{B_0}$ induce an even faster dephasing of the transversal magnetization. Similar to fluctuating fields this static field causes an additive phase and the so called $T2$'-relaxation process. In contrary to the $T2$-relaxation, the static nature of the interfering field offers the possibility to reverse the dephasing within a specially designed NMR experiment. The resultant decaying signal of the combined dephasing is known as the Free Induction Decay (FID), which is also described by an exponential function (Figure 2.6). The comprehensive effect of dephasing in the xy-plane is

characterized by the $T2^*$ parameter:

$$\frac{1}{T2^*} = \frac{1}{T2} + \frac{1}{T2'} \tag{2.13}$$

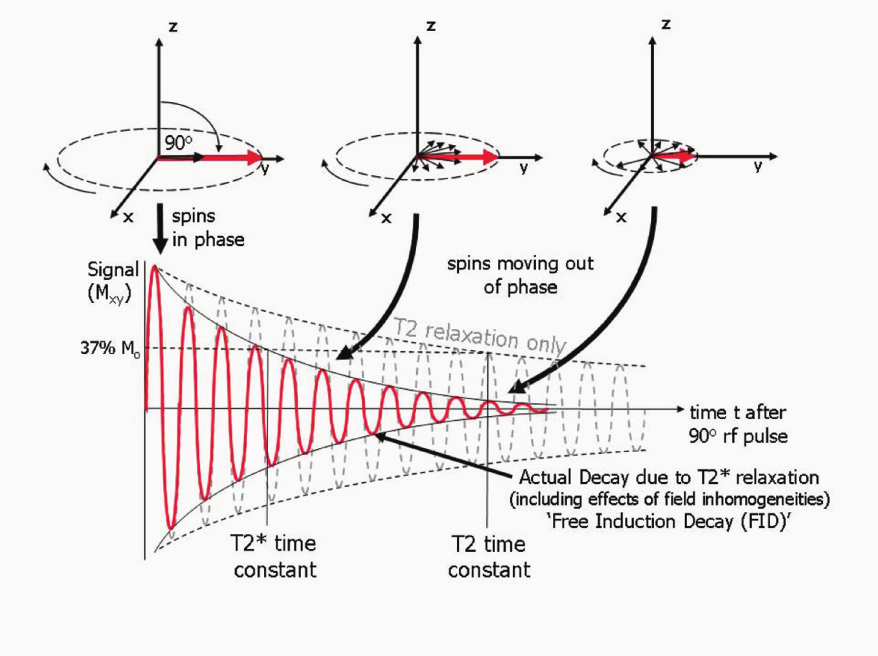

Figure 2.6: $T2$ and $T2^*$ relaxation processes. After an initial RF pulse the transverse magnetization (red arrow) has a maximum amplitude as the population of spins rotate in phase. The amplitude of the transverse magnetization (and therefore the detected signal) decays as the spins start dephasing (shown by the small black arrows). The resultant decaying signal is known as the Free Induction Decay (FID). The overall term for the observed loss of phase coherence (dephasing) is $T2^*$ relaxation, which combines the effect of $T2$ relaxation and additional dephasing caused by local variations (inhomogeneities) in the applied magnetic field. $T2$ relaxation is the result of spin-spin interactions and due to the random nature of molecular motion, this process is irreversible. $T2^*$ relaxation accounts for the more rapid decay of the FID signal, however the additional decay caused by field inhomogeneities can be reversed by the application of a 180° refocusing pulse. Both $T2$ and $T2^*$ are exponential processes with times constants $T2$ and $T2^*$ respectively. This is the time at which the magnetization has decayed to 37% of its initial value immediately after the 90° RF pulse (figure from[24]).

2.1.4.4 Bloch equations

A mathematical description of these processes was first proposed by Felix Bloch [19]. Bloch introduced a set of empirical equations for the temporal evolution of the magnetization vector:

$$\frac{dM_x}{dt} = \gamma(\boldsymbol{M} \times \boldsymbol{B})_x - \frac{M_x}{T2} \tag{2.14}$$

$$\frac{dM_y}{dt} = \gamma(\boldsymbol{M} \times \boldsymbol{B})_y - \frac{M_y}{T2} \tag{2.15}$$

$$\frac{dM_z}{dt} = \gamma(\boldsymbol{M} \times \boldsymbol{B})_z + \frac{M_0 - M_z}{T1} \tag{2.16}$$

In case of a $\theta = \pi/2$ experiment and a perfectly homogeneous field $\boldsymbol{B_0}$ the solutions of the equations above can be easily expressed in the laboratory system:

$$M_{xy}(t) = M_{xy}(0)\exp(i\omega_0 t)\exp(-\frac{t}{T2})$$

$$M_z(t) = M_0 - (M_0 - M_z(0))\exp(-\frac{t}{T1})) \tag{2.17}$$

The M_{xy} component is the relevant measurement variable in an NMR experiment. The oscillating magnetic field is simultaneously detected by an adjacent coil and represents the actual physical quantity detected in NMR and MRI.

2.1.5 Experimental sequence design

The modality and timing in which the RF pulses are applied is called a sequence and it is crucial to select the signal of interest from the investigated tissue. The collection of existing NMR and MRI sequences is very extensive. Here, only the most basic sequences are introduced very briefly.

2.1.5.1 Spin echo sequence

A sequence of pulses to compensate for the reversible dephasing effect due to static inhomogeneities in $\boldsymbol{B_0}$ is called the spin echo sequence. (SE) [25]. After the initial 90° RF pulse the coherent spins in the transversal plane start to dephase due to the local difference in the magnetic field. After a certain time $TE/2$ two spins with a frequency difference of $\Delta\omega$ experience a phase difference of $\Delta\Phi = \Delta\omega\frac{TE}{2}$. The application of 180° pulse causes the phase difference to change sign to $\Delta\Phi = -\Delta\omega\frac{TE}{2}$. Since the local precession frequencies are unaffected by the 180° inversion pulse, after another time period of $TE/2$ the dephasing effect is compensated and all spins are rephased ($\Delta\Phi = 0$.) forming a so called spin echo (Figure 2.7). Hence, the characteristic time parameter TE for this experiment is termed the echo time. The sequence of these two pulses is repeated after a time TR, called the repetition time. The signal decay in this sequence is depending on the T1, T2 and proton density of the examined tissue, which will be discussed further in Subsection 2.2.4 .

2.1.5.2 Gradient echo sequence

The gradient echo sequence (SE) differs from the spin-echo in regard to the flip angle, which is usually below 90,° and the absence of a 180° inversion pulse. As there is no rephasing due to inversion, the spin signal has to be generated in a different manner. As already explained, local differences in the magnetic field cause spin ensembles to dephase. This effect can be employed in order to artificially induce an echo signal. By applying a linear magnetic field gradient a controlled dephasing of the spin system can be obtained. Vice versa, the application of an inverse gradient after a certain time period causes the spins to rephase and form an echo signal. The static magnetic field

inhomogeneities are not compensated for in a gradient echo experiment. Hence, the signal decay is depending on$T2^*$.

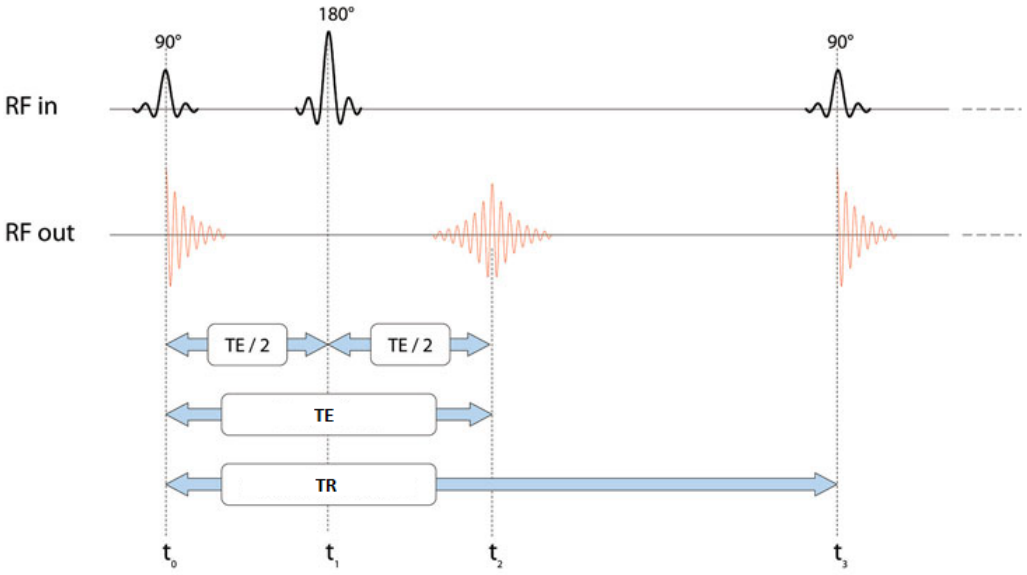

Figure 2.7: Spin echo sequence. The initial 90° RF pulse tips the longitudinal magnetization into the xy-plane. Due to inhomogeneities in the magnetic field coherent spins dephase over time. The application of a 180° pulse after the half echo time $TE/2$ inverts the spins' phases, hence the dephasing effect. After another$TE/2$ the spins are fully rephased forming the spin echo signal. The period between two excitation pulses of 90 ° establishes the TR.

2.1.5.3 Inversion recovery sequence

The Inversion Recovery sequence (IR) is used as subsequence of a variety of MRI sequences and has often the purpose to suppress tissues based on their $T1$ relaxation time, e.g. the short tau inversion recovery (STIR), which is a fat suppression technique [26]. The sequence of IR (Figure 2.8) can be describe as follows: The longitudinal magnetization is inverted by a 180° pulse (inversion pulse). During the TI, the protons belonging to different tissues relax with their characteristic TI time and the longitudinal magnetization M_z grows gradually from the minimum value reached after the 180° pulse to its initial value when only $\boldsymbol{B_0}$ was applied. After the inversion time TI a 90° pulse (readout pulse) is applied. Immediately after the 90° pulse, which rotates the partially relaxed longitudinal magnetization $M_z(TI)$ into the xy-plane, the FID signal is acquired and consequently the recovered longitudinal magnetization is measured.

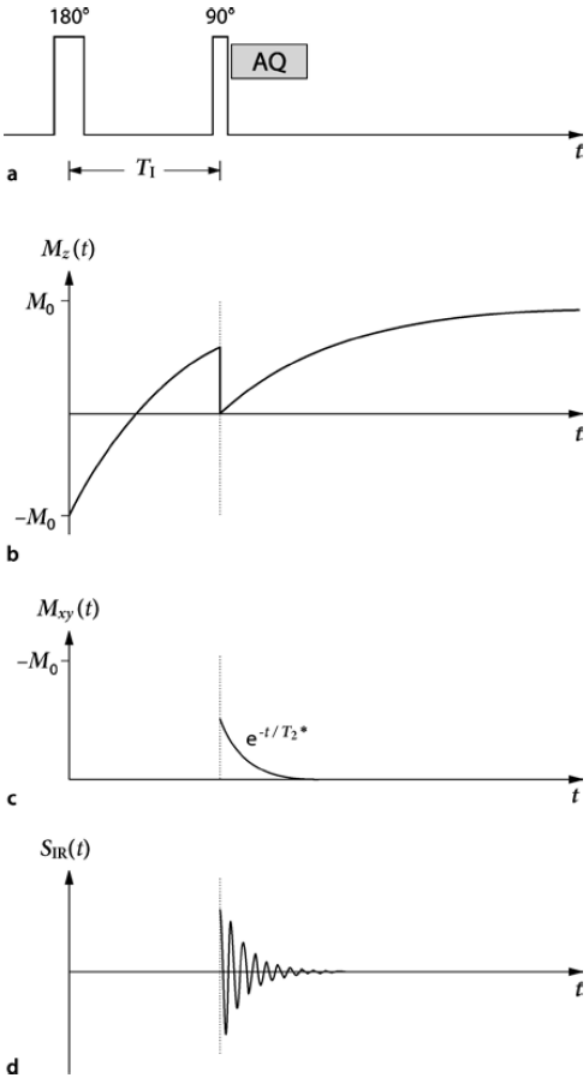

Figure 2.8: Inversion recovery sequence. a) Pulse scheme of the IR sequence (AQ: signal acquisition). b) Initially, the longitudinal magnetization is inverted by the 180° pulse, which is followed after an inversion time TI by a 90° pulse (readout pulse), which rotates the existing longitudinal magnetization $M_z(TI)$ into the xy-plane. After the 90° pulse, the longitudinal magnetization relaxes toward the equilibrium magnetization M_0 . c) Temporal evolution of the transverse magnetizationM_{xy} in the rotating frame. d) Induced MR signal (figure from [13]).

2.2 Magnetic resonance imaging

So far, the signal measured in the described experiments is obtained from the whole measurement object. In the early 1970's Lauterbur [1] and Mansfield [[2] developed a technique called Magnetic resonance imaging (MRI) that opened the possibility to imprint spatial information in the acquired NMR signal. For this achievement they won the Nobel prize in Physiology and Medicine in 2003. MRI employs the effect of NMR to measure the spatial distribution of hydrogen atoms ^1H and other nuclei of non-zero nuclear spin. The following section is intended to give a brief overview of the theoretical background MRI is based on.

2.2.1 Field gradients and spatial encoding

If the object to be imaged, such as the human body, is divided into small cuboidal volume elements (i.e., voxels), the task in MR imaging is to distinguish the signal contributions of the voxels to the detected summation signal from one another and to present them in form of sectional images (tomograms). This can be achieved by superimposing the homogeneous magnetic field $\boldsymbol{B_0}$ by an additional magnetic field with a well-defined dependence on the spatial position, so the Larmor frequency of the MR signal becomes a function of space. In practice, image reconstruction is achieved almost exclusively by means of magnetic gradient fields. These are three additional magnetic fields, whose field vectors point toward the z-direction and whose field strengths depend linearly on the spatial position x, y, z. When applying a gradient field $\boldsymbol{G}$ of the form:

$$\boldsymbol{G} = \frac{\partial B_z}{\partial x}\boldsymbol{e_x} + \frac{\partial Bz_z}{\partial y}\boldsymbol{e_y} + \frac{\partial Bz_z}{\partial z}\boldsymbol{e_z} = G_x\boldsymbol{e_x} + G_y\boldsymbol{e_y} + G_z\boldsymbol{e_z} \tag{2.18}$$

The total magnetic field and hence the Larmor frequency of the affected spins becomes spatially dependent:

$$\omega(\boldsymbol{x}, t) = \gamma B_z(\boldsymbol{x}, t) = \gamma(B_0 + \boldsymbol{x}\boldsymbol{G}(t)) \tag{2.19}$$

The signal dS originating from a spin ensemble in the volume dV located at the position $\vec{x}$ with the local spin density $\varrho(\vec{x})$ is then given by:

$$dS(t) = \varrho(\boldsymbol{x}) \exp(i\gamma(B_0 + \boldsymbol{x}\boldsymbol{G}(t))t) \tag{2.20}$$

Considering this inside the rotating coordinate system, which means physically demodulating the signal at the receiver with the Larmor frequency $\omega_0 = \gamma B_0$, the first part of the exponent vanishes. Volumetric integration of Eq. 2.20 yields

$$S(t) = \int\int\int \varrho(\boldsymbol{x}) \exp(i\gamma(\boldsymbol{x}\boldsymbol{G}(t))t) d\boldsymbol{x} \tag{2.21}$$

This illustrates that the applied gradient dephases the signal. Transverse relaxation was neglected in Eq. 2.20 , so Eq. 2.21 holds true for strong gradients, where the phase spread due to the gradients is much larger than the dephasing due to transverse relaxation. With the definition of the reciprocal, the so-called k-space vector:

$$\boldsymbol{k}(t) = \int d\boldsymbol{k} = \int \frac{\gamma \boldsymbol{G}(t)}{2\pi} dt \tag{2.22}$$

the spin density $\varrho(\boldsymbol{x})$ and the signal in k-space $S(\boldsymbol{k})$ become a Fourier pair:

$$S(\boldsymbol{k}) = \int\int\int \varrho(\boldsymbol{x}) \exp(i2\pi \boldsymbol{k}\boldsymbol{x}) d\boldsymbol{x} \tag{2.23}$$

$$\varrho(\boldsymbol{x}) = \int\int\int S(\boldsymbol{k}) \exp(-i2\pi \boldsymbol{k}\boldsymbol{x}) d\boldsymbol{k} \tag{2.24}$$

By employing three orthogonal linear gradients separately it is possible to manipulate every single component of $\boldsymbol{k}$ and imprint spatial encoding in all three dimensions. While the acquired signal and its Fourier transformation are continuous, sampling in reality is a discrete operation. That is, the signal detected by the detection coil and consequently the k-space are sampled after time intervals Δt or k-space intervals $\triangle k = \gamma G/2\pi \triangle t$, over a non-infinite interval. Equivalent to the continuous case, in reality a discrete version of the Fourier transform is used to reconstruct the real space image from k-space data. The discrete Fourier transform has a periodical spectrum, repeating with the sampling frequency. Consequently, the reconstruction shows identical copies of the magnetization distribution and if the sampled signal contains components with a frequency higher than half the maximal sampling frequency, these repeating copies overlap, which is termed aliasing. Higher frequencies components in the signal can not be resolved and are interpreted as lower frequencies, hence reconstructed as a so-called folding artifact. To avoid this effect the Nyquist sampling theorem [27] has to be satisfied when setting up an NMR experiment.

2.2.2 Slice selection

In most MRI experiments, k-space encoding is not performed in all three spatial directions (which is called a 3D sequence). As shown in Figure 2.9, the magnetization can be excited in a distinct slice of the probe (slice thickness Δz) by a slice-selective pulse. Therefore, an RF pulse with the bandwidth $\Delta\omega$ is combined with the gradient G_z. With slice-selective excitation, 2D gradient encoding can be performed to acquire the image inside the selected slice. Consequently Eq. 2.23 simplifies to

$$S(k_x, k_y) = \int\int \varrho(x, y) \exp(i2\pi(k_x x + k_y y) dx dy \tag{2.25}$$

Since the signal at one point in the k-space, (k_x,k_y), is the Fourier transform of the spin density $\varrho(x, y)$ the reconstruction of the image from the data $[S(k_x,k_y)]$ is an inverse Fourier transform. For Cartesian k-space coordinates, this is usually computed using a fast Fourier transform (FFT) [28]. For non-Cartesian encoding patterns, such as spirals, either the encoding matrix in has to be calculated and inverted numerically, or the k-space are interpolated onto a Cartesian grid with a subsequent Fourier transformation of the data [29].

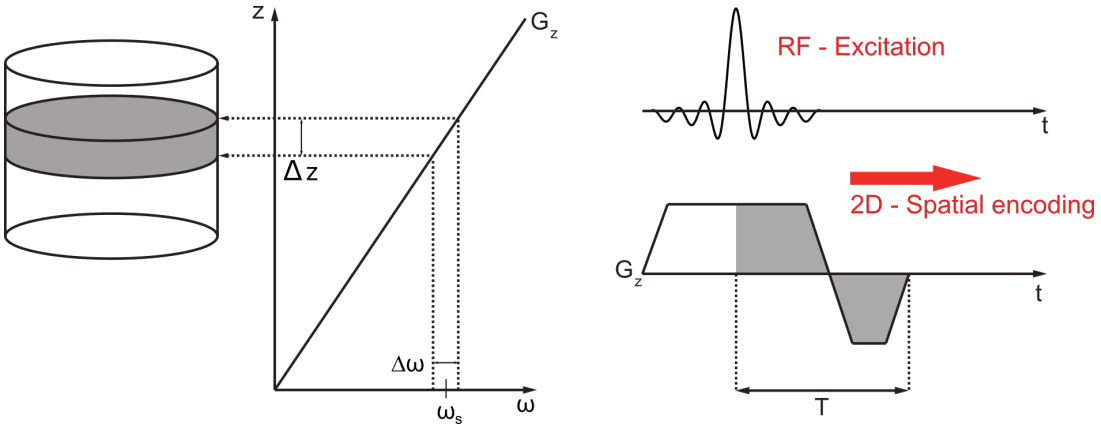

Figure 2.9: Slice selection in MRI. A gradient in z-direction, applied during the excitation pulse irradiated with a central frequency of ω_s and bandwidth $\Delta\omega$ results in a selective excitation of spins contained inΔz (right side). Schematic visualization of the sequence design (left side). The gradient G_z is switched on at the start of the excitation pulse and switched off when the excitation is complete. In order to compensate for the dephasing due to the first gradient a second gradient of opposite sign and half duration is applied.

2.2.3 Basic sequence design for spatial encoding in MRI

The spatial encoding of the xy-plane in many MRI experiments is achieved by phase and readout (frequency) encoding .The concept of phase encoding can easily be realized by applying a magnetic gradient field in, for example, x-direction for a fixed time before the FID signal is detected. Under the effect of this phase-encoding gradient, the phase of the magnetization is changed depending on the gradient. After switching off the phase encoding gradient, the magnetization components of the voxels in the slice precess again with the original, position-independent Larmor frequency around the direction of the $\boldsymbol{B_0}$ — now, however, with position-dependent phase angles. For readout encoding a gradient field in y-direction is applied during the readout, the spatial information is therefore encoded via the frequency. Figure 2.10 demonstrates a schematic sequence design for a basic spin echo sequence as discussed before. It depicts the series of RF pulses, gradients and the acquired signal. As mentioned before, the encoding directions are commonly referred to as slice, phase and readout direction. After the slice selective 90° excitation both phase (green) and frequency (blue) gradients encode the measurement volume in two dimensions, setting the k-value to maximum magnitude (cf. gradients I) for the first measurement (cf. gradients II). Subsequently, the generated echo is sampled by scanning through frequencies $k_{min} : \triangle k : k_{max}$.

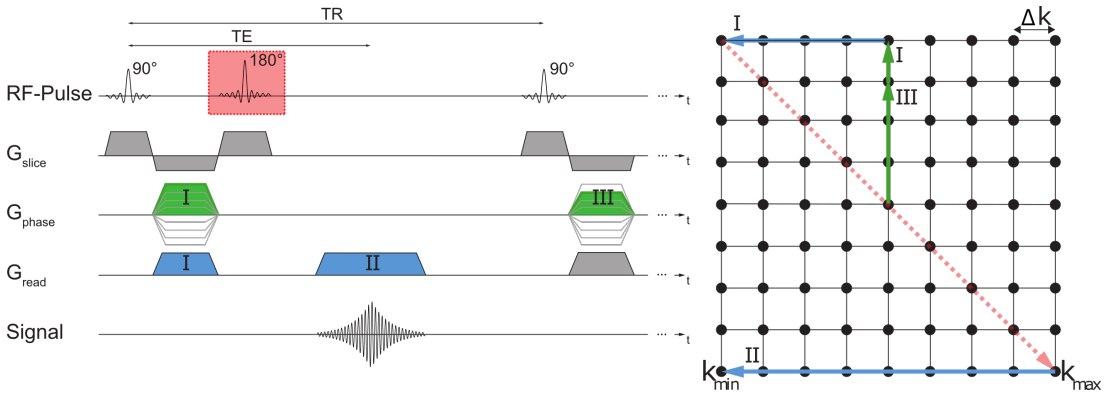

Figure 2.10: Spatial encoding for the spin echo sequence. Right side: After the slice selective 90° RF pulse the spin ensemble is phase encoded and selectively inverted by the 180° pulse. After the echo time TE, the signal is acquired while the frequency is encoded. If necessary the measurement is repeated after the repetitions time TR. Left side: Influence of the applied gradients on the k-space position. The first gradients (I) set phase and frequency to a minimum, while the second gradient (II) shifts through all frequencies during the data acquisition.

2.2.4 Image contrast

One of the strongest benefits of MRI is its flexibility in image contrast. The information possibly contained in the image comprises blood volume, the diffusion coefficient, blood flow, fat content and several other quantities. Primary parameters influencing the contrast of the reconstructed image are already apparent in the basic signal expression of for instance the spin-echo sequence:

$$S(t) = M_0(1 - \exp(-\frac{TR}{T1}))\exp(-\frac{TE}{T2}) \tag{2.26}$$

While the repetition time TR and the echo time TE are parameters set by the MRI user, the spin echo signal depends upon three physical quantities of the tissue: The relaxation times $T1$ and $T2$ and the proton density contained in the equilibrium magnetization M_0. Although most of the tissues in the human body have a similar water or proton density, the signal strength is greatly influenced by the $T1$ and $T2$ relaxation times. This, in turn, influences the intensity of the different tissues in the displayed image. It is important to mention here that diseases can also considerably alter the signal strength of a tissue. The extend to which the image contrast is defined by a certain parameter can be influenced via the set measurement parameters (Figure 2.11). An image contrast primarily defined by for instance the $T1$ time is called $T1$-weighted. In the following, the principle parameter settings for the three main image weightings are presented.

1. Proton density weighted: Long TR ($TR > 5\ T1$) and TE as short as possible: The longitudinal magnetization is completely relaxed after the chosen TR for all contained tissue types. Additionally, due to the short TE almost no transversal relaxation occurs causing the acquired signal to almost exclusively depend upon the proton density of the measurement volume.

2. $T2$-weighted: Long TR and long TE ($TE \approx T2$): Again the longitudinal magnetization is

completely relaxed. However, due to the long echo time, $T2$ relaxation appears, which makes areas of different $T2$ times separable. Hence, the image is $T2$-weighted.

3. $T1$-weighted: Short TR ($TR \approx T1$) and TE as short as possible: Due to differences in $T1$ the signal recovers to a different degree in different areas during the set repetition time. Together with a short echo time, prevent a $T2$ bias of the image contrast, the resulting measurement is primarily $T1$-weighted.

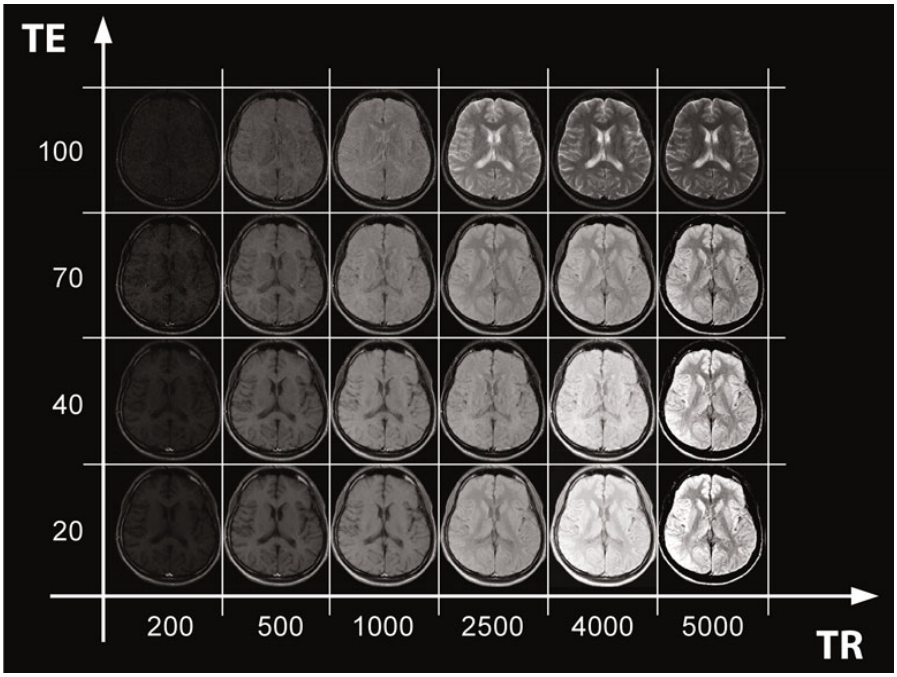

Figure 2.11: MRI contrast modalities. The Development of the contrasts is related to the change of the extrinsic parameters TR and TE. Low values of TR and TE generate an$T1$-weighted image; high values of TR and low TE , lead to an image depending largely on the proton density of the tissue; high values of both parameters lead to an $T2$ -weighted image.

2.2.5 Basic MRI setup

The MRI scanner hardware consists of three main parts. A large magnet that generates the external magnetic field $\boldsymbol{B_0}$. In clinical systems, this is a superconducting coil cooled with liquid helium. Most clinical scanners operate at a field strength of 1.5 T or 3 T; however, human scanners operating at 5 T or 7 T are available. The second part is the gradient system, which contains several coils inside the magnet. They create the time-dependent alternating magnetic fields necessary to spatially encode the signal inside the patient's body. The third system is the transmission and reception system, in which the RF coil is the first and most important part: It excites and receives the signals for MR experiments via electromagnetic induction. The RF coils are typically very close to the patient. In addition to the coil itself, there are several components involved in the receive chain, such as amplifiers, digitizers, and a computer that finally handles the data and performs the

reconstruction to visualize the results. Finally, the patient is lying on a movable bed, in order to minimize the body movements during an MRI scan. A schematic drawing of the structure of an MRI scanner is shown in Figure 2.12.

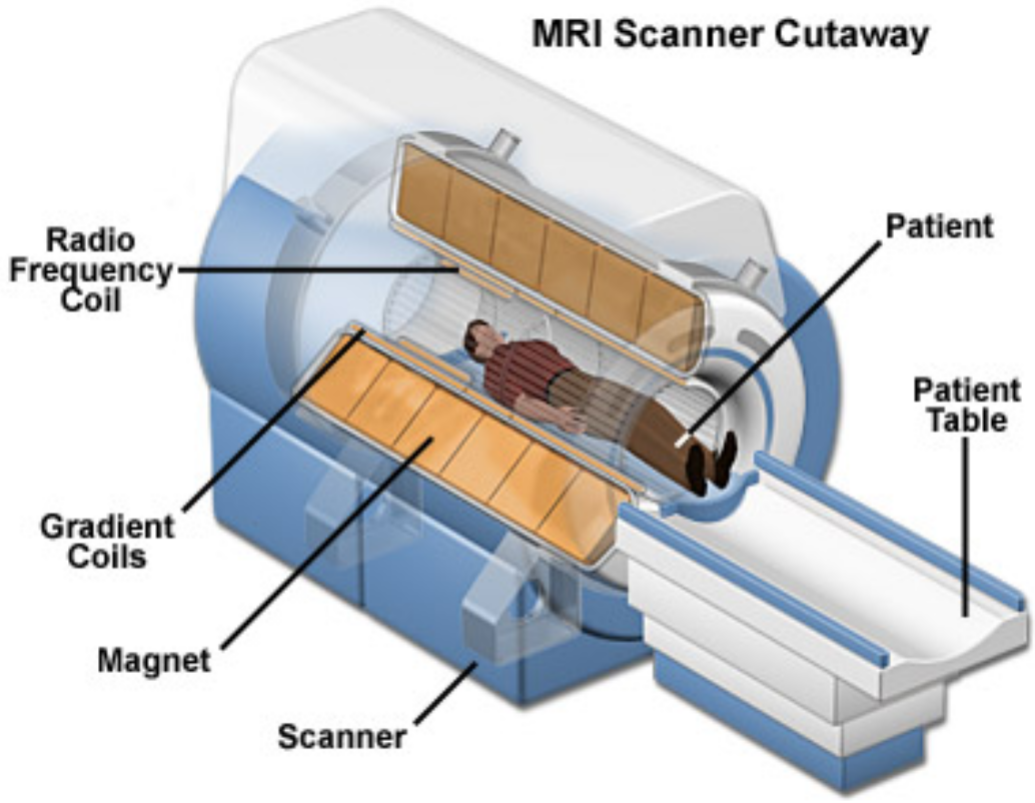

Figure 2.12: Schematic drawing of a whole-body MRI scanner. The most important components are the magnet for the static magnet field, the gradient coils and the radiofrequency coils. (figure from[30]).

2.3 Dynamic contrast-enhanced magnetic resonance imaging

The underlying principle of DCE-MRI is to measure the temporal evolution of the increase of magnetic resonance signal induced in tissue masses, following the injection of an paramagnetic contrast agent (CA). The most commonly used group of contrast agents are gadolinium chelates. The presence of a paramagnetic center affects the timing of the realignment of the magnetization vector of water protons, which represents the signal measured in MRI. More specifically, the introduction of a CA into the tissue leads to a decrease of the *T1* longitudinal relaxation time, and the transverse relaxation time *T2*, increasing the signal acquired with a *T1*-weighted sequence. The effect of the contrast agent on the local magnetic field, can be used to obtain indirect measurements on blood flow, the permeability of the vessels, the vascular volume and the size of the extravascular-extracellular space (EES).

Many studies using DCE-MRI have demonstrated that malignant tumors generally show faster and higher levels of enhancement than is seen in normal tissue [31, 17]. This enhancement characteristic reflects the features of the tumor microvasculature, which usually tends to demonstrate increased proportional vascularity and higher endothelial permeability to the contrast molecules than do normal or less aggressive malignant tissues.

Before the injection of the contrast medium, *T1*-weighted scans are performed in basal conditions to evaluate the effect of contrast enhancement and identify an affected tissue by the tumor. If a large number of healthy cells are present in the tissue, the contrast medium passes into the vessels without leaking into its surroundings. On the contrary, if the tissue is affected by a large number of tumor cells, the contrast medium leaks into the EES. At first the contrast agent accumulates in the extravascular tissue before it diffuses back into the vasculature, therefore, it is drained much more slowly from the blood vessels as in healthy tissue. The analysis of perfusion dynamics of the CA is studied through the evolution of the intensity of the MRI signal as a function of time in a selected region of interest (ROI), where the signal is strongly enhanced by the contrast agent.

2.3.1 Enhancement curve

The perfusion dynamics of a contrast agent is usually described through an time-intensity curve, the so-called enhancement curve (EC) (see Figure 2.13) .

In the EC, the abscissa shows the acquisition time of the signal and the ordinate represents the value of average intensity of the signal. The origin of the abscissa is chosen to coincide with the moment in which the contrast agent begins to perfuse the area of interest.

A first distinction between healthy and pathological tissues can be carried out in a first analysis based on the shape of the enhancement curve. For example, under standard conditions for this kind of analysis, invasive tumor types tend to have a strong initial increase of the signal (typically within 2 minutes after injection of the contrast agent), followed by a plateau or by a wash-out effect, whereas benign lesions exhibit a slower initial signal increase without reaching an effective plateau [32, 33].

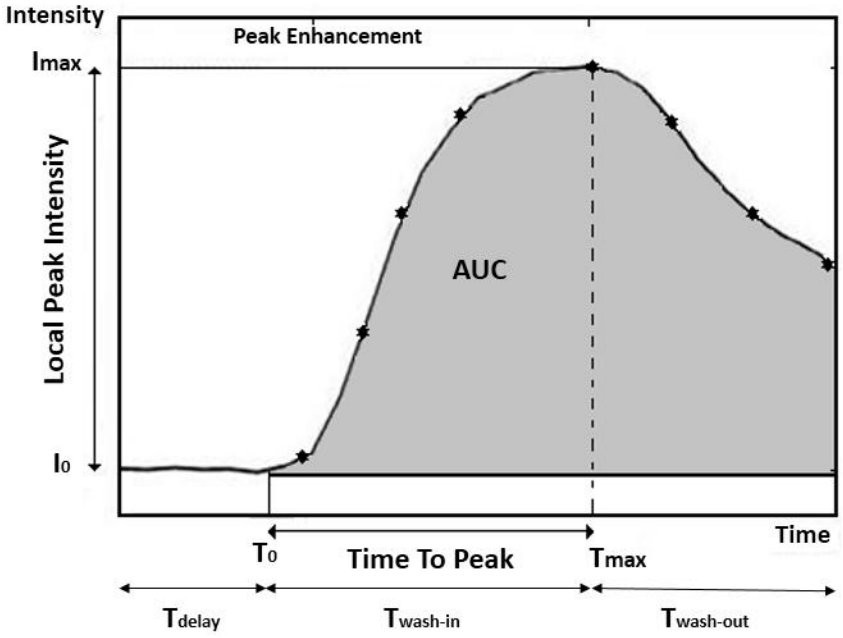

Figure 2.13: Enhancement curve and relevant parameters.

Considering different images $i(T_n)$ belonging to a temporal sequence of N images acquired at the sample times $T_n = T_1, T_2, ..T_N$, it is possible to define the EC as:

$$I(T_n, \bar{x}, \bar{y}) = [i(T_n, \bar{x}, \bar{y}) - i_0(\bar{x}, \bar{y})]\,, \tag{2.27}$$

where $i(T_n, \bar{x}, \bar{y})$ is the intensity of the pixel with the coordinate $(\bar{x}, \bar{y})$ at the sample time T_n and $i_0(\bar{x}, \bar{y})$ is the basal value (i.e., the mean value of the intensity registered before the perfusion of the CA and assumed as background noise).

For each image pixel(x, y) corresponding to a perfused tissue position, the time series of the corresponding values $i(T_n, x, y)$ of all images $i(T_n)$ of the temporal sequence yields the individual local enhancement curve. Hence, the local EC represents how the intensity of the backscattered signal changes as a function of the perfusion of the CA in correspondence to the selected pixel.

2.3.2 Characteristic parameters

As mentioned before, the CE is usually determined for a selected ROI. The average intensity in the ROI can be assessed for each sampling time, as a mean value of the increase in intensity in the specified region. When we define a two-dimensional ROI A that includes M pixels, the average intensity of the ROI is given as:

$$I_A(T_n) = \frac{1}{M} \sum_{(x,y)\in A} [i\,(T_n, x, y) - i_0(x, y)] \tag{2.28}$$

Through a curve-fitting approach, four local parameters (useful for diagnoses) can be calculated from the analysis of the EC of the ROI:

- Area under curve (AUC)
- Local peak intensity (LPT)
- Time to peak (TTP)
- Average Rising (SLOPE)

Before proceeding to the determination of these four parameters, we must determine the global arrival time T_0, which determines the earliest significant global intensity increase in the perfused area after the injection of the contrast agent. It is estimated from the time series of spatially averaged intensities.

2.3.2.1 Area under curve

The area under the curve (AUC) is a parameter, which finds different applications depending on the field in which the data is discussed. In general, this parameter represents the area under the EC starting from T_0. The baseline intensity of each pixel is subtracted prior to averaging the intensities of the pixel in the ROI. The AUC measures the quantity of CA absorbed by the ROI over the entire time series; hence it yields an estimation of blood flow that diffuses into the region:

$$AUC(A) = \sum_{n=1}^{N} \left\{ (T_n - T_{n-1}) \frac{1}{M} \sum_{(x,y)\in A}^{M} \left[i\left(T_n, x, y\right) - i_0(x,y) \right] \right\} \tag{2.29}$$

2.3.2.2 Local peak intensity

The local peak intensity (LPI) parameter is the maximum increment of the intensity with respect to the baseline intensity i_0. This quantity gives an estimation of the maximal quantity of blood that, at a given instant, comes into the region:

$$LPI(A) = \max_{n} \left(\frac{1}{M} \sum_{(x,y)\in A}^{M} \left[i\left(T_n, x, y\right) - i_0\left(x, y\right) \right] \right) \tag{2.30}$$

where the max operator is evaluated with respect to all the frames of the sequence.

2.3.2.3 Time to peak

The time to peak (TTP) parameter is calculated as the duration from the global arrival time T_0 to the time T_{max}, when the maximum contrast agent concentration in the region is reached. Thus, it gives information about whether this region is directly supplied with blood or through revascularisation.

$$TTP(A) = T_{max} - T_0 \quad \text{with} \quad \frac{1}{M} \sum_{(x,y)\in A}^{M} \left[i\left(T_{max}, x, y\right) - i_0\left(x, y\right) \right] \equiv LPI(A) \tag{2.31}$$

2.3.2.4 Average rising

The Average rise (SLOPE) is a measurement of the slope of the curve between T_0 and T_{max}; it is related to the average blood flow perfusion in the ROI and joins temporal and intensity aspects into a single parameter:

$$SLOPE(A) = \frac{LPI(A)}{TTP(A)} \tag{2.32}$$

The four introduced parameters should be evaluated with reference to the whole ROI (and in this case they are estimated averaging the corresponding quantities in the entire ROI) or calculated separately for each single pixel. In the latter case, a useful representation is given by the color maps [34], as represented in Figure... of Chapter....

2.3.3 DCE-MRI artifacts and possible corrections

Motion artifacts remain problematic in magnetic resonance imaging (MRI) and can hamper the precise calculation of the EC in DCE-MRI and thereby a correct diagnosis. Factors that may invalidate the goodness of EC are connected to various sources of patient motion that exist on the relevant time scales. For instance, the heart contracts and deforms as it pumps blood and the diaphragm moves and displaces surrounding tissue and organs during respiration an. Furthermore, peristalsis of organs can also play a role. Abdominal DCE-MRI is particularly affected by respiratory motion in the upper abdominal cavity (kidneys, liver, pancreas and spleen) and since the components of the abdomen consist of non-rigid soft tissues this can also lead to deformation of abdominal organs and structures. Furthermore, patients may also become uncomfortable in the scanner and shift to reposition themselves. Conversely, they may become more comfortable as they become accustomed to the procedure leading to involuntary muscle relaxation (e.g. pectoral muscles). Consequently, motion is present within many of the contexts in which DCE-MRI is utilized.

Standard DCE algorithms implicitly assume that during all the acquisition sequences the ROI does not modify its shape and position. From Eq. 2.27, it is evident that to correctly evaluate DCE, it is necessary to guarantee that there is no movement all along the temporal sequence , i.e., that the spatial location $(\overline{x}, \overline{y})$ remains the same portion in all the images to be compared in the sequence. Unfortunately, due to relatively long time of DCE procedure (up to 10 min), it is impossible to prevent the movements of the patient, breathing, peristalsis, and so forth. This motion artifacts during acquisition can cause inter-frame misalignments, which have a strong impact on the analysis of DCE-MRI data, since apparent intensity changes will be related to a mixture of motion and contrast agent changes. Consequently, the resulting enhancement curves may not reflect the precise contrast-agent dynamics of the tissue under investigation, as the position of the pixel changes over time. To account for this issue, we have to rewrite Eq.2.27 in the following way:

$$I(T_n, \overline{x}, \overline{y}) = [i\,(T_n, x(T_n), y(T_n)) - i_0(\overline{x}, \overline{y})] \tag{2.33}$$

where $x(T_n)$ represents the abscissa and $y(T_n)$ the ordinate of the point $(\overline{x}, \overline{y})$ at the sample time T_n as illustrated in Figure 2.14.

In other terms, there is the need of compensating the displacement of the point $(\overline{x}, \overline{y})$ along the time to correctly calculate DCE. A simple way to reduce the effects of artifacts induced by the movements is to obtain EC with respect to a set of pixels; that is, evaluating an averaging EC in a ROI as demonstrated in Eq. 2.28. In this way, small movements are partially absorbed. Unfortunately, this is a very simple method that considerably reduces the discrimination, and renders it ineffective for early analysis of cancerous lesions. Motion correction in DCE-MRI is challenging because rapid intensity changes can compromise common (intensity based) registration algorithms. To overcome such a drawback, we illustrate a strategy to compensate for artifacts without affecting the discrimination capabilities of the DCE in Chapter 4. To achieve this goal, we exploit the capabilities of active contours to dynamically adopt the morphological changes as a tool for "tracking" the ROI in the different images and, hence, compensate their movements.

It is important to note that the undesired motions during acquisition do not only result in misalignments between the frames of a time sequences, but also cause displacements in the transverse direction relative to the image plane, also referred to as out-of-plane motion. This induces inconsistencies with regard to the shape of the organ profiles within a temporal sequence that can also compromise a correct data analysis. We also present a approach to solve this problem in 4.6.

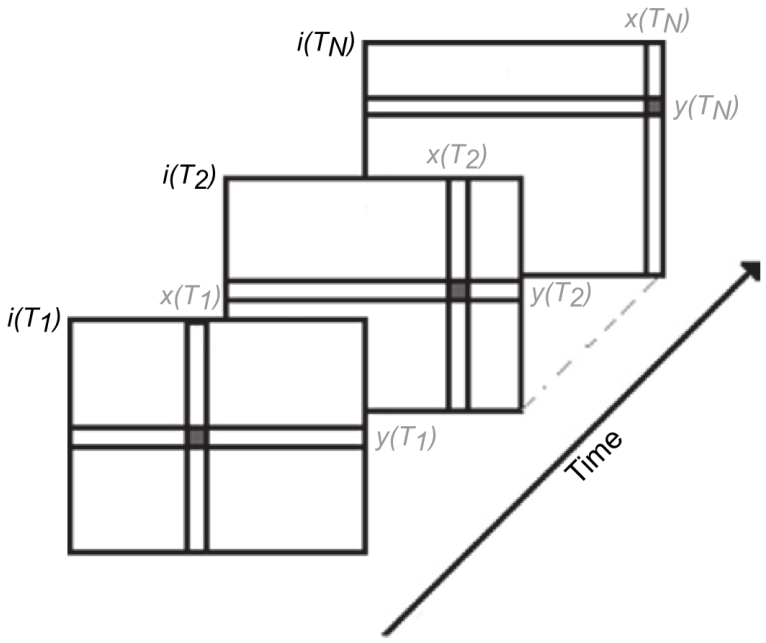

Figure 2.14: Temporal evaluation. The position of the pixel $(\overline{x}, \overline{y})$, which is depicted as the black square, changes on each frame of the sequence.

Chapter 3

Basic concepts of image processing in medical imaging

With the advances in medical imaging techniques and digitalization over the past decades came along a need for computational techniques to process the generated digital image data. Generally speaking, medical image processing methods are applied to enhance an image, to highlight some aspect of the information contained in it, to measure or classify objects within the image or to facilitate future image analysis, either by a user or a machine vision system. Being a highly multidisciplinary research field, medical image processing requires knowledge, theories, methods, and techniques from computer science, engineering, mathematics as well as from general and specialized healthcare domains.

In this chapter, we give an overview of the basic concepts of image processing in medical imaging with focus on the processing steps relevant for this work. Therefore, the chapter is divided into the following sections:

1. Image segmentation
2. Image registration
3. Motion tracking
4. 3D imaging operations

Please note that some of the presented methods cannot be classified exclusively into one of the listed sections, consequently, a partial overlap between the different sections is unavoidable.

For a more extensive description the interested reader is referred to further literature. Comprehensive introductions to image processing in medical imaging can be found in [35, 36]. A detailed description of image processing concepts with focus on the mathematical background is presented in [37]

3.1 Image segmentation

Image segmentation is often described as the process that subdivides an image into its constituent parts and extracts these parts of interest (objects). It is one of the most critical tasks in automatic image analysis because the segmentation results will affect all the subsequent processes of image analysis, such as object representation, image registration, motion tracking and many other possible analysis steps. Image segmentation can be described as the grouping of image pixels, with in some way resembling characteristics, into meaningful and usually connected structures, such as curves (edges) or regions, according to a given set of rules [38].

One way of formally defining image segmentation is given as follows: Assuming an image R with R_i $(i = 1, 2, ..., n)$ being disjoint non-empty regions of R, the segmentation of the image has to fulfill the following conditions [39]:

1. $\bigcup_{i=1}^{n} R_i = R$;
2. for all i and j, $i \neq j$, there exists $R_i \cap R_j = \varnothing$;
3. $P(R_i) = TRUE$, for all $i = 1, 2, ...n$;
4. for all $i \neq j$, there exists $P(R_i \cup R_j) = FALSE$;
5. for $i = 1, 2, ..., n$, the component R_i is connected.

where $\oslash$ is an empty set and $P(R_i)$ is a uniformity predicate for all elements in set R_i.

The first condition implies that the summation of all segmented regions may include all pixels in an image. This means that the segmentation should not terminate until every point is processed. The second condition states that different segmented regions can not be intersecting. The third condition implies that pixels in the same region should have similar properties, whereas the fourth condition states that pixels in different regions should have some different properties. The fifth condition points out that the pixels in a segmented region are contiguous.

From the aforementioned properties, it is clear that a first step for segmentation is the need to identify a suitable predicate of homogeneity for structuring the domain of interest according to certain characteristics. Subsequently, an algorithm can be chosen that is able to efficiently achieve the segmentation process for the defined properties. The result, however, is strictly dependent on both the nature of the applied algorithm and the properties of the image used for the partitioning process (homogeneous predicate). So, in general, there is not an universally good technique, but it is necessary to adapt the procedure to the respective purpose. The applications of segmentation in biomedical images are manifold. Here, the most important applications are listed:

- identifying region of interest (ROI)
- studying anatomical structure
- localizing pathologies
- tracking the progress of diseases
- measuring organ/tumor volumes

- treatment/surgical planning
- cell counting for biomedical studies

Numerous image segmentation techniques are available and there are also various ways of classifying segmentation methods in literature. An overview of segmentation methods classified according to [40, 36] is presented in Table 3.1. In the following subsection, the most common methods are briefly described:

Methods based on pixel levels		Thresholding Region-based methods Edge-based methods
Methods using pattern recognition and machine learning	Supervised classifier	k-nearest neighbor classifier Parzen window classifier Bayes classifier
	Unsupervised classifier	k-means alg. Fuzzy c-mean alg. Expectation-maximization alg.
Model and atlas-based segmentation		Parametric active contour models Geometric active contour models Active shape and appearance models Atlas-based methods
Multispectral segmentation		Gaussian models (Markov random) Variational approach Feature fusion
User interactions in interactive segmentation		Identifying region of interest Providing seeds with predefined labels Controlling topology Correcting segmentation

Table 3.1: Taxonomy of segmentation. According to [40, 36].

3.1.1 Threshold

Thresholding segmentation approaches are based on the pixel values of the image. Some of them are based on the image histogram; others are based on local properties or the local gradient. If the object for detection is clearly separated from the background, then the histogram of the image has at least two dominant distributions of intensities and thus two local maxima. Traditionally, one obvious way to extract the object from the background is to select an intensity threshold value T, which separates the two maxima, e.g. a local minimum between them. The segmentation is then accomplished by scanning the image pixel by pixel and defining each pixel, depending on its intensity value in comparison to the threshold, as object or background (see Figure 3.1). This approach is called global thresholding and creates a binary image $b(x, y)$ from an intensity image $I(x, y)$ according to the simple criterion

$$b(x,y) = \begin{cases} 1 & if\ I(x,y) > T \\ 0 & otherwise \end{cases} \tag{3.1}$$

where T is the threshold.

Instead of just choosing one intensity value, two thresholds can be selected defining a range of intensity values. This might be suitable for images showing a more complex intensity distribution. Global thresholding is a fast and simple computational process. It works well on images that contain objects with uniform intensity values on a contrasting background. However, major drawbacks are that this approach fails if there is a low contrast between the object and the background, the image is very noisy, or if the background intensity varies significantly across the image.

If the threshold depends on local properties of some image regions, for example, local average gray value, thresholding is called local. If the local thresholds are selected independently for each pixel (or groups of pixels), thresholding is called dynamic or adaptive [35].

One common approach for local thresholding is to split an image into subimages that are large enough to contain object and background pixels. Subsequently the thresholds are calculated for each subimage [41].

Another local thresholding method is based on the examination of the image intensities in the neighborhood of each pixel. The local thresholds are selected by using the mean value of the local intensity distribution or other statistics such as the average of the maximum and minimum values [42] or the local gradient magnitude [43].

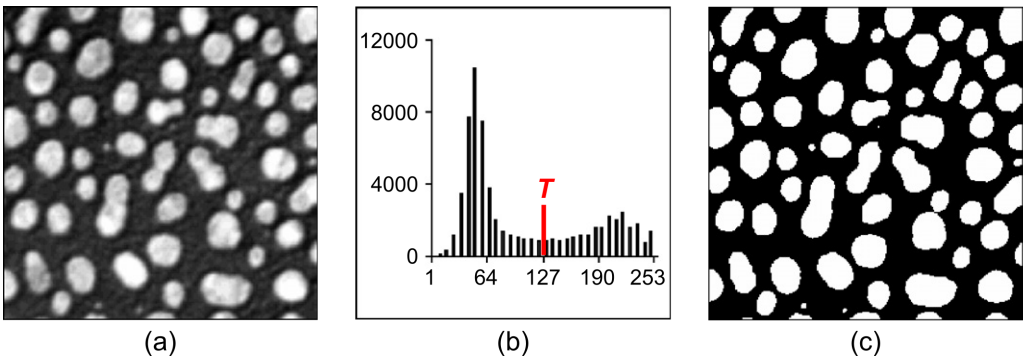

Figure 3.1: Global thresholding. (a) Original image, (b) histogram of image (a) with threshold set at $T = 127$, (c) resulting binary image (modified figure from [35]).

3.1.2 Region-based segmentation

The region approach tries to isolate areas of images that are homogeneous according to a given set of characteristics. Region-based segmentation algorithms can be divided into region growing algorithms, split and merge techniques and watershed algorithms.

3.1.2.1 Region growing methods

Region growing methods [44, 45], also sometimes referred to as growing-and-merging methods are procedures that group pixels or sub regions into larger regions based on a predefined criteria. The basic approach is start with a set of n pixels, the so-called seeds, distributed along the structures of interest. Each seed pixel i is treated as a region R_i with$i = 1, 2, ...n$. In the next step neighboring pixels are examined one at time and added to the growing region if they satisfy similarity criteria relating to their intensity, color or related statistical properties of their own neighborhood. A neighborhood can be 4 or 8 pixels around the center pixel in 2D; 6, 18 or 26 voxels in 3D. The procedure continues until no more pixels can be added. Seeds can be chosen by an operator or provided by an automatic seed finding procedure. The selection of a set of one or more starting points and similarity criteria is often based on the nature of the problem. The choice of homogeneity criteria is crucial for the success of this algorithms, if these criteria are not chosen appropriately, the regions might expand into adjoining areas or merge with regions that do not belong to the object of interest. Another drawback of region growing is that different starting points may not grow into identical regions. Benefits of region growing are their capability of correctly segmenting regions that have the same properties but are spatially separated, and that these algorithms generate connected regions.

3.1.2.2 Split and merge methods

Split and merge methods [46, 47] initially regard the entire image as one big region. In a first step, the region is split iteratively into subsidiary regions, while a homogeneity criterion is not satisfied. For instance, if the homogeneity criterion is not satisfied for the initial region, it is split into four subregions. The homogeneity criterion is then tested on each subregion, and those that do not satisfy the criterion are split into another four subregions. This process continues until all resultant regions satisfy the homogeneity criterion. In a second step, all adjacent regions are reassembled recursively so that resulting regions satisfy the homogeneity criterion, i.e, when a homogeneous region is created, its neighboring regions are checked and the new created region is merged with an existing one if they have identical properties. If the similarity criteria are met by more than one adjacent region, the new region is merged with the most similar one. Split and merge algorithms are computationally fast. Major drawbacks of this techniques are their insensitivity to the image semantics, block-like appearance of the resulting partition due to the region representation as groups of rectangles, and the high sensitivity to image translations.

3.1.2.3 Watershed algorithms

Watershed algorithms [48, 49] are another region-based image segmentation approach relying on image morphology. They require seeds in each object of the image, including the background. An image is considered as a 3D topographic surface (landscape), where the bright pixels represent mountaintops and the dark pixels valleys. A drop of water falling somewhere on a landscape flows downhill until it reaches a river or another body of water. A body of water and the surrounding region from which the water drains into it is called a catchment basin. Obviously, every local

minimum on the landscape will be associated with a catchment basin. Catchment basins are separated topographically from adjacent catchment basins by lines of maximum altitude called watershed lines or simply watersheds. For a drop of water that falls onto a watershed line, it is not defined to which of the adjacent catchment basins it belongs. Those points on the landscape that do not belong to any catchment basin therefore form part of the watershed. The resulting catchment basins then form the regions of an image partition, which are separated by pixels that belong to the watershed. The result of a watershed transform of a gray-scale image is usually represented by a label image in which every pixel that belongs to a given catchment basin is assigned the same label, and the pixels that belong to the watershed are assigned a unique, distinct label. Watershed algorithms are powerful tools for segmentation and that can not only be applied to the original gray-scale image but also to a gradient or binary representation of the image. The major drawback of watershed is that the presence of noise often leads to an oversegmentation of the image.

3.1.3 Edge-based segmentation

The aim of an edge based segmentation is to determine the boundaries of interesting objects by detecting image edges. In a gray level image the edge is a local feature that, with in a neighborhood, separates regions in each of which the gray level is more or less uniform with different values on the two sides of the edge. For a noisy image, it is difficult to detect edges as both edge and noise contain high spatial frequencies, therefore the images are usually filtered to reduce the noise prior to edge detection. Various approaches are available for edge detection, for instance, there methods based on error minimization, maximizing an object function, fuzzy logic, the wavelet approach, the Bayesian approach, morphological operations or genetic algorithms. The majority of these approaches are based in some way on measuring the first derivative (gradient) or the second derivative (Laplace operator) of an image. The Gradient $\boldsymbol{G}$ of an 2D image represented by the function,$f(x,y)$ is defined as a vector:

$$\boldsymbol{G}(f(x,y)) = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix},$$

which points in the direction of the maximum rate of increase of the function $f(x,y)$. The the magnitude of the gradient equals the maximum rate of increase of $f(x,y)$ per unit distance in the direction $\boldsymbol{G}$. It is given by:

$$|\boldsymbol{G}| = \sqrt{G_x^2 + G_y^2}$$

The direction of the gradient is defined as:

$$d = \tan^{-1}\left(\frac{G_y}{G_x}\right)$$

The Laplace operator ∇^2 of an image function $f(x,y)$ is defined as:

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2}$$

The gradient and the Laplace operator are analytical concepts applied to continuous functions. Since images are discrete functions, these concepts have to be approximated. Since derivatives are linear and shift-invariant, gradient calculation is commonly implemented using the concept of image convolution with a convolution kernel. Numerous kernels have been proposed for finding edges, three of the most common are the Roberts, Prewitt and Sobel edge-detection filter kernels. these kernel are discussed in detail in Chapter 4.

The most commonly used edge detection technique is the Canny edge detection algorithm. It was introduce by J. Canny in 1986 [50] with the goal to improve and optimize the already existing edge detection methods. Specifically, he aimed for a lower error rate for the edge detection, a precise localization of the edges and for a single response of the algorithm to each edge in the image. These criteria are implemented as multistage process. In a first step, the image noise is eliminated by a smoothing filter, followed by the calculation of the actual gradient image. Next, for each pixel the direction of the gradient is estimated, and it is classified as an edge point along this direction only if the gradient is greater than the two adjacent pixels; otherwise the value of the pixel is set to zero (nonmaxima suppression). Finally, a double threshold is applied to eliminate the remaining pixels that do not belong to an edge. Pixels that have a gradient less than the low threshold (probably due to noise) are removed and those that exceed the upper threshold are classified as an edge point. If the gradient magnitude is between the two thresholds, the pixel is set to zero unless there is a path from this pixel to a pixel with a gradient above the upper threshold.

3.1.4 Active contour models

Active contour models can be divided into parametric (explicit) and geometric (implicit) models.

In parametric active contour [51] models, also referred to as snake models curves and surfaces are represented explicitly by parameters, i.e. positions of the contour points. During deformation, the points of the contour are moved by applying either an energy minimizing formulation or a dynamic force formulation [52]. A drawback of this approach is the sensitivity to initialization, which may lead to convergence to undesired local minima. Nevertheless, this model has been applied successfully in a wide range of medical image applications and is also used in the developed algorithm of this work. Therefore, it is discussed more detailed in Chapter 4.

Geometric active contour models are represented implicitly as level sets of a scalar function of high-dimensional variables [53, 54]. Instead of evolving the curve with the traditional snakes, this geometric functional evolves in time with respect to the plane. Its major advantage is the ability to automatically handle topological changes and removing the issues of contour parametrization [55].

3.1.5 Fuzzy *c*-means clustering

The fuzzy c-means clustering algorithm [56, 57] belongs to the segmentation techniques based on unsupervised classifiers in which only unlabeled data are used. It is a combination of a conventional k-means clustering [58] and a fuzzy logic system [59]. In contrast to boolean logic, where the truth values of variables may only be the integer values 0 or 1, in fuzzy logic the truth values may be any real number between 0 and 1. It is used to simulate the experience of complex

human decisions and uncertain information. The k-means algorithm is a partitional clustering technique that assigns all patterns to a number of clusters k in each iteration of the algorithm based on minimizing the within-cluster scatter or maximizing the between-cluster scatter. For this approach, the initial clusters centers have to be given and each pattern is assigned to only one cluster, whereas in fuzzy clustering each pattern has a degree of membership to all clusters. Each observer is assigned a fuzzy membership value for a class, and an objective function is then developed based on the fuzzy membership value. The objective function will be minimized iteratively, until convergence is reached, by updating the new fuzzy membership value according to the observers and the number of iterations. The criterion determining the convergence of the objective function is generally defined as when the difference of the values of the objective function between two successive iterations is significantly small. Fuzzy clustering methods have been successfully applied to MRI data segmentation and many other medical segmentation problems. A good overview about these applications is given in [60].

3.1.6 Morphological operations

Mathematical morphology (MM) is a well established theory and methodology for the analysis and processing of geometrical structures, mainly based on the concepts of set theory, topology, stochastic geometry, lattice theory and nonlinear partial differential equations. MM is extensively used for a wide range of image processing tasks. In the context of segmentation, it is a valuable tool for correcting segmented structures especially for the resulting binary structures of thresholding. In this subsection, the basic concepts are presented. For a detailed introduction to MM, the reader is referred to literature, such as the textbooks [61, 62].

Morphological operations were originally applied to binary images, later, these concepts were also extended to gray-scale images. Here, we focus on the concepts for binary images.

We extend the concept of a binary image as given in Eq. 3.1 to the following:

$$b(x,y) = \begin{cases} 1 & if\ (x,y) \in X \\ 0 & otherwise \end{cases}, \tag{3.2}$$

where X is the set of the binary image containing the segmented object, while all $(x,y) \notin X$ are called the complement X^c and usually represent the image background . A morphological operation can generally be described as the interaction of X with a structuring element S . The structuring element is usually small relative to the image.

The basic morphological operations are the following shift-invariant operators.

- *Erosion:* Mathematically erosion can be formulated as $E(X) = \{x : S_x \subseteq X\}$. This means that when the structuring element is placed anywhere in the image and it is fully contained by the set (it is a subset), then the center of the structuring element is part of the eroded set.
- *Dilation:* The mathematical definition of dilation is given by: $D(X) = \{x : S_x \cap X \neq \varnothing\}$. This means that when the structuring element is placed anywhere in the image and it hits the set X, then the center of the structuring element is part of the dilated set.

- *Closing:* The closing operator is defined as: $C(X) = E(D(X))$. When closing is applied to a set; the dilation will expand object boundaries, which will be partly undone by the following erosion. Holes and tubule-like structures in the interior or at the boundaries of objects that are smaller/thinner than the structuring element will be filled up by the dilation, and not reconstructed by the erosion, since these structures no longer have a boundary for the erosion to act upon. Consequently, the closing operation removes small holes and thin cavities.

- *Opening:* The opening operator is given by: $O(X) = D(E(X))$. When opening is applied, the erosion will shrink object boundaries, and this will only be partly reversed by the following dilation. Therefore, the opening operation opens up holes that are near (with respect to the size of the structuring element) a boundary, and removes small objects and protuberances.

The described morphological operations are displayed exemplary in Figure 3.2.

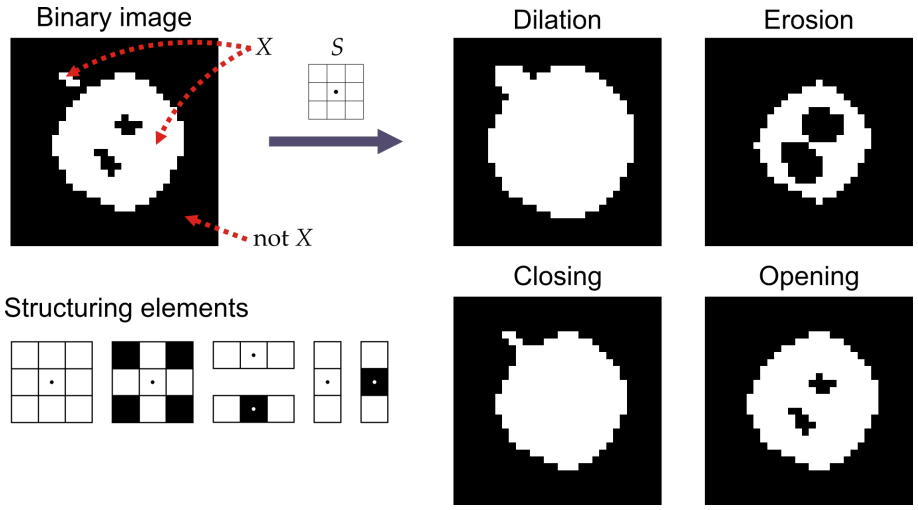

Figure 3.2: Morphological operations. In the upper left a binary image is shown, where the white pixels belong to the set X. In the lower left some examples of structuring elements are displayed. On the right side the basic morphological operations applied to to the binary image are shown (modified figure from [63]).

3.2 Image registration

However, such analysis tasks often require a comparison either between images in different modalities (e.g. MRI and CT) or between the images of single modality taken at different times (e.g. pre-operative vs post-operative) or from different viewpoints. Various factors such as tissue deformations, patient movements or variations in imaging setups make a direct comparison between medical images difficult and necessitate an image registration step. Since these factors are usually unavoidable in most cases, medical image registration has become a core part of almost every medical image analysis procedure. In most medical applications such as the prediction and monitoring of tumor growth, verification of treatments, fusion of images of different modalities, one needs to align images in order to obtain anatomical correspondences between medical images.

Image registration aims at the spatial alignment of two or more images in general. For the sake of simplicity, here, we will consider the co-registration of two images. Let M and F be two images defined in the image domain $\Omega \in \mathbb{R}^N$, which defines a region in space covered by an image with dimensionality N. The image M is called source (or moving) image, which is being registered and F is called the target (or fixed) image. these images are related by a transformation $\boldsymbol{T}$ that is parametrized by θ and operated on M. The goal is then to find the optimal transformation that optimizes an energy function

$$S(F, \boldsymbol{T}_\theta \circ M) + \beta R(\boldsymbol{T}_\theta)$$

where $S(.,.)$ is a similarity measure that quantifies the quality of the alignment. R is used to regularize the transformation to favor any specific property in the solution or to tackle the difficulty associated with the ill-posedness of the problem [64] and β is a coefficient to balancing the two terms. The transformation $\boldsymbol{T}$ is a mapping function of the image domain Ω to itself, which maps point positions to other locations. The transformation $\boldsymbol{T}$ at every position $\boldsymbol{x}$ can be written as a vector field with the displacement or deformation function $\boldsymbol{u(x)}$:

$$\boldsymbol{T}_\theta(\boldsymbol{x}) = \boldsymbol{x} + \boldsymbol{u(x)}$$

Due to the diversity of images to be registered and due to various types of problems in the registration process, it impossible to design a universal method applicable to all registration tasks. Every method should take into account not only the assumed type of geometric deformation between the images but also radiometric deformations and noise corruption, required registration accuracy and application-dependent data characteristics. Nevertheless, the majority of registration methods is divided in the following steps [65]:

1. *Feature detection.* Firstly, salient and distinctive features are chosen and subsequently manually or automatically detected. For further processing, these features can be represented by their point representatives, the so-called control points (CPs).

2. *Feature matching.* In this step, the correspondence between the features detected in the target image and those detected in the source image is established. Various feature descriptors and similarity measures along with spatial relationships among the features are used for that purpose.

3. *Transformation model estimation.* The type and parameters of the transformation model are estimated. The parameters of the transformation model are computed by means of the established feature correspondence.

4. *Image resampling and transformation.* The source image is transformed with the estimated transformation model. Image values in non-integer coordinates are computed by the appropriate interpolation technique.

A variety of image registration techniques have been developed in the last decades for a wide range of applications. They can be classified in terms of many different criteria such as:

- imaging modality
- the dimensionality of the images M and F
- the type of features used for the registration procedure
- the transformation models
- user interaction
- the applied optimization procedure
- the registered objects (for medical imaging: the part of the anatomy).

Here, a brief overview of the basic differences between the various methods is presented. For further information the reader is referred to the [66, 67, 68].

3.2.1 Dimensionality of registration tasks

The dimensionality of the images M and F can be two or three and consequently registration can be performed from 2D to 2D, from 3D to 3D, from 2D to 3D and from 3D to 2D. Additionally, if the images consist of a time series, a fourth dimension may be added. the The simplest way is 2D/2D registration, which means that planar images or objects are aligned. It may be applied to locate, align and compare separate slices of tomographic data [69] or intrinsically 2D images like X-ray projection images or microscopy data. Compared to the 3D/3D case, 2D/2D is less complex by an order of magnitude concerning the number of parameters used and the number of data points involved, thus being usually less computationally extensive and faster. 3D/3D registration establishes correspondences of points in two volumes and is normally applies to the registration of different tomographic data sets. This is commonly used to create volume data with complementary information from different 3D imaging techniques such as PET-MRI [70] or PET/CT [71]. The majority of aforementioned applications allow for off-line registration, therefore, speed issues need only be addressed as constrained by clinical routine. 3D/2D and 2D/3D registration is performed to directly align spatial data to projection data, or to the reverse case. It is widely applied in computer-assisted image-guided intervention [72], where M is the 3D anatomy model (e.g., a pre-operative CT data set) and F is the personalized or intra-intervention image of the respective anatomy, such as an intra-operative X-ray image [73]. Since this registration tasks have to be performed on-line, they are heavily time-constraint and have a strong focus on speed issues connected to the computation of the paradigm and the optimization.

3.2.2 Features for registration

The features used for registration can be divided into extrinsic (not contained in the original data about the anatomy of the patient) and intrinsic features (contained within the data sets).

- *Extrinsic features.* These features are foreign objects introduced in the image space. Commonly, this is achieved by inserting or attaching an artificial device to the patient in a well-defined way. These objects, so called markers or fiducials, are designed to be well visible and

accurately detectable in the respective imaging modalities. Consequently, the registration of acquired images is relatively simple, fast and can usually be automated. Since extrinsic methods by definition cannot include image information about the patient, they normally use rigid transformation models. There are different types of invasive markers, such as bone-implantable markers [74] or stereotactic frames [75]. The latter are screwed rigidly to the outer skull table and are widely used for neurosurgery, since they allow for a high accuracy of registration. Since the major drawback of this markers is the invasiveness of the method, non-invasive markers are preferred whenever their application is feasible. Non-invasive extrinsic markers, for instance, are skin surface fiducials [76], which are glued to the skin of the patient. However, these non-invasive fiducials are, by nature, more mobile than invasive markers, and they may lead to an increase in fiducial localization errors. While extrinsic methods can generally provide registration results with high accuracy in a fast way, there are also some major disadvantages of this methods. First of all, these methods have an prospective character, i.e., preparations have to be made in the pre-acquisition phase and often require invasive interventions. Furthermore, fiducials may be decoupled with the anatomical information of the data set, thereby introducing decoupling correspondence errors in the registered images.

- *Intrinsic features.* These features are based on the image formation as generated by the patient. Depending on the available data different kind of features can be extracted. A commonly used intrinsic registration method is landmark-based registration. Landmarks can be anatomical, i.e., salient and precisely locatable points of the morphology of the visible anatomy [77], or geometrical, such as corners or local curvature extrema [78]. Landmark based registration is a versatile method that can theoretically be applied to any kind of image. They are mostly used to find rigid or affine transformations. While anatomical landmark can be relatively fast and provide good results, a typical drawback is that user interaction is usually required. Another widely applied registration technique is based on segmented features. For this approach, anatomical structures are extracted from both images by segmentation and are registered based on a rigid transformation models [79, 80] or deformable models [81, 82]. The third main category of intrinsic registration techniques is based on the voxel intensities and they are also referred to as area-based methods. In contrast to the aforementioned approaches it operates directly on the image gray values without prior data reduction by the user or segmentation. This can be done by using the full image content throughout the registration process [83] or by reducing the image gray value content to a representative set of scalars and orientations [84]. Area-based methods do not necessarily detect any features, but rather put their emphasis on the feature matching step, therefore there are sometimes called template matching. To find a correlation between the chosen areas or the entire images, different methods can be used, such as cross-correlation in the image space [83, 85], Fourier-domain based methods [86] and mutual information methods [87].

3.2.3 Transformation models

There are many different transformation models used for registration that are also often referred to as mapping functions. Sometimes, these mapping functions are divided into global and local

mapping functions [65]. While global models use all control points for estimating one set of the mapping function parameters valid for the entire image, the local mapping functions treat the image as a composition of patches and the function parameters depend on the location of their support in the image. This approach leads to the tessellation of the image, usually a triangulation, and to the defining of parameters of the mapping function for each patch separately. Transformation models can also by divided in the following main categories:

- *Rigid and affine transformation.* . For rigid objects, only translation and rotation operations are considered, thus the transformation $\boldsymbol{T}$ is not a function of the position $\boldsymbol{x}$. This type of transformations usually provide a global alignment of the data sets in a fast way as fewer parameters are involved. It is normally used for coarse registration [86]. A widely used and efficient method is the interactive closest point algorithm [88]. Affine mapping functions allow additionally for scaling and shearing. Consequently, they involve more parameters to be decided in the registration. In many situations, affine transformation would be sufficient for the alignment of objects [89].

- *Deformable Transformation.* Many biomedical image registration techniques utilize deformable (or non-rigid) transformations. In this case, the transformation $\boldsymbol{T}$ is a function of the position $\boldsymbol{x}$ and consequently the transformation is local. Deformable models can be can be categorized according to their theoretical basis into three groups [90]. The first group is based on physical models of materials including linear elasticity [91, ?], viscous fluid flow [92, 93] and optical flow [94, 95]. The second group originates from interpolation and approximation theory and includes methods based on radial basic functions [96], splines [97, 98] and wavelets [99, 100]. The third group are methods that constrain the transformation according to some desirable mathematical properties, such as inverse consistency [101], topological preservation [102] and diffeomorphic transformations [103, 104].

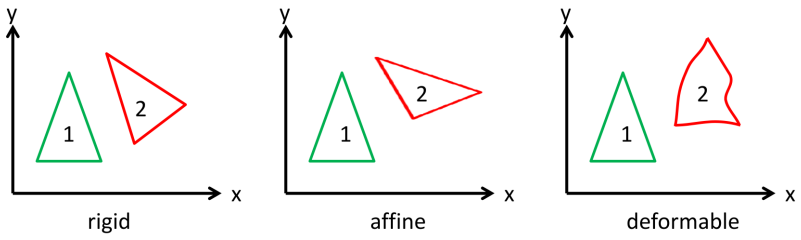

Figure 3.3: Transformation types. While rigid transformation does not change the scale of the objects, affine transformations may change the scale and the aspect ratio. On the other hand, deformable transformations are local and can change the shape of the object.

3.3 Motion tracking

Motion tracking (or motion estimation) in medical imaging aims at deriving the trajectory over time of moving organs and body parts, and in certain cases, the trajectory of a medical instrument. Motions of the anatomical objects (e.g., internal organs), surgical instruments and imaging devices

are very common and almost inevitable during most medical imaging procedures. Motion artifacts can be a significant source of error, especially for imaging modalities with relatively long acquisition times, e.g., in magnetic resonance imaging (MRI). These artifacts can cause serious problems in diagnostic decision, therapy planning, and image-guided intervention, such as misdiagnosis or mistakes in surgeries. Therefore, it is a crucial task in medical image analysis to develop precise, reliable, and fast tracking of deformable organs from medical image sequences and to use the gained information to compensate artifacts in the image reconstruction process.

Tracking should be distinguished from object detection and segmentation, as those techniques aim at estimating object position and orientation in continuous images. However, detection and tracking are not totally unrelated, since motion tracking basically combines the detection and correspondence matching steps, i.e., the object's positions and orientations in the next frame are predicted instead of detected, using information derived from the current (or previous) frames.

In its simplest form, the purpose of motion tracking is to estimate the path of an object over time by finding its position in each frame of the video sequence (e.g.; a series of MRI images acquired over a certain period of time). In other words, a tracker assigns consistent labels to the tracked objects in different frames of a video. Moreover, depending on the tracking domain, a tracker can further provide object-centric information, such as orientation, area, or shape of an object.

Tracking objects can be a complex process due to the following issues [36] :

- Noise in images
- Complex object motion
- Nonrigid or articulated nature of objects
- Loss of information caused by projection of the 3D object on a 2D image
- Partial and full object occlusions
- Scene illumination changes
- Real-time processing requirements.

There are a variety of approaches for object tracking that have been proposed for medical image processing and analysis. These approaches can be distinguished from each other primarily based on the way they address the following questions:

1. Which object representation is suitable for the tracking?
2. Which image features should be used?
3. How to model the shape and motion of the object?

There are a large number of object representation and tracking methods, which attempt to answer these questions for a variety of scenarios. The proposed solutions depend on the context in which the tracking is performed, the techniques used to implement them, and the objectives to be achieved.

In order to answer the questions above, the following overview about motion tracking is divided into three parts:

1. Object representation for tracking
2. Feature selection for tracking
3. Tracking techniques

3.3.1 Object representation for tracking

The first step of object tracking is to define a suitable representation of the object. Objects can be represented according to their intrinsic characteristics , such as shape or geometric configuration, and extrinsic, such as intensity, brightness or color. First the object shape representations commonly employed for tracking are introduced and then the joint shape and appearance representations are described.

- *Points.* The object is represented by a point which is the centroid or a set of points [105, 106]. This representation is suitable for tracking objects that have small regions in an image.
- *Primitive geometric shapes.* The object shape is represented by primitive geometric shapes such as ellipses or rectangles [107]. This method is used to approximate rigid bodies assuming that the distance between the different points representing the object remain unchanged during its motion. However, this approach is not very suitable for representing non-rigid objects that are deforming during motion.
- *Object silhouette and contour.* Contour representation defines the boundary of an object. The silhouette of an object is defined as the region inside the contour. Silhouette and contour representations are suitable for tracking complex non-rigid shapes whose pattern can change in transition from one frame to the next [108].
- *Articulated shape models.* Articulated objects are composed of body parts that are held together by joints [109]. The relationships between them are described by appropriate kinematic motion models. To represent an object of this type, you can model its individual parts using cylinders, ellipses.
- *Skeletal models.* Object skeletons can be extracted by applying medial axis transform to the object silhouettes. The skeletal models are commonly used as a shape representation for recognizing objects [110, 111], and such a representation can be used to model both articulated and rigid objects.

There are different ways to represent the appearance features of objects. It is important to mention that shape representations can also be combined with the appearance representations for tracking. Some common appearance representations in the context of tracking algorithms are given here:

- *Probability densities of object appearance.* In this method, probability function is used to estimate the probability densities of object appearance features, such as color or texture. These probability densities can either be parametric, e.g., a Gaussian [112] or a Gaussian mixture model [113] or non-parametric, e.g., Parzen windows [114] or histograms [107]. The probability densities of object appearance features can be computed from the image regions specified by the shape models, e.g., interior regions of an ellipse or a contour.

- *Templates.* Templates are obtained using simple geometric shapes or silhouettes and it carries spatial information about the intrinsic properties of an object within an image. The limitation of template based models is that they only encode appearance of the objects from one view. Thus, they are only suitable for tracking objects whose pose does not vary considerably.

- *Active appearance models.* These are models that take into account both the shape of the object and its extrinsic characteristics [115]. Through the use of suitable landmarks positioned (which can reside on the object boundary or inside the object region) it is possible to identify the shape, intensity and color. Active appearance models require a training phase where both the shape and its associated appearance is learned from a set of samples using, for instance, the principal component analysis.

- *Multiview appearance models.* These models are used when objects are replicated from different views within a single time sequence[116] . One limitation of this approach is that the appearance information of all views of the tracked objects is required ahead of time.

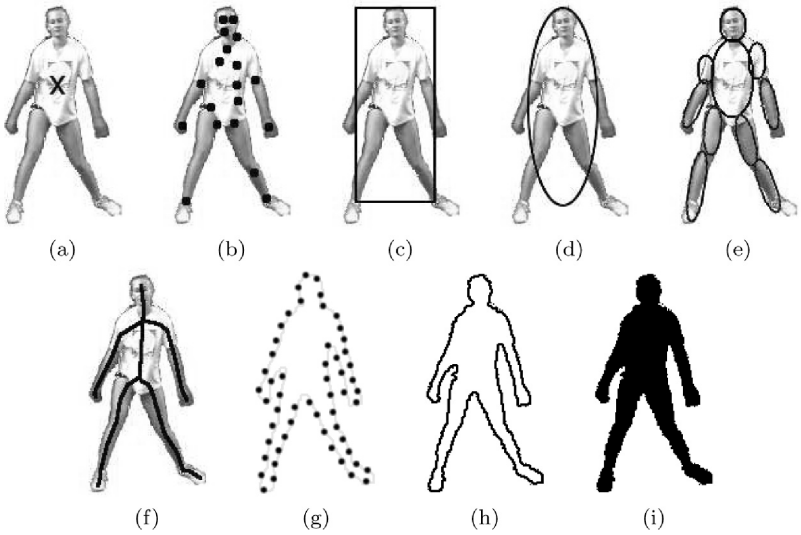

Figure 3.4: Object representations. (a) Centroid, (b) multiple points, (c) rectangular patch, (d) elliptical patch, (e) part-based multiple patches, (f) object skeleton, (g) complete object contour, (h) control points on object contour, (i) object silhouette (figure from [117]).

3.3.2 Feature selection for tracking

Selecting the right features plays a critical role in tracking. In general, the most desirable property of a visual feature is its uniqueness so that the objects can be easily distinguished in the feature space. Feature selection is closely related to the object representation. The objectives of feature selection are multifold. The details of common visual features are described in the following according to [117]:

1. *Color.* The apparent color of an object is affected primarily by two factors, namely, the spectral power distribution of the light source and the surface reflectance properties of the

object [118]. Among all the features, color is one of the most widely used. However, these color spaces are sensitive to noise.

2. *Edges.* Edge detection is used to identify strong changes in image intensities generated by object boundary. Edges features are less sensitive to illumination changes compared to color features. An evaluation of the edge detection algorithms is provided by [119].

3. *Optical flow.* Optical flow is commonly used as a feature in motion-based segmentation and tracking applications. It is a dense field of displacement vector which defines the translation of each pixel in a region. Its use allows to assign to each pixel belonging to the current frame, a motion vector pointing to the position of the same pixel in a subsequent reference frame. It is computed using the brightness constraint, which assumes brightness constancy of corresponding pixels in consecutive frames [94, 120].

4. *Texture.* Texture is a measure of the intensity variation of a surface, which quantifies properties such as smoothness and regularity. In contrast to color, texture requires a processing step to generate the descriptors, e.g., gray-level cooccurence matrices [121], wavelets [122] or steerable pyramids [123]. The texture features are less sensitive to illumination changes compared to color as edge features.

In most cases, the features are chosen manually by the user depending on the application domain. Nowadays, users pay large attention to the problem of automatic feature selection, which can be divided into[117, 124]:

1. *Filter methods [125].*The filter methods rely on feature selection based on a general criteria, and are able to find a good feature subset independently of the model selection step. The benefit of this approach is that they can be easily scaled to very high-dimensional data sets while being computationally simple and fast. Furthermore, they are independent of the classification algorithm. However, they ignore the interaction within the classifier.

2. *Wrapper methods* [126, 127]. The wrapper methods incorporate the model hypothesis search in the feature subset search. Consequently, the advantages are the interaction between feature subset search and model selection, and the possibility to take into account feature dependencies. A common drawback is that they are more likely over fitting than filter techniques and require intensive computations.

3. *Embedded methods* [128, 129]. The embedded methods can be considered as a search in the combined space of feature subsets and model hypothesis. The benefit of this approach is the interaction with the classification model, meanwhile being far less computationally intensive than wrapper methods.

3.3.3 Object tracking

Object tracking is closely related to object detection and feature selection. In a tracking scenario, the trajectory of an object is generated by locating its position in every frame of the video. Object tracker may also provide the complete region in the image that is occupied by the object at every

time instant. The tasks of detecting the object and setting up correspondence across frames can either be performed separately or jointly. When performed separately, the object regions in every frame are obtained through object detection algorithms, and then the tracker corresponds to objects across frames. When performed jointly, the object region and correspondence are jointly estimated by iteratively updating object location and region information obtained from previous frames. In both tracking approaches, the objects are represented using the shape and/or appearance models described in Section 3.3.1. In medical imaging, most of the objects to be tracked are nonrigid, hence, silhouettes or contours are the most descriptive representations, and both parametric and non-parametric models can be used to specify their motion.We now briefly introduce the main tracking categories and discuss them more detailed in the following subsections:

1. *Point tracking.* Tracking can be elaborated as the correspondence of detected objects represented by points across frames. Indeed, point tracking methods can be divided into deterministic [105, 130] and statistical methods [131, 132]. The deterministic methods use qualitative motion to constrain the problem, while probabilistic methods consider the object measurement and uncertainties to establish correspondence.

2. *Kernel tracking.* Kernel tracking is performed by computing the motion of the object, represented by a primitive object region across the frame. The object motion is usually in the form of parametric motion or the dense flow field computed in subsequent frames. We divide the kernel tracking methods into two categories: templates and density-based appearance models [107, 133, 134] , and multiview appearance models [135].

3. *Silhouette tracking.* Objects may have complex shapes that cannot be well described by simple geometric shapes. Silhouette-based methods can provide a more precise shape description for such objects. We divide silhouette tracking methods into two categories: shape matching [136, 137] and contour tracking [138, 139].

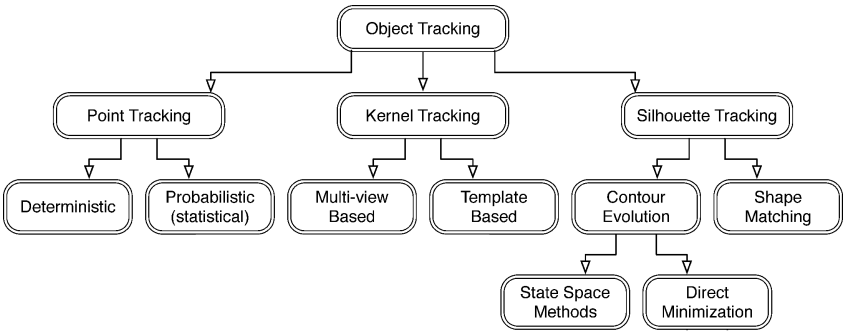

Figure 3.5: Taxonomy of tracking methods (figure from [117]).

3.3.3.1 Point-based tracking

Point-based tracking can be formulated as the correspondence of detected objects that are represented by points across frames of the image sequence. Point correspondence is a complex problem, especially in the presence of occlusions, misdetections, entries, and exits of objects. Recognition

of the points can be done in relatively simple way, by thresholding. Potentially good points for tracking are those that have distinctive local characteristics, such as brightness, contrast or texture, e.g. points associated with high-local curvature information such as corners and edges. As mentioned before, Point-based tracking methods can be divided into two broad categories, namely, deterministic and statistical methods.

1. *Deterministic methods.* For this methods constraints are used, such as qualitative motion heuristics [105], to obtain the correspondence of the points. Commonly, a cost is defined associating each object in frame $t-1$ to a single object in frame t using a set of motion constraints. Minimization of the correspondence cost is formulated as a combinatorial optimization problem. A solution, which consists of one to one correspondences among all possible associations, can be obtained by optimal assignment methods, such as the greedy search method [140]..The correspondence cost is usually defined by using a combination of the following constraints, however, it should be noted here that these constraints are not specific to the deterministic methods, and they can also be used in the context of point tracking using statistical methods:

   (a) *Proximity* assumes that the location of the object would not change notably from one frame to the next, as shown in Figure 3.6 (a)

   (b) *Maximum velocity* defines an upper bound on the object velocity and limits the possible correspondences to the circular neighborhood around the object, as shown in Fig. (b).

   (c) *Small velocity change* assumes that the direction and speed of the object does not change drastically, as shown in Figure 3.6 (c).

   (d) *Common motion* constrains the velocity of objects in a small neighborhood to be similar, as shown in Figure 3.6 (d).

   (e) *Rigidity* assumes that the distance between different point on the actual object will remain unchanged, as shown in Figure 3.6 (e).

   (f) *Proximal uniformity* is a combination of the proximity and the small velocity change constraints.

2. *Statistical method*s. They explicitly take the object measurement and uncertainties into account to establish correspondence. They use the space approach to model the object properties such as position, velocity, and acceleration. Measurements usually consist of the object position in the image, which is obtained by a detection mechanism. Bayesian methods are the most commonly used statistical technology for motion tracking in medical images, and are especially used for motion parameter estimation and object tracking. For instance, Schaap et al. [141] used a Bayesian framework for tracking of carotid arteries and Peressutti et al. [142] combined the motion model with real-time 3D echocardiography images based on a Bayesian framework to estimate cardiac motion.

One important issue of point-based tracking is the handling of missing or noisy observations. Medical images inevitably contain noise and the object motion can undergo random perturbations. To address these problems, computational point-based tracking methods often use a combination of

motion-based constraints. Statistical correspondence methods solve these tracking problems by taking the measurement and the model uncertainties into account during object state estimation.

Point-based tracking methods are suitable for tracking small or single objects, which can be represented by a small number key points. In the context of tracking larger or multiple objects represented by a large number of points, automatic clustering and tracking of points that belong to the same object becomes an important problem. This is due to the need to distinguish between multiple objects and, between objects and the background.

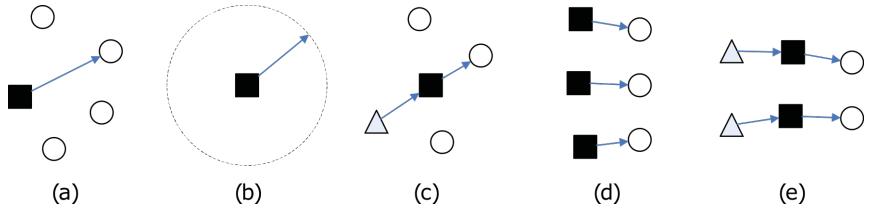

Figure 3.6: Different motion constraints. (a) Proximity; (b) maximum velocity; (c) small velocity change; (d) common motion; (e) rigidity constraint (figure from [36]).

3.3.3.2 Kernel-based tracking

Kernel Tracking is the most commonly used when the object in motion follows the parametric motion model. The applied algorithms differ in terms of the appearance of the object to be tracked, the number of objects to be tracked and most significantly the motion model used to estimate the next state of the object along the path.

These tracking methods can be divided into two subcategories based on the used appearance representation, namely, template-based and multiview-based tracking methods.

1. *Template-based tracking.* This method is closely related to region-based tracking because a template is basically a model of the image region (and, therefore, of the corresponding object). This approach is widely used due to its relative simplicity and low computational costs. In particular it is for 2D tracking applications in MRI, such as optical flow models [143] and B-spline models [144].

2. *Multiview-based tracking.* This tracking method is mainly applied for 3D medical image analysis. When applying the 2D methods discussed before, there is always the possibility that the appearance model is no longer valid, when the object view changes dramatically during tracking, thus the object track might be lost. This problem can be solved by learning different views of the object offline and use this information for tracking. A typical application for multiview-based tracking is MRI-HARP imaging with cardiac motion tracking [145].

Kernel-based tracking techniques overcome some limitations of traditional point-based tracking techniques. The advantages are as follows:

- They can track single or multiple objects.
- They have the ability to handle occlusion.
- They have general applicability.

The major drawback of kernel-based tracking methods is that parts of the objects may be left outside of the defined shape, while parts of the background may reside inside it. This phenomenon can be observed for rigid objects when the object pose changes and nonrigid objects when local motion results in changes in object appearance. In such cases, the object motion estimated by maximizing model similarity may not be correct.

3.3.3.3 Silhouette-based tracking

In medical images, the objects of interest often have complex shapes that cannot be well described by few points or simple geometric shapes. Silhouette-based methods provide an accurate shape description for these objects and are the most convenient approach for non-rigid object tracking like hands, vessels and similar body parts. The goal of the silhouette-based tracking is to identify the object region in the successive frame by computing the model obtained from the previous frames. Silhouette trackers can be categorized based on the used approaches into shape matching and contour tracking techniques:

1. *Shape matching method* are similar to the template-based tracking method described in Subsection 3.3.3.2. An object silhouette and its associated model are searched in the current frame and the search is performed by computing the similarity of the object with the model generated from the hypothesized object silhouette based on the previous frame. It is assumed that the shape changes from one image to the next can be described by simple translations, hence non-rigid object motions are not explicitly handled. To compensate for non-rigid object motions and to overcome problems with viewpoint and lightning condition changes, the object model is reinitialized in every frame after the object is located. The object model used for this method is usually an edge-based representation as performed by Huttenlocher et al. [136]. Another approach to match shapes is to find corresponding silhouettes detected in two consecutive frames. This kind of silhouette matching, can be considered similar to point matching with the main difference being the object representations and the object models used. In particular, silhouette matching uses the complete object region, whereas point matching uses points. In addition, silhouette matching makes use of an objects appearance features, whereas point matching uses only motion and position-based features.

2. *Contour tracking methods* are iterative procedures that evolve an initial contour in the previous frame to its new position in the current frame. Contour tracking can either be done using *state space models* or by the *minimization of the contour energy functional*.

   (a) *State space model.* In state space models, the state of an object is represented by the shape and motion parameters of the contour, which is updated in each frame by maximizing the contour's posteriori probability. This posterior probability in turn is dependent on the prior state and the likelihood function is defined in terms of the distance of the contour from the identified edges. A variety of different versions of this approach has been developed such as the tracking with so-called Kalman snakes [146]. It is important to mention that all the state space models represent the contours using explicit representation, for example, parametric splines. Explicit representations do not allow topology

changes such as region split or merge.

(b) *Minimization of the contour energy functional.* In the context of contour evolution, there is an analogy between the segmentation methods discussed in Subsection 3.1.4 and the contour evolution techniques described here. Both the segmentation and tracking methods minimize the energy functional either by greedy methods or by gradient descent. The contour energy is defined in terms of temporal information in the form of either the temporal gradient (optical flow) [139] or appearance statistics generated from the object and the background regions [108]. The methods based on direct minimization of energy functional can use implicit representations and allow topology changes.

Silhouette-Based Tracking is applied when tracking of the complete region of an object is required. The most important advantage of tracking silhouettes is their flexibility to handle a large variety of object shapes. Deformable models are widely used in medical image analysis and offer a unique and powerful approach to image analysis that combines geometry, physics, and approximation theory. In medical images, the general shape, location, and orientation of objects are known and this knowledge may be incorporated into the deformable model in the form of initial conditions, data constraints, constraints on the model shape parameters, or into the model fitting procedure. One of the most important applications for deformable models in MRI are cardiac motion analyses [147, 148]].

In summary, it can be said that deformable models overcome many of the limitations of traditional low-level image processing techniques and offer an attractive approach to tackle an array of challenging problems centered on the motion of anatomic structures from medical images because these models are able to represent the complex shapes and broad shape variability of anatomical structures. The main advantages can be summed up to the following three points:

- They are insensitive to initial conditions and noisy data.
- They can exploit multiple image attributes and high-level or global information to increase the robustness of shape recovery.
- They can easily incorporate additional constraints and any other a priori anatomic knowledge of object shape and motion.

In this thesis, we developed a system for tracking of deformable organs and tissue structures using the active contour model. Therefore, a detailed description of this method is presented in the following chapter.

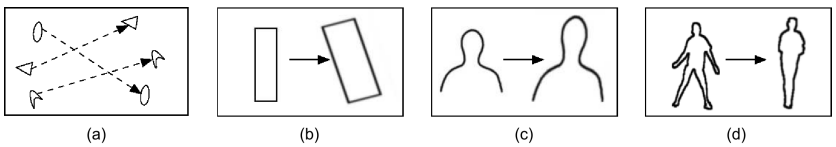

Figure 3.7: Different tracking approaches. (a) Multipoint correspondence, (b) parametric transformation of a rectangular patch, (c, d) Two examples of contour evolution (figure from [117]).

3.4 3D imaging operations

3D imaging is a branch of medical imaging that generates volumetric data of an object. The most common techniques in clinics are X-ray computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), single-photon emission tomography (SPEC) and ultrasound imaging (US). It is important to mention here that data obtained from a 3D imaging technique do not necessarily have three dimensions. For, instance a slice of an MRI data set is a 2D image and a time series of volumetric data is called a 4D data set. Furthermore, the expressions "3D data", "3D data set" or "3D image" are not exclusively used for volumetric data sets in the classical sense, but can also refer to a time series of 2D images, such as MRI slices or projection images from digital radiography. In the following section a 3D data set, i.e., a three-dimensional array of voxels is also termed scene. Furthermore, an object may refer to a physical object, such as an organ or other anatomical or pathological structures, or to a conceptional object, such as an isodose surface in radiation treatment planing. An object system refers to a collection of physical or conceptional objects.

The main purpose of 3D imaging in medicine is to produce qualitative and quantitative information about an object under study. The qualitative information regarding the appearance of the visualization, and are information that the specialist, doctor or radiologist can obtain simply by examining the visualized 3D data. The quantitative information mainly relates to the morphology and function of the object or object system. They can be quantitative information, for example, geometric information, such as the volume or perimeter of the region of interest, or information related to a parameter that characterizes the region.

3D imaging is a crucial tool in many medical applications, such as diagnosis, treatment planing or computer-assisted image-guided intervention.

In General, three-dimensional imaging operations can be broadly categorized into 4 groups [149]:

- *Preprocessing*: In this step the object system is defined to create a geometric model of the objects of interest
- *Visualization*: These operations are focused on viewing and comprehending the structure of the object system.
- *Manipulation*: The goal of these operations is to alter objects or the relationship among objects in the object system, for example, to virtually simulate surgery procedures.
- *Analysis*: The aim of these operations is to retrieve quantitative data of the object system.

In the following subsections, the categories above are briefly presented with focus on introducing concepts that have not been discussed in the previous sections and which are necessary to understand the content of this thesis. For a more detailed introduction into 3D image operations please see [149, 150].

3.4.1 Preprocessing

The goal of preprocesssing operations is to use a set of scenes as input and to directly generate geometric object models or alternatively another set of scenes, which facilitates the creation of

object models.

The most commonly applied preprocessing operations can be classified in:

- *Volume of Interest (VOI):* Defining a volume of interest basically converts a given scene into another scene, which usually consists of a smaller number of voxels. Its purpose is to reduce the amount of data by specifying a region of interest and a range of intensity of interest.
- *Filtering*: Filtering is a conversion of a given scene into another scene. Its aim is to enhance wanted (object) information and suppress unwanted information, such as noise, background or objects that are not of interest. Filters can be divided into suppressing filters, such as smoothing filters to suppress noise, and enhancing filters, such as edge filters.
- *Interpolation*: This operation converts a given scene to another scene. The goal of interpolation is to change the sampling (the level of discretization sampling) of the input scene. Interpolation is necessary to convert the non-isotropic discretization of the input scene to an isotropic or to a desired level of discretization, to represent longitudinal scene acquisitions of a patient in a registered common scene coordinate system, to represent multimodality scenes in a registered common scene coordinate system or to re-section the input scene. 3D interpolation methods calculate the intensity of a voxel in the output scene on the basis of the neighboring voxels in the input scene. There are many different approaches regarding the determination of neighborhoods of a voxel and how the neighboring voxel intensities are used to estimate the intensity of the voxel in the output scene. The most common approaches are nearest neighbor interpolation, linear interpolation and cubic interpolation methods.
- *Segmentation*: For a given set of scenes, segmentation generates computer models of object information contained in the scenes. The goal of segmentation is to identify and delineate objects, as already discussed in Section 3.1.
- *Registration*: In the context of 3D imaging operations, registration takes two scenes or two objects as input and calculates a transformation that is applied to the second scene/object in order to match as accurately as possible to the first. Its purpose is to combine and compare scene/object information from multiple scenes/modalities/protocols, to determine changes between the scenes/objects, such as growth, motion, displacements of objects, and to facilitate object identification (see Section 3.2).

3.4.2 Visualization

3D visualization is a method of extracting meaningful information from volumetric data using interactive graphics and imaging. It aims at a representation of 3D data that facilitates the visual perception of object information and allows for comprehending objects in their full dimensionality. Visualization methods can be divided into scene-based and object-based approaches.

3.4.2.1 Scene-based visualization

Scene-based visualization methods create renderings directly from given scenes. These methods can be subdivided into slice mode and volume mode.

- *Slice mode.* These methods differ on what is considered to constitute a slice (axial, coronal, sagittal, oblique/curved slices) and how this information is visualized (montage, roam through/fly through, gray scale/pseudo color). A simple example for scene-based visualization in slice mode is displaying a slice of a tomographic data set, such as an MR or CT slice. For multislice sets, the images can be displayed one at a time in a window that allows the user to move through the image stack or as a montage that shows different slices at the same time, such as orthographic views.

- *Volume mode.* In this mode information is displayed as surfaces, interfaces or intensity distribution by using surface rendering, volume rendering or maximum intensity projection (MIP). In all cases, a projection technique is required to get from the three-dimensional data to a two-dimensional representation on a viewing plane. Two methods of projection are typically used in converting higher dimensional object information to their 2D renditions: ray casting [151]which consists of tracing a line perpendicular to the viewing plane into the scene of domain, or voxel projection [152] which consists of directly projecting voxels encountered along the projection line from the scene onto the viewing plane.

  1. *Surface rendering.* Initially, the surface to be rendered is determined, usually by gray-level thresholding, i.e., it is defined whether a voxels is on the surface or not by assigning an opacity value of 100% or 0%. The actual rendering process consists of three steps: projection, hidden part removal and shading. These steps are necessary to give the surface rendering a 3-D appearance. Additional cues for three-dimensionality may be provided by techniques such as stereoscopic display, motion parallax by rotation of the objects, shadowing, and texture mapping. Here, we describe these cues very briefly according to [153]:

     – Hidden part removal is used to mimic a realistic perspective. If another object is blocking the view of an object, then it is not seen.

     – Shading supports the impression that the surface has a shape and angle by making the sides of an object appear darker than the surface perpendicular to the viewer.

     – Stereoscopic displays simulate how the observer normally perceives three-dimensional objects in reality, i.e., each eye receives a slightly different view of the object, and these views are combined in the brain to give the impression of depth. This is accomplished in visualization by generating a left-eye view and a right-eye view of the object and ensuring that the respective view is seen only by the corresponding eye. For this purpose 3D-glasses based on color filters or shutters are used.

     – The impression of three-dimensionality can be reinforced by making the object appear to rotate, so that the shading changes as the object moves relative to the light source.

     – Casting shadows can further enhance the impression of a three-dimensional surface.

     – Mapping texture on a surface can be used to give a more realistic appearance of the materials resembled by the data. Furthermore, the texture is often designed to make the surface look shiny in order to enhance the 3-D appearance.

2. *Volume rendering*: Volume rendering is very similar to surface rendering, but instead of defining the object surface in absolute terms, a "fuzzy" object is used. In this case, each voxel has a opacity value ranging from 0% to 100% that indicates how much it resembles a voxel that is on the surface of interest. The opacity value determines the transmission, emission and reflection from that voxel when the rendering is performed. The fuzzy approach is very useful for medical imaging, where partially volumed voxels are common. The partial volume effect arises because each voxel may represent more than one tissue type, and so have a value that does not clearly belong to one tissue or another. Another feature associated with volume rendering that distinguishes it from surface rendering is that the full set of data remains available, meaning that rather complex visualizations including surfaces, slices and objects can be generated. Like surface rendering, volume rendering consists of three basic step: projection, hidden part removal and shading. While the principles underlying projection are identical to those described for surface renderings, hidden part removal is much more complicated in volume rendering [154, 151]. The shading operation is also more complicated since it has to take into account transmission, reflection and emission. Additional cues for 3D-appearance, such as the ones discussed for surface renderings, can also be applied to volume rendering.

3. *Maximum intensity projection (MIP).* Maximum intensity projection is achieved by projecting only the brightest voxel that occurs along the projection line [155, 156]. It is the simplest of all 3D rendering techniques. It works most efficiently when the object of interests are the brightest features in the image and have a simple 3D morphology. Furthermore, it is important that there are only a small number of bright objects in the scene. MIP is widely used, especially for vascular applications such as contrast-enhanced CT [157] and magnetic resonance angiography [158] . Its main advantage is that it does not necessarily require segmentation.

3.4.2.2 Object-based visualization

In object-based visualization, initially the objects of interest are explicitly defined and subsequently rendered. This approach is mainly used in difficult segmentation situations or when the amount of data of the entire scene is very large. In this cases it would be impractical or very time-consuming to perform scene-based rendering. The general representation of object-based data are the same as for scenes. While object-based surface rendering requires hard object descriptions as input, volume rendering and maximum intensity projections rely on a fuzzy object description.

Surface rendering of objects consists of the same basic steps as the scene-based case. While the basic concepts of projection, hidden-part removal and shading are identical, many methods specially suitable for object-based surface rendering have been developed. For instance, there are surface description methods for objects, such as voxels [152, 159], voxel faces [160] and triangles [161, 162]. The three-dimensional perception can be improved by the aforementioned additional cues for 3D appearance.

Object-based volume rendering is based on the same principles as the scene-based case. Since the

object description usually consists of a smaller amount of data than the original scene and additional information for increasing computation speed can be stored as part of the object description, volume rendering based on fuzzy objects requires less computing power and can be performed online even on personal computers.

For object-based maximum intensity projection, the initial scene is converted via fuzzy segmentation to a new scene, in which the objects of interest have a higher intensity than the non-object regions. Subsequently, the new scene is rendered via scene-based MIP as described before.

3.4.3 Manipulation

Manipulation operations are mainly used to create a second object system from a given object system by changing objects or their relations. The main goal of these operations is to simulate surgical procedures on the basis of patient-specific scene data and to develop aids for interventional and therapy procedures. The commonly manipulations can be divided into rigid and deformable manipulations.

Rigid manipulations cut, separate, add, subtract, move, and mirror objects and their components that can be defined in both hard [163] or fuzzy fashions [164]. In rigid manipulation, the user interacts directly with an object-based surface or volume rendering of an object system, hence the manipulations must be performed online.

Deformable manipulations include stretching, compressing, bending and so on. It is used to mechanically model soft-tissue structures, such as muscles, tendons or ligaments. Another application is facial plastic surgery for which facial expression animations are modeled [165]. Deformable manipulations are considerably more complex than rigid manipulations and require more computing power.

3.4.4 Analysis

The main goal of analysis operations is to create a quantitative description of the morphology, architecture, and function of the object system from a given set of scenes or an object system.

In scene-based analysis, quantitative descriptions are based directly on the intensities of the 3D data set and include ROI statistics, density, activity, perfusion, and flow. Structure information derived from a scene acquired with another modality is often used to guide the selection of regions for these measurements.

In object-based analysis, quantitative description is obtained from an object on the basis of morphology, architecture, change over time, relationships with other objects in the system, and changes in these relationships. Examples of measurements obtained in this way include distance, length, curvature, area, volume, kinematics, kinetics, and mechanics. Since object information in images is fuzzy, the challenge is to develop algorithms based on fuzzy mechanics and morphometry theories in order to analyze the object information in a suitable manner.

Chapter 4

Methodology and implementation of motion tracking algorithm

As previously mentioned in Chapter 2, the organs and lesion examined via the DCE procedure can change their shapes and positions over time due to patient movements. The basic idea behind the presented work is to register the images of a time sequence in such a way that information can be obtained from the same pixel throughout the time sequence in order to retrieve accurate information of the analyzed anatomical structures. For this purpose, we identify automatically the ROI contour for each image of the sequence via an active contour model, calculate the displacement in the different images and use a rigid rototranslational transformation to realign the images. In the following, we will present a detailed mathematical description of the applied active contour model, followed by a delineation of the individual methodological steps of our method.

4.1 Active contour model

Since deformable models were first introduced by Kass et al in 1987[51], they have become a very powerful,versatile, dynamic technique applied in many image processing fields, such as image segmentation, edge detection, shape modeling, and tracking of medical image structures. Deformable or active contour models can generate a continuous boundary of an object of interest in spite of possible large shape variations, the presence of image noise and inhomogeneities, and discontinuous object boundaries due to occlusions. An active contour model (AC) is a curve in a 2D image or a surface in a 3D image that evolves to outline a goal object. Its evolution is guided by internal and external forces, which are defined so that the deformable model will eventually conform to the object boundary. The internal forces, coming from the curve/surface itself, are designed to keep the evolved curve smooth and unified. The external forces, depending on the image, propagate the evolution towards the object boundary. Generally, all the parametric deformable models are initialization-sensitive, and different initializations may produce different results. In the majority of models, the initial contour or surface should be placed close to the goal object boundary.

The deformable model geometry usually allows a wide coverage of the shape of the object

through the use of geometric representations that involve many degrees of freedom. The model remains manageable since the degrees of freedom are generally not allowed to evolve independently, as their evolution is governed by physical principles that give to the model a significantly intuitive behavior from a geometric point of view. The name deformable models derives mainly from the use of the theory of elasticity on a physical level; the physical interpretation sees the deformable models such as elastic bodies that naturally respond to the forces and the constraints applied to them. To confirm this, the deformation energy function of the model is defined in terms of geometric degrees of freedom. The deformation energy gives rise to both elastic forces inside the model and to attractive forces outside the model itself: with increasing of internal energy the model deforms and at the same time is affected by the loss of those terms that limit the elasticity and symmetry of the model. External potential energy functions are defined in terms of the data of interest to which the model refers: these potential energies give rise to external forces that deform the model in such a way that it best fits the data in question. In the following, we describe the mathematical principles that are the basis of the functioning of deformable models.

4.1.1 Mathematical formulation

Active contour models (or snakes) are deformable contours represent a special case of the general multidimensional deformable model theory [166] and have been used in many image analysis applications, including the image-based tracking of rigid and nonrigid objects, . In their basic forms, the mathematical formulation is derived from the theory of optimal approximation involving a functional. A traditional snake is a parametric contour embedded in the image plane $(x, y) \in \Re^2$. The contour is represented as a parametric curve:

$$V(s) = [X(s), Y(s)] \tag{4.1}$$

where $X(s)$ and $Y(s)$ are the coordinate functions and $s \in [0, 1]$ is the parametric domain. This curve is generally obtained using B-splines functions [167, 168, 169, 170] that allow to depict a continuous and smoothing parameterized contour.

B-splines are a particular, computationally convenient representation for spline functions. In the B-spline form, a spline function $V(s)$ is constructed as a weighted sum of $N_B$basis functions (hence, B-splines) $B_n(s)$, $n = 1, ..., N_B$. The basic idea of the B-splines is the possibility to obtain a contour as a combination of a set of curves as a replication of a basis function. For each basis function, $B_j(s)$, a control point $\boldsymbol{q_j} = (q_j^x, q_j^y)^T$ must be defined and the curve is a weighted vector sum of those control points, written as follows:

$$V(s) = \sum_{j=1}^{N_B} \mid B_j(s) \cdot \boldsymbol{q_j} \mid \tag{4.2}$$

where the vectors $\boldsymbol{q_j}$ contain the coordinates of control points, $B_j(s)$ is the polynomial basis function, s is the curve parameter, and $V(s) = [X(s); Y(s)]$ is a point in the 2-D image.

By convention, B-spline basis functions are constructed in such a way that they sum to 1 at all points:

$$\sum_{j=1}^{N_B} \mid B_j(s) \cdot \boldsymbol{q_j} \mid = 1 \tag{4.3}$$

In the presented approach, we use a simple quadratic B-spline. The first B-spline basis function $B_0(s)$ has the following form:

$$B_0(s) = \begin{cases} \frac{s^2}{2} & \forall s \in [0,1) \\ \frac{3}{4} - (s - \frac{3}{4}) & \forall s \in [1,2) \\ \frac{(s-3)^2}{2} & \forall s \in [2,3) \\ 0 & otherwise \end{cases} \tag{4.4}$$

and the rest are simply translated copies of the first one.

$$B_{n,d}(s) = B_0(s-n)$$

Figure 4.1:

As can be seen from the graph above, each B-spline is different from zero only in a specific range depending on the order to which d belongs:

$$B_{n,d}(s) \neq 0 \, \forall s \in [n, n+d)$$

One may rewrite the sum in Eq. 4.2 taking into account that active contours are closed lines and by using the arithmetic modular, as follows:

$$V(s) = \sum_{i=j-2}^{j} \mid B_0(s-j) \cdot \boldsymbol{q_{1+mod(i,N_B)}} \mid \tag{4.5}$$

By varying the number and the location of the control points it is possible to change the contour smoothness and its ability to fit a generic figure.

As mentioned before, the final goal of an AC is to move the contour in such a way that an energy measure is minimized. It is possible to associate each AC with an energy function [51, 171, 172], which represents a measurement:

$$E^{(i)}_{snake} = E^{(i)}_{internal} - E^{(i)}_{external} \tag{4.6}$$

where the first term is represents the internal energy (or smoothness energy) of the snake, and the second term expresses the external energy computed over the entire contour (at the *ith* position of the contour). The negative sign in front of the second integral implies that we want to maximize the sum of the image gradient magnitude over the entire contour.

The internal energy is used to control the rate of stretch and to prevent discontinuity in the contour and can be defined as follows:

$$E_{int} = \frac{1}{2}\int_0^1 \left[\alpha\left(\left|\frac{dX}{ds}\right|^2 + \left|\frac{dY}{ds}\right|^2\right) + \beta\left(\left|\frac{d^2X}{ds^2}\right|^2 + \left|\frac{d^2Y}{ds^2}\right|^2\right)\right] ds \tag{4.7}$$

where α and β are weighting parameters that control the snake's tension and rigidity respectively. The first term in Eq. 4.7 is related to the contour elasticity, thus called elastic energy, while the second term specifies the contour's strength and resistance against sudden changes, thus referred to as bending energy:

$$E_{elastic} = \int_0^1 \alpha\left(\left|\frac{dX}{ds}\right|^2 + \left|\frac{dY}{ds}\right|^2\right) ds \tag{4.8}$$

$$E_{bending} = \int_0^1 \beta\left(\left|\frac{d^2X}{ds^2}\right|^2 + \left|\frac{d^2Y}{ds^2}\right|^2\right) ds \tag{4.9}$$

The discrete version of the two energy terms is given by:

$$E_{elastic} = \alpha\sum_{i=1}^{N}\left[(X_i - X_{i-1})^2 + (Y_i - Y_{i-1})^2\right] \tag{4.10}$$

$$E_{bending} = \beta\sum_{i=1}^{N}\left[(X_{i+1} + 2X_i + X_{i-1})^2 + (Y_{i+1} - 2Y_i + Y_{i-1})^2\right] \tag{4.11}$$

where N is the total number of points that compose the snake, while X_i and Y_i are the coordinates of each point.

The second term in Eq. 4.6 represents the external energy, which moves a snake towards the feature of interest (e.g., borders) in an image. This value will be minimal when the snake arrives at the features of interest. It can be defined as follows:

$$E_{ext} = \int_0^1 \kappa\left(f\left[X(s), Y(s)\right]\right) ds \tag{4.12}$$

$$f(x,y) = |\nabla(G_\sigma(x,y) * I(x,y))| \tag{4.13}$$

$$G_\sigma(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(\frac{-(x^2+y^2)}{2\sigma^2}\right) \tag{4.14}$$

where $I(x,y)$ is a gray-level image and $f(x,y)$represents the absolute value of the gradient of the gray-level image previously convoluted with a Gaussian function $G_\sigma(x,y)$.

The Gaussian filter is used to blur the image, thereby making the surface uniform and eliminating spikes and ripples due to noise. The gradient is applied to find the edges of the image in order to attract the snake to the edges [51, 172]. It is important to underline here that we introduced a scaling factor κ that multiplies the modulus of the gradient. This is a novelty with respect to [51] and reduces the convergence time of the algorithm.

The contour and its energy function depend only on the location of N_B control points; hence, the problem is to find the best contour that approximates the edge of the ROI. From a mathematical point of view, this can be reduced to the minimization of the energy function. When we insert Eq.

4.7 and Eq. 4.12 into Eq. 4.6 the energy function of the snake is given by:

$$E_{snake} = \frac{1}{2}\int_0^1 \left\{ \alpha \left(\left| \frac{dX}{ds} \right|^2 + \left| \frac{dY}{ds} \right|^2 \right) + \beta \left(\left| \frac{d^2X}{ds^2} \right|^2 + \left| \frac{d^2Y}{ds^2} \right|^2 \right) \right\} ds - \int_0^1 \kappa f\left[X(s), Y(s)\right] ds \tag{4.15}$$

Minimizing the energy function describes the process of finding the optimum location of the snake where the energy is minimal. This results in the following equations:

$$\frac{\delta E_{Snake}}{\delta X} = -\alpha \frac{d^2X}{ds^2} + \beta \frac{d^4X}{ds^4} - \kappa \frac{\partial f}{\partial x} = 0 \tag{4.16}$$

$$\frac{\delta E_{Snake}}{\delta Y} = -\alpha \frac{d^2Y}{ds^2} + \beta \frac{d^4Y}{ds^4} - \kappa \frac{\partial f}{\partial y} = 0 \tag{4.17}$$

The equations 4.16 and 4.17 are known as Euler equations. In general, there is no closed form solutions for those equations. Therefore, it is necessary to resort to the use of numerical techniques to solve them. The gradient descend method (GDM) allows unconstrained minimization based on the use of search direction opposite to the gradient. In other words, when assuming a point moving on a potential surface, it is obvious that the point moves in the opposite direction of the gradient of the potential surface and stops where the gradient magnitude is zero, consequently, the velocity of the point is proportional to the negative gradient of the potential surface. In analogy, the GDM dictates that, with respect to a time, the velocity of the snake will be proportional to the negative gradient of Eq. 4.15 and is given by:

$$\frac{\partial X}{\partial \tau} = \alpha \frac{d^2X}{ds^2} - \beta \frac{d^4X}{ds^4} + \kappa \frac{\partial f}{\partial x} \tag{4.18}$$

$$\frac{\partial Y}{\partial \tau} = \alpha \frac{d^2Y}{ds^2} - \beta \frac{d^4Y}{ds^4} + \kappa \frac{\partial f}{\partial y} \tag{4.19}$$

A physical interpretation of the equations is given by a force balance equation in the form of:

$$\boldsymbol{F_{int}} + \boldsymbol{F_{ext}} = 0 \tag{4.20}$$

Starting from an initial snake location, the resulting force drives the snake and when the resulting force is zero, the snake reaches its equilibrium position and we obtain the desired snake location. The Equations 4.18 and4.19 are combinations of internal and external forces, whose contributions are properly weighted by the adimensional parameters α, β, κ.

In analogy to the energy terms, the internal forces can be classified as elastic and bending forces that prevent the contour from stretching and bending and they are classified as:

- Elastic forces

$$F_x(elastic) = \alpha \cdot \frac{d^2X}{ds^2}$$

$$F_y(elastic) = \alpha \cdot \frac{d^2Y}{ds^2}$$

- Bending forces

$$F_x(bending) = \beta \cdot \frac{d^2X}{ds^4}$$

$$F_y(bending) = \beta \cdot \frac{d^2Y}{ds^4}$$

The external forces attract the snake toward the edge of the ROI by varying the intensities of the pixels of the boundary and neighborhood:

$$F_x(external) = \kappa \cdot \frac{\partial f}{\partial x}$$

$$F_y(external) = \kappa \cdot \frac{\partial f}{\partial y}$$

Applying the GDM to a discrete grid, i.e., a digital image, one obtains the following iterative equations, which are the discrete versions of Eq. 4.18 and Eq. 4.19:

$$X^{k+1} = X^k + \left(F^k_{x(elastic)} - F^k_{x(bending)} + F^k_{x(external)}\right)\Delta \tag{4.21}$$

$$Y^{k+1} = Y^k + \left(F^k_{y(elastic)} - F^k_{y(bending)} + F^k_{y(external)}\right)\Delta \tag{4.22}$$

where $X^{k+1}$and Y^{k+1}collect the coordinates of the new contour while $X^k$and Y^k collect the same quantity for the current one; k is the iteration index and Δ is the algorithm step. The Equations 4.21 and 4.22 produce a sequence of contour point locations that gradually minimize the energy of the snake; the process is generally stopped when the next iteration does not introduce a significant change of the total energy. The crucial point to guarantee the rapid convergence of the iterative procedure is the to select initial contour points (i.e.,X_0, Y_0) belonging to a region that is sufficiently close to the actual contour.

4.2 Outline of the developed algorithm

Starting from the considerations above, we propose to face the problem of tracking the region of interest in DCE-MRI measurements by two consecutive steps. First, we precisely identify the contour of the ROI (organ or lesion) in a particular image. Subsequently, we establish a correspondence among the positions of such ROIs in the different images that compose a sequence.

To depict automatically a contour inside a medical gray-scale image is a challenging task due to the very small difference in contrast and brightness between the ROI and the background . For a particular ROI, the contour recognition can be performed manually by a human expert or by the support of some suitable graphical tools. In the case of our algorithm, the ROI in one image is manually selected and the consecutive ROIs are determined automatically.

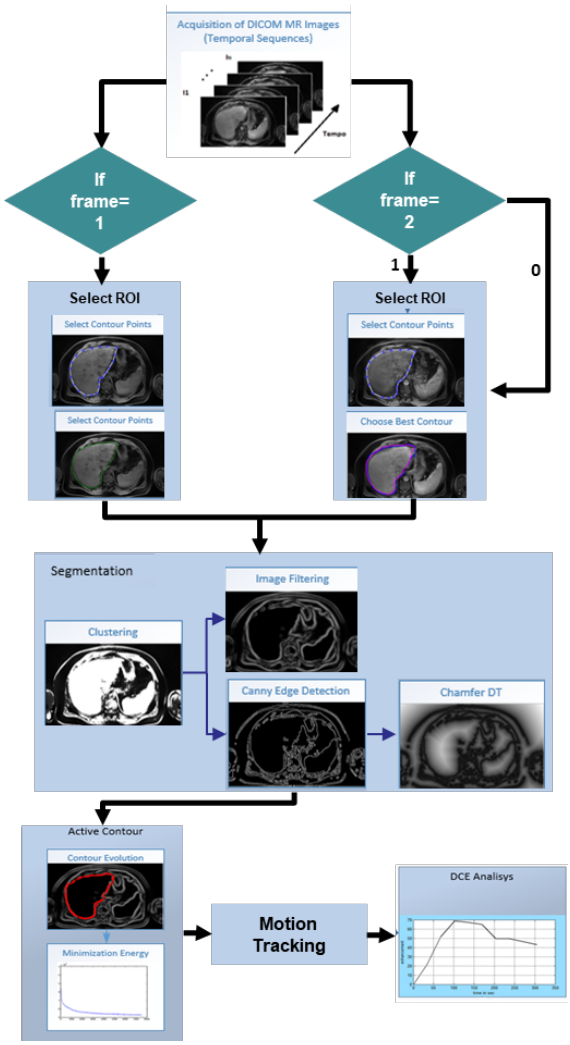

Figure 4.2: Flow chart of the motion tracking algorithm.

The proposed algorithm has been implemented MATLAB R2013a and it is composed by the following steps:

1. With the help of an expert, the ROI on the first image is selected ;
2. Use the AC algorithm to identify the optimal contour of the ROI by iterating the Equations 4.21 and 4.22;
3. Evaluate the center of mass and principal axis of the ROI;
4. Project the optimal contour on the next image in the temporal sequence, this will represent the starting location of the control points on which the AC procedure for the next image will be applied;
5. Use the AC algorithm, with the modification described later, to identify the optimal contour on the current image;

6. Evaluate the center of mass and principal axis of the actual ROI and calculate the rigid motion displacement with respect to the first image;

7. Compensate the motion of each point inside the ROI accordingly

8. repeat 4-5-6-7 until the end of the sequence.

In the following, we illustrate how the active contours have been implemented and how the algorithm can be used to compensate motion artifacts in a DCE-MRI time series and thereby allowing a more accurate analysis of the enhancement curves.

4.3 ROI selection

Since common DCE-MRI images display a larger area than the area occupied by the organs of interest, it is reasonable to perform a preprocessing step by selecting a ROI containing only information about the organ under investigation. To realize the segmentation of the anatomical structure of interest, it is necessary to initialize a contour that contains the ROI. Since this task is most efficiently performed manually, we ask a medical doctor to define a contour of the ROI on one of the images of the sequence (if necessary, the first one) by clicking several times on the contour of the ROI. The selected points are then interpolated using B-splines to obtain an initial contour that will subsequently be subject to optimization via active contours.

4.4 Segmentation

4.4.1 Image filtering and edge detection

As explained in Subsection 4.1.1, active contours are computer-generated curves that move within the image to find object boundaries. The movement of the contour is controlled by internal and external forces. The external energy requires the computation of the gradient of the image previously convoluted with a Gaussian smoothing filter. The gradient of the image (or the edge map) is commonly calculated by a convolution of the image with a gradient operator, which typically consists of horizontal and vertical gradient kernels. To find the best result for the edge map we tested the following operators and compared the resulting images:

- *Roberts operator.* The Roberts edge detection was introduced by L. G. Roberts in 1963 [173]. It performs a simple, quick to compute, 2-D spatial gradient measurement on an image. This method emphasizes regions of high spatial frequency which often correspond to edges. Pixel values in every point in the output represent the estimated complete magnitude of the spatial gradient of the input image at that point. The kernels can be applied separately to the input image, to produce separate measurements of the gradient component in each orientation (G_x and G_y). These can then be combined together to find the absolute magnitude of the gradient at each point and the orientation of that gradient. The Roberts operator consists of a pair of 2×2 convolution kernels.

$$G_x \approx \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \; G_y \approx \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

- *Sobel operator.* The Sobel or Sobel-Feldman operator was introduced by I. Sobel and G. Feldman in 1968[174] . The Sobel operator performs a 2-D spatial gradient measurement on an image and thereby highlights regions of high spatial frequency that correspond to edges. In general it is used to find the estimated absolute gradient magnitude at each point in an gray-scale image. The operator consists of a pair of 3×3 convolution kernels:

$$G_x \approx \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \; G_y \approx \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

- *Prewitt operator.* The Prewitt edge detection was proposed by J.M.S. Prewitt in 1970 [175]. The Prewitt operator is a discrete differentiation operator which is used to compute an approximation of the gradient of the image intensity function and is an appropriate way to estimate the magnitude and orientation of an edge. Although differential gradient edge detection needs a rather time consuming calculation to estimate the orientation from the magnitudes in the x- and y-directions the compass edge detection obtains the orientation directly from the kernel with the maximum response. This gradient based edge detector is estimated in the 3×3 neighborhood for eight directions. It consist of the following two kernels:

$$G_x \approx \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \; G \approx \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

In Figure 4.3, the results for the application of the three described filters to an exemplary image is displayed. From our results of the filter comparison, we concluded that the Sobel edge detection operator produces the best results for the given task and was used to compute the external energy, whose discrete version is restored below:

$$E_{external} = \kappa \sum_{i=1}^{N} f(X_i, Y_i) \tag{4.23}$$

where N are the snake points and f represents the Gaussian smoothing and the gradient computation.

As explained in Subsection2.3.3, the ROI in n DCE-MRI measurements changes its intensity due to the CA perfusion over time. This implies that the application of the gradient operator does not always provide crisp and sharp images. To compensate for this effect a clustering technique was applied when necessary. By using a fuzzy c-means clustering algorithm, previously described in Subsection 3.1.5, a cleaner image and a sharper contrast is achieved that allows an easier tracking of the reference's edges by the snake.

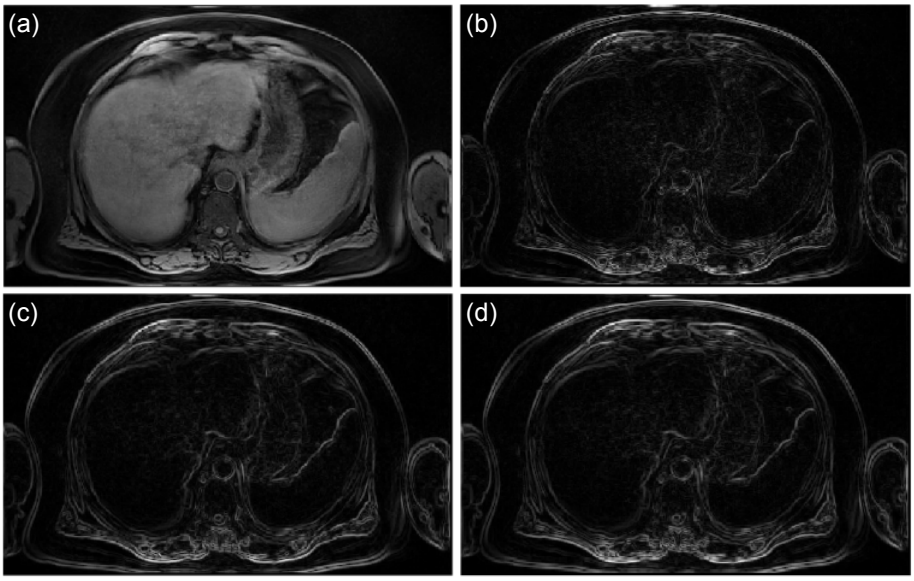

Figure 4.3: Comparison of edge detection filters. (a) Original image; (b) Roberts filter; (c) Sobel filter; (d) Pewitt filter.

4.4.2 Chamfer distance transform and modification of the energy functional

First results with traditional snake formulations showed that the gradient-based edge potential force field f has a limited capture range, because the magnitude of the external forces decreases quite rapidly away from the object boundary. The location of the initial contour must be set close to the object boundary, otherwise, it will potentially evolve to the local minima due to the limited capture range. Increasing the standard deviation σ of the Gaussian in Eq. 4.13 can enlarge this range, but the localization of boundaries will become less precise and distinct.

Among the methods proposed to address this problem is the concept of distance potential forces (DPFs). The basic idea is to increase the capture range of the external force fields and to guide the contour toward the desired boundary. L.D. Cohen proposed an external force model that dramatically increases the range of capture of a traditional snake [171]. These external forces are the negative gradient of a potential function that is calculated using an Euclidean distance or an approximation of this distance . This DPF is an additional contribution to the external force, making it converge toward the edge of interest, whether the region is homogeneous or not. We customized the DPF (normally computed only on binary images) by considering the mean intensity of pixels of the ROI and added its contribution to the external force. The value of the distance map at each pixel is obtained by calculating the distance between the pixel and the closest boundary point. This calculation can be performed by used the Euclidean distance transform. However, this transform is a computationally expensive operation, therefore we applied the Chamfer distance transformation [176], which is a very good approximation of the Euclidean distance transformation without having its high computational cost. By defining the potential energy function based on the distance map, one can obtain a potential force field that has a large attraction range. We can

calculate the distance map $D(x, y)$in the following way:

$$D(x, y) = min_{(p,q)\in\{(a,b):I(a,b)=1\}} [d(x, y; p, q)] \tag{4.24}$$

where $d(x, y; p, q)$ is a distance metric between two locations (x, y) and (p, q) of the image $I(x, y)$.The result of the transform is a gray-scale image that looks similar to the input image, except that the intensities of points are changed to show the distance to the closest boundary from each point. To create the edge map used as input image for the computation of the distance transform, the original image was preprocessed with the fuzzy c-means clustering algorithm to create a crisper and sharper contrast. Subsequently, the edge map was calculated with the Canny edge detection algorithm [50], as it provided the best performance both visually and quantitatively based on the measures such as mean square distance, error edge map, and signal-to-noise ratio. A detailed description of the Canny edge detection method is given in Subsection 3.1.3.

After the computation of $D(x, y)$, the additional term for the external forces are given by :

$$F_{x(external)} = -\delta\frac{\partial D}{\partial x} \tag{4.25}$$

$$F_{y(external)} = -\delta\frac{\partial D}{\partial y} \tag{4.26}$$

where δ is a scaling factor (in analogy to κ for the classical external forces) that is used to reduce the convergence time of the algorithm. Including this new terms into the external forces leads to new coordinates of the contour that are more accurate than those obtained by simple using the contribution of the gradient as described in Equation 4.23.

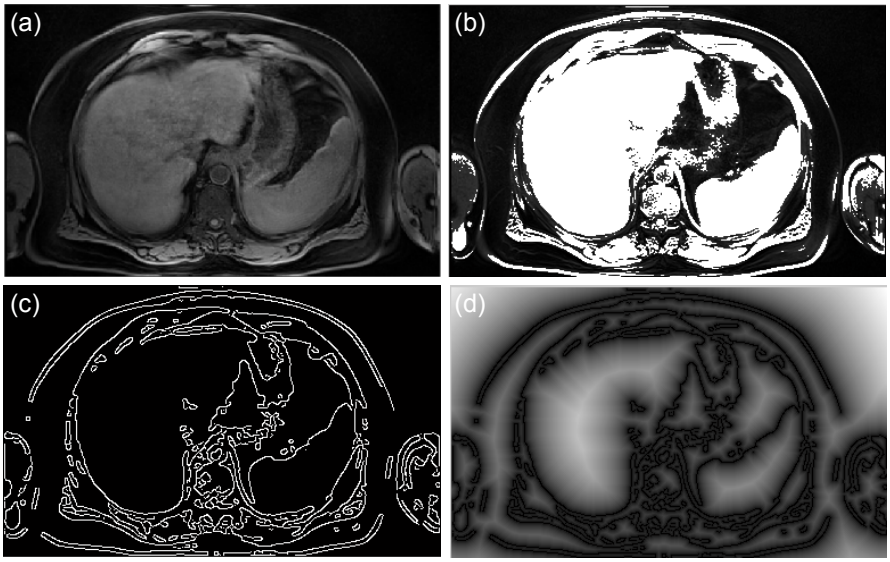

Figure 4.4: Procedure for calculation of the distance map. (a) Original image. (b) Clustered image generated with a fuzzy C-means clustering algorithm. (b) Edge map calculated from the clustered image with the Canny edge detection algorithm. (c) Distance map computed from the edge map by using the Chamfer distance transformation.

4.4.3 Minimization of the energy functional

Once the distance map and magnitude of the gradient image are defined, we deform the initial contour by applying forces obtained by minimizing the energy functional as described above (recall that the external force is given by the sum of the force calculated on the gradient image and that calculated on the distance map).

The coordinates of the snake are updated for each iteration during the evolution of the algorithm. When the algorithm has reached an alleged equilibrium position, the following criteria are checked in order to ensure a correct convergence. It is verified that the distance between the current and the previous contour does not exceed a predefined limit. Similarly, the minimized energy terms of the previous and the current contour are compared.

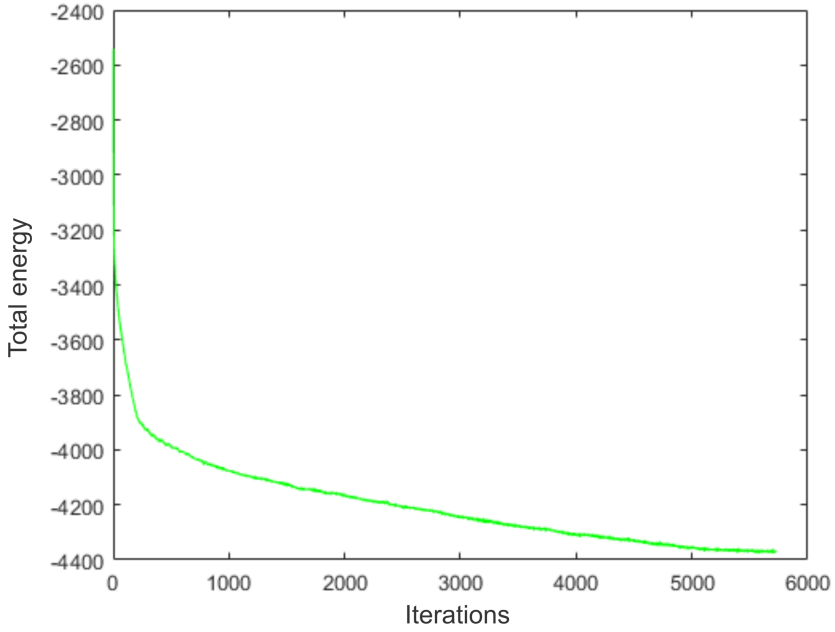

Figure 4.5: Energy function of the active contour model. The total energy E_{snake} is depicted dependent on the number of iterations of the algorithm.

As depicted in Figure 4.5 the energy function is constantly decreasing and its values are negative. The fact that the function decreases is obvious, since the algorithm evolves with the aim of minimizing the snake's energy content. The negative values are caused by the strong contribution of the external forces to the total energy, which is fundamental for the correct evolution of the active contour.

The algorithm is extended to the whole temporal sequence so that the snake is able, with the best possible approximation, to reproduce the contour of the ROI in every single image. In the transition from one slice to the next, the final contour obtained in the previous frame is used as a template in the next frame, i.e., the previous contour is used as the initial contour of the algorithm

in the next frame. This optimizes the edge tracking by the snake, since the contour that evolves from one slice to the next does not undergo any large variations as the tracked organs perform only small displacements or are subject to slight deformations. Consequently, this procedure allows a more accurate and efficient segmentation of the ROI.

4.5 Motion tracking

After using the active contour model to identify the contour of the ROI in each image, the next step is to compensate the displacement of the ROIs caused by motion artifacts. For this purpose, we use the assumption that the ROI can be represented by a rigid body and apply a rototranslational transformation[177, 178]. In order to perform a rototranslation, it is necessary to calculate the center of mass and the rotation angle (the principal axes of inertia), which requires the knowledge of the moments of the first and second order [179]. The moments are calculated by applying the following general formula:

$$m_{pq} = \sum_{y=1}^{M}\sum_{x=1}^{N} x^p y^q I(x,y) \tag{4.27}$$

where$I(x,y)$ is the considered image and N and M, respectively, are the number of rows and columns that compose the image matrix.

The first-order moments are used for the assessment of the center of mass coordinates $x_c$and y_c:

$$x_C = \frac{\sum_{y=1}^{M}\sum_{x=1}^{N} xI(x,y)}{\sum_{y=1}^{M}\sum_{x=1}^{N} I(x,y)} \tag{4.28}$$

$$y_C = \frac{\sum_{y=1}^{M}\sum_{x=1}^{N} yI(x,y)}{\sum_{y=1}^{M}\sum_{x=1}^{N} I(x,y)} \tag{4.29}$$

The second-order moments are used for the assessment of the principal axis of inertia. The orientation of the principal axes with respect to the first slice reference framework is given by

$$\varphi = 0.5 \arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right) \tag{4.30}$$

where φ is the angle between the reference axis and the principal axes of inertia and μ_{11}, μ_{20}, μ_{02} are the central moments given by:

$$\mu_{20} = \sum_{y=1}^{M}\sum_{x=1}^{N} (x - x_c)^2 I(x,y) \tag{4.31}$$

$$\mu_{02} = \sum_{y=1}^{M}\sum_{x=1}^{N} (y - y_c)^2 I(x,y) \tag{4.32}$$

$$\mu_{11} = \sum_{y=1}^{M}\sum_{x=1}^{N}(x - x_c)(y - y_c)I(x,y) \tag{4.33}$$

By computing, the coordinates of the center of mass and the rotation angles in each image, it is possible to reconstruct the movements associated with the ROI in terms of rototranslation (Figure 4.6). On the image on the left, we projected the point P_n belonging to the ROI in the image on the right in order to illustrate the erroneous calculation that would be induced in the absence of the motion compensation. Indeed, the effective corresponding point of P_n is the point $P_{n\to 0}$, which is obtained when applying the motion correction. Mathematically, the generic point of the coordinate (x_n^p, y_n^p) inside the ROI in the *nth* image corresponds to the point in the first image with the following coordinates:

$$P_{n\to 0} = \begin{pmatrix} x_{n\to 0}^p \\ y_{n\to 0}^p \end{pmatrix} = R_{T_n}^{T_0} P_n + \Delta_{T_n}^{T_0} \tag{4.34}$$

where $R_{T_n}^{T_0}$ is the rotation matrix, and $\Delta_{T_n}^{T_0}$ is the contribution in translation given by the difference between the x and y components of the centers of mass of the ROI in the *nth* image and first image respectively.

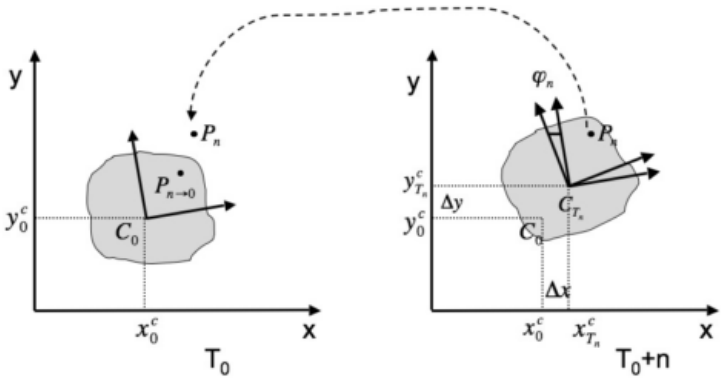

Figure 4.6: Center of mass displacement and rotation angles.

In order to reevaluate the EC curve with the compensation for the motion of the ROI included, Eq. 2.27 can be rewritten as:

$$I(T_n, x, y) = \left[i\left(T_n, R_{T_n}^{T_0}(x_\alpha, y_\alpha)\right) + \Delta_{T_n}^{T_0}(x_\alpha, y_\alpha)\right] - i_0(x, y) \tag{4.35}$$

where(x, y) is the center of mass location on the first image, i_0 is the baseline intensity of the ROI, (x_α, y_α) is the coordinates of the point in the *nth* image, whose projection on the first image is the point (x, y).

4.6 Compensation of out-of-plane motion

As mentioned before, the motion artifacts do not only result in misalignments between the frames of a time sequences, but also cause displacements in the transverse direction relative to the image plane, so- called out-of-plane motion. This induces inconsistencies with regard to the shape of the organ profiles within a temporal sequence that can also compromise a correct data analysis. In this section, we briefly present an approach, how this inconsistencies can be corrected. The presented out-of-plane motion compensation is based on the motion tracking algorithm that was described in detail in the previous sections of this chapter. We basically extended this algorithm to 4D by also taking into account the adjacent spatial planes that do not belong initially to the selected time series. An overview of the performed steps of the algorithm are shown in Figure 4.7.

We start with a particular temporal sequence with the plane position z_0 that is selected manually by the radiologist. The contour of the first frame is manually segmented and used as starting point for the AC algorithm. The final contour of the first frame found by the AC algorithm is then used as a template to find the best matching spatial plane in the next frame. For this purpose the images of the planes z_{-1},z_0,z_{+1} of the next frame are processed as follows. First a clustering is performed, followed by the creation of an edge map and finally the Chamfer transform is used to generate a distance map. The contour of the previous frame is then superimposed on the distance maps of the three planes and the root-mean-square (RMS) of the pixels covered by the contour is calculated for each plane. This value basically represents the average distance of the contour from the edges in the respective image and estimates how good the contour fits to the ROI in the respective plane. Consequently, the image with the lowest RMS is defined as the best match and is assigned to the initial time series. This procedure is repeated frame-by-frame until the new time series is completed. Afterwards, the previously described motion tracking algorithm is applied and the final result is time series that is corrected for both, the inter-frame and the out-of-plane motion artifacts.

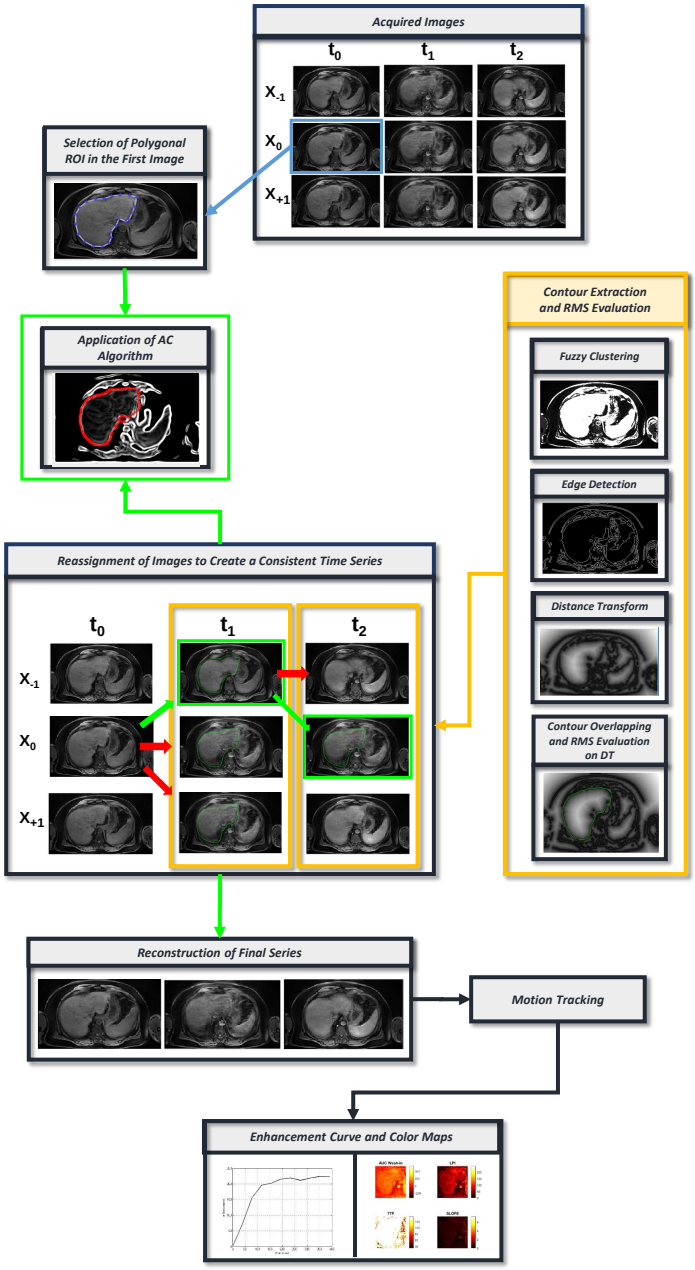

Figure 4.7: Flow chart of the out-of-plane motion compensation algorithm.

Chapter 5

Methodology and implementation of 3D analysis and visualization tool

Three-dimensional imaging methods are of crucial importance in medical diagnostics due to their evident advantage of providing more realistic anatomical and physiological information compared to 2D methods that cannot adequately represent the three-dimensional nature of the organs and structures in the human body.

The DCE-MRI procedure is usually performed on a temporal series of MRI slices, hence, normally, only 2D results are provided from this methodology, such as the Contrast Enhancement (CE) curve and the 2D color maps related to the DCE parameters. To overcome the limits of two-dimensional diagnosis, we extend the DCE analysis to the whole organ/lesion in order to gather 3D information of the entire ROI, namely, the Enhancement surface and 3D color maps of the characteristic parameters. The aim of this project is to provide the medical doctor with a fast, reliable and easy-to-handle tool that provides three-dimensional vision of the ROI and additional quantitative data, thereby allowing a deeper understanding of the underlying perfusion processes and a more accurate diagnosis, e.g., by highlighting the spatial areas with higher vascularization in the 3D volume of the organs of interest.

To generate the required 3D information from DCE-MRI data, we split our approach into two main steps, namely, the analysis and reconstruction of the 3D data. In the first step, we use different segmentation techniques to extract the 3D representation of the organs of interest, which are subsequently used for the reconstruction and visualization. Furthermore, we define a rectangular ROI in every image of the 4D data set and perform the perfusion analysis for each time series separately. In the second step, we perform volume and surface reconstructions to obtain the enhancement surface of the perfusion analysis and volume representations and 3D color maps of the examined organs.

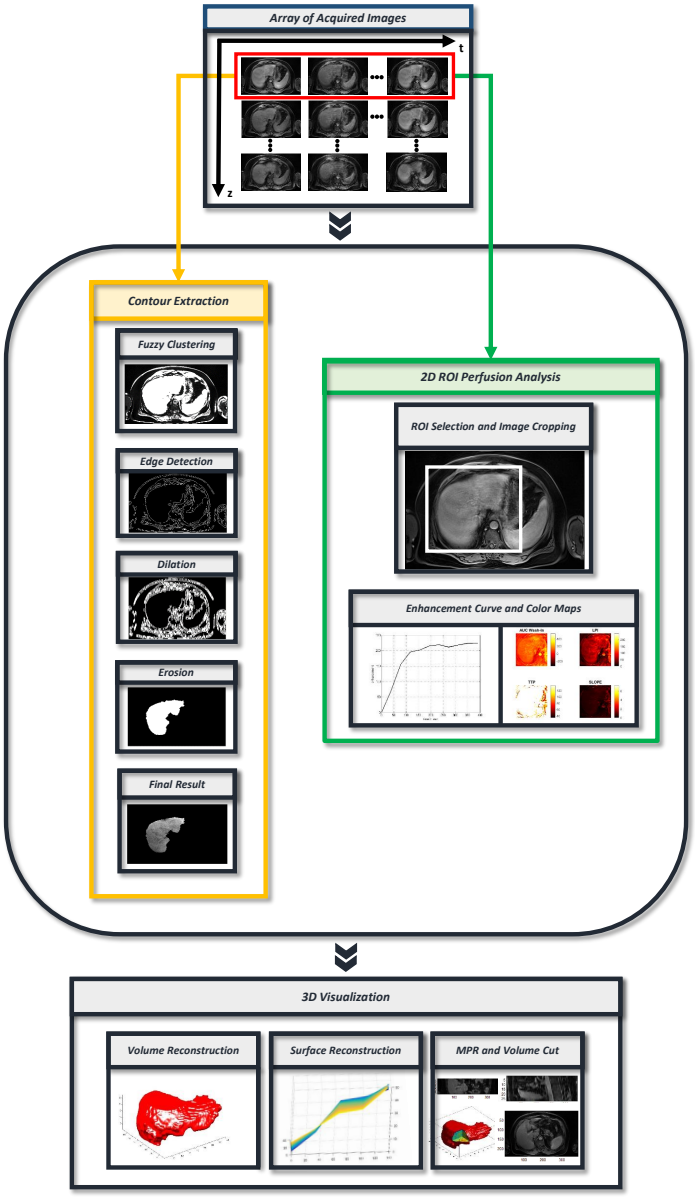

Figure 5.1: Flow chart of the 3D analysis and reconstruction tool.

The algorithm can be divided into four processing blocks:

1. Spatial and temporal arrangement of the images;
2. Segmentation and perfusion analysis;
3. Three-dimensional reconstruction of the volume of interest (VOI);
4. Reconstruction of orthogonal planes and VOI cutting.

5.1 Spatial and temporal arrangement of the images

The set of images to be processed consists of multi-planar temporal sequences of DCE-MRI data stored as individual 2D slices in the DICOM format. Each 2D slice can be assigned to a plane coordinate z and a time coordinate t. In order to facilitate and accelerate the data access of our algorithm, the images are loaded and arranged into a matrix structure. We create a data array of size $n \times m$, where each entry is a single 2D slice, each column represents a spatially growing order of n spatial planes (z_i with $i = 1...n$) at a given time point t_j, and each row represents a temporally growing order for m time points (t_j with $j = 1...m$) for a specific spatial plane z_i. So an image $I(z_i, t_j)$ is a 2D image form the spatial plane with the position z_i acquired at the time t_j. This is illustrated in the exemplary matrix shown in Table 5.1. The correct arrangement of the images in the array is crucial for the accurate analysis and reconstruction in the following steps. To guarantee this, we access the meta data of the DICOM files that contain additional information about the measurement and patient, such as the patient ID or the spatial location of the 2D slice within the data set. We readout the spatial position and the time stamp of each 2D slice and use this information to sort the images into the array.

$I(z_1, t_1)$	$I(z_1, t_2)$	...	$I(z_1, t_m)$
$I(z_2, t_1)$	$I(z_2, t_2)$	...	$I(z_2, t_m)$
...	...	...	...
$I(z_n, t_1)$	$I(z_n, t_2)$	...	$I(z_n, t_m)$

Table 5.1: Matrix for image sorting of 4D DCE-MRI sequences. Each column represents a spatially growing order of n spatial planes at a given time point t_j.Each row represents a temporally growing order for m time points for a specific spatial plane z_i.

5.2 Segmentation and perfusion analysis

The processing phase preceding the three-dimensional reconstruction can be subdivided into segmentation of the organs of interest and the perfusion analysis in a cropped region of the images.

Both procedures begin with a rough manual selection of the pathological area of interest in the first image of the first time series by the user in order to accelerate the further processing steps.

5.2.1 Segmentation

The basic idea for the segmentation step was to develop a fast algorithm that is able to recognize a region of interest in each image of the 4D data without further user intervention and extract the contour of the ROI as a binary template from the rest of the image. Subsequently, this template can be applied like a mold to the original gray value images to retrieve the desired structure of interest that can be used for further analysis or reconstruction (see contour extraction in Figure 5.1). The algorithm works in two subsequent steps:

1. Initial segmentation by an unsupervised fuzzy c-means clustering algorithm to perform a first cleaning and clustering of the image.

2. An edge-based segmentation algorithm that applies morphological operations to retrieve the final contour.

The fuzzy c-means algorithm (FCM), previously described in Subsection 3.1.5, is implemented in MATLAB and classifies the image by grouping similar data points in the feature space into a predefined number of clusters. To achieve the purpose of fast presegmenting and image cleaning, the applied FCM algorithm clusters the input image in two regions. The output of the algorithm comprises the coordinates of the cluster centroids and an array U that specifies for each pixel coordinate the degree of membership to a cluster. In the next step, we convert the we converted the fuzzy values associated with the degree of membership in U to "hard" integer values,i.e, by applying a threshold, we substitute all values below the threshold by 0 and all values equal or above the threshold by 1, thereby creating a binary representation of the clusters. This information is then used to extract the gray values of the ROI from the original image and is stored for every image of the data set.

Subsequently, the extracted ROI in gray-value representation is used as input for the edge-based segmentation. At first, an intensity check of the different clusters is performed to determine the best cluster for the creation of the edge map. Furthermore, the algorithm asks the user to denominate the organ to be segmented in order to proceed with an optimized segmentation procedure for the respective organ. In particular, a Canny edge operator, previously described in Subsection 3.1.3 is used for the edge map generation of the kidney, whereas a LoG operator is applied for the liver. The LoG operator is a combination of a Gaussian smoothing filter and the Laplacian operator that calculates the second derivative [180].

The created edge maps are further processed by morphological operations, which were previously described Subsection 3.1.6. Generally speaking, we apply a series of morphological operations on the edge maps, such as dilation, erosion, or hole-filling operations. For some of the operations the images are temporarily converted to binary images. The applied kernels and operations also differ for the respective organ, for instance, the liver is dilated with the "periodic line" kernel, whereas the kidney is dilated with the "line" kernel. An overview of the most important applied structure elements is shown in Figure 5.2. The goal of the applied morphological operations is to retrieve a fast, continuous and clean segmentation of the organs of interest without including any surrounding features or areas. The segmented structures are stored as binary templates and are consecutively applied as a mold to the original grayvalue image or the color maps.

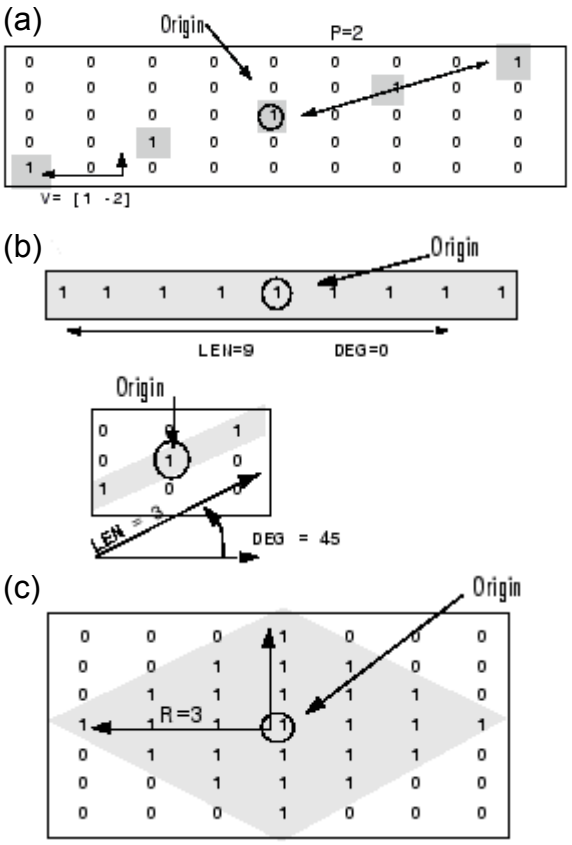

Figure 5.2: Different structure elements used for morphological operations. (a) Periodic line kernel. (b) Line kernel. (c) Diamond kernel.

5.2.2 Perfusion analysis

The standard 2D perfusion analysis model used for DCE-MRI data is described in detailed in Section 2.3. Its main purpose is to calculate an enhancement curve that depicts the dynamics of the contrast agent perfusion in the selected ROI. Furthermore, it is possible to retrieve four local parameters (AUC, TTP, LPI, SLOPE) from the EC which are typically displayed as four color-coded maps.

During the development phase of our algorithm , we performed the perfusion analysis for the segmented organs of interest and for a rectangular cropped region in the original images. We found that the latter procedure provided more reliable results than the analysis of the segmented data, consequently, the results from the rectangular ROI analysis are used in the following for the visualization. We analyze all time series for a ROI that preserves the 3D structure of the results and save the generated curves and associated parameters for all spatial planes. By combining the perfusion results and the templates of the segmented organs, it is now possible to create 3D color maps of the characteristic parameters, as depicted for the liver and the AUC parameter in Figure 5.3

Extending the curve-fitting approach for the EC to the perfusion analysis of all spatial planes

allows for the creation of an enhancement surface, as illustrated in Figure 5.4. By evaluating the surface, it is possible to gain additional volumetric information about the perfusion process. Furthermore, new characteristic parameters can be introduced, such as the volume under the surface, and existing parameter can be redefined, e.g., the time to peak can be defined as the time to reach the maximum intensity in the volume of interest (VOI).

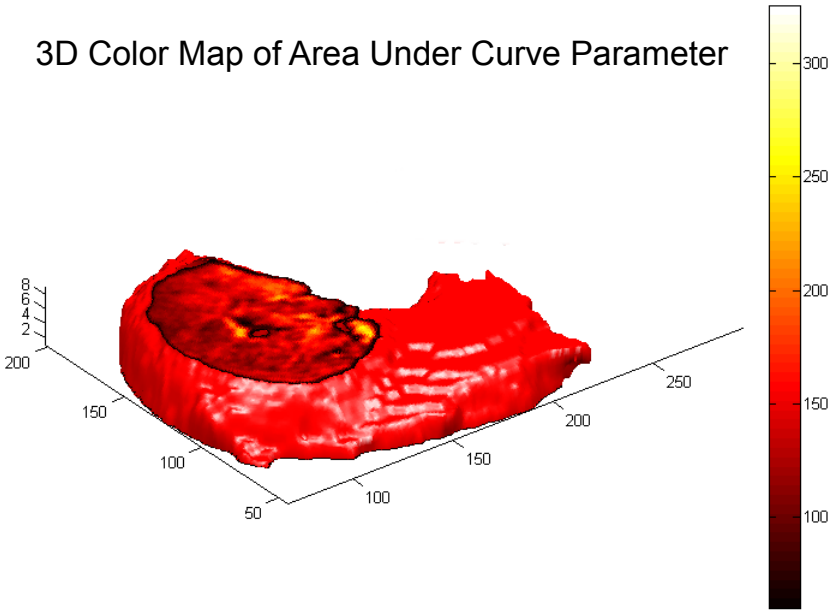

Figure 5.3: 3D color map of liver. The area under curve parameter was retrieved from the perfusion analysis of the 4D DCE-MRI data of the liver and is visualized by color coding.

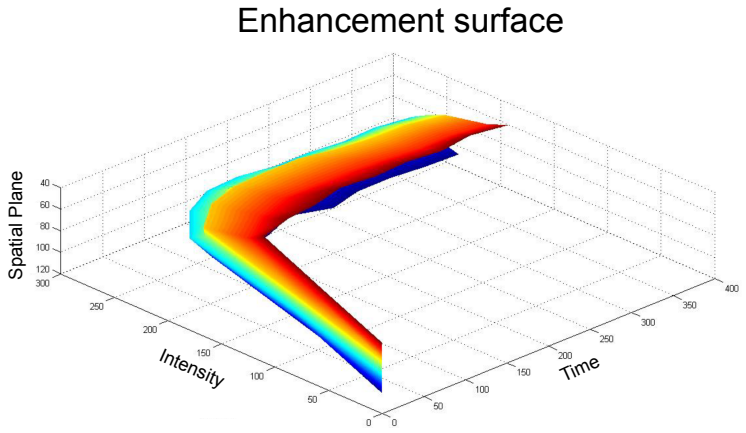

Figure 5.4: Enhancement surface. The depicted surface contains the information of the enhancement curves of all spatial planes and can be used to extract additional volumetric information.

5.3 3D reconstruction

In this section, we briefly describe how the organs and color maps are reconstructed from the array containing the segmented images. To enable fast 3D reconstruction and visualization, we use the implemented Graphics functions in MATLAB .

As mentioned before, by applying the binary segmentation templates of all spatial planes to the corresponding images representing either the intensity distribution or a characteristic parameter of the perfusion analysis, we can create 3D arrays containing the volumetric information in the respective representation. The reconstruction module takes these 3D arrays as input and generates a surface rendering of the array content.

Surface rendering techniques usually require the extraction of contours (edges) and faces that define the surface of the structure to be visualized. Subsequently, the surface is represented by a mosaic of connected polygons. In medical applications, isosurfaces are mostly used to visualize three-dimensional volume data gathered from medical-imaging devices, e.g. the visualization of internal organs, bones, or other structures from volumetric data generated with MRI or CT. In our case, we applied the built-in "isosurface" function of MATLAB, since it requires low computation power and can provide fast visualization results. The output structure of the function contains faces, edges and vertices of the isosurface which are then patched together with the MATLAB "patch" function to create filled polygons composing the visualized object. Furthermore, isocaps can be added to the rendering to provide a visual context for the isosurface. Isocaps show a cross-sectional view of the interior of the isosurface for which the isocap provides an end cap. The user can adjust the isovalue for the isosurface and find the appropriate isovalue from the histogram. Additionally, the isocaps properties can be adjusted to provide the optimal visualization.

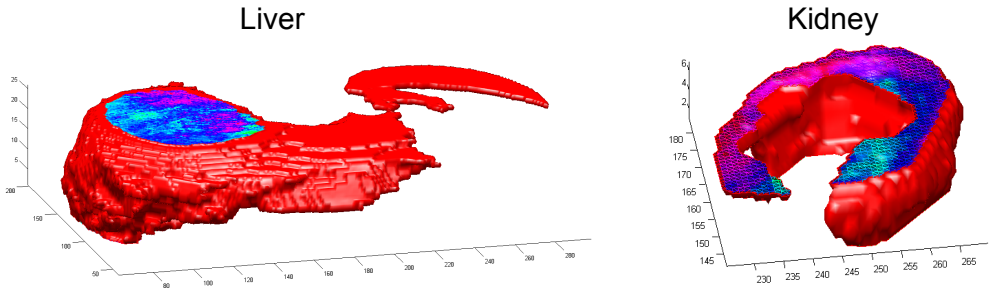

Figure 5.5: Liver and kidney isosurface rendering. The renderings were generated with the MATLAB isosurface function and show the intensity distributions of the organs.

5.4 Multiplanar reconstruction

A main benefit of displaying a anatomical structure as a 3D representation is that we can specify how 'deep' to search beneath the surface for a better visualization of the inner structure of the organ/lesion. To show the desired internal area of the VOI, it is necessary to perform a precise

cutting step that is defined by cutting coordinates. In the case of straight cutting along the coordinate axis of the 3D volume, the control points for the cutting coordinates can be retrieved from the view of the anatomical planes through multiplanar reconstruction (MPR). Multiplanar reconstruction provides efficient computation of images and involves generating images at right angles relative to a stack of axial slices so that the orthogonal perspectives can be viewed as well. Anatomical planes are planes used to transect the human body, in order to describe the location of structures or the direction of movements . The three most commonly used planes are described in the following:

- *Transversal plane.* This plane, also known as axial or horizontal plane, is an xy-plane, parallel to the ground, which (in humans) separates , the head from the feet or in other words, the superior from the inferior. In MRI images this is type (see Figure 5.6). In MRI imaging, this is usually the plane orientation in which the images are generated, as explained in Subsection 2.2.2.

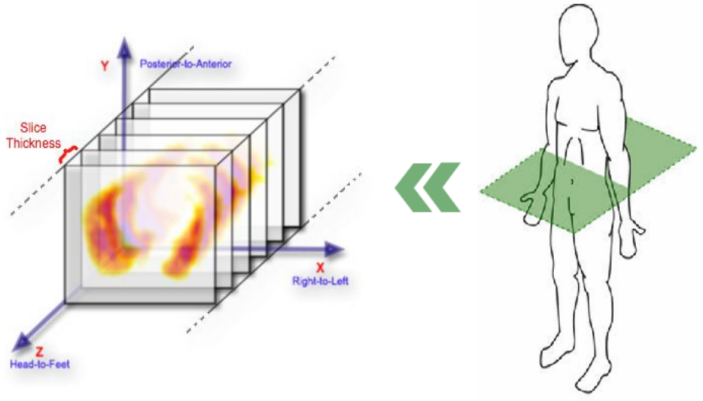

Figure 5.6: Transversal plane. An xy-plane that separates the head from the feet.

- *Coronal Plane.* It is an xz-plane, perpendicular to the ground, which (in humans) separates the anterior from the posterior, the front from the back, the ventral from the dorsal. Here, the image stack is rotated so that the z-axis becomes the vertical axis and a stack of images is reconstructed using parallel planes of voxels arranged from the patient's anterior to posterior surface, as illustrated in Figure 5.7.

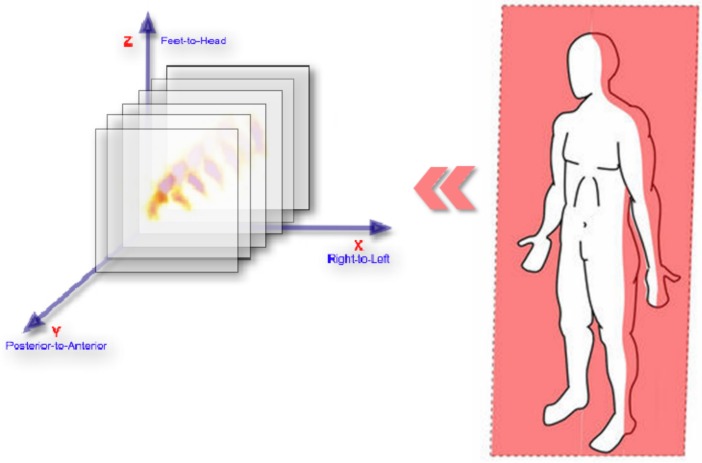

Figure 5.7: Coronal plane. An xz-plane that separates the anterior from the posterior.

- *Sagittal Plane.* It is an yz-plane, perpendicular to the ground, which separates left from right. Sagittal reconstructions are possible through additional rotations of the image stack so that a patient's left-to-right slice sequence (see Figure 5.8).

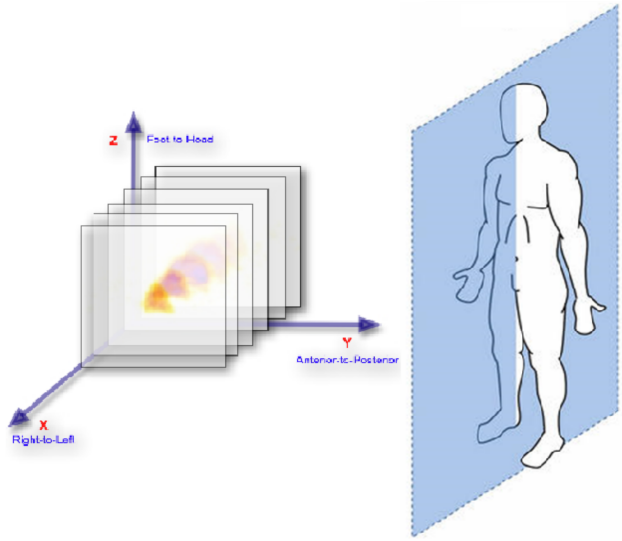

Figure 5.8: Sagittal plane. An yz-plane that separates left from right.

We propose a simple and fast approach for the cutting procedure of volumetric structures in 3 steps:

1. Based on a sequence of images, we reconstruct a three-dimensional model and enable the simultaneous visualization of axial, coronal and sagittal planes in a completely interactive way.

2. We use the position of the mouse to project cross-lines pointers on all slices. The Cross lines on each image represent lines of intersection of the respective plane with the two other orthogonal planes. Moving one of the slices, the others will follow to make sure that the center of each of the slices intersect in the same point in space. By left dragging the mouse on any plane the user can explore the volume.

3. The selected coordinates will then be used as a reference for the cutting of the three-dimensional structures.

This cutting approach can be used for 3D color maps of the perfusion parameters as well as for surface renderings of the intensity distribution.

5.5 Graphical user interface

We implemented a dedicated graphical user interface (GUI) for our program to facilitate the execution and interaction processes by the user. The idea was to design a GUI as intuitive and simple as possible, so that medical doctors, who may have not a strong background in informatics and programming, can use the program without further instructions. Here, we briefly introduce the main features of the GUI without going too much into detail.

The main window of the GUI consists of three panels (see Figure 5.9):

- The upper left panel shows the surface rendering of the entire selected data of the volume of interest in order to give the user an overview of the loaded data and to facilitate further operations

- The lower left panel displays the surface rendering of the cropped volume of interest. The selection of the coordinates for the volume cutting is explained in Section 5.4.

- On the right, the control panel for the MPR visualization is displayed, which simultaneously shows the transversal, coronal and sagittal planes. Moreover, this panel includes fields for the numerical input of 3D coordinates and buttons for the selection of different visualization options.

By clicking on the start button, the program execution is started and the user is prompted to select the data to be imported. The user has the options to import an entire 4D data set or to only select a certain range of time series in order to accelerate the following computations. Initially, the intensity values of the segmented data are loaded and the surface renderings and MPR slices are displayed either in gray value or color representation. Using the set button we can change the reference point and, consequently, the cutout coordinates in case of depth analysis.

The View button loads the non-segmented data set without the DCE analysis. Consequently, the multi-plane view is extended to all planes of the imported sequences and the left panes show the rendering of the full volumetric data. Displaying full spatial data can be very helpful to provide the user with additional information from the volume surrounding the segmented structures. In some cases, this step is crucial to understand the results of the perfusion analysis and thus ensuring a correct diagnosis.

Clicking on the 3D Map button opens a new window dedicated exclusively to displaying colorimetric 3D maps of the characteristic parameters of the perfusion analysis (see Figure 5.10. The user can choose between the four parameters extracted from the enhancement curves and different color maps for the color coding. Furthermore, either the full volume of interest or the cropped form of the VOI can be displayed.

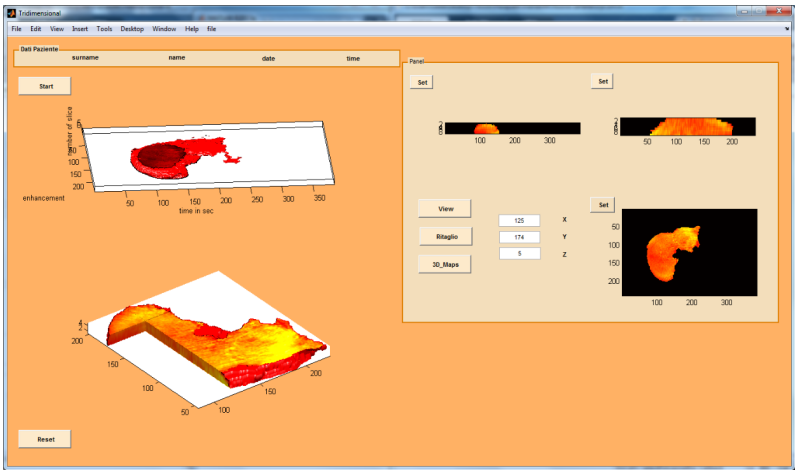

Figure 5.9: Main window of the GUI. The upper left panel shows the surface rendering of the entire selected data of the volume of interest. The lower left panel displays the surface rendering of the cropped volume of interest. The right panel shows the control panel for the MPR visualization.

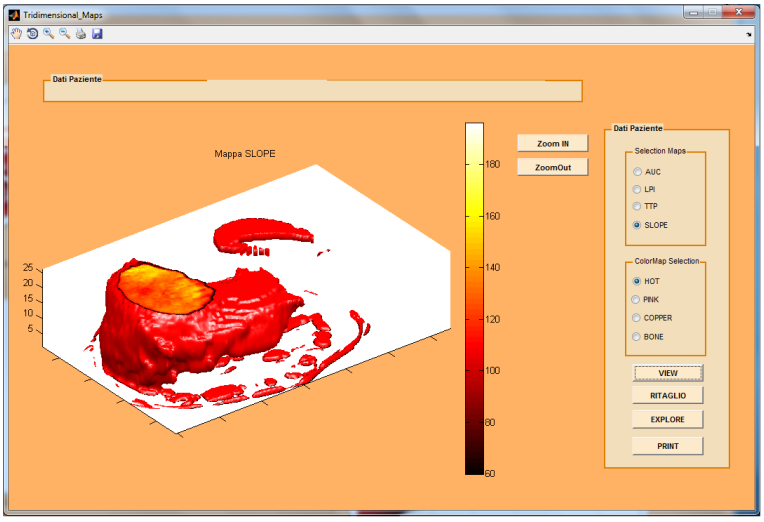

Figure 5.10: GUI window for 3D visualization of characteristic parameters. The user can choose the desired EC parameter and the used color map for visualization.

Chapter 6

Application of motion tracking algorithm to clinical DCE-MRI data

The motion tracking algorithm was developed with the aim to compensate for motion artifacts in DCE-MRI time sequences an to improve quality of postconstrastographic analysis. This a challenging task since it requires the tracking of soft-tissue structures that are not only affected by motion artifacts, but also by changes in intensity generated by the contrast agent (CA) perfusion.

To demonstrate the applicability and suitability of our technique, we applied the algorithm to four different organs from the abdominal region, namely, liver, spleen, kidney, and bowel wall. These organs were chosen, since they are particularly affected by undesired movements. In order to illustrate the effect of our method we performed the following two analyses on the DCE-MRI data of the four organs:

• Standard DCE perfusion analysis without any motion correction. We refer to this analysis as DCE in the following.

•Perfusion analysis after the application of the motion correction algorithm that is based on rototranslation and active contour segmentation. We refer to this analysis as AC in the following.

We generated two different types of results for both analysis methods, namely, the enhancement curves and the color maps of the characteristic parameters. Subsequently, we compared these results to quantify the effect of our algorithm. A major conceptional difference between the two types of results is that the enhancement curves are calculated from the mean intensity of the ROI, whereas the color maps are generated from a pixel-wise analysis. The anatomical and physiological meaning of the color maps is briefly described in the following:

- **AUC:** The area under curve map depicts the quantity of CA absorbed by each pixel over the entire time series and thereby provides information about the vascularization of the diseased tissue. A high brightness in a region of the color map depicts areas of greater CA concentration usually indicating increased local vascularization due to the presence of the lesion.

- **LPI:** The local peak intensity map shows the maximum intensity reached in each pixel estimating the maximal quantity of blood that is flowing through the pixel at a particular time

point of the temporal sequence. The brightness of the pixels is proportional to the quantity of blood.

- **TTP:** The time to peak map provides an assessment of the time period for each pixel to reach its peak intensity. Thus, it gives information about whether this pixel is directly supplied with blood or through revascularisation. Here, a higher brightness indicates a longer diffusion time.

- **SLOPE:** The slope map provides information about the average blood flow perfusion speed in the tissue. The brighter areas represent regions of faster perfusion indicating abnormal tissue vascularization.

6.1 Adaption and optimization of the algorithm for different organs

As explained in Chapter 4, the edge detection of the ROI is related to the minimization of the energy function E_{snake}, given by the weighed sum of the internal ($E_{internal}$) and external ($E_{external}$) energy components. The external energy requires the computation of the gradient of the image, convoluted with a Gaussian smoothing filter. The effect of the Gaussian controlled by the dimensions of corresponding mask and the value chosen for the standard deviationσ. The latter is the only parameter that modifies the transfer function, consequently, it defines how much the surrounding pixels affect the new intensity value to the central pixel. The size of the convolution mask depends mainly on the size of the image characteristics we want to highlight (or remove).

For the estimation of the optimal values of these parameters, several trials run were performed for each organ. A series of different combinations of the filter parameters were used as input for the algorithm and the optimal values were found by comparing the corresponding final results of the segmentation process. Representative images showing the application of different filter parameters, namely, the filter dimensions and the standard deviation of the Gaussian are presented in Figure 6.1 and Figure 6.2.

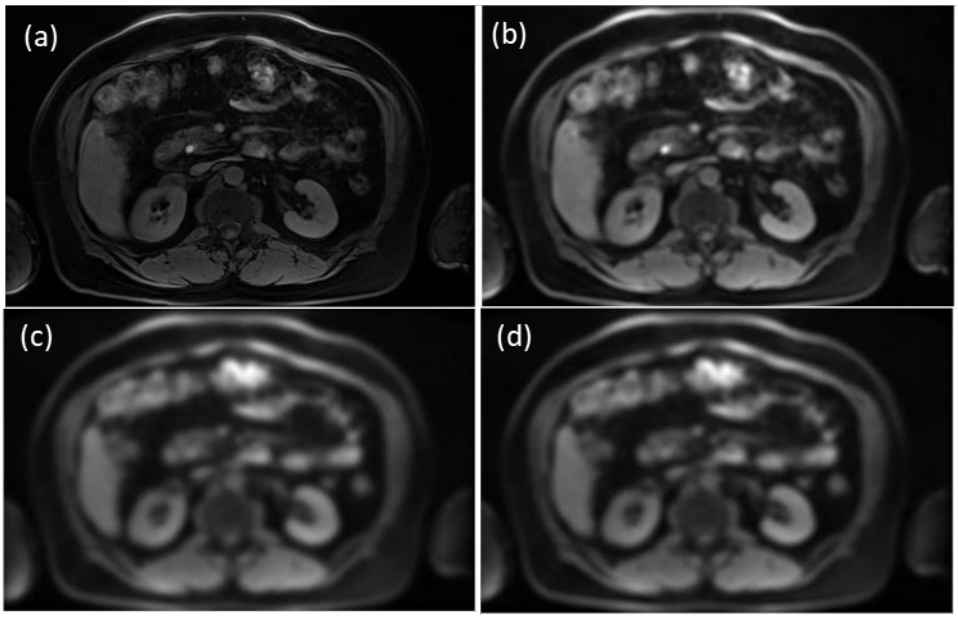

Figure 6.1: Comparison of different sizes of the Gaussian convolution mask for a standard deviation of 1 pixel. (a) Original image. (b) Filtered image with mask size of 5 x 5 pixels. (c) Filtered image with mask size of 10 x 10 pixels. (d) Filtered image with mask size of 15 x 15 pixels.

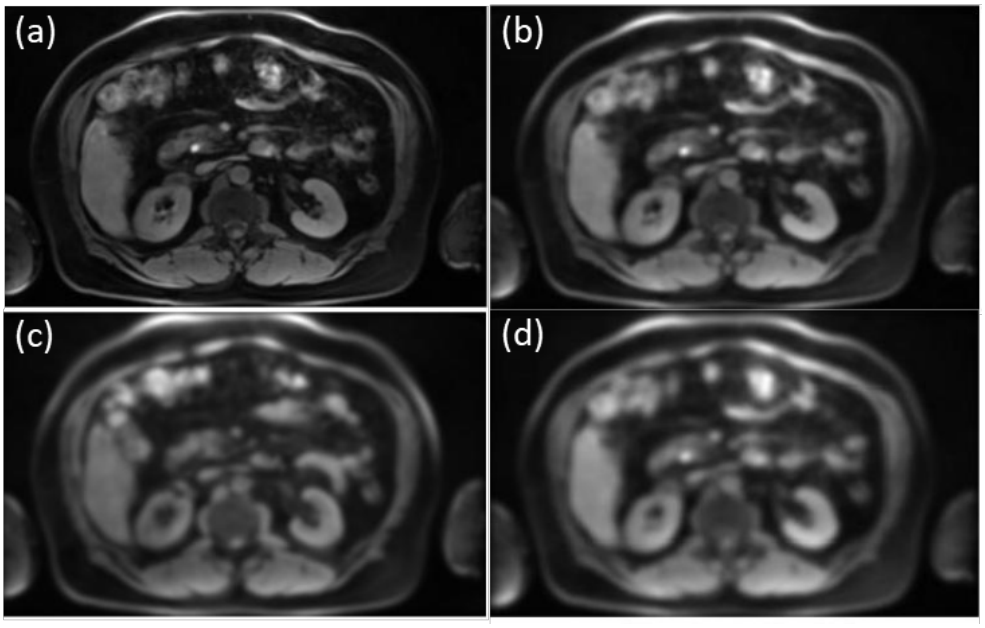

Figure 6.2: Comparison of different standard deviations of the Gaussian convolution mask with a size of 15 x 15 pixels. (a) Image filtered with a Gaussian with a standard deviation of σ=1 pixel. (b) Image filtered with a Gaussian with a standard deviation of σ=2 pixels. (c) Image filtered with a Gaussian with a standard deviation of σ=2.5 pixels. (d) Image filtered with a Gaussian with a standard deviation of σ=3 pixels.

For the kidney case, a 15 x 15 pixel mask proved to be the most effective solution in order to reduce irrelevant image details that can cause problems in the segmentation process. For the given filter size of 15x15 pixels, a standard deviation of 2.5 pixels provided the best results for the subsequent segmentation. The respective optimized values for the other organs are listed in Table 6.1:

	Kidney	***Bowell Wall***	***Liver***	***Spleen***
Filter Dimension	15x15	15x15	30x30	15x15
σ	2.5	3	3	2.5

Table 6.1: List of the optimized parameters describing the Gaussian filter for the different organs.

For an optimized application of the algorithm to the different organs, it is necessary to identify the ideal weighting factors for the different components of the energy functional for each organ separately. During extensive test runs of different combinations of the weighting factors, we identified the following problems related to the the contour evolution. In case of a too dominant contribution of the elastic energy term (weighting factor α) the contour may collapse, on the other hand, if the bending term (weighting factor β) is dominating, the snake is unable to contract and correctly track the edges of interest. Moreover, if the external energy term (weighting factor κ for the gradient term and δ for the distance map term) is given too much weight, the snake can neither contract nor reproduce the curvature of the object.

From the previous considerations, it is clear that it is of great importance to find a set of weighting factors that suitably combine the effects of the different energy terms. Therefore, several trial runs were performed for each organ and the the optimal sets of weighting factors were determined by comparing the final contours, as demonstrated for the kidney in Figure 6.3.

For the case of the kidney, the best result was achieved when using $\alpha = 0.3000$, $\beta = 0.0200$ and $\kappa = 0.0100$. The respective values for all organs are shown in Table 6.2:

	Kidney	***Bowell wall***	***Liver***	***Spleen***
α	0.3	0.3	0.3	0.3
β	0.02	0.02	0.03	0.03
κ	0.01	0.01	0.015	0.025
δ	0.05	0.05	0.05	0.05

Table 6.2: Weighting factors for the different terms that contribute to the total energy functional.

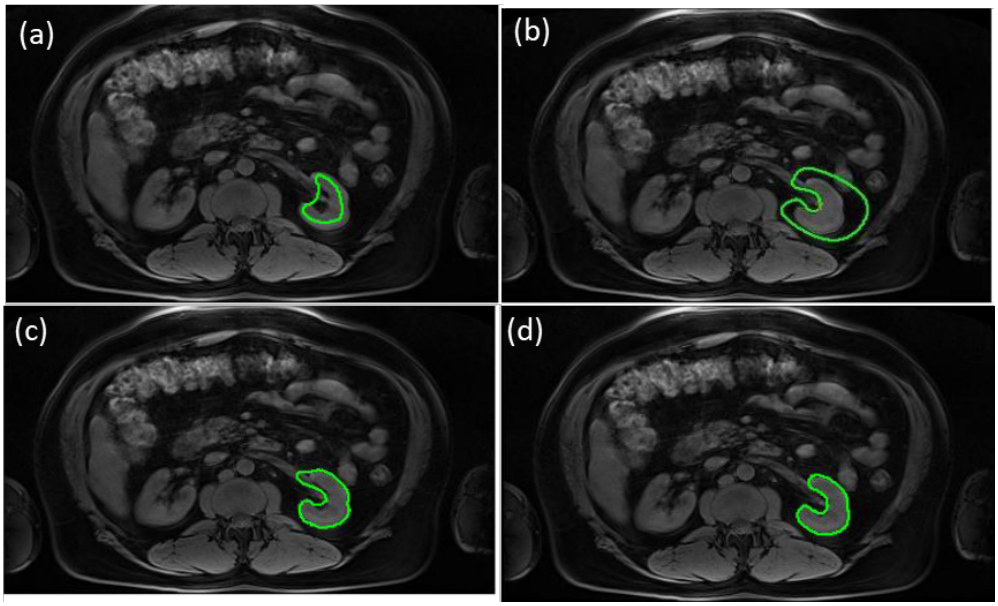

Figure 6.3: Comparison of different sets of weighting factors for the kidney segmentation. (a) Resulting contour for $\alpha = 0.1000$, $\beta = 0.0010$, $\kappa = 0.0001$. (b) Resulting contour for $\alpha = 0.0001$, $\beta = 0.1000$, $\kappa = 0.0001$. (c) Resulting contour for $\alpha = 0.5000$, $\beta = 0.0400$, $\kappa = 0.0200$. (d) Resulting contour for $\alpha = 0.3000$, $\beta = 0.0200$, $\kappa = 0.0100$.

In addition to carefully selecting the weighting factors, we also introduced two correction mechanisms to prevent the aforementioned problems and guarantee a reliable and precise contour extraction. The correction methods are described in the following:

- Contour collapse: This problem occurs in particular during the transition from one frame to the next. In this case, it is possible that the final contour of the last frame that is used as the initial contour in the next frame is too far away from the edges in the new image, which results in collapsing of the snake in the worst case. To tackle this problem, a check of the size of the ROI was implemented, since the ROI, although deformable, should not have large variations in size. For this purpose, the area of the contour of the previous frame is determined and the contour of the current frame is constrained to a size of at least 80% of the previous one. If it falls below this limit, the contour is reinitialized as the average between the contour initialized in the previous frame and the contour of the current frame in the iteration before it has collapsed. This results in a slightly expanded boundary allows for the continuation of contour evolution.

- Contour smoothing: Since the final contour of a frame is often rough-edge and jagged, we apply a resampling and interpolation step to smooth the contour before using it as initial contour of the next frame. First, the perimeter of the contour is calculated and divided into uniform intervals. Subsequently, the contour is resampled with the previously calculated sampling steps and it is interpolated using B- splines of the new control points.

In order to test the robustness of our algorithm, we analyzed numerous time sequences from different

spatial planes for each organ and compared the results for the DCE and AC method. In the following section, representative examples of these results are shown for each organ.

6.2 Case study: Bowel wall and kidney

We present the results of the bowel wall and the kidneys in a combined form, since both organs are relatively small and the applied algorithm is very similar. The value for the weighting factors are identical and the the parameters for the Gaussian filter only differ slightly in the σ value. Furthermore, the algorithm does not include the clustering step in both cases, since the problem of delineating the contour is less pronounced in small organs. If the initial contour is defined close to these organ, the snake will tend to converge, while maintaining a good level of stability throughout the segmentation of the entire sequence. This is demonstrated in Figure 6.4.

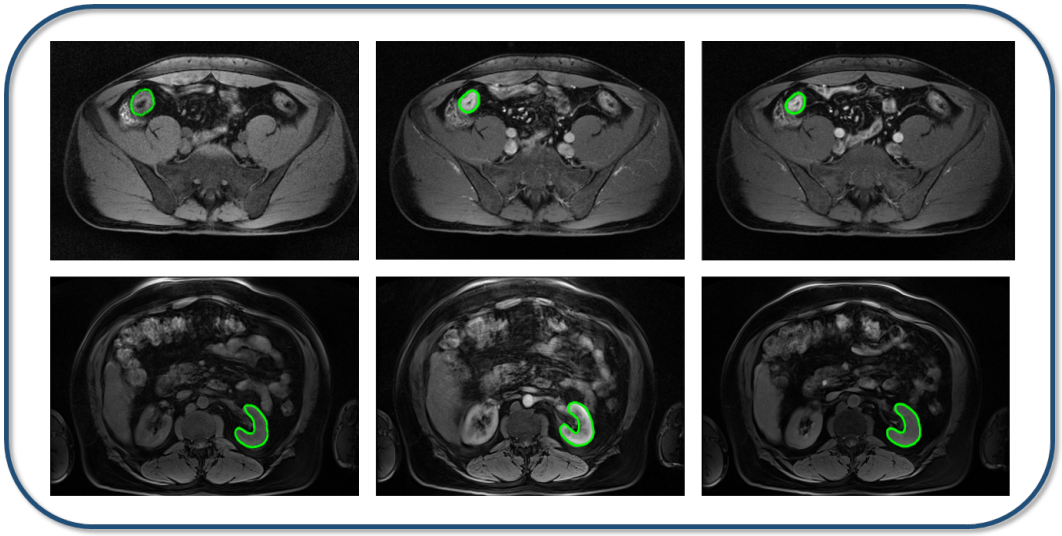

Figure 6.4: Contour evolution for kidney and bowel wall. The upper row shows the contour evolution of the bowel wall, the lower row the snake evolution of the kidney . The left column displays the first frame, the middle row depicts the central frame and the right row the last frame of the time sequence.

6.2.1 Bowel wall

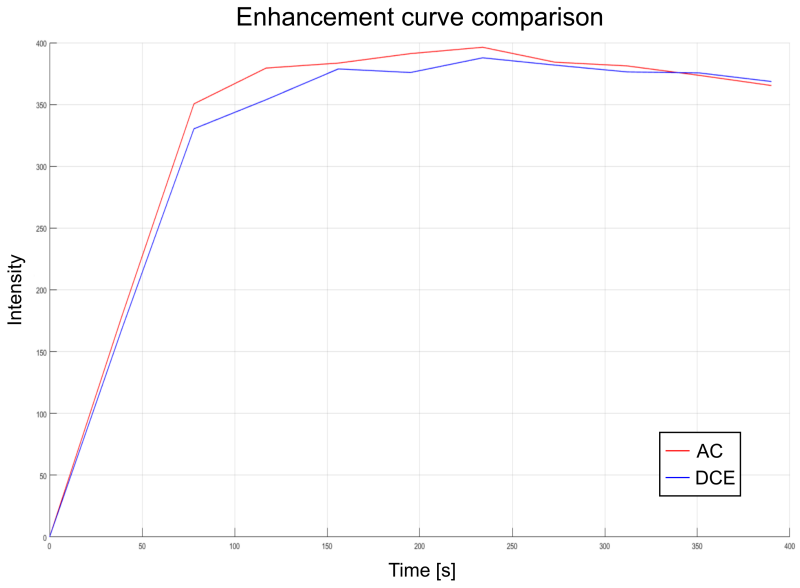

Figure 6.5: Enhancement curve comparison for the bowel wall.

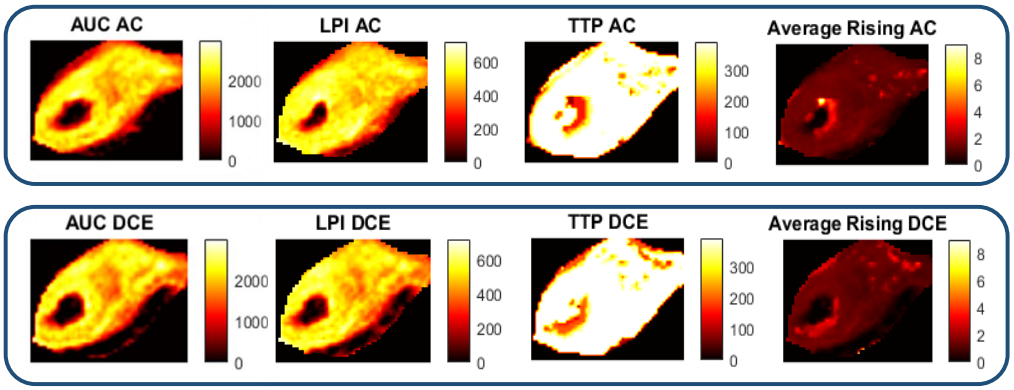

Figure 6.6: Color map comparison for the bowel wall.

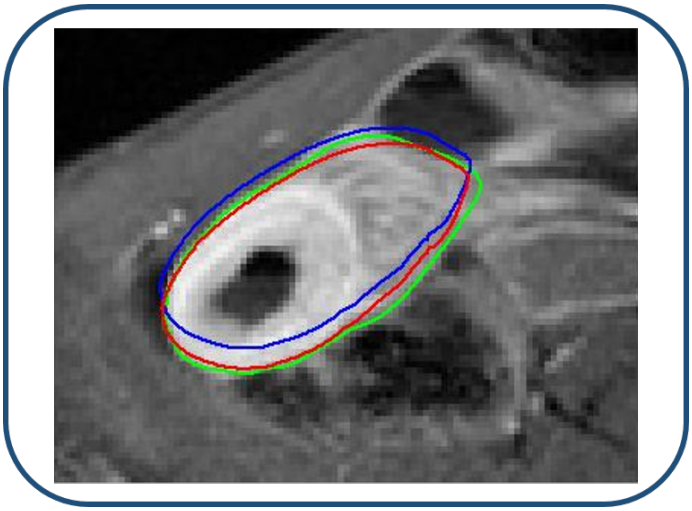

Figure 6.7: Contour comparison for the bowel wall. The green line is the final contour of the current frame. The red line is the final contour of the first frame projected on the current frame and the blue line is the rototranslated final contour of the first frame projected on the current frame.

#Img	1	2	3	4	5	6	7	8	9	10	11
y_b	118	117	117	117	118	118	119	119	119	119	119
x_b	84	83	84	82	82	81	82	82	82	82	82
φ	-1.1510	-1.0795	-1.1288	-1.0762	-1.0663	-1.0603	-1.0623	-1.0565	-1.0617	-1.0531	-1.0466

Figure 6.8: Center of mass coordinates (y_b and x_b) and rotation angle for the rototranslation of the bowel wall.

6.2.2 Kidney

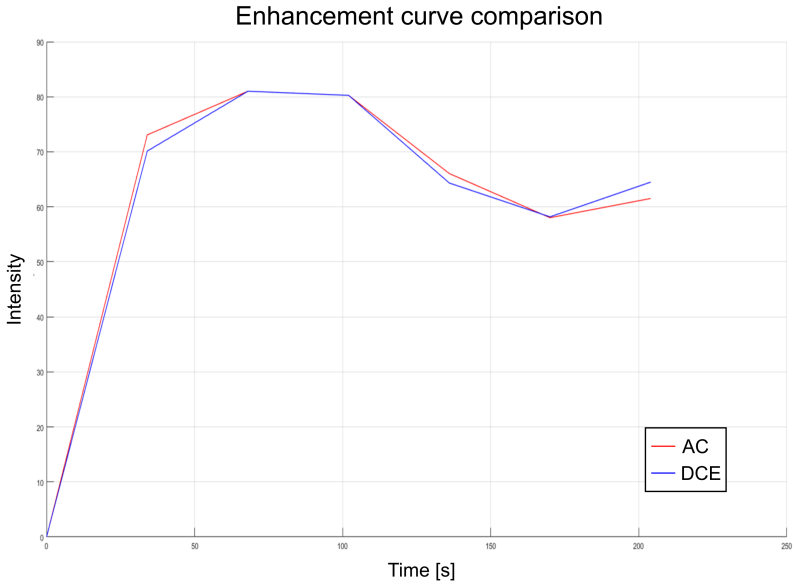

Figure 6.9: Enhancement curve comparison for the kidney.

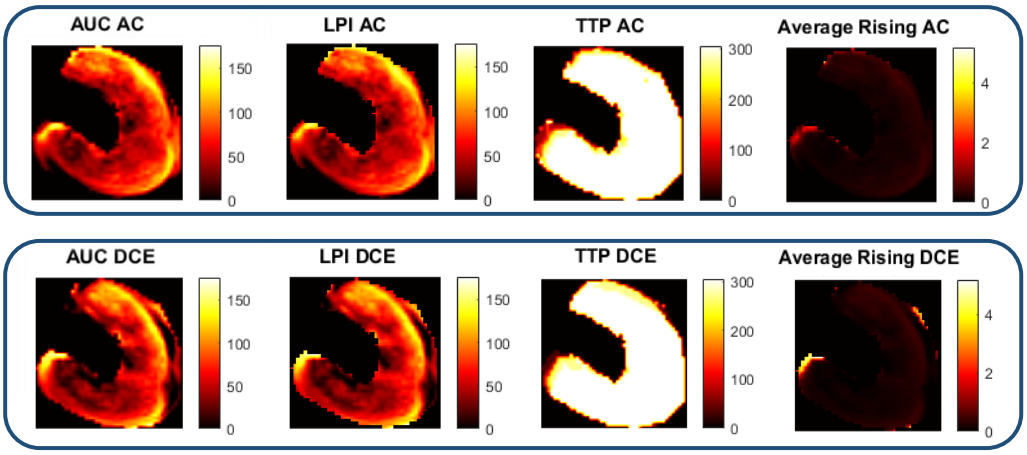

Figure 6.10: Color map comparison for the kidney.

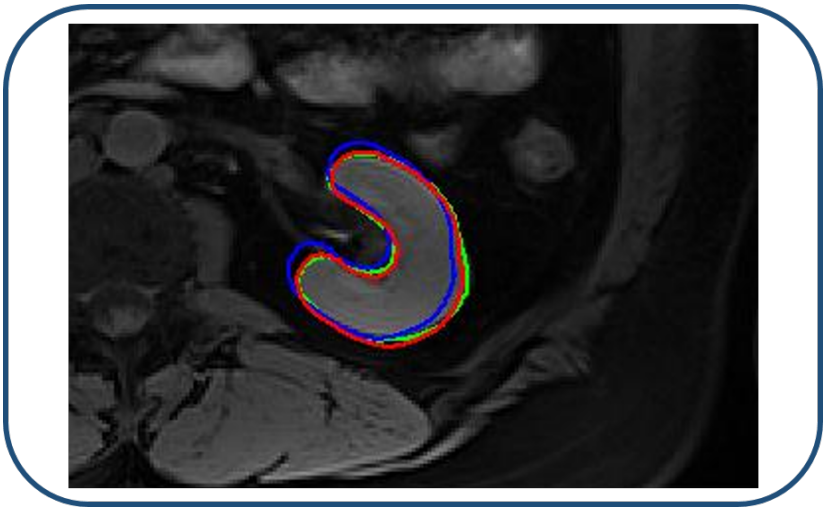

Figure 6.11: Contour comparison for the kidney. The green line is the final contour of the current frame. The red line is the final contour of the first frame projected on the current frame and the blue line is the rototranslated final contour of the first frame projected on the current frame.

#Img	1	2	3	4	5	6	7	8	9	10
y_b	165	162	163	163	163	165	165	165	165	165
x_b	253	254	253	254	253	254	254	254	254	255
φ	-0.0605	0.1668	-0.0412	-0.0363	0.0194	0.0505	-0.0048	-0.0690	-0.0749	-0.1470

Figure 6.12: Center of mass coordinates (y_b and x_b) and rotation angle for the rototranslation of the kidney.

6.3 Case study: Liver and spleen

In analogy to bowel wall and kidney, we also present the results of liver and spleen in a combined form, since both organs are relatively large and similarly processed. The values for the applied weighting factors are identical except for κ, while the size of the Gaussian filter is much larger for the liver due to the size of the organ. For both organs, we apply the clustering step, since this is necessary for the correct delineation of the contour. If the initial contour is defined close to to liver or spleen, the snake will also maintain a good level of stability throughout the segmentation of the entire sequence. This is demonstrated in Figure 6.13.

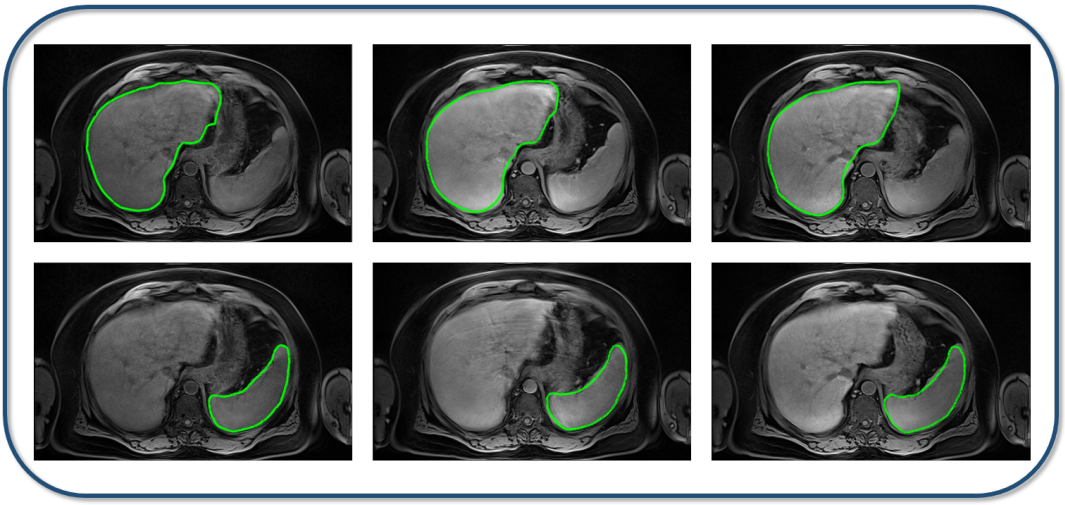

Figure 6.13: Contour evolution for liver and spleen. The upper row shows the contour evolution of the liver, the lower row the snake evolution of the spleen. The left column displays the first frame, the middle row depicts the central frame and the right row the last frame of the time sequence.

6.3.1 Liver

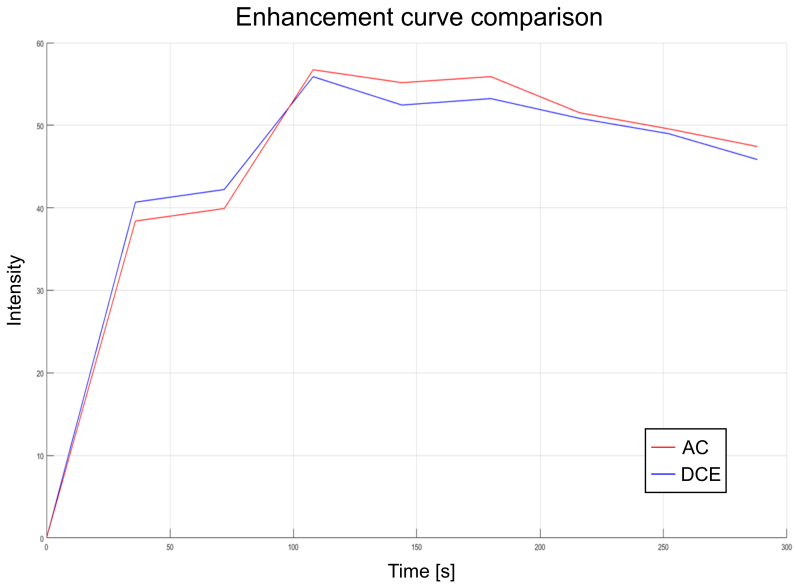

Figure 6.14: Enhancement curve comparison for the liver.

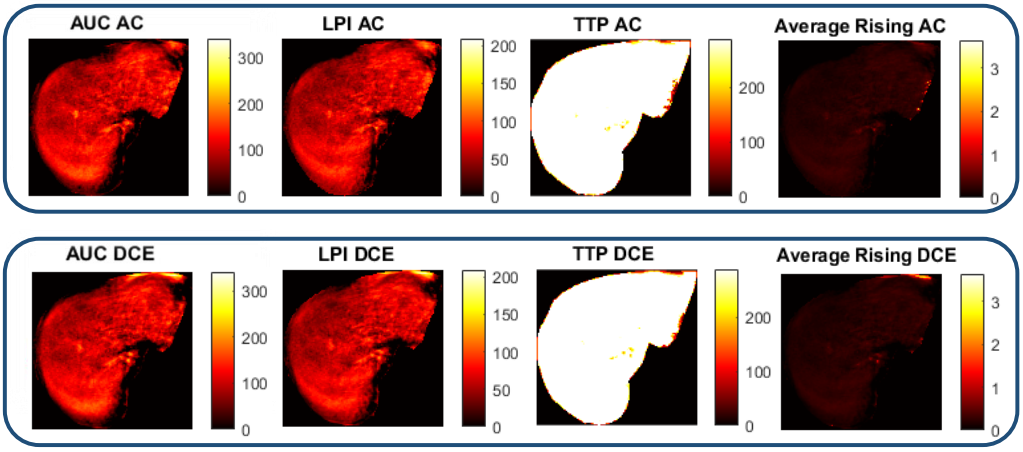

Figure 6.15: Color map comparison for the liver.

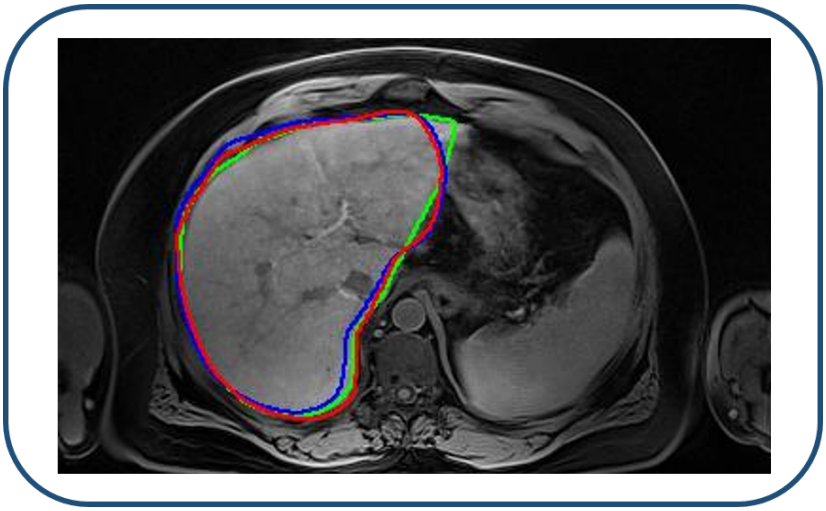

Figure 6.16: Contour comparison for the liver. The green line is the final contour of the current frame. The red line is the final contour of the first frame projected on the current frame and the blue line is the rototranslated final contour of the first frame projected on the current frame.

#Img	1	2	3	4	5	6	7	8	9
y_b	110	113	110	115	116	115	114	114	114
x_b	140	140	140	139	137	139	140	140	141
φ	-0.7449	-0.7667	-0.8044	-0.6829	-0.6118	-0.6592	-0.6616	-0.6634	-0.6675

Table 6.3: Center of mass coordinates (y_b and x_b) and rotation angle for the rototranslation of the liver.

6.3.2 Spleen

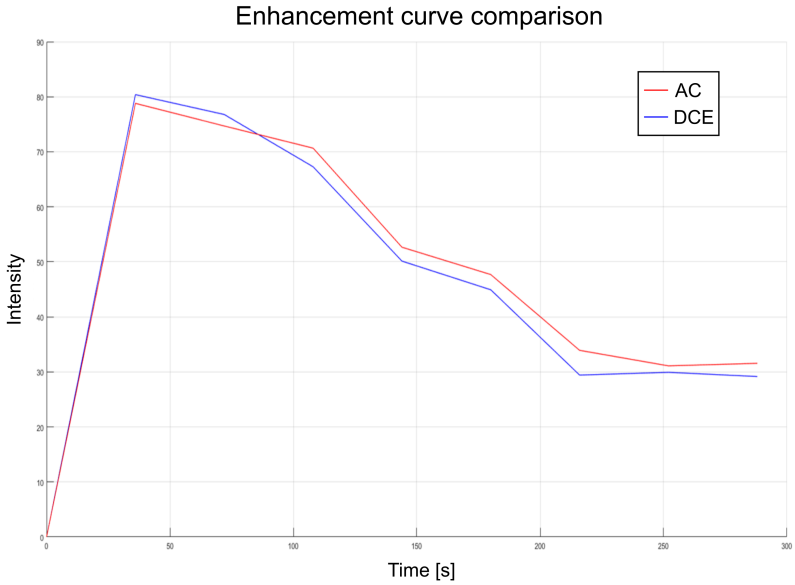

Figure 6.17: Enhancement curve comparison for the spleen.

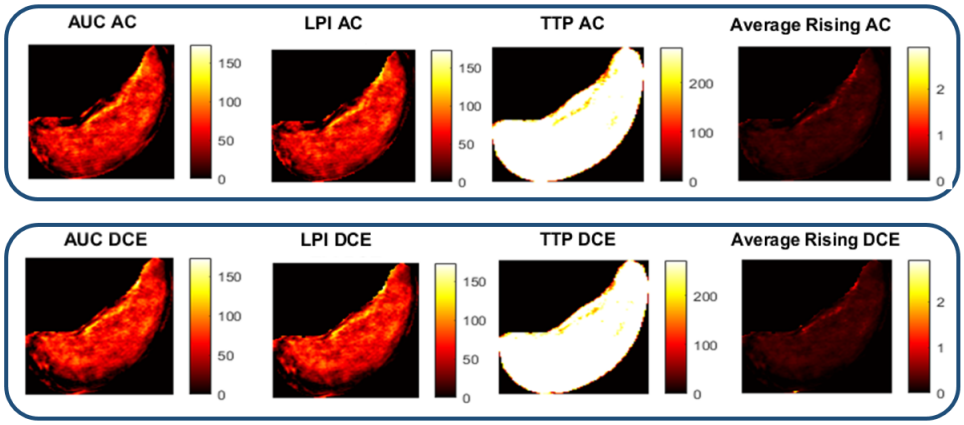

Figure 6.18: Color map comparison for the spleen.

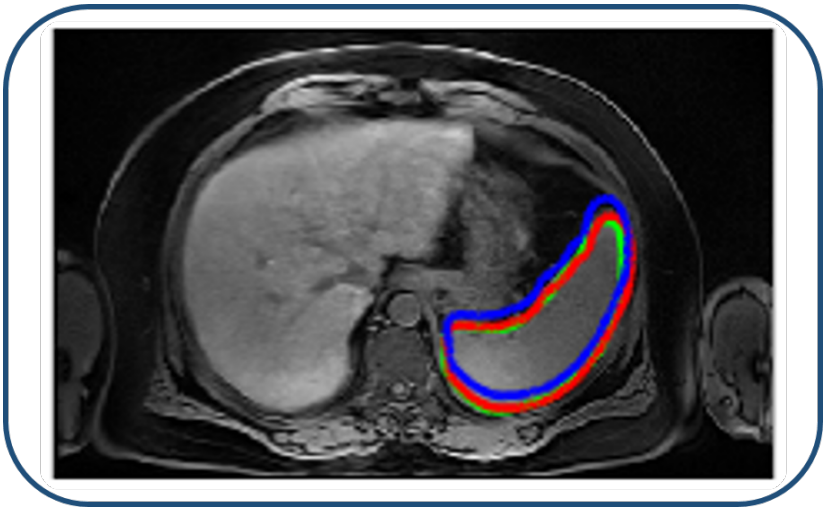

Figure 6.19: Contour comparison for the spleen. The green line is the final contour of the current frame. The red line is the final contour of the first frame projected on the current frame and the blue line is the rototranslated final contour of the first frame projected on the current frame.

#Img	1	2	3	4	5	6	7	8	9
y_b	167	165	166	165	165	167	166	166	167
x_b	262	258	258	259	259	257	258	257	257
φ	-0.8483	-0.8984	-0.9459	-0.8581	-0.8570	-0.8737	-0.8587	-0.8671	-0.8981

Table 6.4: Center of mass coordinates (y_b and x_b) and rotation angle for the rototranslation of the spleen.

6.4 Ground truth

A common problem in image processing is to objectively quantify the performance of a particular processing method. A widely used approach is the comparison of the processing result with a ground truth obtained from synthetic data or by manual human interaction. The quantification is usually obtained by using a supervised measure that computes a similarity/dissimilarity value between the ground truth and the processing result. For image segmentation methods, the ground truth is the perfect segmentation of the structure to which the segmentation method is applied. The most common method for ground-truth definition for the segmentation of natural images remains manual labeling by humans.

For this purpose, we manually generated ground-truth contour maps for an entire temporal sequence of the four organs. In the following, we compare them with the results of our active contour algorithm. In this context, it is important to mention that the manual definition of a

ground truth by humans is never completely objective and the results of such an analysis depend highly on the person who segmented the data.

To retrieve a quantitative index of the performance of our algorithm, we employ one of the most popular descriptors, Pratt's Figure of Merit (FoM) [181], which is a measure to express the closeness of a detected contour map to an ideal contour map. This distance measure ranges from 0 to 1, where 0 corresponds to a perfect segmentation. The Figure of Merit is defined as:

$$FoM = 1 - \frac{1}{max\left(|G_T|, |D_C|,\right)} \cdot \sum_{p \in D_C} \frac{1}{1 + k \cdot d^2_{G_T}(p)}$$

where:

- $d^2_{G_T}(p)$ represent the minimal euclidean distance between p and D_C;
- k is a scaling constant
- G_T is the reference contour map (corresponding to the ground truth) and D_C is the detected contour map of an original image.

Since we apply the FoM descriptor to all iterations of our algorithm, D_C corresponds in our case to the contour of each iteration in each image.

The second applied descriptor is a root-mean-squared (RMS) distance-based error measure, based on Chamfer matching [182]. The idea of Chamfer matching is to apply the Chamfer distance transform to the reference contour map G_T to generate a distance map, where the pixel values give the shortest distance to the contour. Subsequently, the detected contour D_C is superimposed on the distance map. Therefore, the pixel values of the distance map that correspond to the points composing the contour D_C can be used to estimate the average distance between the detected contour and the ground truth. A normalized value of the RMS average is computed as the descriptor of the Chamfer matching:

$$CM = \frac{1}{N} \sqrt{\frac{1}{m} \sum_{i=1}^{m} d_i^2}$$

where d is the distance value, m the number of edge points and N a normalization factor.

In the following, we will first show a qualitative visual comparison of the contours retrieved from the AC algorithm and the ground truth contours for an entire time series of the analyzed organs. Subsequently, we present the calculated descriptors in dependence on the iterations of the AC method for different frames.

6.4.1 Bowel wall

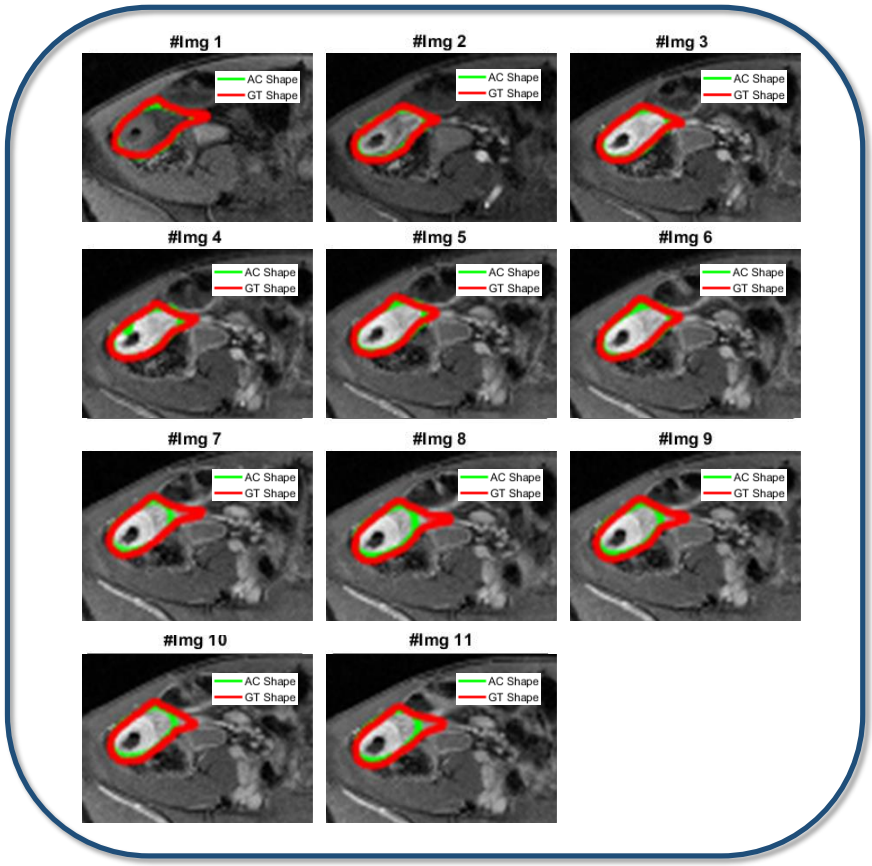

Figure 6.20: Comparison of detected contour (green) and ground truth contour (red) for a temporal sequence of the bowel wall.

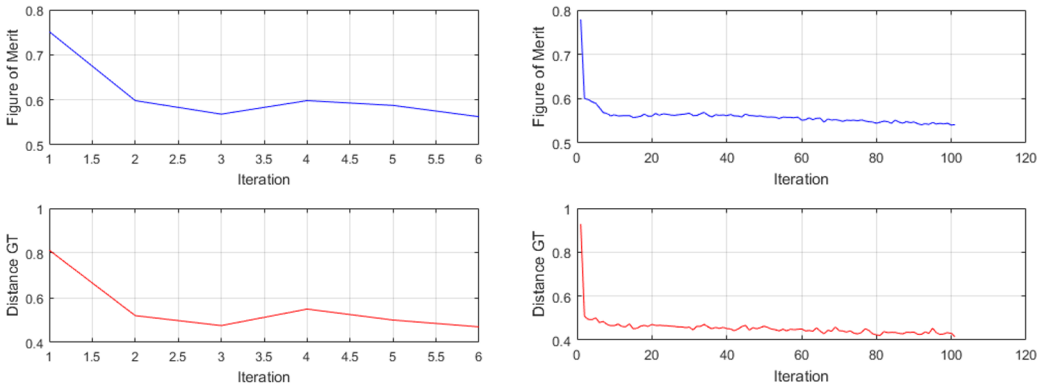

Figure 6.21: Descriptors of the similarity between the ground-truth contour and the detected contour in dependence on the iterations of the AC algorithm for the bowel wall. The upper row shows Pratt's Figure of Merit for two different frames, the lower row shows the Chamfer matching descriptor for two different frames.

6.4.2 Kidney

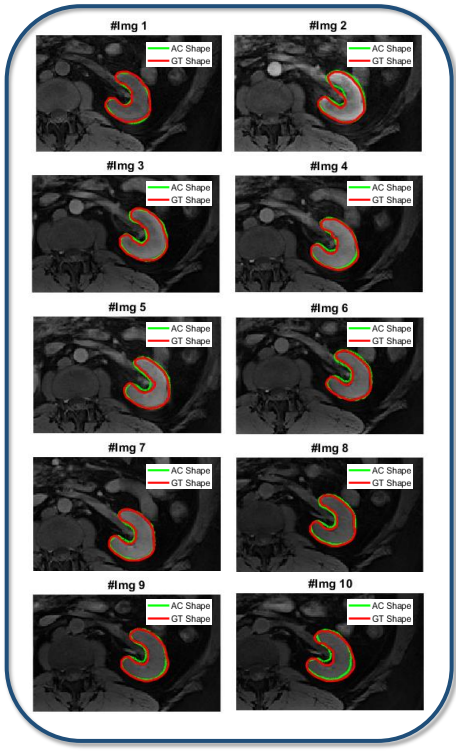

Figure 6.22: Comparison of detected contour (green) and ground truth contour (red) for a temporal sequence of the kidney.

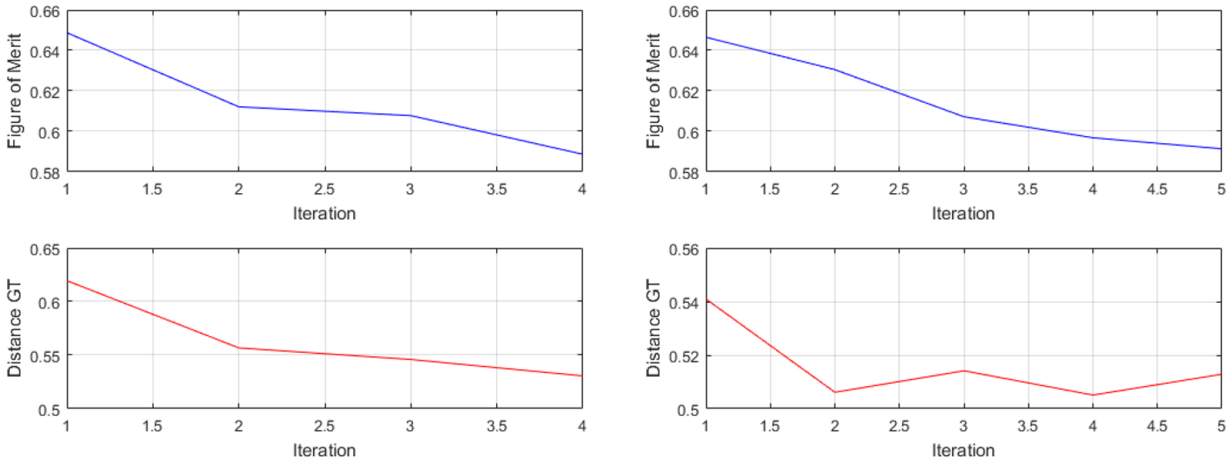

Figure 6.23: Descriptors of the similarity between the ground truth contour and the detected contour in dependence on the iterations of the AC algorithm for the kidney. The upper row shows Pratt's Figure of Merit for two different frames, the lower row shows the Chamfer matching descriptor for two different frames.

6.4.3 Liver

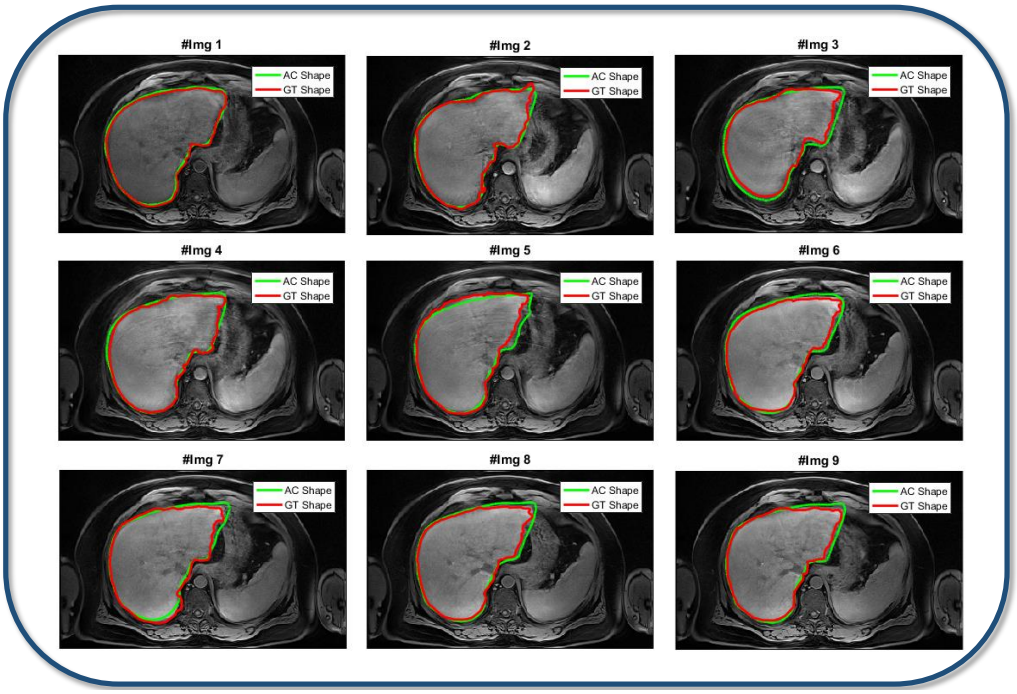

Figure 6.24: Comparison of detected contours (green) and ground truth contours (red) for a temporal sequence of the liver.

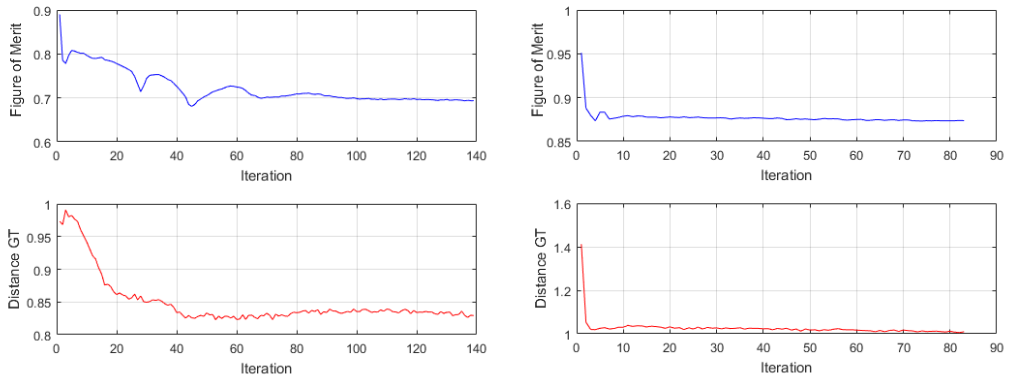

Figure 6.25: Descriptors of the similarity between the ground truth contour and the detected contour in dependence on the iterations of the AC algorithm for the liver. The upper row shows Pratt's Figure of Merit for two different frames, the lower row shows the Chamfer matching descriptor for two different frames.

6.4.4 Spleen

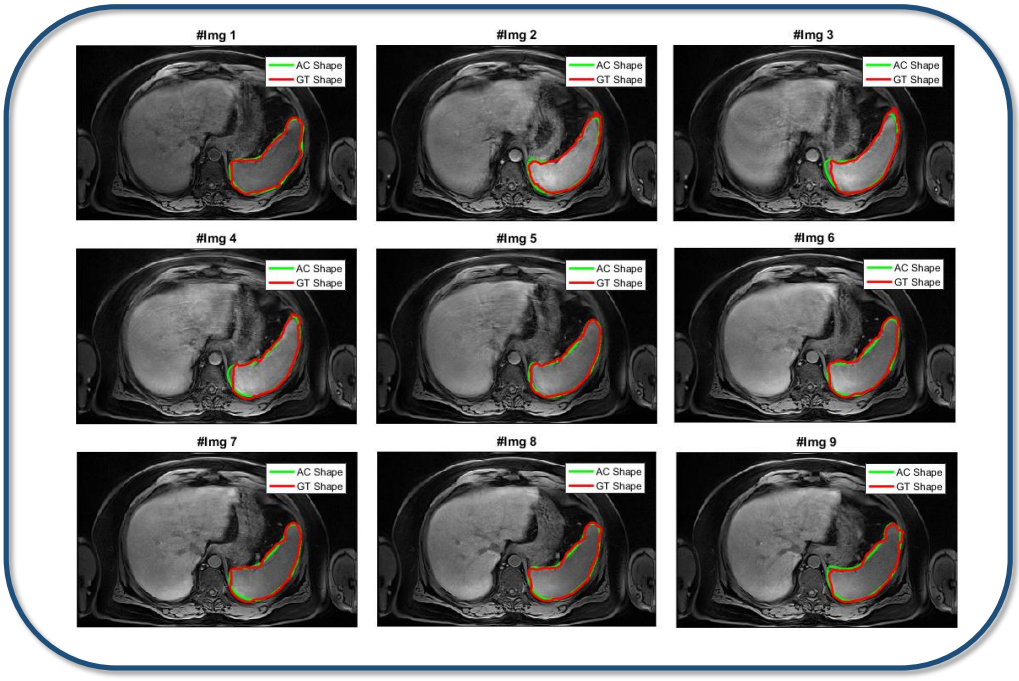

Figure 6.26: Comparison of detected contours (green) and ground truth contours (red) for a temporal sequence of the spleen.

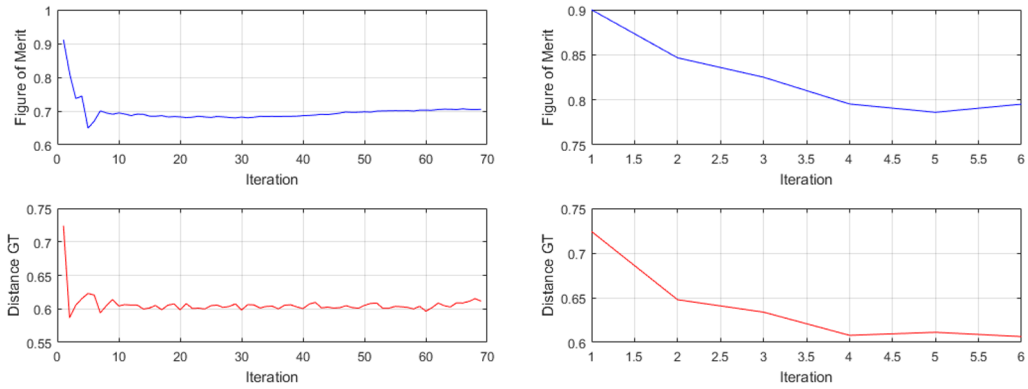

Figure 6.27: Descriptors of the similarity between the ground truth contour and the detected contour in dependence on the iterations of the AC algorithm for the spleen. The upper row shows Pratt's Figure of Merit for two different frames, the lower row shows the Chamfer matching descriptor for two different frames.

6.5 Discussion of the results

From the shown results, we can draw several conclusions about the performance of our algorithm and its effect on the analyzed data.

In general, we can state that our segmentation algorithm works relatively stable and provides accurate contours for small (Figure 6.4) and large organs (Figure 6.13). This was quantitatively and qualitatively confirmed by the comparison of our segmentation results with a manually generated ground truth. For the small organs, the algorithm converges generally faster than for the larger organs (Figures 6.21, 6.23, 6.25, 6.27), as the initial contour is more often relatively close to the edges in image. On the other hand, convergence problems related to snakes that got stuck between two edges were also observed more often for smaller organs, due to higher probability of encountering two edges that are close to each other. However, in most cases the described correction mechanisms could solve this problems and allowed for a precise tracking of the desired organ boundaries.

The effect of motion correction on results of the perfusion analysis is depending on the relative displacement of the organ or analyzed ROI in relation to its size. From this consideration, it is obvious that the motion correction affects the color maps more strongly than the enhancement curves, since the former are computed pixel by pixel and the latter are retrieved from intensity values averaged over the entire organ or ROI.

The degree of movement of the organs can be estimated from the figures, in which we compared the contour of a given frame with the projected superimposed contours from the first frame (Figures 6.7, 6.11,6.16, 6.19.) From these figures, it is evident that the bowel wall shows the largest relative displacement of all analyzed organs. This is not surprising, since the cross section of the bowel wall is relatively small in the analyzed data and the organ is affected by strong peristaltic movements. The large relative displacements of this organ are reflected in a considerable improvement of the results from the motion-corrected analysis when compared to the uncorrected analysis. As explained before, the effect of motion compensation is particularly pronounced in the color maps of the characteristic

parameters, which can be seen in Figure 6.6. The relative displacement of the other organs was considerably smaller for the analyzed measurements, therefore, the effect of motion correction is less pronounced, but still clearly visible.

In conclusion, we can state that our motion compensation algorithm provides reliable results that can improve the perfusion analysis of different organs. The benefit of the application of the method depends strongly on the movement and the size of the organ and is expected to be higher for small organs that are affected by strong motions.

Chapter 7

Summary and discussion

The increasingly important role of medical imaging in the diagnosis and treatment of disease has created a wide range of challenges regarding the processing of medical imaging data. Image processing methods like segmentation, registration, motion tracking and visualization are crucial for the analysis and understanding of the acquired data and thereby for the reliable and accurate diagnosis by the physician or other health personnel.

This work is focused on the improvement of the analysis and visualization of data dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI), which has become popular imaging technique in the field of oncology. This method enables the analysis of blood perfusion dynamics in vessels by measuring the temporal evolution of the MRI signal in the tissue or organ of interest after the injection of an paramagnetic contrast agent. The acquired signal is analyzed for a selected region of interest (ROI) in each frame of the time series to generate an enhancement curve (EC), which reflects the tissue's response to the arrival of the CA by time-dependent enhancement values of the detected intensity. By analyzing ECs, a set of parameters can be extracted, which are usually visualized by 2D color maps and provide important information for clinical and oncological diagnosis, in particular for the characterization of vessel morphology and the discrimination between healthy and malignant tissue. A major problem regarding the correct determination of the enhancement curves arises from movement and deformations of the analyzed region of interest over the course of the examination. These motion artifacts are caused by breathing, heart contraction, organ peristalsis or shifting of the patient and are not corrected in the standard analysis procedure for DCE-MRI data, which can lead to a wrong or inaccurate estimation of the enhancement curve.

7.1 Motion tracking algorithm

The major contribution of this work is the development of a motion tracking algorithm for the compensation of undesired motion during the DCE-MRI procedure (Chapter 4). Using an active contour (AC) model, the deformation of the ROI is determined for each image in the sequence. The AC method relies on the minimization of an energy functional composed of internal and external energy terms. While the internal energy ensures the smoothness of the contour, the external energy

causes the snake to be attracted to the edges of the ROI. In standard AC algorithms, the external energy merely includes the gradient information of the image, leading to a limited capture range of the snake and thereby to an inaccurate segmentation. To circumvent this effect we added an additional term to the energy function, i.e., we included a distance map calculated with the Chamfer transform into the external energy to increase the capture range. This results in a considerably more accurate detection of the contour of the ROI by the snake and thus in a better segmentation. After the segmentation, the center of mass and the principal axes for each ROI are calculated and subsequently a rototranslational transformation is applied to the ROIs in order to correct for misalignment between the individual frames. We applied the algorithm to four abdominal organs, namely, liver, spleen, kidney, and bowel wall, since the DCE-MRI analysis of these organs is particularly affected by motion artifacts. We corrected and analyzed numerous sequences from different spatial planes and compared the retrieved results to those generated from the uncorrected time sequences . The results of the motion-corrected analysis demonstrated a significant improvement of the accuracy of the enhancement curves and the characteristic parameters when compared to the uncorrected results, in particular for organs that are affected by large movement relative to their size. Furthermore, we used a ground truth approach to quantify the performance of our AC segmentation method. The results from this evaluation showed that our AC model can generate accurate segmentations of small and large organs.

In order to further improve and optimize the developed motion compensation method, we discuss the currently existing weaknesses and issues of the algorithm and propose approaches to solve this problems.

A drawback of the used parametric AC model is its sensitivity to initialization, which may lead to convergence to undesired local minima. The modification of the energy functional compensates partly for this effect, but there are still cases where the AC looses the edges of the object and collapses in the worst case. We presented methods to tackle this problem by optimizing the weighting factors for the individual energy terms and by implementing additional constraints and control mechanisms (Chapter 6). However, there is still a trade-off between avoiding the collapse of the snake and finding the accurate and smooth contour for the ROI. Consequently, it is desirable to further optimize the used AC model by implementing more advanced control mechanisms and optimization methods. Alternatively, the used parametric model could be substituted by a geometric AC model [53, 54]. These models are based on curve evolution theory and the level set method. Instead of evolving the curve with the plane-like snakes, this geometric functional evolves in time with respect to the plane. A geometrical model has several advantages over a parametric model, such as computational simplicity and the ability handle topological changes. Most importantly, the contour can start far from the boundary and will still converge to boundary concavities [55].

The estimation of the trajectory of the ROI by assuming movements associated with a rigid body clearly improves the results compared to the uncorrected case and has the advantage of being computationally simple and fast. However, the internal organs of the human body do not conform to a rigid approximation. A first improvement for the future, which could be implemented in a relatively easy way, is to relax the assumptions of rigid body motion by introducing an element of distortion to be taken into account in the tracking. The most suitable approach is the

implementation of a nonrigid transformation [90], which also accounts for soft tissue deformation during imaging. While this approach promises the most accurate results, it is also a complex and challenging task that poses several additional problems for the implementation.

The undesired motions during acquisition do not only result in misalignments between the frames of a time sequences, but also cause displacements in the transverse direction relative to the image plane, thereby generating inconsistencies in the shape of the organ profiles in the same temporal series. In this thesis, we presented a first approach to account for this misalignments by including the slices from the adjacent spatial planes into our algorithm in order to reassign images that belong to the same temporal sequence (Section 4.6). The method segments the ROI of a particular frame via the AC model and uses the generated contour to find the best match of the three considered slices of the consecutive frame by superimposing the contour on the distance map representation of the slices. This is done frame-by frame in order to create time sequences without inconsistencies in the organ profiles. The algorithm was successfully applied to a first DCE-MRI data set of a pathological breast, however, the algorithm has not been extensively tested and validated yet, therefore it is too early to draw conclusions about the reliability and general applicability of this method.

7.2 3D analysis and visualization tool

Three-dimensional imaging methods as well as the analysis and visualization of the volumetric data are of crucial importance in medical diagnostics. They provide the physicians with a more realistic perspective of the examined structures and may allow for the understanding and analysis of complex anatomical structures and physiological processes that cannot be sufficiently represented by two-dimensional approaches.

In DCE-MRI examinations, the analysis is usually performed on a single temporal series of MRI slices, consequently the provided results are naturally two-dimensional, such as the Contrast Enhancement (CE) curve and the color maps displaying the characteristic parameters. In this thesis, we developed an innovative tool that overcomes the limits of two-dimensional diagnosis. To achieve this goal, we developed and implemented suitable segmentation algorithms and visualization operations. Our tool provides the possibility to analyze the entire set of temporal sequences from a DCE-MRI examination and visualize the gathered volumetric results in form of an enhancement surface and 3D color maps of the characteristic parameters. We implemented different reconstruction techniques, such as multiplanar reconstruction and volume cutting, to allow a flexible visualization of the 3D data. Furthermore, our tool is optimized to avoid long computation times and can provide the results almost on-the-fly. To facilitate user interaction we embedded the program in a dedicated graphical user interface that allows intuitive execution of the various processing steps. we developed implemented suitable. In conclusion, we provide the medical doctors with a fast, reliable and easy-to-handle tool that provides three-dimensional vision of the ROI and additional quantitative data. The tool enables a deeper understanding of the examined anatomical structures and the underlying perfusion processes and allows for a comprehensive and more accurate diagnosis.

Future improvements of the tool may be the following:

A main objective for the development of this tool was a fast generation of the results. Consequently, it was not possible to include the motion compensation algorithm into the analysis of the time sequences, since the large amount of data and the computational effort for the contour segmentation would have increased the execution time drastically. Therefore, an improvement of the tool involves the development of a fast motion correction algorithm that might not be as accurate as the described AC algorithm, but could still improve the reliability and accuracy of the analysis.

Furthermore, it may be beneficial to implement the optional possibility to create a volume rendering of the volume of interest, which allows to use a transparency threshold. While this feature would require more extensive computations than the currently implemented isosurface rendering, it could also further improve the three-dimensional understanding of the examined structures.

7.3 Conclusion

In conclusion, it can be stated that this work provides important contributions to the improvement of the analysis and visualization of DCE-MRI data.

The developed motion tracking algorithm is a valuable method to generate more reliable and accurate results for the analysis of perfusion dynamics of the contrast agent. This was demonstrated by applying the algorithm to several internal organs. Possible improvements of the developed method in the future comprise further optimization of the contour segmentation of the ROI, an improved transformation model and the additional correction for out-of-plane motion.

The 3D analysis and visualization tool can be of great benefit for the clinical personell that works with DCE-MRI data. It can improve their understanding of the perfusion processes and provide new quantitative volumetric information. Suggested additional features for the tool are a motion tracking correction and a volume rendering mode.

Bibliography

[1] Lauterbur, P. C. Image formation by induced local interactions: Examples employing nuclear magnetic resonance. *Nature* **242**, 190 – 191 (1973).

[2] Mansfield, P. & Grannell, P. K. Nmr 'diffraction' in solids? *Journal of Physics C: Solid State Physics* **6**, L422 (1973).

[3] Le Bihan, D., Turner, R., Douek, P. & Patronas, N. Diffusion mr imaging: clinical applications. *AJR. American journal of roentgenology* **159**, 591–599 (1992).

[4] DeCharms, R. C. Applications of real-time fmri. *Nature Reviews. Neuroscience* **9**, 720 (2008).

[5] Prasad, P. V., Edelman, R. R. & Epstein, F. H. Noninvasive evaluation of intrarenal oxygenation with bold mri. *Circulation* **94**, 3271–3275 (1996).

[6] Kauczor, H.-U., Surkau, R. & Roberts, T. Mri using hyperpolarized noble gases. *European radiology* **8**, 820–827 (1998).

[7] Damadian, R. Tumor detection by nuclear magnetic resonance. *Science* **171**, 1151–1153 (1971).

[8] Hanahan, D. & Weinberg, R. A. The hallmarks of cancer. *cell* **100**, 57–70 (2000).

[9] Carmeliet, P. & Jain, R. K. Angiogenesis in cancer and other diseases. *nature* **407**, 249 (2000).

[10] Knopp, M. V., Giesel, F. L., Marcos, H., von Tengg-Kobligk, H. & Choyke, P. Dynamic contrast-enhanced magnetic resonance imaging in oncology. *Topics in Magnetic Resonance Imaging* **12**, 301–308 (2001).

[11] Ernst, R. R., Bodenhausen, G., Wokaun, A. *et al.* Principles of nuclear magnetic resonance in one and two dimensions (1987).

[12] Callaghan, P. T. *Principles of nuclear magnetic resonance microscopy* (Oxford University Press on Demand, 1993).

[13] Reiser, M. F., Semmler, W. & Hricak, H. *Magnetic Resonance Tomography* (Springer, 2008).

[14] Bernstein, M. A., King, K. F. & Zhou, X. J. *Handbook of MRI pulse sequences* (Elsevier, 2004).

[15] Brown, R. W., Cheng, Y.-C. N., Haacke, E. M., Thompson, M. R. & Venkatesan, R. *Magnetic resonance imaging: physical principles and sequence design* (John Wiley & Sons, 2014).

[16] Jackson, A., Buckley, D. L., Parker, G. J. *et al. Dynamic contrast-enhanced magnetic resonance imaging in oncology* (Springer, 2005).

[17] Padhani, A. R. Dynamic contrast-enhanced mri in clinical oncology: Current status and future directions. *Journal of Magnetic Resonance Imaging* **16**, 407–422 (2002).

[18] Rabi, I. I., Millman, S., Kusch, P. & Zacharias, J. R. The molecular beam resonance method for measuring nuclear magnetic moments. the magnetic moments of $_3Li^6$, $_3Li^7$ and $_9F^{19}$. *Phys. Rev.* **55**, 526–535 (1939).

[19] Bloch, F., Hansen, W. W. & Packard, M. The nuclear induction experiment. *Phys. Rev.* **70**, 474–485 (1946).

[20] Purcell, E. M., Torrey, H. C. & Pound, R. V. Resonance absorption by nuclear magnetic moments in a solid. *Phys. Rev.* **69**, 37–38 (1946).

[21] Gerlach, W. & Stern, O. Der experimentelle nachweis der richtungsquantelung im magnetfeld. *Zeitschrift für Physik* **9**, 349–352 (1922).

[22] Zeeman, P. The Effect of Magnetisation on the Nature of Light Emitted by a Substance. *nature* **55**, 347 (1897).

[23] Stanisz, G. J. *et al.* T1, T2 relaxation and magnetization transfer in tissue at 3T. *Magnetic Resonance in Medicine* **54**, 507,512 (2005).

[24] Ridgway, J. P. Cardiovascular magnetic resonance physics for clinicians: part i. *Journal of Cardiovascular Magnetic Resonanc* **12**, 71 (2010).

[25] Hahn, E. L. & Maxwell, D. E. Spin echo measurements of nuclear spin coupling in molecules. *Phys. Rev.* **88**, 1070–1084 (1952).

[26] Fleckenstein, J. L. *et al.* Fast short-tau inversion-recovery mr imaging. *Radiology* **179**, 499–504 (1991).

[27] Nyquist, H. Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers* **47**, 617–644 (1928).

[28] Cooley, J. W. & Tukey, J. W. An algorithm for the machine calculation of complex fourier series. *Mathematics of Computation* **19**, 297–6301 (1986).

[29] Fessler, J. & Sutton, B. Nonuniform fast fourier transforms using min-max interpolation. *Trans. Sig. Proc.* **51**, 560–574 (2003).

[30] National high magnetic field laboratory, http://nationalmaglab. org/education/magnet-academy/learn-the-basics/stories/ mri-a-guided-tour/.

[31] Padhani, A. Dynamic contrast-enhanced mri studies in human tumours. *The British journal of radiology* **72**, 427–431 (1999).

[32] Daniel, B. L. *et al.* Breast disease: dynamic spiral mr imaging. *Radiology* **209**, 499–509 (1998).

[33] Kuhl, C. K. *et al.* Dynamic breast mr imaging: are signal intensity time course data useful for differential diagnosis of enhancing lesions? *Radiology* **211**, 101–110 (1999).

[34] Russo, V. & Setola, R. Dynamic contrast enhancement: Analysis's models and methodologies. In *Handbook of Research on Advanced Techniques in Diagnostic Imaging and Biomedical Applications*, 392–406 (IGI Global, 2009).

[35] Bankman, I. *Handbook of medical image processing and analysis* (academic press, 2008).

[36] Lim, J.-H., Ong, S.-H. & Xiong, W. *Biomedical Image Understanding: Methods and Applications* (John Wiley & Sons, 2015).

[37] Scherzer, O. *Handbook of mathematical methods in imaging* (Springer Science & Business Media, 2010).

[38] Petrou, M. & Petrou, C. *Image processing: the fundamentals* (John Wiley & Sons, 2010).

[39] Zhang, Y.-J. An overview of image and video segmentation in the last 40 years. *Advances in Image and Video Segmentation* 1–15 (2006).

[40] Rastgarpour, M. & Shanbehzadeh, J. Application of ai techniques in medical image segmentation and novel categorization of available methods and tools. In *Tools, Proceedings of the International MultiConference of Engineers and Computer Scientists 2011 Vol I, IMECS 2011, March 16-18, 2011, Hong Kong* (Citeseer, 2011).

[41] Chow, C. & Kaneko, T. Automatic boundary detection of the left ventricle from cineangiograms. *Computers and biomedical research* **5**, 388–410 (1972).

[42] Davies, E. R. *Machine vision: theory, algorithms, practicalities* (Elsevier, 2004).

[43] Kittler, J., Illingworth, J. & Föglein, J. Threshold selection based on a simple image statistic. *Computer vision, graphics, and image processing* **30**, 125–147 (1985).

[44] Zucker, S. W. Region growing: Childhood and adolescence. *Computer graphics and image processing* **5**, 382–399 (1976).

[45] Adams, R. & Bischof, L. Seeded region growing. *IEEE Transactions on pattern analysis and machine intelligence* **16**, 641–647 (1994).

[46] Chen, P. & Pavlidis, T. Image segmentation as an estimation problem. *Computer Graphics and Image Processing* **12**, 153–172 (1980).

[47] Fukada, Y. Spatial clustering procedures for region analysis. *Pattern Recognition* **12**, 395–403 (1980).

[48] Beucher, S. & Lantuéjoul, C. Use of watersheds in contour detection (1979).

[49] Meyer, F. & Beucher, S. Morphological segmentation. *Journal of visual communication and image representation* **1**, 21–46 (1990).

[50] Canny, J. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence* 679–698 (1986).

[51] Kass, M., Witkin, A. & Terzopoulos, D. Snakes: Active contour models. In *Proc. 1st Int. Conf. on Computer Vision*, vol. 259, 268 (1987).

[52] Xu, C., Pham, D. L. & Prince, J. L. Image segmentation using deformable models. *Handbook of medical imaging* **2**, 129–174 (2000).

[53] Caselles, V., Catté, F., Coll, T. & Dibos, F. A geometric model for active contours in image processing. *Numerische mathematik* **66**, 1–31 (1993).

[54] Malladi, R., Sethian, J. A. & Vemuri, B. C. Shape modeling with front propagation: A level set approach. *IEEE transactions on pattern analysis and machine intelligence* **17**, 158–175 (1995).

[55] Cremers, D., Rousson, M. & Deriche, R. A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape. *International journal of computer vision* **72**, 195–215 (2007).

[56] Dunn, J. C. A fuzzy relative of the isodata process and its use in detecting compact well-separated clusters (1973).

[57] Bezdek, J. C., Ehrlich, R. & Full, W. Fcm: The fuzzy c-means clustering algorithm. *Computers & Geosciences* **10**, 191–203 (1984).

[58] MacQueen, J. *et al.* Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, 281–297 (Oakland, CA, USA., 1967).

[59] Zadeh, L. A. Fuzzy sets. *Information and control* **8**, 338–353 (1965).

[60] Bezdek, J., Hall, L., Clark, M., Goldgof, D. B. & Clarke, L. Medical image analysis with fuzzy models. *Statistical Methods in Medical Research* **6**, 191–214 (1997).

[61] Serra, J. *Image analysis and mathematical morphology, v. 1* (Academic press, 1982).

[62] Soille, P. *Morphological image analysis: principles and applications* (Springer Science & Business Media, 2013).

[63] Meijering, E. & van Cappellen, G. Biological image analysis primer, https://imagescience.org/meijering/publications/download/biap2006.pdf (2006).

[64] Sotiras, A., Davatzikos, C. & Paragios, N. Deformable medical image registration: A survey. *IEEE transactions on medical imaging* **32**, 1153–1190 (2013).

[65] Zitova, B. & Flusser, J. Image registration methods: a survey. *Image and vision computing* **21**, 977–1000 (2003).

[66] Goshtasby, A. A. *2-D and 3-D image registration: for medical, remote sensing, and industrial applications* (John Wiley & Sons, 2005).

[67] Hill, D. L., Batchelor, P. G., Holden, M. & Hawkes, D. J. Medical image registration. *Physics in medicine and biology* **46**, R1 (2001).

[68] Maintz, J. A. & Viergever, M. A. A survey of medical image registration. *Medical image analysis* **2**, 1–36 (1998).

[69] Graf, F., Kriegel, H.-P., Schubert, M., Pölsterl, S. & Cavallaro, A. 2d image registration in ct images using radial image descriptors. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011* 607–614 (2011).

[70] Woods, R. P., Mazziotta, J. C., Cherry, S. R. *et al.* Mri-pet registration with automated algorithm. *Journal of computer assisted tomography* **17**, 536–546 (1993).

[71] Goerres, G. W. *et al.* Pet-ct image co-registration in the thorax: influence of respiration. *European journal of nuclear medicine and molecular imaging* **29**, 351 (2002).

[72] Van Dye, J., Batista, J. & Bauman, G. S. Accuracy and uncertainty considerations in modern radiation oncology. *The Modern Technology of Radiation Oncology* **3**, 361–412 (2013).

[73] Markelj, P., Tomaževič, D., Likar, B. & Pernuš, F. A review of 3d/2d registration methods for image-guided interventions. *Medical image analysis* **16**, 642–661 (2012).

[74] Maurer, C. R. *et al.* Registration of head volume images using implantable fiducial markers. *IEEE transactions on medical imaging* **16**, 447–462 (1997).

[75] Lemieux, L., Kitchen, N. D., Hughes, S. W. & Thomas, D. G. Voxel-based localization in frame-based and frameless stereotaxy and its accuracy. *Medical physics* **21**, 1301–1310 (1994).

[76] Barnett, G. H., Miller, D. W. & Weisenberger, J. Frameless stereotaxy with scalp-applied fiducial markers for brain biopsy procedures: experience in 218 cases. *Journal of neurosurgery* **91**, 569–576 (1999).

[77] Hill, D. L. *et al.* Registration of mr and ct images for skull base surgery using point-like anatomical features. *The British journal of radiology* **64**, 1030–1035 (1991).

[78] Thirion, J.-P. Extremal points: definition and application to 3d image registration. In *CVPR*, 587–592 (1994).

[79] Guéziec, A. & Ayache, N. Smoothing and matching of 3-d space curves. In *Computer Vision ECCV 92*, 620–629 (Springer, 1992).

[80] Jiang, H., Robb, R. A. & Holton, K. S. New approach to 3-d registration of multimodality medical images by surface matching. In *Proceedings if the SPIE: The international society for optical engineering*, 196–196 (SPIE INTERNATIONAL SOCIETY FOR OPTICAL, 1992).

[81] Gueziec, A. Large deformable splines, crest lines and matching. In *Computer Vision, 1993. Proceedings., Fourth International Conference on*, 650–657 (IEEE, 1993).

[82] MacDonald, D., Avis, D. & Evans, A. C. Multiple surface identification and matching in magnetic resonance images. In *Visualization in Biomedical Computing 1994*, 160–169 (International Society for Optics and Photonics, 1994).

[83] Junck, L., Moen, J. G., Hutchins, G. D., Brown, M. B. & Kuhl, D. E. Correlation methods for the centering, rotation, and alignment of functional brain images. *Journal of nuclear medicine: official publication, Society of Nuclear Medicine* **31**, 1220–1226 (1990).

[84] Alpert, N., Bradshaw, J., Kennedy, D. & Correia, J. The principal axes transformation - a method for image registration. *Journal of nuclear medicine* **31**, 1717–1722 (1990).

[85] Munch, B. & Ruegsegger, P. 3-d repositioning and differential images of volumetric ct measurements. *IEEE Transactions on medical imaging* **12**, 509–514 (1993).

[86] De Castro, E. & Morandi, C. Registration of translated and rotated images using finite fourier transforms. *IEEE Transactions on pattern analysis and machine intelligence* 700–703 (1987).

[87] Viola, P. & Wells, W. M. Alignment by maximization of mutual information. In *Computer Vision, 1995. Proceedings., Fifth International Conference on*, 16–23 (IEEE, 1995).

[88] Besl, P. J., McKay, N. D. *et al.* A method for registration of 3-d shapes. *IEEE Transactions on pattern analysis and machine intelligence* **14**, 239–256 (1992).

[89] Reddy, B. S. & Chatterji, B. N. An fft-based technique for translation, rotation, and scale-invariant image registration. *IEEE transactions on image processing* **5**, 1266–1271 (1996).

[90] Holden, M. A review of geometric transformations for nonrigid body registration. *IEEE transactions on medical imaging* **27**, 111–128 (2008).

[91] Bajcsy, R. & Kovačič, S. Multiresolution elastic matching. *Computer vision, graphics, and image processing* **46**, 1–21 (1989).

[92] Christensen, G. E., Rabbitt, R. D. & Miller, M. I. 3d brain mapping using a deformable neuroanatomy. *Physics in medicine and biology* **39**, 609 (1994).

[93] Bro-Nielsen, M. & Gramkow, C. Fast fluid registration of medical images. In *Visualization in Biomedical Computing*, 265–276 (Springer, 1996).

[94] Horn, B. K. & Schunck, B. G. Determining optical flow. *Artificial Intelligence* **17**, 185 – 203 (1981).

[95] Thirion, J.-P. Image matching as a diffusion process: an analogy with maxwell's demons. *Medical image analysis* **2**, 243–260 (1998).

[96] Fornefett, M., Rohr, K. & Stiehl, H. S. Radial basis functions with compact support for elastic registration of medical images. *Image and vision computing* **19**, 87–96 (2001).

[97] Bookstein, F. L. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on pattern analysis and machine intelligence* **11**, 567–585 (1989).

[98] Rueckert, D. *et al.* Nonrigid registration using free-form deformations: application to breast mr images. *IEEE transactions on medical imaging* **18**, 712–721 (1999).

[99] Amit, Y. A nonlinear variational problem for image matching. *SIAM Journal on Scientific Computing* **15**, 207–224 (1994).

[100] Wu, Y.-T., Kanade, T., Li, C.-C. & Cohn, J. Image registration using wavelet-based motion model. *International Journal of Computer Vision* **38**, 129–152 (2000).

[101] Christensen, G. E. & Johnson, H. J. Consistent image registration. *IEEE transactions on medical imaging* **20**, 568–582 (2001).

[102] Noblet, V., Heinrich, C., Heitz, F. & Armspach, J.-P. 3-d deformable image registration: a topology preservation scheme based on hierarchical deformation models and interval analysis optimization. *IEEE Transactions on image processing* **14**, 553–566 (2005).

[103] Glaunes, J., Trouvé, A. & Younes, L. Diffeomorphic matching of distributions: A new approach for unlabelled point-sets and sub-manifolds matching. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, vol. 2, II–II (Ieee, 2004).

[104] Janssens, G., Jacques, L., de Xivry, J. O., Geets, X. & Macq, B. Diffeomorphic registration of images with variable contrast enhancement. *Journal of Biomedical Imaging* **2011**, 3 (2011).

[105] Veenman, C. J., Reinders, M. J. T. & Backer, E. Resolving motion correspondence for densely moving points. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**, 54–72 (2001).

[106] Serby, D., Koller-Meier, E. & Gool, L. V. Probabilistic object tracking using multiple features. In *Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04) Volume 2 - Volume 02*, ICPR '04, 184–187 (2004).

[107] Comaniciu, D., Ramesh, V. & Meer, P. Kernel-based object tracking. *IEEE Trans. Pattern Anal. Mach. Intell.* **25**, 564–575 (2003).

[108] Yilmaz, A., Li, X. & Shah, M. Contour-based object tracking with occlusion handling in video acquired using mobile cameras. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**, 1531–1536 (2004).

[109] Sundaresan, A. & Chellappa, R. Multicamera tracking of articulated human motion using shape and motion cues. *IEEE Transactions on Image Processing* **18**, 2114–2126 (2009).

[110] Ballard, D. & Brown, C. *Computer Vision* (Prentice-Hall, 1982).

[111] Ali, A. & Aggarwal, J. K. Segmentation and recognition of continuous human activity. In *IEEE Workshop on Detection and Recognition of Events in Video*, 28– (2001).

[112] Zhu, S. C. & Yuille, A. Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **18**, 884–900 (1996).

[113] Paragios, N. & Deriche, R. Geodesic active regions and level set methods for supervised texture segmentation. *International Journal of Computer Vision* **46**, 223–247 (2002).

[114] Elgammal, A., Duraiswami, R., Harwood, D. & Davis, L. S. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. *Proceedings of the IEEE* **90**, 1151–1163 (2002).

[115] Edwards, G., Taylor, C. & Cootes, T. Interpreting face images using active appearance models. In *International Conference on Face and Gesture Recognition*, 300–305.

[116] Moghaddam, B. & Pentland, A. Probabilistic visual learning for object representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **19**, 696–710 (1997).

[117] Yilmaz, A., Javed, O. & Shah, M. Object tracking: A survey. *ACM Comput. Surv.* **38** (2006).

[118] Paschos, G. Perceptually uniform color spaces for color texture analysis: an empirical evaluation. *IEEE Transactions on Image Processing* **10**, 932–937 (2001).

[119] Bowyer, K., Kranenburg, C. & Dougherty, S. Edge detector evaluation using empirical roc curves. In *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, vol. 1, 359 Vol. 1 (1999).

[120] Black, M. J. & Anandan, P. The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding* **63**, 75 – 104 (1996).

[121] Haralick, R. M., Shanmugam, K. & Dinstein, I. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics* **SMC-3**, 610–621 (1973).

[122] Mallat, S. G. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**, 674–693 (1989).

[123] Greenspan, H. *et al.* Overcomplete steerable pyramid filters and rotation invariance. In *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 222–228 (1994).

[124] Saeys, Y., Inza, I. & Larranaga, P. A review of feature selection techniques in bioinformatics. *bioinformatics* **23**, 2507–2517 (2007).

[125] Yu, L. & Liu, H. Efficient feature selection via analysis of relevance and redundancy. *J. Mach. Learn. Res.* **5**, 1205–1224 (2004).

[126] SIEDLECKI, W. & SKLANSKY, J. On automatic feature selection. *International Journal of Pattern Recognition and Artificial Intelligence* **02**, 197–220 (1988).

[127] Inza, I., Larranaga, P., Etxeberria, R. & Sierra, B. Feature subset selection by bayesian network-based optimization. *Artificial Intelligence* **123**, 157 – 184 (2000).

[128] Duda, R. O., Hart, P. E. & Stork, D. G. *Pattern classification* (Wiley, New York, 1973).

[129] Weston, J., Elisseeff, A., Schölkopf, B. & Tipping, M. Use of the zero-norm with linear models and kernel methods. *Journal of machine learning research* **3**, 1439–1461 (2003).

[130] Salari, V. & Sethi, I. K. Feature point correspondence in the presence of occlusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**, 87–91 (1990).

[131] Broida, T. J. & Chellappa, R. Estimation of object motion parameters from noisy images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 90–99 (1986).

[132] Streit, R. L. & Luginbuhl, T. E. Maximum likelihood method for probabilistic multihypothesis tracking. In *SPIE's International Symposium on Optical Engineering and Photonics in Aerospace Sensing*, 394–405 (1994).

[133] Shi, J. *et al.* Good features to track. In *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on*, 593–600 (1994).

[134] Tao, H., Sawhney, H. S. & Kumar, R. Object tracking with bayesian estimation of dynamic layer representations. *IEEE transactions on pattern analysis and machine intelligence* **24**, 75–89 (2002).

[135] Black, M. J. & Jepson, A. D. Eigentracking: Robust matching and tracking of articulated objects using a view-based representation. *International Journal of Computer Vision* **26**, 63–84 (1998).

[136] Huttenlocher, D. P., Noh, J. J. & Rucklidge, W. J. Tracking non-rigid objects in complex scenes. In *Computer Vision, 1993. Proceedings., Fourth International Conference on*, 93–101 (IEEE, 1993).

[137] Sato, K. & Aggarwal, J. K. Temporal spatio-velocity transform and its application to tracking and interaction. *Computer Vision and Image Understanding* **96**, 100–128 (2004).

[138] Isard, M. A. blake," condensation-conditional propagation for visual tracking. *Int. J. Computer Vision* (1998).

[139] Bertalmio, M., Sapiro, G. & Randall, G. Morphing active contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**, 733–737 (2000).

[140] Resende, M. G. & Ribeiro, C. C. Greedy randomized adaptive search procedures: Advances, hybridizations, and applications. In *Handbook of metaheuristics*, 283–319 (Springer, 2010).

[141] Schaap, M. *et al.* Bayesian tracking of tubular structures and its application to carotid arteries in cta. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007* 562–570 (2007).

[142] Peressutti, D. *et al.* A novel bayesian respiratory motion model to estimate and resolve uncertainty in image-guided cardiac interventions. *Medical image analysis* **17**, 488–502 (2013).

[143] Gupta, S. N. & Prince, J. L. On variable brightness optical flow for tagged mri. In *Information Processing in Medical Imaging*, vol. 3, 323–334 (Kluwer, 1995).

[144] Huang, J., Abendschein, D., Davila-Roman, V. G. & Amini, A. A. Spatio-temporal tracking of myocardial deformations with a 4-d b-spline model from tagged mri. *IEEE Transactions on Medical Imaging* **18**, 957–972 (1999).

[145] Pan, L., Prince, J. L., Lima, J. A. & Osman, N. F. Fast tracking of cardiac motion using 3d-harp. *IEEE transactions on Biomedical Engineering* **52**, 1425–1435 (2005).

[146] Terzopoulos, D. Tracking with kalman snakes. *Active vision* (1992).

[147] Park, J., Metaxas, D., Young, A. A. & Axel, L. Deformable models with parameter functions for cardiac motion analysis from tagged mri data. *IEEE Transactions on Medical Imaging* **15**, 278–289 (1996).

[148] Pham, Q., Vincent, F., Clarysse, P., Croisille, P. & Magnin, I. A fem-based deformable model for the 3d segmentation and tracking of the heart in cardiac mri. In *Image and Signal Processing and Analysis, 2001. ISPA 2001. Proceedings of the 2nd International Symposium on*, 250–254 (IEEE, 2001).

[149] Udupa, J. K. & Herman, G. T. *3D imaging in medicine* (CRC press, 1999).

[150] Höhne, K. H., Fuchs, H. & Pizer, S. M. *3D imaging in medicine: algorithms, systems, applications*, vol. 60 (Springer Science & Business Media, 2012).

[151] Levoy, M. Display of surfaces from volume data. *IEEE Computer graphics and Applications* **8**, 29–37 (1988).

[152] Frieder, G., Gordon, D. & Reynolds, R. A. Back-to-front display of voxel based objects. *IEEE Computer Graphics and Applications* **5**, 52–60 (1985).

[153] Berry, E. *A practical approach to medical image processing* (Taylor & Francis, 2007).

[154] Udupa, J. K. & Odhner, D. Shell rendering. *IEEE Computer Graphics and Applications* **13**, 58–67 (1993).

[155] Schreiner, S., Paschal, C. B. & Galloway, R. L. Comparison of projection algorithms used for the construction of maximum intensity projection images. *Journal of computer assisted tomography* **20**, 56–67 (1996).

[156] Brown, D. G. & Riederer, S. J. Contrast-to-noise ratios in maximum intensity projection images. *Magnetic resonance in medicine* **23**, 130–137 (1992).

[157] Napel, S. *et al.* Ct angiography with spiral ct and maximum intensity projection. *Radiology* **185**, 607–610 (1992).

[158] Hertz, S. M. *et al.* Comparison of magnetic resonance angiography and contrast arteriography in peripheral arterial stenosis. *The American journal of surgery* **166**, 112–116 (1993).

[159] Reynolds, R. A., Gordon, D. & Chen, L.-S. A dynamic screen technique for shaded graphics display of slice-represented objects. *Computer Vision, Graphics, and Image Processing* **38**, 275–298 (1987).

[160] Artzy, E., Frieder, G. & Herman, G. T. The theory, design, implementation and evaluation of a three-dimensional surface detection algorithm. *Computer graphics and image processing* **15**, 1–24 (1981).

[161] Wyvill, G., McPheeters, C. & Wyvill, B. Data structure for soft objects. *The visual computer* **2**, 227–234 (1986).

[162] Chuang, J.-H. & Lee, W.-C. Efficient generation of isosurfaces in volume rendering. *Computers & graphics* **19**, 805–813 (1995).

[163] Trivedi, S. S. Interactive manipulation of three dimensional binary scenes. *The visual computer* **2**, 209–218 (1986).

[164] Odhner, D. & Udupa, J. K. Shell manipulation: Interactive alteration of multiple-material fuzzy structures. In *Medical Imaging 1995*, 35–42 (International Society for Optics and Photonics, 1995).

[165] Terzopoulos, D. & Waters, K. Techniques for realistic facial modeling and animation. In *Computer Animation 91*, 59–74 (Springer, 1991).

[166] Terzopoulos, D. On matching deformable models to images. In *Topical meeting on machine vision*, vol. 12, 160–167 (1987).

[167] Stammberger, T., Rudert, S., Michaelis, M., Reiser, M. & Englmeier, K.-H. Segmentation of mr images with b-spline snakes: A multi-resolution approach using the distance transformation for model forces. In *Bildverarbeitung für die Medizin 1998*, 164–168 (Springer, 1998).

[168] Stammberger, T., Eckstein, F., Michaelis, M., Englmeier, K.-H. & Reiser, M. Interobserver reproducibility of quantitative cartilage measurements: comparison of b-spline snakes and manual segmentation. *Magnetic resonance imaging* **17**, 1033–1042 (1999).

[169] Lingrand, D. & Montagnat, J. Levelset and b-spline deformable model techniques for image segmentation: a pragmatic comparative study. In *Scandinavian Conference on Image Analysis*, 25–34 (Springer, 2005).

[170] Pottmann, H., Leopoldseder, S. & Hofer, M. Approximation with active b-spline curves and surfaces. In *Computer Graphics and Applications, 2002. Proceedings. 10th Pacific Conference on*, 8–25 (IEEE, 2002).

[171] Cohen, L. D. On active contour models and balloons. *CVGIP: Image understanding* **53**, 211–218 (1991).

[172] Xu, C. & Prince, J. L. Snakes, shapes, and gradient vector flow. *IEEE Transactions on image processing* **7**, 359–369 (1998).

[173] Roberts, L. G. *Machine perception of three-dimensional solids*. Ph.D. thesis, Massachusetts Institute of Technology (1963).

[174] Sobel, I. & Feldman, G. A 3x3 isotropic gradient operator for image processing. *Talk at the Stanford Artificial Project* 271–272 (1968).

[175] Prewitt, J. M. Object enhancement and extraction. *Picture processing and Psychopictorics* **10**, 15–19 (1970).

[176] Butt, M. A. & Maragos, P. Optimum design of chamfer distance transforms. *IEEE Transactions on Image Processing* **7**, 1477–1484 (1998).

[177] Gyftakis, S., Agouris, P. & Stefanidis, A. Image-based change detection of areal objects using differential snakes. In *Proceedings of the 13th annual ACM international workshop on Geographic information systems*, 135–142 (ACM, 2005).

[178] Pardàs, M. & Sayrol, E. Motion estimation based tracking of active contours. *Pattern recognition letters* **22**, 1447–1456 (2001).

[179] Gupta, S. N., Solaiyappan, M., Beache, G. M., Arai, A. E. & Foo, T. K. Fast method for correcting image misregistration due to organ motion in time-series mri data. *Magnetic Resonance in Medicine* **49**, 506–514 (2003).

[180] Marr, D. & Hildreth, E. Theory of edge detection. *Proceedings of the Royal Society of London B: Biological Sciences* **207**, 187–217 (1980).

[181] Abdou, I. E. & Pratt, W. K. Quantitative design and evaluation of enhancement/thresholding edge detectors. *Proceedings of the IEEE* **67**, 753–763 (1979).

[182] Liu, M.-Y., Tuzel, O., Veeraraghavan, A. & Chellappa, R. Fast directional chamfer matching. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, 1696–1703 (IEEE, 2010).

List of Figures

Publications

- Setola, R., Montoni, L.M. & Zobel, B.B. Motion artifacts compensation in DCE-MRI framework using active contour model. In *Emerging trends in applications and infrastructures for computational biology, bioinformatics, and systems biology: systems and applications,* Chapter 27, 391-409 (Elsevier 2016)

- Montoni, L.M. , Setola, R. Novel tool for 3D analysis and visualization in the DCE framework (in progess)