MRI in Radiation Therapy Planning: The Use of Ultrashort Echo Time Imaging

Hans Martin Kjer

Herlev Hospital

Kongens Lyngby 2012
IMM-MSc-2012

Technical University of Denmark
Informatics and Mathematical Modelling
Building 321, DK-2800 Kongens Lyngby, Denmark
Phone +45 45253351, Fax +45 45882673
reception@imm.dtu.dk
www.imm.dtu.dk IMM-MSc-2012

Summary (English)

Background: Radiation therapy is a technique used in the treatment of a number of cancers. The radiation therapy planning process is currently based on a CT-scan of the patient because these images are geometrically accurate and contain the information of the electron densities required for calculating dose plans.

Radiation therapy based only on MRI would be advantageous in many aspects. The superior soft tissue contrast would for example make the delineations of tumours and organs at risk more accurate. It requires that the information from CT can be estimated from MRI. Conventional MRI sequences lack the ability to visualise compact bone, which hinders proper conversion of MR images into CT images. Visualisation of compact bone is however made possible with a newly installed MRI sequence where images are obtained with ultra short echo times (UTE). The objective of this report is to explore the potential of this sequence for enabling MRI based radiation therapy.

Materials and Methods: It should be established how UTE acquisitions should be conducted at the department's 1 Tesla open MR-scanner dedicated for radiation therapy. A cut-off knee from a calf was scanned with a range of varying parameters. Regions of muscle and bone tissue were manually annotated on the images. Optimal parameters could be investigated by calculating a contrast-to-noise ratio.

Strategies for estimation of CT from MRI should further be explored. The overall chosen strategy was to segment the MRI into different tissue groups and assign an appropriate bulk electron density to each of them with . Four different approaches were investigated and compared to each other. The evaluation was

based both on the geometric and dosimetric accuracy compared to CT, and was tested on both data from the knee-phantom and on human head anatomy with data from a single patient.

Results: It was shown that the first echo time should be minimized, the second echo time should be placed close to 4 ms, and that the flip angle of the sequence was optimal at 25 degrees.

A classification method based on Markov Random Fields was shown to have the overall best performance of the compared methods.

Conclusion: Proper acquisition parameters for UTE imaging were established and a relatively successful segmentation approach was implemented. There is room for improvement both concerning image acquisition and segmentations, but the results are promising and good enough for conducting further studies into radiation therapy based only on MRI.

Summary (Danish)

Baggrund: Stråleterapi er en teknik der indgår i behandlingen af en række kræftformer. Planlægningen af et stråleterapiforløb er i dag baseret på en CT-skanning af patienten, fordi disse billeder er geometrisk præcise og indeholder den information omkring vævets elektrondensitet der muliggør beregninger af stråledosisplaner.

Stråleterapi baseret på MRI ville have en række fordele. F.eks. vil den overlegne bløddelskontrast gøre indtegningen af tumor og risikoorganer mere præcis på stråleplanerne. Det kræver dog, at man kan estimere CT-informationen fra MRI. Konventionelle MRI sekvenser kan ikke visualisere kompakt knogle, hvilket gør det svært at konvertere MR billeder om til CT-billeder på ordentligt vis. Det er imidlertid blevet muligt med en nyligt installeret MRI sekvens, hvor der optages med ultrakorte ekko tider (UTE). Denne sekvens' potentiale for at kunne bruges til MRI baseret stråleterapi skal afdækkes.

Materialer og Metoder: Det skulle undersøges hvordan optagelsen med UTE skal foregå på afdelingens 1 Tesla åbne MR-skanner dedikeret til stråleterapi. Det blev gjort ved at skanne et afskåret kalveknæ med en række varierende parametre. Regioner af muskel og knoglevæv blev manuelt annoteret på billederne. Ved beregning af et kontrast-støj forhold kunne det undersøges hvilke parametre der er optimale.

Metoder til at estimere CT fra MRI skulle desuden undersøges. Den valgte overordnede strategi gik ud på at MRI segmenteres ind i forskellige vævsgrupper som tildeles en passende gennemsnitlig elektrondensitet. Fire forskellige måder at gøre dette på blev undersøgt og vurderet i forhold til hinanden. Vurderingen

var baseret både på den geometriske og dosimetriske akkurathed i forhold til CT-billeder, og blev undersøgt på både data fra kalve-fantomet og på menneskelig hoved anatomi med data fra en enkelt patient.

Resultater: Det blev vist at den første ekko tid bør vælges så kort som muligt, den anden ekko tid bør placeres i omegnen af 4 ms og sekvensens flip vinkel bør være på 25 grader.

En klassifikationsmetode baseret på Markov Random Fields viste sig at klare sig bedre på alle punkter end de andre testede metoder.

Konklusion: Det blev etableret hvordan UTE kan optages på fornuftig vis og en relativ god segmenteringsmetode blev implementeret. Der kan arbejdes på forbedringer både med hensyn til billedoptagelse og segmentering, men resultaterne er så lovende, at videre studier omkring stråleterapiplanlægning udelukkende baseret på MRI kan foretages.

Preface

This thesis was created in the period from February to July 2012 in a collaboration between the Technical University of Denmark, Department of Informatics and Mathematical Modelling and Copenhagen University Hospital, Herlev, Department of Oncology (R).

The work corresponds to 30 ECTS points and is part of the requirements for acquiring an M.Sc. in Medicine & Technology at the Technical University of Denmark (DTU) and Copenhagen University (KU).

Supervisors:
Jens M. Edmund, PhD, DABR
Copenhagen University Hospital, Herlev
Department of Oncology (R)

Rasmus R. Paulsen, Associate Professor, PhD
Technical University of Denmark
Department of Informatics and Mathematical Modelling

Included in the thesis

A H. M. Kjer, R. H. Hansen, R. R. Paulsen, J. M. Edmund. "Optimized Acquisition Parameters for MRI Only RT Using Ultra-short Echo Times", American Society for Therapeutic Radiology and Oncology (ASTRO), 54th annual meeting, 2012.

B H.M. Kjer. "MRI Mimicry of CT". Poster presentation: Visionday 2012, DTU, Denmark.

C J. Edmund, H. M. Kjer, R. H. Hansen, "Auto-segmentation of bone in MRI-only based radiotherapy using ultra short echo time", Radiotherapy and Oncology, volume 103, no. Supplement 1, p. S75, 2012.

Abstract **A** was accepted for a poster presentation at ASTRO 2012. Abstract **C** was accepted for an oral presentation at ESTRO 2012. The contribution of the undersigned to **C** was primarily gathering of data for the presentation.

Lyngby, 27-July-2012

Hans Martin Kjer

Acknowledgements

First, I would like thank the Department of Oncology at Herlev Hospital for enabling this project, providing facilities and access to the MRI-scanner. A special thanks to the radiologic technologists at the department, in particular Annette Kahlen and Pernille Toft-Jensen for helping me gather all the data, and teaching and guiding me through all the practical things concerning MRI. On that notion, I would also like to thank Rasmus H. Hansen for his help with both theoretical and practical MRI questions and issues that were encountered during the project.

Additionally, a great thanks to both of my supervisors, Rasmus R. Paulsen and Jens M. Edmund for their guidance and counselling within their respective fields.

Finally, I would like to thank my colleagues and fellow students for friendly company, much appreciated support and constructive feedback throughout the project.

Contents

List of Figures

List of Tables

List of Acronyms

APPA	Anterior-Posterior, Posterior-Anterior
ASTRO	American Society for Therapeutic and Radiation Oncology
CNR	Contrast-to-Noise Ratio
CT	Computed Tomography
DRR	Digitally Reconstructed Radiograph
DTU	Technical University of Denmark
DVH	Dose Volume Histogram
EQD_2	Equivalent Dose in 2 Gy fractions
ESTRO	European Society for Therapeutic Radiology and Oncology
FID	Free Induction Decay
FOV	Field Of View
GRE	Gradient Echo
HU	Hounsfield Unit

IMRT	Intensity Modulated Radiation Therapy
IQR	Interquartile Range
LINAC	Linear Accelerator
MR	Magnetic Resonance
MRI	Magnetic Resonance Imaging
MRF	Markov Random Field
MU	Monitor Unit
NSA	Number of Signal Acquisitions
OAR	Organs At Risk
PDF	Probability Density Function
PRV	Planned Risk Volume
PET	Positron Emission Tomography
PTV	Planned Target Volume
ROI	Region Of Interest
RF	Radio Frequency
RT	Radiation/Radio Therapy
SNR	Signal-to-Noise Ratio
TE	Echo Time
TR	Repetition Time
TPS	Treatment Planning System
UTE	Ultras short Echo Time
dUTE	difference/dual Ultra short Echo Time
VMAT	Volumetric Modulated Arc Therapy
VOI	Volume of Interest

Chapter 1

Introduction

Radiation Therapy (RT) is a common approach for the treatment of a variety of cancers, and it can be a standalone treatment or a supplement to other treatments. The principle is to damage cancer cells on a cellular level by the use of ionizing radiation. In the western world the most common way of producing and delivering the radiation is with linear accelerators (LINACs).

In principle, all cancer cells can be killed if the dose is high enough. Healthy tissue will also be irradiated causing radiation damage with various side-effects and increased risk of developing secondary cancers years after the treatment. The goal is thus to maximize the dose to the tumour volume while minimizing the dose to healthy tissue - in particular sparing organs at risk (OAR), which are especially vulnerable tissues. To accomplish this goal, an individual treatment plan is made. Currently, the planning process is based on a Computed Tomography (CT) scan. The tumour volume and OARs are delineated on the images, and advanced computer programs are employed to do dose calculations and find optimal plans. Various technological improvements and techniques have been developed to ensure more effective treatments and research constantly expands the knowledge within this field.

One such improvement is to include Magnetic Resonance Imaging (MRI). The soft-tissue contrast of MRI is superior compared to CT (Figure 1.1 left vs. right). This allows for more accurate and precise delineations of the tumour

and OARs[1]. MRI is further capable of a variety of imaging techniques that could be valuable in for instance the assessment of treatment response [2].
To transfer the information gained from MRI to CT, the two volumes have to be registered. This can potentially introduce a systematic error[3]. By increasing the use of supplementary MRI in the RT work-flow, the seriousness of such registration errors will increase, especially if multiple MR images are taken during the course of treatment for response evaluation. Further, it is costly and an added burden for the patient to have them scanned with both modalities.

There is a clear incentive for moving towards RT based solely on MRI. A prerequisite for so-called MRI-only RT is that the information gained from CT can be sufficiently extracted from MRI. This is troublesome to realize due to some inherent MRI issues. First of all, MRI is geometrically less accurate and, secondly, there is no relationship to electron density which then has to be estimated. One approach is to deform a CT atlas to the MR image[4, 5]. Another approach is to segment the MR image into different tissues and then assign them a bulk electron density[6]. The most important part of this bulk strategy is to correctly segment the bone structures. They have an impact on the accuracy of the dose calculations due to their high attenuation compared to soft tissue. Further, Digitally Reconstructed Radiographs (DRR) of the bones are used for set-up verification at the LINACs. Bone segmentation is unfortunately also the most challenging part of the bulk strategy, since signal from cortical bone cannot be acquired with conventional MRI. Manual segmentations of the bone is a possibility but not feasible in real clinical practice[7, 8].

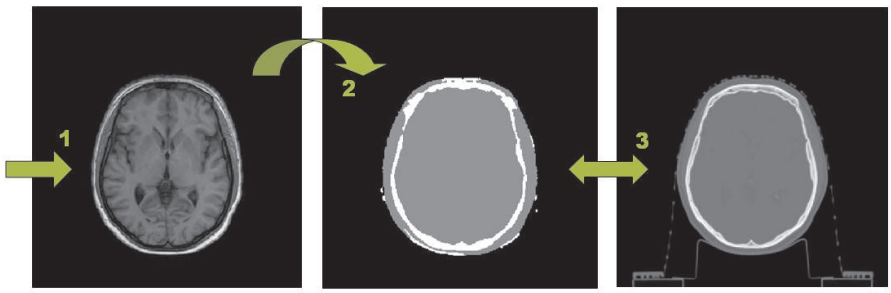


Figure 1.1: Similar transaxial slice of human head. The challenges and objective for the project. **1:** MR image acquisition, **2:** Bulk segmentation of tissues, **3:** Evaluation and comparison to true CT. **Left:** MRI T1-weighted. **Middle:** Bulk segmented MRI/CT-estimate. **Right:** CT.

A key element in realization of automatic MRI bone segmentation is to use a

MRI sequence where images are obtained with Ultra short Echo Times (UTE). The signal from fast relaxing tissues and materials, such as cortical bone, is received with this sequence, which enable ways of auto-segmenting bone structures[9, 10, 11, 12, 13, 14]. The UTE sequence was provided and installed by Philips on some of the MR-scanners at Herlev Hospital prior to this project.

1.1 Objectives

UTE imaging has not been available at Herlev Hospital until recently and there are no protocols and procedures on how to use the sequence. The objective of this report is explore the use of the sequence and test its potential for making MRI-only RT treatment plans. More specifically its ability to automate bone segmentations. This poses several questions that can be addressed as three separate parts (Figure 1.1):

1. How should UTE images be acquired on a 1 Tesla open MR-scanner to give the best bone segmentations? It is important to establish appropriate settings, practical procedures, understand limitations etc.
2. Which bone segmentation approach should be used? How does the performance of various methods compare?
3. How close is the estimation to true CT? In order to asses the clinical potential it is important to evaluate the segmentation's geometric accuracy and the dosimetric agreement with CT-based treatment plans.

The objectives will be investigated mostly with phantom based studies but also with a patient study. A bone segmentation strategy has been implemented, and its performance will be compared to other proposed methods.

The report is structured in the following manner:
A theory chapter that introduces MRI and UTE, classification methods and general concepts used throughout this report. The Materials & Methods in Chapter 3 describe the data acquisition and processing that was done to investigate the objectives. The corresponding results are presented in Chapter 4 and then discusses in Chapter 5. Finally a conclusion in Chapter 6 and an appendix containing three independent scientific projects that was done in relation to and during the course of this project; An abstract accepted for the American Society for Therapeutic and Radiation Oncology (ASTRO) annual meeting 2012; A poster that was presented at the Visionday conference 2012 at the Technical University of Denmark (DTU); And further an abstract accepted

for the annual meeting at the European Society for Therapeutic Radiology and Oncology (ESTRO) 2012.

1.2 Previous Work

A number of papers regarding the three specified objectives have already been published. The work described below serves as the foundation for the strategies pursued in this report.

1.2.1 UTE imaging

Although the UTE sequence is not readily available on a standard clinical MRI-system, it has in fact been present in the literature for over a decade. Previously it was just considered for the standard diagnostic purpose of visualising pathological processes[15, 16]. However the images are generally of low quality compared to other MRI sequences, and concerning the ability to visualise bone, then x-ray and CT perform much better because of their direct visualisation of tissue attenuation properties. Supposedly, it is the reason why UTE did not manage to become part of the standard sequences available. However, the use of MRI has developed since then. MRI is becoming increasingly important in RT[2]. Clinical Positron Emission Tomography (PET)/MRI systems are emerging, but in order to do acceptable attenuation correction maps for PET reconstructions a problem identical to the one in MRI-only RT is faced - how to do CT-estimates from MRI[6]. Since UTE imaging shows a promising potential in this regard, a great incentive for using it has risen.

The experience with UTE imaging that was gathered earlier is valuable. A review article from Robson et al.[17] from 2003 gives an insight into the basic physics and technical aspects concerning the sequence. The paper also introduces a couple of methods for visually enhancing bone in an image, one of which is the difference UTE (dUTE). This strategy appears to be the most widely adopted strategy for UTE imaging currently, and also the strategy employed in this report. No specific parameter settings (echo times, flip angle etc.) can be directly derived from the paper. Going through other papers that have adopted the dUTE strategy[9, 10, 11, 12, 13, 14] there is no consistent choice of UTE acquisition parameters.
We build upon this work by investigating the use of a quantitative Contrast-to-Noise Ratio (CNR) to establish optimal acquisition parameters for our MRI-scanner.

1.2.2 Classification strategies

There are several methods in the literature for trying to estimate a CT image from MRI. Getting correct classification of air and the soft-tissues is less challenging than segmentation of bone. The strategies can broadly be categorised as either an atlas-based or voxel-based approach[6].

Atlas based strategies have been reported by Hofmann et al.[4] and Dowling et al.[5] without using the UTE sequence. The basic principle is to generate both a MRI-atlas and a co-registered CT-atlas from numerous patients. Given a new patient the MRI is deformed to fit the MRI-atlas. The same deformation field is applied to the CT-atlas which results in the desired CT-estimate (dubbed pseudo-CT).
Dowling et al.[5] report results from a study with 26 prostate patients. Segmentations of the pelvic bone with an average Dice coefficient around 0.8 was achieved. However these strategies have difficulties coping with atypical patient anatomy, which is observed in the same study. Two patients were excluded from the study, one simply because of a Body Mass Index much lower than the average.

In this project we consider a voxel-based strategy. The general principle is to segment the MR image into different tissue groups (typically air, soft tissue and bone) and assign them a bulk electron density. Several authors within the last few years have reported various voxel-based strategies using UTE imaging[9, 10, 11, 12, 13, 14]. Prior to these papers UTE voxel-based strategies had not been considered for bone segmentations according to a review article on PET/MRI attenuation correction strategies from 2009 by Hofmann et al.[6].

Thörnqvist et al.[10] demonstrates an approach using standard image filtering techniques combined with masks determined from manual set thresholds to segment cortical bone using boolean logic. The project was made with RT applications in mind. Data was obtained from 4 human volunteers with a 3 Tesla scanner. The focus was head anatomy and visual results were presented. By missing CT information of the 4 volunteers no direct comparison to CT could be done. The average segmented bone volume (530 ± 157 cm^3) were however found to be comparable to a volume segmented from CT-data of 12 patients (555 ± 84 cm^3).

Keereman et al.[11], Catana et al.[12] and Berker et al.[14] use variations of another strategy for bone classification. Berker et al. also include a MRI Dixon sequence. They use various different image analysis techniques (filtering, morphological operations, region growing) and a mathematical operation to calculate a map where voxels with a fast intensity decay are enhanced. A threshold is

used to determine when a decay is considered fast enough to be bone. The drawback of this approach is that intensity relaxation in MRI is not only governed by tissue properties but also inhomogeneities and variations in the magnetic field. These papers consider applications within PET/MRI and the use of 3 Telsa scanners. They favour making their performance assessment in PET relevant manners to demonstrate the clinical impact/feasibility. This makes sense but it makes comparisons to RT approaches difficult. Other than PET relevant evaluations, they employ visual and qualitative descriptions. Keereman et al. reports 5 patient cases where 85%-95% of voxels are correctly classified within a Volume of Interest (VOI) compared to CT. Similarly Berker et al. reports 81.1% overall correct classification based on 6 patients. Details on the bone accuracy are not specified, and the overall classification results are not useful numbers for comparisons in this study.

Johansson et al.[13] use more advanced image analysis techniques for bone classification. A CT and two sets of dUTE MRI images are obtained for a number of patients. The intensities in the MRI images are linked to the intensities of the registered CT-images using a Gaussian mixture model. Given a new MRI image the calculated model parameters are used for estimating a map of CT-values (dubbed substitute-CT). The method appears to yield accurate and robust results, but the paper reports no numbers or values that other methods can easily be compared to. The main drawback of this approach is that registered CT images are required in training of the model. The consequence is a method that lacks flexibility in the sense that the training can only be applied to images acquired in a very specific way.

For a classification strategy in this report, we initially did something very similar to Thörnqvist et al.[10]. This method was used and presented in the abstract for ASTRO (Appendix A). In the final phase of our investigation we implemented an approach with similarities to Johansson et al.[13], but without including the CT information. We only employ UTE images, and consider CT images as something to compare results to.
Here, we test a method that to our knowledge has not been reported previously. In an effort to compare performance of various approaches, we have further implemented versions of the logical masking method proposed by Thörnqvist et al.[10], and the R2-mapping approach proposed by Keereman et al.[11]. These approaches are based solely on UTE images as well, making it easy to do relatively fair comparisons of performance.

1.2.3 Dosimetric Evaluation

A prerequisite for pursuing MRI-only RT planning using bulk density assigned MRI is to know that it can eventually result in clinical acceptable plans. Other authors have already investigated this[7, 8, 18, 19] by comparing dosimetric agreements between CT-based plans and $\mathrm{CT_{Unit}}$, $\mathrm{CT_{Bulk}}$, $\mathrm{MRI_{Unit}}$, $\mathrm{MRI_{Bulk}}$ plans. Here *Unit* means that all tissue is assigned a uniform density corresponding to water, and *Bulk* refers to bulk assigned densities to soft-tissue and bone. The bone segmentations on MR images have in these cases been manually segmented or transferred from CT.
The conclusion from these studies is that MRI-based treatment plans can be clinically acceptable, and that *Bulk* plans are generally more dosimetric accurate than *Unit* plans.

Our contribution in this regard is to evaluate dosimetric agreement for bulk MRI plans where bone has been automatically estimated. The expectation is that the result will be more dosimetric accurate than $\mathrm{MRI_{Unit}}$ plans, but less accurate compared to $\mathrm{MRI_{Bulk}}$.

Chapter 2

Theory

2.1 Magnetic Resonance Imaging

A full description of general MRI theory is beyond the scope of this report. However, the UTE sequence is a rather special kind of sequence, and there are some important points that should be stated. For this report the work of Nishimura[20] and Hanson[21] has served as the theoretical foundation for understanding the basics of MRI.

Cortical bone has a T2 relaxation of ≈0.5 ms at 1.5 Tesla[17]. This is an ultra fast decay rate compared to the echo times (time from excitation to signal readout) employed with conventional MRI sequences. This is why cortical bone is not normally seen on MR images, the bone signal is decayed before signal readout. In order to achieve echo times short enough to pick up the signal from cortical bone, all time consuming steps of the acquisition have to be considered.

• **Excitation:** The radio-frequency (RF) excitation pulse is normally assumed to be instantaneous because the tissue relaxation is negligible during the pulse duration. This assumption does not hold for imaging of tissues with short T2. The increased signal gained by spending more time reaching a larger flip angle is simply lost due to relaxation. The implication is that flip angles should be kept small, which excludes the possibility of manipulating short T2 tissues with

180° pulses required in Spin Echo sequences.

• **Post-excitation:** The signal should be sampled as quickly as possible after excitation.
- There is a hardware lower limit called the coil tune delay. It is the time it takes to switch coils form transmit to receive mode[22].
- There is no time to recall and sample an echo. Instead the Free Induction Decay (FID) is sampled. Although the UTE sequence is not in principle a Gradient Echo (GRE) sequence, it is still referred to as one[17], since it shares some similarities and drawbacks of GRE sequences.
- The signal has to be sampled even during the ramp up of the readout gradients. Usually the ramp up shape is irrelevant and considered ideal. If the correct shape and timing are not taken into account, the consequence is that samples in k-space are slightly misplaced, which results in blurring and spatial distortions of the image[23]. Adjustments are made with a mathematical correction during image reconstruction. The correction is governed by a single parameter called the *trajectory delay*, τ. It is affected by the echo time, the bandwidth and the choice of Field of View (FOV). It is therefore necessary to calibrate the value of τ for different settings.
- The fastest sampling of the k-space is achieved with a readout direction that is radially outward from the centre of k-space. The consequence is that the lower frequencies (centre of k-space) are more densely sampled than higher frequencies. This corresponds to a low pass filtering of the image, which is further enhanced due to signal loss during signal readout. When the higher frequencies are sampled a degree of signal loss has already occurred[22].

• **Artefacts:** Similarly to GRE sequences the UTE sequence suffers from the following artefacts.
- Susceptibility: UTE imaging concerns tissues that decay rapidly. However the decay is not only governed by intrinsic tissue properties (T2) but also by sources that de-phase the signal due to variations in the local magnetic field (T2*). Such susceptibility artefacts are especially pronounced at air-to-tissue boundaries.
- Chemical shift: Water and fat have slightly different Larmor frequencies. At certain echo times (≈3.4 ms at 1 Tesla) the two signals are out of phase. Voxels with a mix of water and fat will as a consequence have no measurable signal.

• Bandwidth should preferably be broad[17]. Although this lowers the Signal-to-Noise ratio (SNR) it reduces chemical shift artefacts and read-out times.

• Other tissues have short T2s. Short T2 tissue is present as a minority in most tissues, but is the majority in for example ligaments, tendons, cortical bone. These tissues will show a similarly behaviour in UTE images[15].

2.2 UTE imaging strategy

By acquiring a single UTE echo (Echo Time, TE $\approx$ 0.1 ms) all tissues should have a measurable signal. This allows a separation of cortical bone and air, which is not possible using conventional echo times. A distinction between cortical bone and other tissues is, however, difficult unless something extra is added to the sequence. There are several approaches [17], but the one used for this project is the difference UTE or dual UTE. A second echo (Echo 2) is acquired shortly after the first (Echo 1). Because tissues with short T2 will have experienced a larger signal loss in the time between the echoes, they will appear bright in a subtraction image (Echo 1 - Echo 2). This principle is demonstrated in Figure 2.1, and the following can be observed:

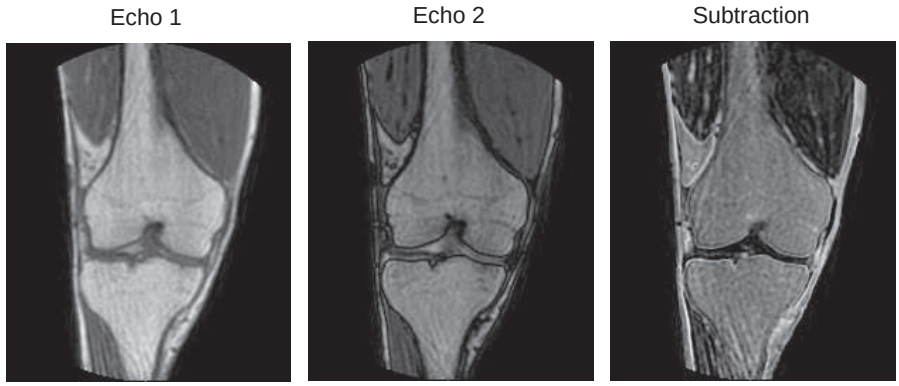


Figure 2.1: UTE echoes and subtraction image of human knee joint

- **Air** has low intensity in both echoes. The intensity is not necessarily zero, in particular close to the tissue interfaces. This can partly be contributed to the partial volume effect, in the sense that signal is picked up from a voxel that contains a mix of tissues and the recorded intensity is then some kind of average.
- **Cortical bone** behaves as expected. Signal is present in the first echo due to the ultrashort TE, but it is gone in the second. This results in a high intensity in the subtraction image, which is visible as a thin bright contour around the bone.
- **Spongy bone** has high intensity in both echoes. A significant signal loss is still observed resulting in a bright appearance in the subtraction image as well.

- **Muscle tissue** appear with very similar intensity in both echoes and is clearly suppressed in the subtraction image. Notice that voxels at the interface with other tissues have lost most of their intensity in echo 2. This can be caused by chemical shift or susceptibility artefacts, and results in bright contours at tissue interfaces - similar to cortical bone. Distinguishing between signal loss due to intrinsic tissue properties (T2) or T2* effects is not possible using dUTE.
- **Other soft tissues** such as ligaments and the skin behave much like spongy bone. Tendons will show the same behaviour [15]. Distinguishing between these tissues is expected to be difficult.
- **Noise:** Although not shown, some voxels have a greater intensity in echo 2 which would result in a negative intensity in the subtraction image. Only noise and artefacts can cause such a gain in intensity.

The information can also be displayed in an intensity map (Figure 2.2). A specific voxel is interpreted as a 2-D point (x,y) using the intensities from the echoes (I_{E1}, I_{E2}). The spatial information in the images is lost, but the different tissue species are expected to occupy different regions in accordance with the above stated observations. The degree of separation between tissue groups has an impact on the performance of the classification methods and the accuracy of the CT-estimates. The actual scaling and appearance of the intensity map is influenced by almost everything - the magnetic field strength, choice of coil, acquisition parameters and imaged anatomy.

The intensity map, the color coding and the straight line showing no signal loss between echoes in Figure 2.2 will be used throughout the theory chapter.

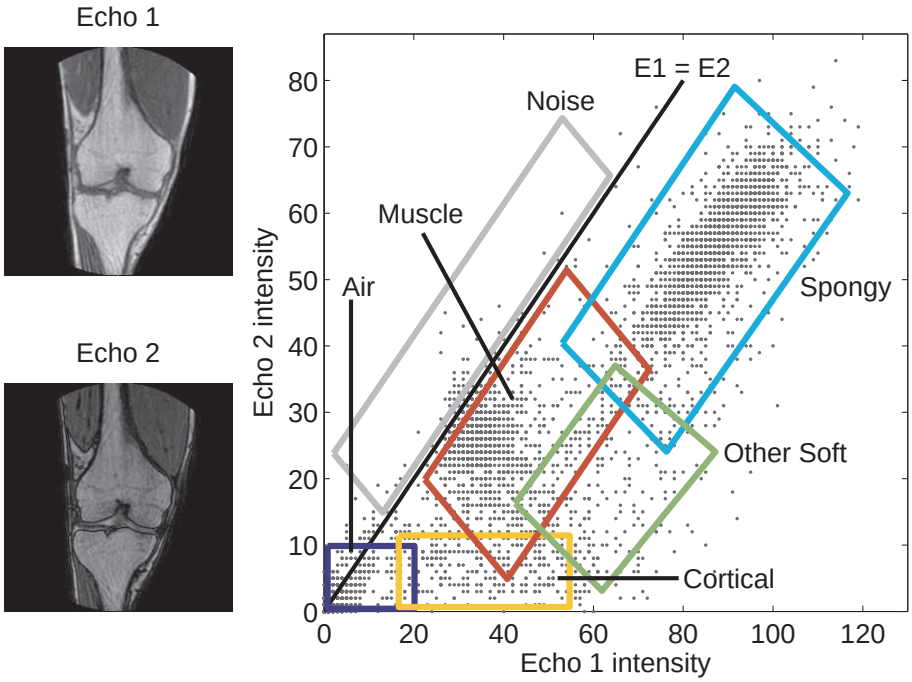


Figure 2.2: UTE intensity map. Distinction between air and cortical bone is only possible due to the UTE. Color coding: Air (dark blue), cortical bone (yellow), muscle tissue (red), other soft tissues (ligament, tendons etc) (green) and spongy bone (light blue).

2.3 Contrast to Noise Ratio

The acquisition parameters have an impact on the separation of the different tissue classes on the intensity map in Figure 2.2. A metric is needed for a quantitative evaluation of how well the tissue types are distinguished. The ability to identify a certain tissue group is dependent on two things - how well is the group defined and how different is it from other groups.

The former is typically defined as the Signal-to-Noise Ratio (SNR). In image analysis the SNR is often defined as the ratio of the mean pixel intensity to the standard deviation of a region of interest (ROI). There are variations to this depending on the application. However, just because an image has a high SNR, it does not imply that different tissue regions are easily distinguished between. This is illustrated in Figure 2.3 with an artificial image containing two separate rectangular ROIs. The image to the left has high SNR in both regions, however the contrast is much better in the image to the right in spite of low SNR. When the contrast is also an important aspect of the metric, then the contrast-to-noise ratio (CNR) is more appropriate. This metric takes the signal difference of the ROIs into account:

$$\mathrm{CNR} = \frac{|S_A - S_B|}{\sigma} \tag{2.1}$$

where S_A and S_B are the signal strength of the two regions, and σ is the noise. For the remainder of this report the signal strength is defined as the median pixel intensity and the noise is defined as the root of the added squared interquartile ranges (IQR) of the two regions ($\sigma = \sqrt{\mathrm{IQR}_A^2 + \mathrm{IQR}_B^2}$).

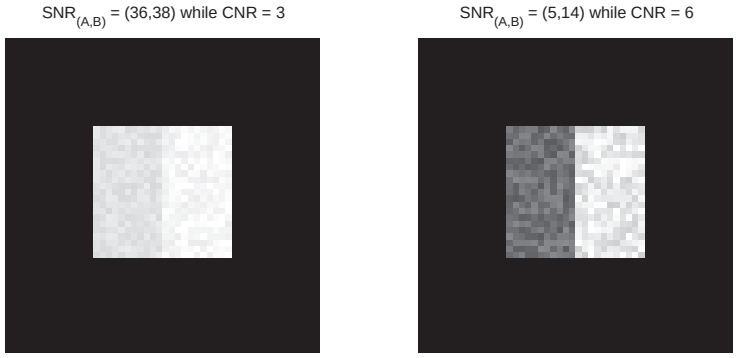


Figure 2.3: Two images each containing two homogeneous regions (A and B) with added noise of identical strength.

2.4 Classification

The output of the MRI dUTE sequence is two image volumes as illustrated in Figure 2.4. The task is to create a volume where all voxels are assigned to a group that represents a specific Hounsfield Units (HU). A minimum of 3 groups are considered - air (-1000 HU), soft tissue (0 HU) and bone (500-2000 HU). This image volume is the CT estimate, $\widetilde{CT}$.
$\widetilde{CT}$ is obtained by using a voxel classifier. Given the information in the image and perhaps some a priori knowledge, the classifier is a method for assigning each voxel into one of K different classes. This will be referred to as the label image/volume. Each class or label corresponds to a tissue type that appears distinctively different in the MR images. The main differences between the $\widetilde{CT}$ and the label image is that the soft tissue class in the $\widetilde{CT}$ may be comprised of several different tissue classes in the label image. For instance, white and grey matter can appear very different on MRI, and it may be most appropriate to have a separate label for each of them. Their CT-numbers are, however, similar and should appear identical on the $\widetilde{CT}$.

There are many different classification approaches that could be considered. The methods that are tested in this project are described in the following sections. The various strategies are applied to the artificial UTE images and intensity map shown in Figure 2.4.

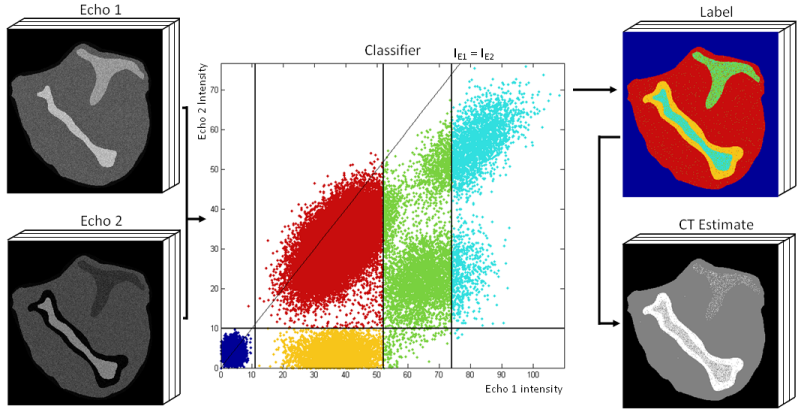


Figure 2.4: Illustration of the classification process. Color coding as in Fig. 2.2. The decision boundaries (thick black lines) of the classifier divides the intensity map into regions and each voxel is assigned the corresponding label of that tissue group. Each label is then assigned an appropriate electron density (or CT-number).

2.4.1 Classification using R2-mapping

The approach with this strategy is to estimate the relaxation time constant, T2, for each voxel. The strategy was proposed by Keereman et al.[11] and developed for head anatomy. The intensity in a voxel should ideally drop exponentially:

$$I(t) = I_0 \exp(-\mathrm{R2} \cdot t)$$

where $\mathrm{R2} = 1/\mathrm{T2}$. The intensities of a voxel in the two echo images (I_{E1}, I_{E2}) can be considered two measurements of the function. The solution for R2 is then:

$$\mathrm{R2} = \frac{\ln(I_{E1}) - \ln(I_{E2})}{\mathrm{TE}_2 - \mathrm{TE}_1} \tag{2.2}$$

This calculation can be done for every voxel and plotted as an image with R2 values (R2-map). R2 is a tissue dependent factor and cortical bone has a higher R2 than soft tissues since it decays much faster. A threshold based on measurements of the T2 of cortical bone is used to determine when the decay has been fast enough to be classified as cortical bone.
Echo 1 is used for classification of air. A region growing algorithm is applied to define all air around the scanned object based on a manual set threshold, Th_1. Voxels within the object are classified as air only if the intensity is lower than another manual set threshold, Th_2 ($Th_1 > Th_2$).
In comparison to the UTE intensity map from Figure 2.4, this method defines its decision boundaries as shown in Figure 2.5.

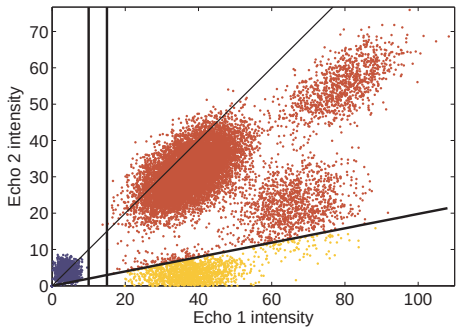


Figure 2.5: R2 classifier boundaries. Color coding as in Figure 2.2.

.

2.4.2 Classification using masks and boolean logic

This approach is based on the a priori knowledge regarding the echo images as described earlier. The method was proposed by Thörnqvist et al.[10] and was developed for head anatomy.

- Only voxels of air will ideally have zero intensity in echo 1 .
  - Lemma 1: Everything above a relatively low threshold is tissue.

- The signal from the cortical bone has ideally decayed to nothing in echo 2.
  - Lemma 2: Everything above a threshold is not cortical bone.

- The short T2 of the bone results in a high intensity in the subtraction image.
  - Lemma 3: Everything below a relatively high threshold is not bone.

The three appropriate thresholds are found manually and defines three masks (M1, M2, M3). Classification is done using boolean logic as follows:

- Air = $\neg$M1.

- Soft tissue = M1 $\wedge$ M2 $\wedge$ M3.

- Bone = M1 $\wedge$ $\neg$M2 $\wedge$ $\neg$M3.

where $\neg$ is negation and $\wedge$ is the conjunction. In an effort to reduce the noise from air-tissue interfaces an edge detection filter is applied to echo 2. The detected edges are widened and used for a reduction of the mask that defines bone. In relation to the intensity map from Figure 2.4, the decision boundaries of this approach are placed as illustrated in Figure 2.6.

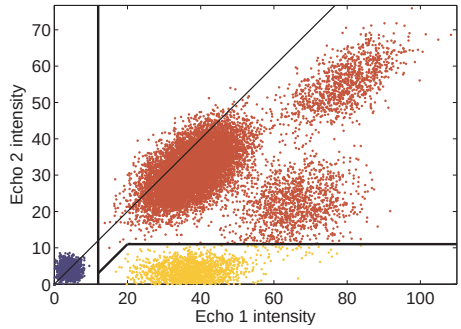


Figure 2.6: Logic classifier boundaries. Color coding as in Figure 2.2.

2.4.3 Classification using Bayesian statistics

Bayes' theorem is a very generalized statement that has found applications in a variety of fields. It basically states that when confronted with some new objective information, the initial (the prior) state of knowledge or belief should be updated to a new and improved state (the posterior). It is contributed to Thomas Bayes, but was mostly developed by Laplace[24]. The concept has also found its way into image analysis and image classification as described below[25, 26].

The basic assumption is that the acquired intensities in an image volume actually comes from a mixture of K normal distributions. The measured intensity in each voxel is interpreted as a random sample drawn from one these distributions. Samples are drawn with different frequencies which represents the prior probabilities. If the image contains, for example, 38% background/air, then 38% of the voxels would have been random samples from that distribution.

The K Guassian distributions are described by a n-dimensional mean value, $\mu_{\mathbf{k}}$, and a n-by-n covariance matrix, $\mathbf{\Sigma_k}$. The dimensionality is 2 in this report, since only intensities from the two UTE echoes are ever considered. However, this method is easily expanded to include intensities from other MR images as well.
Suppose that the parametric properties ($\mu_{\mathbf{k}}$ and $\mathbf{\Sigma_k}$) of the K Gaussian distributions and the priors are somehow known. Given a voxel with intensity, $\mathbf{x} = [x_1, \ldots, x_n]$, it is then possible to estimate the posterior probability, $P(k|\mathbf{x})$, that it was drawn from the k'th distribution using Bayes' rule:

$$P(k|\mathbf{x}) = \frac{P(\mathbf{x}|k)P(k)}{P(\mathbf{x})} \tag{2.3}$$

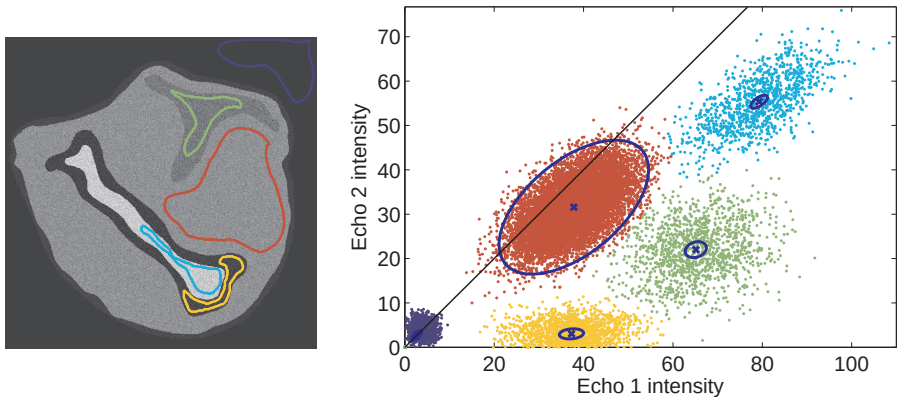
On the right hand side the term $P(k)$ is the prior probability as mentioned above. It is independent of the intensity, and in its most basic form it serves as a scaling. The term $P(\mathbf{x})$ is a normalising constant. The term $P(\mathbf{x}|k)$ is the conditional probability and it states the probability that the intensity value $\mathbf{x}$ is in the k'th class. Still assuming that the classes are normally distributed the conditional probability is approximated by evaluating the intensity in the Probability Density Function (PDF):

$$f_{\mathbf{x}}(x_1, \ldots, x_n) = \frac{1}{(2\pi)^{n/2}|\mathbf{\Sigma_k}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{x} - \mu_{\mathbf{k}})^T \mathbf{\Sigma_k}^{-1}(\mathbf{x} - \mu_{\mathbf{k}})\right) \tag{2.4}$$

The posterior probability is evaluated for each of the K classes, and the voxel is classified according to the class with the highest posterior probability. The nor-

malising term in Equation 2.3 ensures that the sum of the posterior probabilities equals to one.

The process is repeated for all voxels, and the resulting intensity map can be seen in Figure 2.7(b).



(a) Manual training (b) Bayes classifier

Figure 2.7: Manual training of the classifier (left) and the labelled intensity map and estimated Gaussian distributions representing tissue classes (right). Color coding as in Figure 2.2.

2.4.3.1 Training

Previously, it was assumed that the properties of the Guassian distributions were somehow known. Training is the process where the parametric properties are defined and it can either be manual (supervised) or automatic.

Supervised The manual training approach is illustrated in Figure 2.7(a). A representative ROI for each tissue class is annotated, the $\mu_{\mathbf{k}}$ and $\mathbf{\Sigma_k}$ are calculated and an appropriate electron density is assigned to the class.

The annotation of a bone ROI on UTE images would for example contain N_k voxels providing the observations:

$$\mathbf{x}_k = (I_{E1}, I_{E2}) = \begin{pmatrix} I_{E1-1} & I_{E2-1} \\ \vdots & \vdots \\ I_{E1-N_k} & I_{E2-N_k} \end{pmatrix}$$

The mean point or class centre is then calculated as

$$\mu_{\mathbf{k}} = (\mu_{E1}, \mu_{E2}) = (\mathbf{E}[I_{E1}], \mathbf{E}[I_{E2}])$$

The covariance is a 2-by-2 matrix whose element (i,j) is defined as

$$\mathbf{\Sigma_k^{i,j}} = \mathrm{cov}(I_{Ei}, I_{Ej}) = \mathbf{E}[(I_{Ei} - \mu_{Ei})(I_{Ej} - \mu_{Ej})]\frac{N_k}{N_k - 1} \tag{2.5}$$

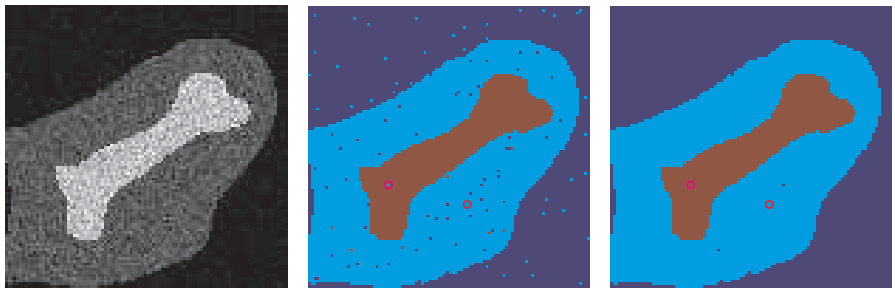
and $\mathbf{\Sigma_k^{i,j}}$ should be evaluated for $i = 1:2$ and $j = 1:2$.

Automatic There are different algorithms that can automatically estimate the underlying Gaussian distributions of a set of points. For this report an implementation of the Expectation Maximization algorithm[27] for Matlab was found. A detailed explanation of the algorithm is beyond the scope of this report, but it should be stated that this particular implementation uses a random sampled subset of the points. The consequence is that two separate trainings can end up with two different results. The algorithm will converge to similar results if enough points are included and if the convergence criterion is small enough.
Using an automatic training presents another challenge. There is no immediate way of knowing what tissue is represented by which class, and therefore no knowledge of the correct electron density. The assignment of electron densities is done separately after the training is completed. The established class centres, $\mu_{\mathbf{k}}$, are evaluated according to the a priori knowledge presented earlier. A class is assigned a density corresponding to air if $\mu_{\mathbf{k}}$ is smaller than some threshold. A class is defined as spongy bone if $\mu_{\mathbf{k}}$ is larger than a threshold in echo 1 intensity. Similarly, if $\mu_{\mathbf{k}}$ has a R2-value (Equation 2.2) larger than a threshold it is considered cortical bone. Otherwise classes are assumed to be soft-tissues.

2.4.4 Classification using Markov Random Fields

Markov Random Fields (MRF) is a very general term with applications in a variety of fields. The term Markov originates from probability theory where it is used to characterise a certain type of stochastic process or system. The system contains multiple states, and if the transition from the current state to the next is dependent only the current state - it is said to full-fill the Markov property. The concept of Markov Random Fields is the same, just expanded into multiple dimensions, and in image analysis it is thus a way of incorporating that pixels/voxels are spatially dependent on each other.

The approach chosen in this project can be seen as an expansion of the Bayesian classifier. In the Bayes approach all voxels were treated independently. It is a



(a) Medical image (b) Bayes (c) MRF

Figure 2.8: Left: A artificial medical image showing the bone tissue as high intensity, the muscle as tissue mid level intensity and the dark background represents air. **Middle:** The result of Bayes classification. The two encircled pixels have identical intensities. **Right:** The result with the MRF prior added to the classifier. The marked pixel in bone has changed classification due to the influence of the local neighbourhood.

simplifying and naive assumption considering the nature of medical images. A certain tissue type will always be part of a smaller or larger connected region (organ). The concept of using MRF is shown in Figure 2.8. An artificial medical image containing air, muscle and bone tissue is shown in Figure 2.8(a). Figure 2.8(b) shows the result of Bayes classification. The two encircled pixels share the same intensity, giving them a [0, 51, 49]% posterior probability of belonging to respectively air, muscle and bone. Hence they are both classified as muscle tissue, but when considering the neighbourhood it is most reasonable to assume that the pixel embedded in bone is not classified correctly. The MRF approach provides a method for changing the posterior probability of each individual pixel by looking at the surrounding neighbourhood. If many neighbours belong to the same tissue type it would strengthen the posterior probability and vice versa. The result of using MRF is shown in Figure 2.8(c), and the posterior probability of the two encircled pixels are now respectively [0, 99, 1]% and [0,1,99]%.

In a more mathematical sense, the knowledge of spatial dependency is incorporated into the prior, $P(k)$, of Bayes' theorem as an energy function, $E(k)$:

$$P(k) = \frac{1}{Z}\exp(-E(k))$$

Where Z is a normalising term. The energy function can be specified in many ways. For this report the chosen energy function favours labels which are spatially clustered, in the sense that $E(k) \to 0$ if a high number of neighboring voxels are classified in the same manner, and $E(k) \to \infty$ otherwise. Because

voxels can now influence each other the prior is no longer considered on a global scale. It can be shown that it is possible to define the prior for the k'th class for each individual voxel, i[26]:

$$P_i(k) = \frac{\pi_k \exp\left(-\beta \sum_{j \in \Re_i} (1 - q_j(k))\right)}{\sum_{k'} \pi_{k'} \exp\left(-\beta \sum_{j \in \Re_i} (1 - q_j(k'))\right)} \tag{2.6}$$

The first term, π_k, is the constant frequency prior from the Bayes classifier. The denominator is for normalising the prior probabilities to ensure that $\sum_k P_i(k) = 1$. The constant β is the weighting of neighbourhood influence. If $\beta = 0$ then the equation will reduce to the constant prior used in Bayes. The term $q_j(k)$ is the current posterior probability of voxel j belonging to class k. The summation is then the summing of probabilities of not belonging $(1 - q_j(k))$ to class k over the voxels, j, in a neighbourhood, $\Re_i$, around voxel i. In a soft-sense it is the counting of the number of neighbours current classified as something else. In 2-D images the neighbourhood is considered as the 4 or 8 nearest pixels. In 3-D volumes the nearest 6, 18 or 26 voxels can be defined as the neighbourhood.

Every time the posterior probability of the voxel i is updated using the MRF prior it will have an impact on the neighbouring voxels are vice versa. All voxels should be updated several times to ensure convergence, and neighbouring voxels cannot be updated simultaneously. Further, the order of which voxels are updated can affect the final result. This iterative nature of the approach has a drawback. The computational time is significantly increased compared to the other methods. The principle on Markov Random Fields can also be illustrated on the intensity map (Figure 2.9) previously introduced.

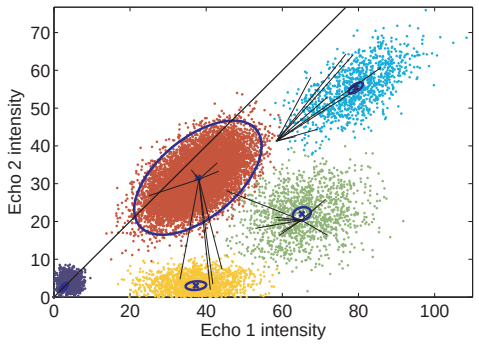


Figure 2.9: The principle of Markov Random Fields. Three pixels and their eight spatial neighbours are linked. A voxel can be assigned to another class based on how its spatial neighbours are defined.

2.5 Geometric Evaluation

The performance of the classifiers can be evaluated on how correctly the bone was segmented compared to a corresponding CT-scan. Geometric evaluation is in this report done quantitatively with the Dice similarity coefficient and qualitatively with Digitally Reconstructed Radiographs (DRRs).

2.5.1 Dice coefficient

The Dice similarity coefficient originates from ecology studies[28], but is a general method, that in this case is used for comparison of segmented volumes. Given the intersection volume (A $\cap$ B) and the individual volumes A and B, the Dice coefficient is calculated as:

$$\text{DICE} = \frac{2(A \cap B)}{A + B} \tag{2.7}$$

The coefficient can be anything between 0 (no overlap) to 1 (perfect overlap). In the specific case for this report A would be the bone volume from a CT image, while B would be the bone volume from the segmented MR image. CT is considered the ground truth, and the interpretation is that the higher the Dice coefficient the better the MRI segmentation. Failing to get a Dice coefficient of 1, which is always the case, the Dice coefficient itself does not reveal whether this is caused by a low intersection volume or by over-segmentation of bone on the MRI. This is illustrated in Figure 2.10, where two very different segmentations have the same Dice coefficient. Therefore a further clarification to the geometric evaluation can be to report the fractions of CT missed and MRI falsely classified volume.

2.5.2 Digitally Reconstructed Radiograph

The 2-D x-ray image of an object from a given angle can be approximated from a CT-scan of the object. The so-called DRR (Figure 2.11) is used for set-up verification at the LINACs prior to the treatments. A corresponding image can be obtained with the LINAC. A manual registration between the two images can fine tune the position of the patient to match the expected position in the dose plan. Two orthogonal DRRs are acquired in order to cover for all three dimensions. Only high dense materials are seen on these x-ray images, so matching is most often based on bone structures.

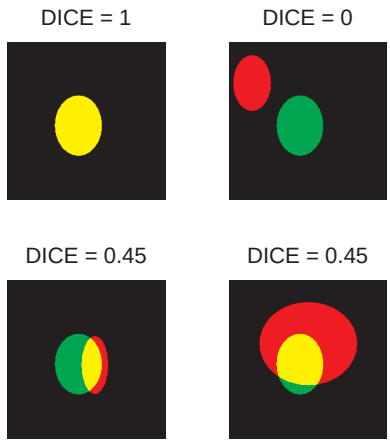


Figure 2.10: Illustration of Dice coefficient. The same Dice coefficient can have different causes.

It is crucial that decent DRRs can be generated from MRI-scans in order to realize MRI-only RT. This requires truthful bone segmentations. A visual comparison of the DDR generated from CT and from bulk segmented MRI is therefore a clinical relevant evaluation method.

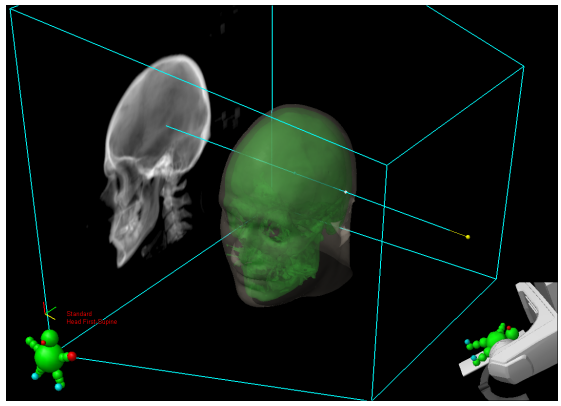


Figure 2.11: Illustration of DRR generation

2.6 Dosimetric Evaluation

The performance of the classifier can further be evaluated by comparing dose plans calculated on CT- and on bulk segmented MR images. Realization of MRI-only RT requires minimal deviation from the CT based dose plan, and it is further important to evaluate if the segmented *Bulk* plans are more accurate than the MRI *Unit* plans. A full description the of dose planning process is beyond the scope of this report, but some relevant concepts are shortly introduced below.

2.6.1 Dose Volume Histograms

The output of a dose plan (Figure 2.12) is a rather complex set of 3-D informations. Visual browsing through slices can reveal hot-spots or other regions that require special attention. It is, however, a large amount of spatial information, and in order to do a more quantitative evaluation and make comparisons to other plans the information is reduced to something simpler. The dose volume histogram (DVH) is a way of representing the 3-D dose distributions in 2-D graphs[29]. Similar to standard image histograms, the DVH is created by counting the number of voxels in a defined structure (target volume for example) that receive a dose that falls within a small dose-range (bin). It is most common to show the DVH in its cumulative form (Figure 2.12 top right), so that the bin-height represents how many voxels that receive that dose or higher. Unless otherwise stated DVH refers to the cumulative DVH in this report.

Although the spatial information is lost in the DVH, analysis of the 2-D graphs can reveal other useful information. The gradient of the slope in the DVH represents the dose homogeneity in the volume. Specific DVH points have been shown to be of clinical relevance for the outcome of the treatment and the occurrence of side-effects.

The comparisons in this report are based on the similarity of the DVHs and deviations in specific DVH points. The CT based plan is considered the truth and as such the MRI bulk segmented plans should get as close as possible to the CT plan.

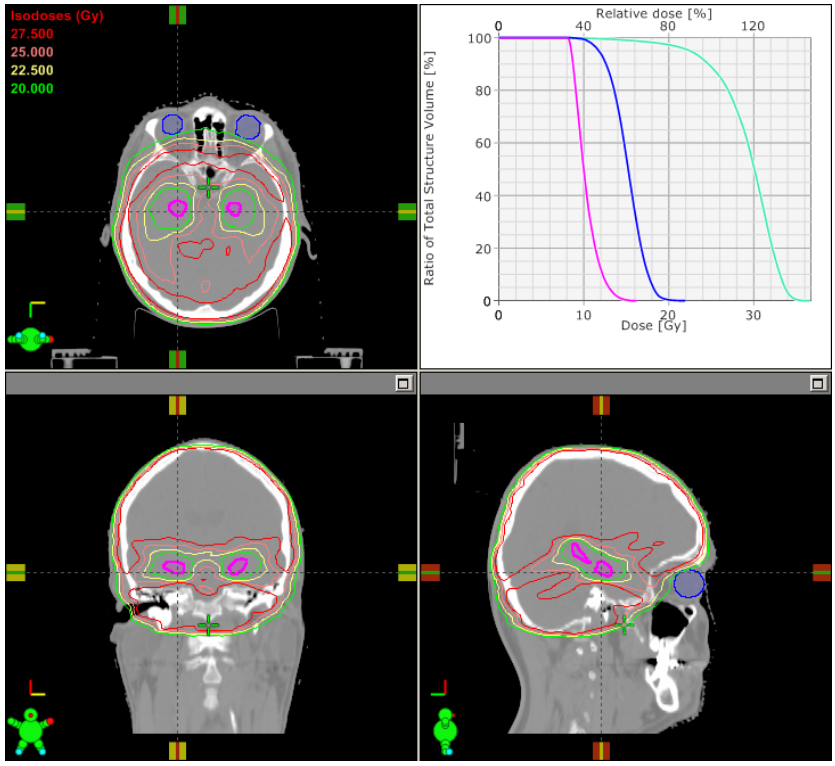


Figure 2.12: Illustration of dose plan and DVH. Color coding of structures: Hippocampus (pink), Eyes (blue) and Brain(light blue)

2.6.2 Equivalent dose in 2 Gy fraction

The doses presented in clinical dose plans are rarely given in a single treatment. RT is more effective when it is fractionated, but the fractioning makes intercomparisons of dose plans a bit troublesome. A 20 Gy dose delivered in 4 fractions of 5 Gy is not biologically equivalent to a 20 Gy dose delivered from 10 fractions of 2 Gy. An approach for standardising reported doses is to use a mathematical model of isoeffective dose relationships to calculate the equivalent dose in 2 Gy fractions (EQD_2)[30, p. 126]:

$$\mathrm{EQD}_2 = D\frac{d + \alpha/\beta}{2 + \alpha/\beta} \tag{2.8}$$

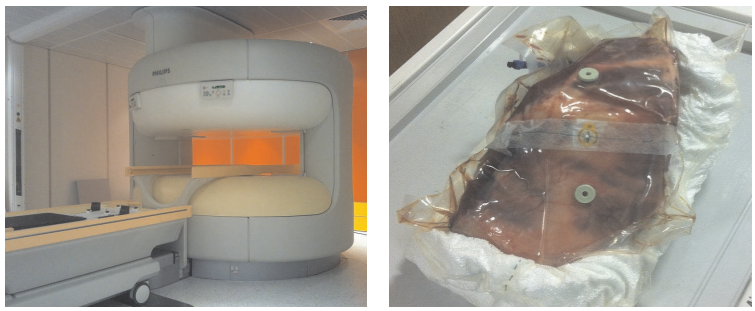
Where D is the total dose, d the dose rate and α/β is a tissue dependent parameter that describe the sensitivity to the dose rate. A lower α/β equals a higher sensitivity. The unit of EQD_2 is Gy, but for clarification it can be reported as $\mathrm{Gy}_{\alpha/\beta}$.

The dosimetric evaluation in this project involves a head patient. One of the OARs in the brain is the hippocampus, and a study has shown that a EQD_2 greater than $7.3\,Gy_2$ to 40% of hippocampus correlates with long-term memory impairment[31]. The doses from the calculated dose plans are converted to EQD_2 by measuring the hippocampus DVH point $D = D_{40\%}$. Given the number of fractions, n, the dose rate can be calculated ($d = D/n$). α/β is assumed to be 2, and the EQD_2 can then be calculated using Equation 2.8.

Chapter 3

Methods & Materials

The details concerning data acquisition and processing are described in the following sections. Unless otherwise stated all MR images were acquired with a Philips Panorama 1 Tesla open MRI-system, which is shown in Figure 3.1(a). All CT-scans were obtained with a Philips Brilliance Big Bore CT.



(a) Panorama (b) Bovine phantom

Figure 3.1: MRI system and bovine phantom

All obtained data is logistically considered as belonging to three separate studies.

The first study referred to as **UTE calibration** concerns the measurements

required for establishing appropriate values for the trajectory delay, τ, which was introduced in Chapter 2. The scanned object was a 3-D Philips phantom that is otherwise used for measurements of geometric distortions.

The objective of the **parameter study** is to determine which parameters UTE images should be obtained with on the 1T open MRI. This is done by measuring a CNR for UTE images acquired with varying parameters. The study is based on images of a bovine phantom. The phantom (Figure 3.1(b)) is a cut-off knee of a calf, vacuum packed and set-up in a vacuum fixation.

Some of the bovine data was further used in the **classification study**. The purpose here is manifold. First of all, it must be investigated if the CNR measure correlates with the performance of the classifier and the dosimetric accuracy of the segmentations. Secondly, the performance of various segmentation methods is investigated both on bovine data and images of a patient. The compared methods are the Bayesian classifier (*Bayes*), the Bayesian classifier with Markov Random Fields (*MRF*), classification based on measured R2-values[11] (*R2map*) and finally the classifier using logical masks[10] (*Logic*).

3.1 Image Acquisition

Table 3.1 presents an overview of the various scanning parameters used for the UTE scans in the three studies. The Field of View (FOV) and voxel size are both isotropic meaning that the listed values are identical in all three dimensions. The Number of Signal Acquisitions (NSA) is a parameter that control how many times the same measurement is obtained, and the final outcome is then an average of the acquisitions. It is a way of increasing the SNR but also the total scan time. The bandwidth was kept as wide as possible throughout all scans.

3.1.1 Phantom

The phantom used for UTE trajectory delay calibration was primarily imaged using the Head-coil. The Large-coil was employed when large FOVs (>300 mm) were investigated, and the Flex-coil was the only coil able to test for the currently lowest achievable TE1. Only the first UTE echo is of interest and the second echo was therefore not obtained. The trajectory delay was investigated for FOV sizes of 190:10:250 (from 190 mm to 250 mm with increments of 10

Table 3.1: MRI UTE acquisition parameters for various scans. * represents a parameter variation

	Calibration phantom	Bovine phantom				Patient Head
		TE1	TE2	α	Cross	
TE_1[ms]	0.14	*	0.11	0.11	0.11	0.09
TE_2 [ms]	–	3.6	*	3.2	*	3.6
α [°]	15	25	10	*	*	25
TR [ms]	6	6.9	9.3	6.5	9	9
FOV [mm]	*	220	220	220	220	220
Voxel size [mm]	1.5	1.6	1.6	1.6	1.6	1.5
NSA	2	1	1	1	1	1
Coil	*	Head				Flex

mm) and additionally 350 mm and 450 mm. The effect of changing the first echo time was further investigated with TE_1 of 0.09, 0.11 and 0.14 ms.

3.1.2 Bovine phantom

The bovine phantom was scanned only once with CT. The MRI data included in this report contains variations of the first echo time (TE1), the second echo time (TE2), the flip angle (α) and variation of both flip angle and second echo time (Cross). Each of the 4 data sets were obtained at four separate sessions. The head coil was used because the phantom fitted nicely into it and because it allowed for a quick and very reproducible set-up. Except for the TE2-session each measurement was repeated a few times to get a sense of the reproducibility and the acquisition order was randomized to negate unforeseen sources of variation such as gradual heating.

- **TE1**: Echo times of [0.11, 0.16, 0.21, 0.26] ms were investigated with four repetitions.
- **TE2**: Only a single measurement of the echo times [1.74, 2.1, 2.5, 2.8, 3.1, 3.4, 3.7, 4.0, 4.3, 4.5, 5.0, 6.0] ms.
- $\boldsymbol{\alpha}$: Flip angles of 10:5:50° was obtained with three repetitions.
- **Cross**: Corresponding measurements of [10,17,25]° and [1.74, 2.7, 3.6, 4.6, 5.6] ms was made with 2 repetitions.

3.1.3 Patient

Patients were scanned in accordance with in-house protocols. Appropriate fixation devices were used in both MRI and CT. All data was anonymized using Conquest DICOM server version 1.4.16.

- **Head patient** was scanned using the Flex-coil and with the following sequences: A T1-weighted, a T2-weighted and an UTE sequence. The details concerning the non-UTE sequences are irrelevant in this project and therefore not stated.

3.2 Image Processing

An overview of the data processing step are shown in Figure 3.2.

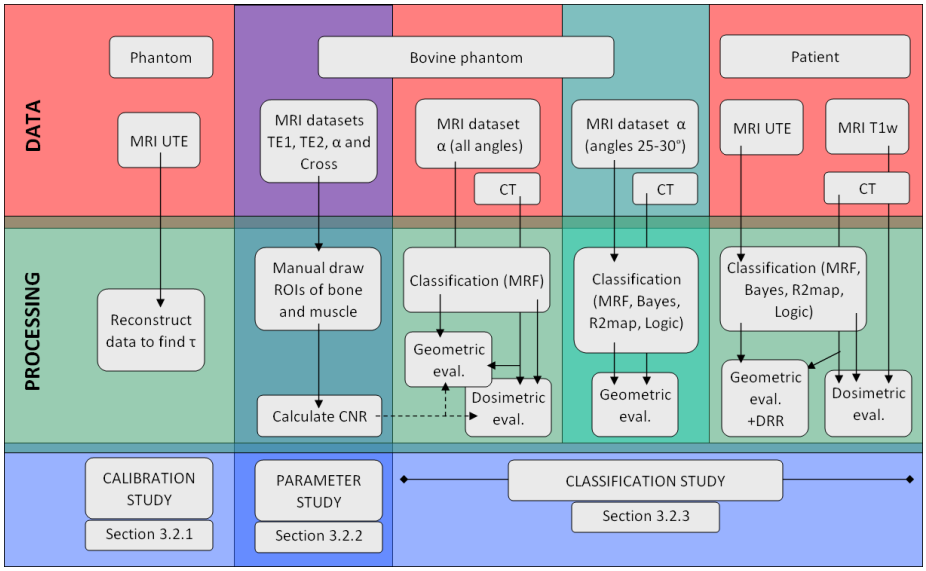
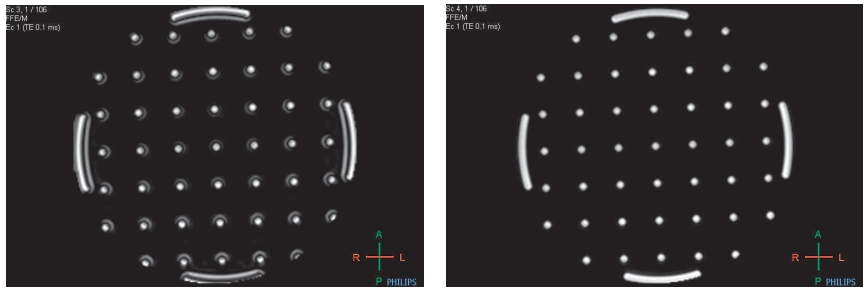


Figure 3.2: Data processing overview

3.2.1 UTE Calibration

A qualitatively visual assessment of each phantom scan was made in order to find an acceptable trajectory delay value, τ, for the given FOV & TE1.

The raw data was saved allowing reconstructions with other τ-values than those used during image acquisition. Reconstructions was tested with a precision of 0.5 μs. An example of an image with non-optimal and correct τ-value are shown in Figure 3.3 for comparison. When τ is non-optimal blurred artefacts appear especially at the four elongated radial markers. When τ is too large the blur appears towards to the centre and vice versa. The effect is only present on the first echo.



(a) Echo 1, τ = 5 μs (b) Echo 1, τ = -1 μs

Figure 3.3: Example of visual assessment of τ-values. Clear blurring is observed with τ = 5 μs.

3.2.2 Parameter study

The echo images were loaded and the subtraction image calculated. All voxels in the subtraction image with values below zero were truncated to zero, since voxels can only have a higher intensity in the second echo due to noise or artefacts.

One dataset from each of the four sessions was randomly selected. Bone and soft tissue ROIs were drawn manually on multiple slices, as illustrated in Figure 3.4, and the CNR was calculated (Equation 2.1). The ROIs could be transferred to all the other datasets from the same session, since they were registered to each other.

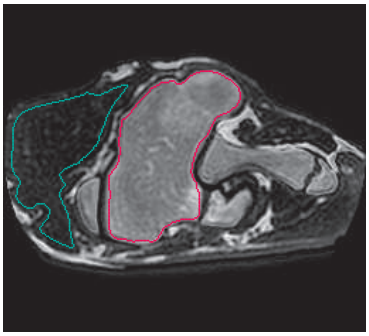


Figure 3.4: Illustration of a manually drawn muscle ROI (green) and bone ROI (red) in one slice.

3.2.3 Classification study

The following is applicable to all the classification studies. The specific details are described in the individual sections below. All classifications was done using Matlab. For geometric and dosimetric evaluation data was exported in DICOM format and imported into Eclipse v10.0 (Varian Medical Systems) treatment planning system (TPS). The CT-data is considered to be the truth. The registration of CT and MR was done using an initial manual rigid alignment, which was fine tuned (still rigidly) using an iterative algorithm.

Geometric evaluation was performed on all datasets by defining and measuring the following volumes:

- BoneAgree = BoneCT ∩ BoneMR
- BoneMiss = BoneCT - BoneAgree
- BoneFalse = BoneMR - BoneAgree

BoneCT refers to the bone volume defined on the CT images, and that was then transferred to the MR images. The *BoneMR* was the volume that was classified as bone on the segmented MRI.

Dosimetric evaluation was done by copying a CT-based plan onto the bulk segmented MR datasets. Relevant structures defined on the MRI were assigned an appropriate CT-number (the standard calibration curve of the TPS was used for converting CT-numbers into electron densities). The plan was then recalculated using the fixed monitor units (MU) noted from the CT-plan. Relevant DVHs were exported and specific DVH points of interest were noted.

3.2.3.1 Classification performance vs. CNR

The data considered here are all the bovine phantom images from the session with varying flip angles (total of 27 data sets - 9 different flip angles with 3 repetitions of each).
Each data set was segmented into air, soft tissue, compact bone and spongy bone using the *MRF*-classifier ($\beta = 0.2$, 5 iterations). A automatic separate training (8 classes) was performed for each data set.
The registration was performed between the original echo 1 MRI (flip angle 30°) and the CT. All MRI datasets share the same registration.
The structures used for geometric and dosimetric evaluation were defined as:

- BodyMR: All segmented tissue. Was assigned to 0 HU.
- BoneMR: The segmented bone tissue. Was assigned to 501 HU, which was a rough representative estimate based on the CT image.
- TissueCT: CT voxels [0 HU to 241 HU], cropped to the MR body.
- BoneCT : CT voxels [241 HU to 2005 HU].

The CT-based treatment plan consisted of two opposing fields. An Anterior-Posterior and a Posterior-Anterior (APPA) field. A 2 Gy dose should be delivered in 1 fraction to the isocenter of the phantom which required a LINAC output of 109/109 Monitor Units (MU). Beside calculating dose plans for all the bulk segmented MRIs, an additional plan was evaluated where the entire BodyMR volume was assigned to 0 HU corresponding to a $\mathrm{MRI}_{\mathrm{Unit}}$ plan.

3.2.3.2 Comparison of classification methods using bovine data

The data considered here are 6 datasets from the session with varying flip angles that had similar CNR scores (the data with flip angles 25 and 30°).

Each data set was segmented into air, soft tissue and cortical bone using both *Bayes*, *MRF*, *R2map* and *Logic* classifiers (a total of 24 datasets). A separate automatic training (8 classes) was done for each data set, but was applied to both *Bayes* and *MRF* ($\beta = 0.2$, 5 iterations). Thresholds for the *Logic* classifier was set manually on one dataset and afterwards applied to all the others.

The registration was performed between one *MRF* bulk segmented image and the CT. All the MRI datasets share the same registration.

The structures used for geometric evaluation was defined as:

- BoneCT: CT voxels [700 HU to 2005 HU]. The lower limit of cortical bone was a visual estimate.
- BoneMR: The segmented (cortical) bone tissue.

3.2.3.3 Comparison of methods using the head patient data

The UTE images were segmented into air, soft-tissue and bone using both *Bayes*, *MRF*, *R2map* and *Logic* classifier. The same automatic training (7 classes) was shared for *Bayes* and *MRF* ($\beta = 0.7$, 10 iterations). The three thresholds for *Logic* was determined manually.

The CT volume and the MR volume did not cover the same field of view. CT extended to include several vertebrae and most of the mouth, while MRI covered less. Combined with the more complex anatomy the definitions of structure volumes were then a bit more comprehensive:

- BodyMR: All voxels classified as tissue, including all air cavities. The outer contour expanded to include ear canal and nasal cavities. Was assigned to 0 HU.
- CavityMR: All voxels classified as air within BodyMR and part of a connected component larger than 0.5 cm^3. Was assigned a CT-number corresponding to air.
- BoneMR: All voxels classified as bone. Was assigned to 971 HU [32].
- BoneCT: CT bone voxels, but cropped to be within BodyMR. This was to exclude the excess CT volume from the geometric evaluation.
- Brain: The planned target volume (PTV) segmented from the T1 weighted MR image
- Hippocampus: Planned Risk Volume (PRV) segmented from the T1 weighted MR image.
- Eyes: PRV segmented from the T1 weighted MR image.

The CT-based plan used for dosimetric evaluation was a pre-clinical Volumetric Modulated Arc Therapy (VMAT) plan. A dose of 25 Gy should be delivered

in 10 fractions using two arcs. The LINAC output was 142 and 323 MUs respectively. Beside calculating dose plans for the four bulk segmented MRIs, an additional plan was evaluated where the entire BodyMR volume was assigned to 0 HU.

Chapter 4

Results

4.1 UTE Calibration

Purpose: Establish appropriate trajectory delay, τ, values to avoid blurring artefacts.

Observation: The relative small changes in TE_1 that were tested showed not to have a markedly effect on τ. Decent values for a range of FOVs are listed in Table 4.1. It is noted that τ decreases with increasing FOV. It was found that the amount of blurring in images were not sensitive to small deviations of τ. For example, a choice of $\tau = -1$ μs for at FOV of 210 mm would also keep blurring at a minimum. With this kind of tolerance it is reasonable to linear interpolate trajectory delays for FOVs not listed.

FOV [mm]	190	200	210	220	230	240	250	350	450
$\boldsymbol{\tau}$ **[$\boldsymbol{\mu}$s]**	-1	-1	-1.5	-1.5	-1.5	-2	-2	-3	-4

Table 4.1: Trajectory delay settings

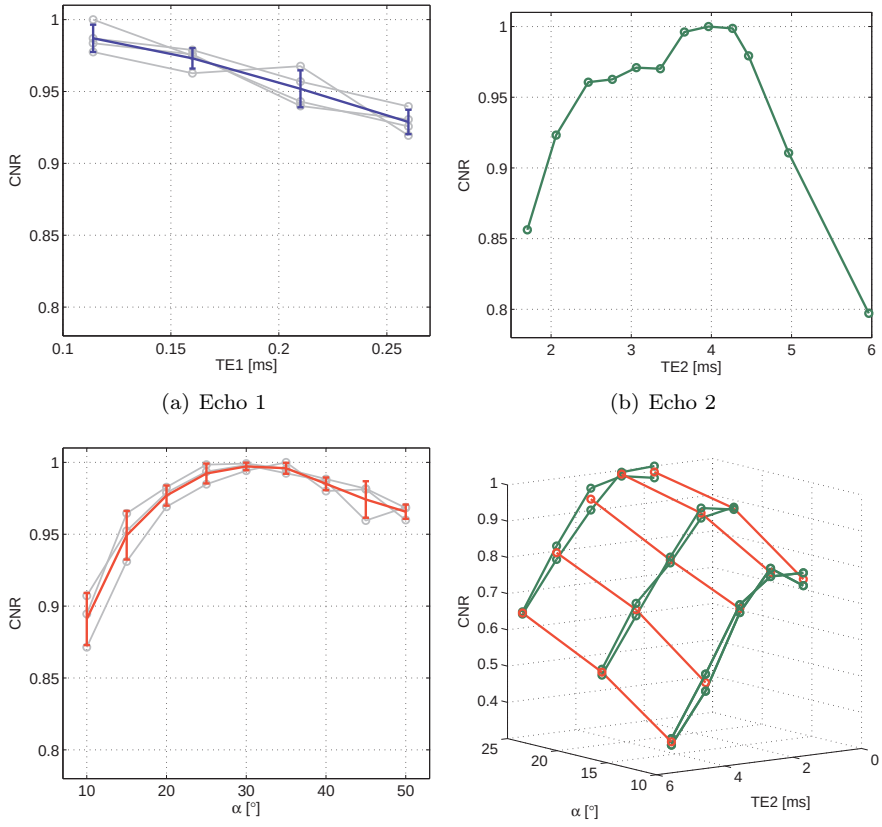
4.2 Parameter study

Purpose: Investigate how controlled variations of acquisition parameters influence the CNR. Parameters that maximize CNR are considered optimal.

Observation: CNR scores plotted against imaging parameters are displayed in Figure 4.1. The CNR score of each individual repetition is further shown on Figure 4.1(a) and 4.1(c), to give an impression of the variation between measurements. The following is observed from the figures:

- CNR drops linearly when the first echo time is increased.
- Variations of the second echo time results in a parabolic-like behaviour with a maximum around 4 ms.
- CNR increases with an almost logarithmic behaviour until $25 - 30^\circ$ where it begins to drop of.
- The CNR behaves similarly for all combinations of flip angle and second echo time. No cross-effect is observed.

When the cross effect measurements were acquired the phantom had markedly decayed. This resulted in a slightly different CNR behaviour concerning variations in TE2. Qualitatively the "parabolic" shape is reproduced, but the maxima were observed at lower TE2 than previously.



(a) Echo 1

(b) Echo 2

(c) Flip angle

(d) Cross effect

Figure 4.1: **a:** Varied first echo time with 4 repetitions for each setting.
b: Varied second echo time with 1 repetitions.
c: Varied flip angle with 3 repetitions.
d: Variations of both flip angles and second echo time with 2 repetitions for each set of parameters values. Points of the red curves are the mean of the green points.

4.3 Classification study

4.3.1 Classification performance vs. CNR

Purpose: Having observed that flip angles of $25 - 30°$ yields the highest CNR, it should be explored if the high CNR correlates with higher geometric and dosimetric agreement.

The results of geometric evaluation are shown in Figure 4.2. The Dice coefficient peaks in accordance with the flip angles that yielded high CNR (Figure 4.1(c)). It is further observed that the fraction of falsely classified MR volume steadily drops for increasing flip angles. The fraction of missed bone drops significantly at the optimal flip angles.

The dosimetric evaluation is shown in Figure 4.3 and 4.4. All the data points were calculated as the average of the three datasets with same flip angle. A single MRI_{Unit} plan was also evaluated. The absolute difference from the corresponding data points in the CT-based plan are plotted. A similar shape is observed in all figures - the deviation is minimized when CNR is high. The effect is less clear with the soft-tissue median dose and DVH, but still observable. The MRI_{Unit} plan deviates more than any of the bulk segmented plans, regardless of the flip angle.

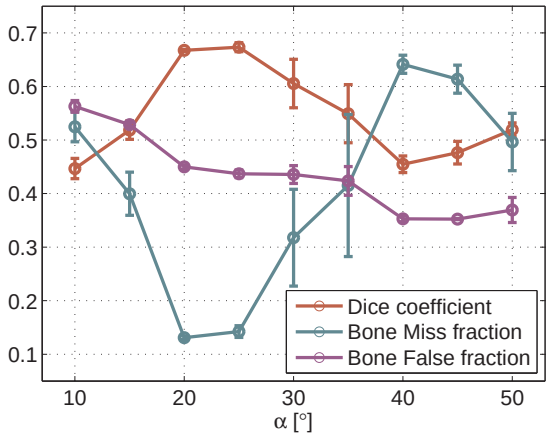


Figure 4.2: The geometric evaluation correlated with flip angles (CNR). The shown points are the average of the three measurements with half a standard deviation on both sides.

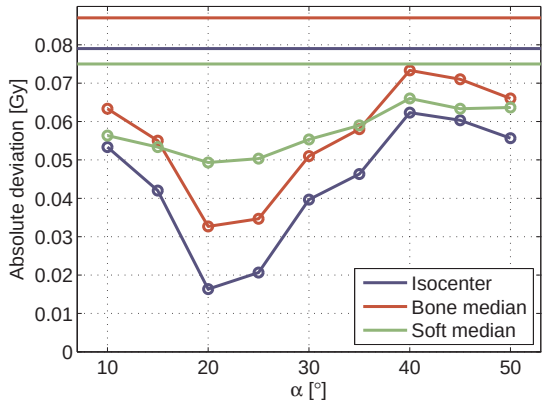
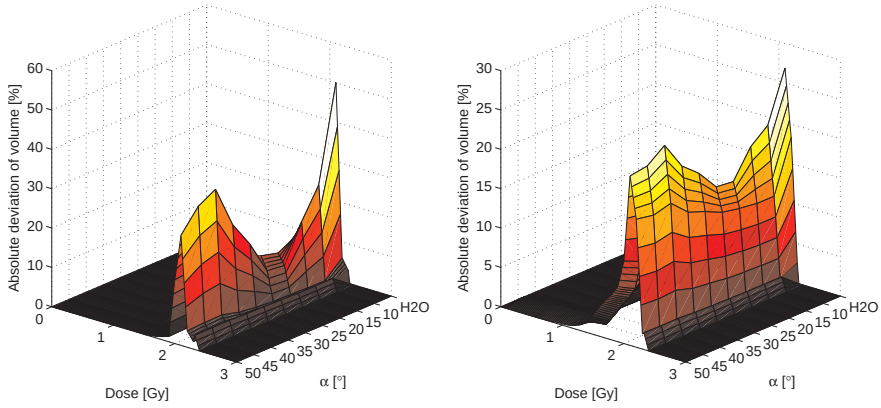


Figure 4.3: The median and isocenter point dose deviation from CT-based plan correlated with flip angle (CNR). The horizontal lines shows the corresponding deviation between CT and $\mathrm{MRI}_{\mathrm{Unit}}$ plan



(a) Bone DVH

(b) Soft tissue DVH

Figure 4.4: All DVHs subtracted from their corresponding DVH in CT-based plan. The $\mathrm{MRI}_{\mathrm{Unit}}$ plan is marked as flip angle 'H20'.

4.3.2 Comparisons of methods using bovine phantom

Purpose: Compare the ability of the four classifiers to segment cortical bone from the bovine phantom data that was observed to have to highest CNR.

Six datasets were classified and geometrically evaluated. The data presented in Table 4.2 is the measured volume averaged over the six datasets. The data is further presented graphically in Figure 4.5. The Dice coefficient and bone miss/false fractions were calculated for each data set separately and then averaged, and not based on the average reported volumes.

It is observed that *MRF* segmentation performs slightly better than *Bayes*. *Logic* masking performs better than *R2-mapping*, but is still worse than the *Bayes*.

Method	**Bayes**	**MRF**	**R2map**	**Logic**
CT Bone [$\mathbf{cm}^3$]	165.30	165.30	165.30	165.30
MR Bone [$\mathbf{cm}^3$]	143.94	145.56	157.79	89.23
Bone Agree [$\mathbf{cm}^3$]	58.49	63.20	26.64	33.06
Bone Miss [$\mathbf{cm}^3$]	94.43	92.81	129.22	124.51
Bone False [$\mathbf{cm}^3$]	80.31	77.81	127.29	52.61
Dice coef.	0.38	0.41	0.16	0.26

Table 4.2: The averaged structure volumes and Dice coefficients for bovine phantom (6 datasets)

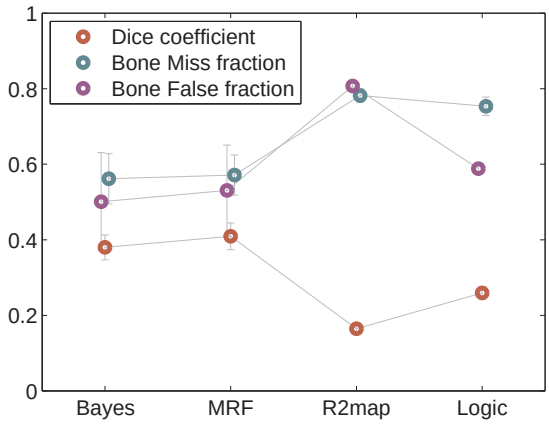


Figure 4.5: The geometric evaluation of bovine phantom data segmented with four different approaches.

4.3.3 Comparisons of methods using patient data

Purpose: The performance of the four classifiers were tested on head anatomy data of a single patient.

The geometric evaluation is presented in Table 4.3 and Figure 4.6. Compared to the segmentation of the phantom a similar result is observed. *MRF* performs slightly better than *Bayes* which in turn is slightly better than the *Logic* approach. The DRRs are shown in Figure 4.8. The apparent size difference between CT and MR images is not due geometric distortions. The CT DRRs are viewed with a different scaling because they cover a larger anatomical area.

The dosimetric evaluation is presented in Figure 4.7. The DVH is shown for the CT-plan, the bulk segmentation plan from each of the four methods and the MRI_{Unit} plan. The deviation from the CT-plan for each structure is shown below the figure for better visual assessment of the differences. Specific DVH dose points are further presented in Table 4.4. The percentage difference from the corresponding CT point, ΔCT, are also listed.

The largest difference between the CT-plan and the MRI-based plans is observed for the hippocampus. All the methods overestimate the dose, but do so in a similar behaviour.
For the PTV, *MRF* and *Bayes* show a performance close to CT and markedly better than the other plans both concerning median dose and DVH deviation. It is noted that *R2map* behaves closely to *Water* (MRI_{Unit} plan).
The most degree of different behaviour in between the MRI-plans is observed at the Eyes volume. The median dose for *MRF*, *Bayes* and *R2map* is in good agreement with *CT*. However, *R2map* has a very different DVH behaviour and the $D_{2\%}$ dose is underestimated with 11%.

Method	**Bayes**	**MRF**	**R2map**	**Logic**
CT Bone [$\mathbf{cm}^3$]	695.37	694.04	688.71	692.91
MR Bone [$\mathbf{cm}^3$]	561.91	615.93	208.00	476.51
Bone Agree [$\mathbf{cm}^3$]	334.75	376.45	134.77	292.48
Bone Miss [$\mathbf{cm}^3$]	333.01	299.59	534.11	376.67
Bone False [$\mathbf{cm}^3$]	206.35	221.30	63.70	168.09
Dice coef.	0.53	0.57	0.30	0.50

Table 4.3: The measured structure volumes and calculated Dice coefficients for patient data.

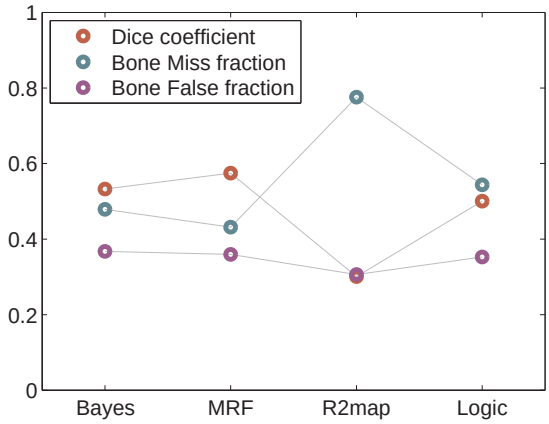


Figure 4.6: The geometric evaluation of patient data segmented with four different approaches.

Plan	CT	Bayes	MRF	R2map	Logic	Water
PTV, D_{Median} [Gy]	30.07	30.05	30.10	30.44	30.29	30.41
ΔCT [%]	-	0.07	-0.10	-1.23	-0.73	-1.13
PTV, $D_{\text{D2\%}}$ [Gy]	34.24	34.36	34.42	34.74	34.60	34.69
ΔCT [%]	-	-0.35	-0.53	-1.46	-1.05	-1.31
Eyes, D_{Median} [Gy]	15.22	15.24	15.26	15.27	15.36	15.34
ΔCT [%]	-	-0.13	-0.26	-0.33	-0.92	-0.79
Eyes, $D_{\text{D2\%}}$ [Gy]	18.76	18.39	18.45	16.68	18.62	18.62
ΔCT [%]	-	1.97	1.65	11.09	0.75	0.75
Hippocampus, $\text{EQD2}_{40\%}$ [Gy_2]	7.82	8.56	8.56	8.46	8.49	8.43
ΔCT [%]	-	-9.46	-9.46	-8.18	-8.57	-7.80

Table 4.4: Notable dosimetric points for patient data, and the percentage deviation from CT.

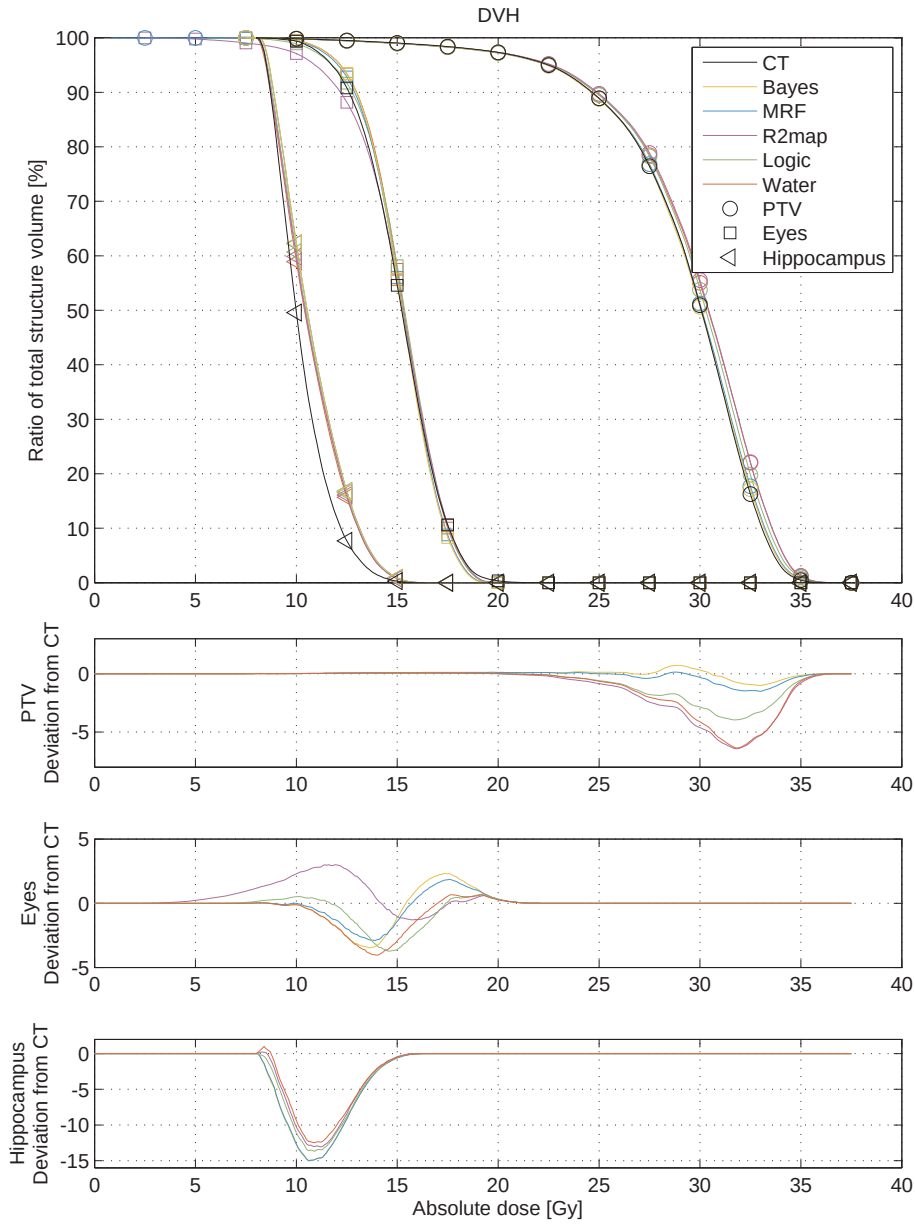


Figure 4.7: The DVHs from pre-clinical plan belonging to head patient. The deviation from CT DVH in percent of structure of volume receiving a given dose are shown below.

Figure 4.8: Head anatomy DRRs. Marked distance measures: 13.78 cm (*CT*) and 13.80 cm (*R2map*)

Chapter 5

Discussion

5.1 UTE Calibration

Values for trajectory delays were established for a range of clinical useful FOVs. It deserves to be restated that also the bandwidth and TE1 has an impact on τ. It seems reasonable to keep the bandwidth as wide as possible in UTE scans, and this parameter was therefore never changed in this report, eliminating the need for exploring the bandwidth effect on τ. A small interval of TE1s were investigated (0.09 to 0.14 ms), and it was found not to have an impact on the amount of blurring. Results from the parameter study showed that TE1 should be minimized, thus negating the need for investigating appropriate τ-values for higher TE1. Should it be possible to decrease TE1 below 0.09 ms in the future, it might be appropriate to investigate the effect on τ.

The listed values in Table 4.1 were used for all subsequent UTE scans. The values appears to work fine, but are only based on a visual assessment of blurring in a single or few slices and only with an accuracy of 0.5 μs. The importance of this parameter should however not be understated. Spatial distortions and blurring is the consequence of using an incorrect τ, which was demonstrated in Figure 3.3 and in Figure 5.1.

Spatial distortions between the two echoes are clearly visible (Figure 5.1) .

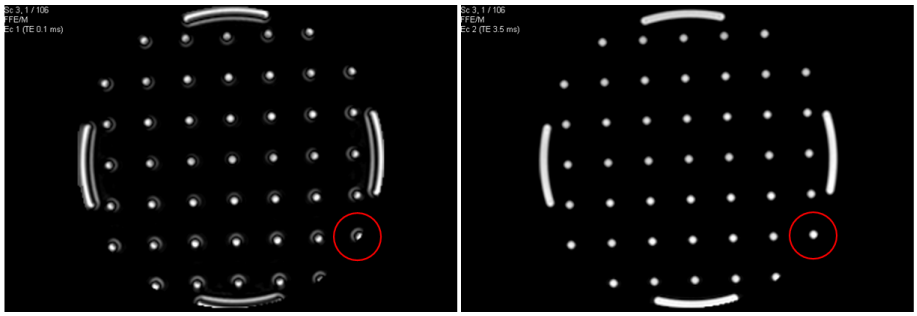


Figure 5.1: Effect of wrong τ-value. Echo 1 (left) and echo 2 (right).

The segmentation methods assume that the voxel position in the echo images are identical. By using the correct τ-values distortions will be minimal and will probably cause a little noise at worst, but if τ is wrong it can potentially be a serious problem for the classifiers.

Blurring is apparent in certain voxels in echo 1, but it is gone in echo 2. The blurred voxels have a behaviour identical to the expected behaviour of cortical bone and thus a potential source of misclassifications.

A more thoughrough study of trajectory delays could be considered. Perhaps using a quantitative metric instead of a subjective visual assessment. This could potentially allow τ-values to be established with a higher accuracy than 0.5 μs. Further is should be considered to use a larger phantom when testing large FOVs. It can be seen from Figure 5.1 that the blur increase with radial distance from centre. With the relatively small phantom used, it was not possible to see blur far from the centre when the largest of the FOVs were tested.

5.2 UTE Parameters

First Echo Time The results from the CNR measurements on the bovine phantom suggest that TE1 should be minimized. The TE1 effect on CNR can easily be demonstrated on an intensity map with some artificial UTE (Figure 5.2). The data was created similarly to the data used in Figure 2.4. Here it can be seen how a larger TE1 compresses the intensity map towards the y-axis, and thus ruins the contrast between tissues.

The CNR curve in Figure 4.1(a) is approximately linear. Extrapolating the curve to $\mathrm{TE}_1 = 0$ ms will result in a CNR very close 1, which theoretically

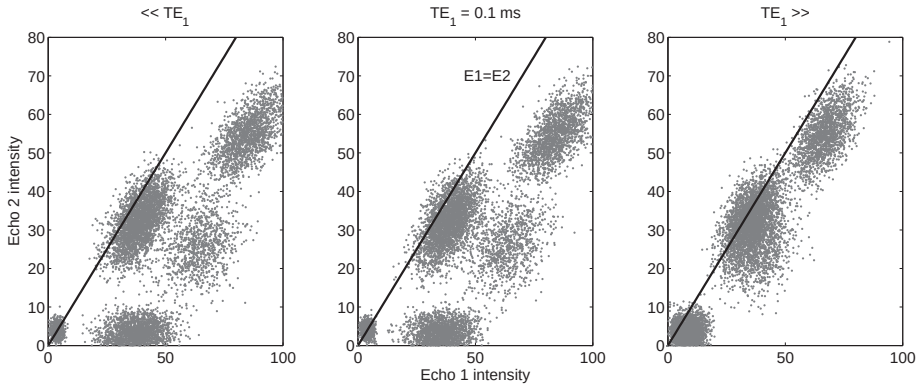


Figure 5.2: The effect of TE1 shown on intensity maps of artificial data

would be the optimum. Linear extrapolation to higher TE1s is however not appropriate. This would suggest that bone contrast remains adequate even to rather large TE1 values. This is because the CNR measure primarily reflects the contrast between spongy bone and muscle tissue, since the ROIs mainly consisted of these tissues. The most important aspect of choosing TE1, is to ensure that cortical bone can be separated from air. Although this is not reflected in the current CNR measure, minimizing TE1 is the result that is in accordance with theory.

The lower limit to TE1 is due to hardware limitations. The coils are manufactured with a time delay from switching between transmit to receive mode. The delay is there to let remaining energy of the RF excitation pulse ring down and to tune the coil to receive mode[22]. From a UTE point of perspective it would be interesting to see if this delay could be decreased even further.

Second Echo Time The results suggest that there is an optimum second echo time around 4 ms. This can again be demonstrated using the UTE intensity map (Figure 5.3). The T2 (or T2*) determines how fast a tissue group will descend towards the x-axis. Naturally, this will result in the existence of an optimum. If the CNR measure was based only on cortical bone and muscle tissue, this optimum would likely had been found at the point when cortical bone intensities approach zero in echo 2. The CNR measure in this study is based largely on spongy bone, and overall best separation was found with an echo time of $\approx$4 ms for the 1-T scanner.

This result presents a dilemma. The chemical shift of water-fat at 1 Tesla is strongest at $\approx$3.4 ms. Signal from voxels with a mix of water and fat will cancel

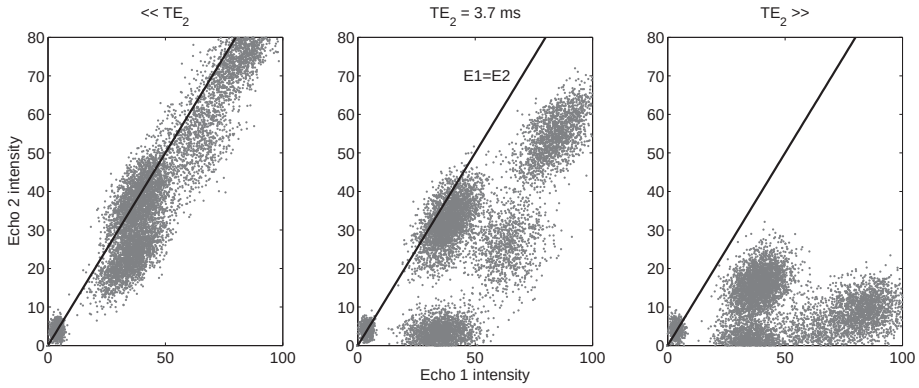


Figure 5.3: The effect of TE2 shown on intensity maps of artificial data

out to a large degree. In effect, they have a short T2* and will behave like cortical bone voxels resulting in misclassifications. However, the first in-phase echo time for water and fat is 6.9 ms, which is found to be much too long in the sense of optimal CNR (Figure 4.1(b)). The chemical shift issue appears to be lesser of two evils. By including other MRI sequences into the classifier in the future it might be possible to make a distinction between water-fat cancelled voxels and cortical bone voxels.

Flip angle The observed result can be visualised with the intensity maps of the actual measured bovine data in Figure 5.4. Choosing the flip angle too low brings spongy bone and soft tissue too close to each other. Increasing the flip angle appears to shift the spongy bone group outwards, but at the cost of an increase in the variance. The ideal (high CNR in equation 2.1) is to have spatially separated clusters (large difference in the nominator), with as little variance as possible (low denominator). An optimum flip angle was found to be 25-30°.

5.3 Classification

5.3.1 Classification performance vs. CNR

Having established a set of parameters that yield a high CNR, it should be shown than it in fact correlates with the classifiers' ability to perform better segmentations. It was observed that the Dice coefficient was highest for 20-25°,

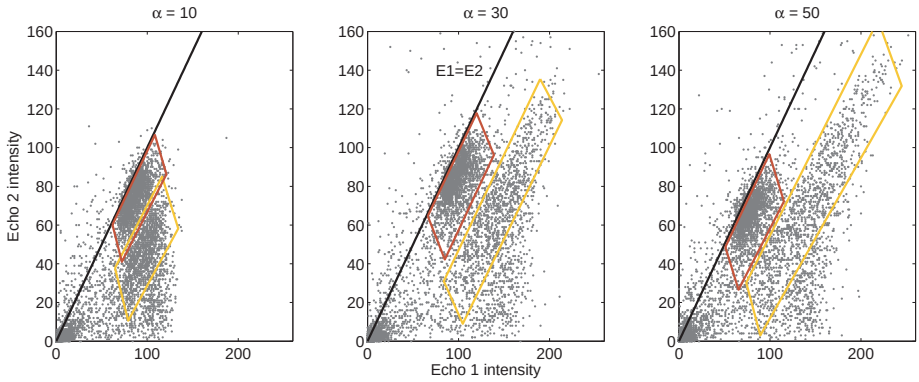


Figure 5.4: The effect of flip angle shown on intensity maps of bovine phantom data. Approximate regions of spongy bone (yellow) and muscle tissue (red) are marked

mostly due to a much lower miss fraction. The fraction of false classification steadily decreases with increasing flip angles. This behaviour can be explained with the observations from Figure 5.4. Spongy bone clearly separates itself more from other tissues with increasing flip angle which results in fewer false classifications. However, the increased variance results in a lot more missed bone.

A high variance in the geometric results was observed regarding flip angles 30-35°. This is contributed to the variability of the training used. The randomness has probably caused it to place the Gaussian distributions differently. However, the tendency is clear, a higher CNR gives better segmentations (higher Dice coefficient). The same conclusion was reached in the abstract in Appendix A, where the datasets and segmentation strategy was different.

The remaining thing that should be established is whether higher Dice coefficient correlates with better dosimetric accuracy. The results suggest that it does. All investigated points and DVHs deviated less from CT-based plan when flip angles were 20-25°. This conclusion is also in accordance with the one presented in the abstract in Appendix A. Further, it was observed that all segmentations regardless of α deviated less from CT than the $\mathrm{MRI}_{\mathrm{Unit}}$ plan.

The above stated behaviour was only shown for the *MRF* classifier, but is reasonable to assume that it would be similar for the other classifiers used in this report. The *R2map* and *Logic* classifiers have very similar decision boundaries and the approach used in the abstract (Appendix A) are very similar to the *Logic* classifier.

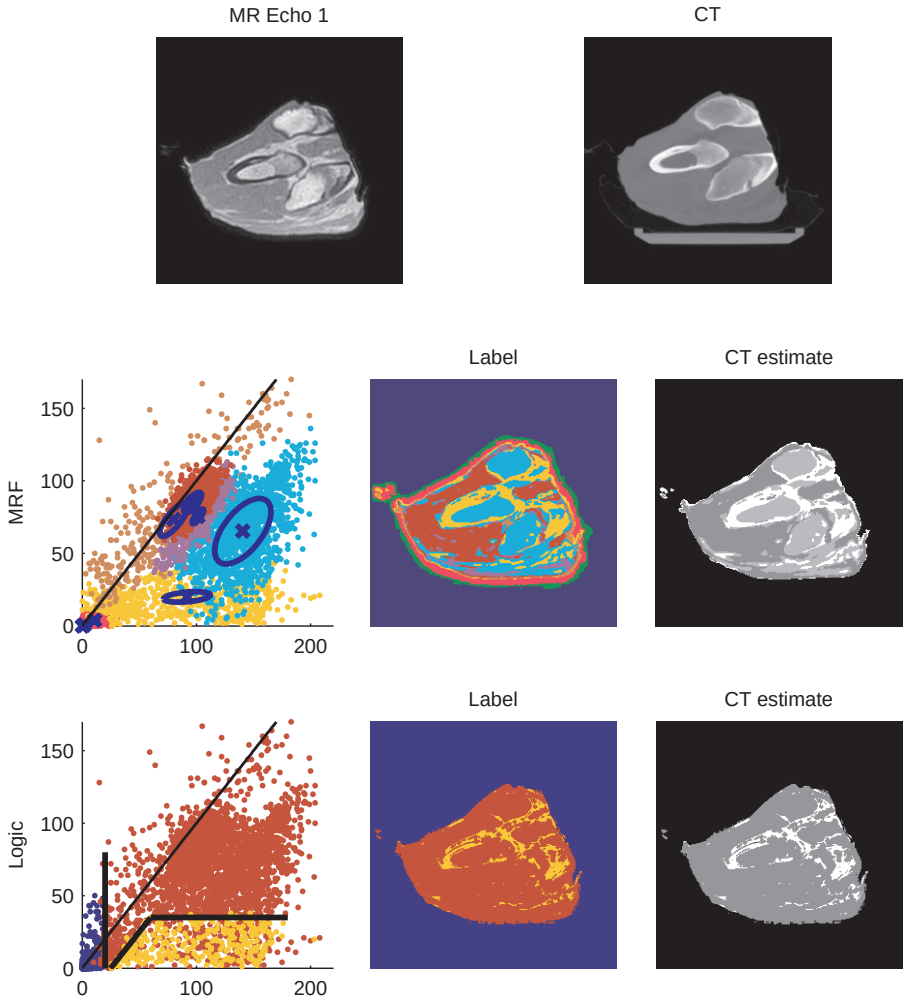


Figure 5.5: Bovine phantom, comparison of CT, MRI, and CT estimates of *MRF* and *Logic* classifiers. Color coding: Air (dark blue), cortical bone (yellow), soft tissue (dark red), spongy bone (blue), more soft tissues (brown and purple) and more air (red and green).

5.3.2 Comparisons of methods

The bovine phantom Datasets with the highest CNR were used to test the performance of the 4 classification strategies. Only geometric results were compared. A dosimetric comparison seems more meaningful with the data from the head patient. Especially because the *Logic* and *R2map* classifiers were developed for this particular anatomy. They were not designed to deal with spongy bone, which is what a high percentage of the bovine phantom consists of.

A visual presentation of the CT-estimates from the *MRF* and *Logic* classifier are shown in Figure 5.5. The susceptibility artefacts at air-tissue boundaries clearly causes misclassifications. The edge reduction scheme in the *Logic* classifier appears to reduce this to a great extent, and it is the reason why some voxels are labelled with a different colour in the intensity map (Figure 5.5 lower left) than the decision boundaries would otherwise determine.
Ligaments/tendons are easily identified on the MR image, and appears to be misclassified as either spongy bone or cortical bone on the CT-estimates. This was expected since these tissues behave similarly according to theory. If a larger percentage of the phantom had been ligaments or tendons, then the *MRF/Bayes* training might had been able to classify it as a separate tissue group. This flexible training is clearly an advantage that *MRF/Bayes* has over the other methods, however it is dependent on the imaged anatomy. Cortical bone is something that is never present in large amounts, which means that the training of the *MRF/Bayes* classifier sometimes fails to recognise it as a separate tissue.

The results showed that *MRF* had the best performance (Dice = 0.41), but only slightly better than *Bayes* (Dice = 0.38). The drawback of using *MRF* is illustrated by comparing the computational time of the classifiers (Figure 5.6). The test was done on square artificial UTE 2-D images. The *R2map* classifier shows long computation times due to the region growing algorithm used for defining the air outside the object. The *MRF* classifier is shown for seven full iterations using 5 classes. Better implementations could increase speed for both algorithms, which is something that would be valuable to explore with studies in the future.

Patient data The geometric comparison comes to the same conclusion as with the bovine phantom, *MRF* has the best performance with a Dice coefficient of 0.57. The atlas-based approach presented by Dowling et al.[5] gave segmentations of the pelvic bone with a Dice coefficient of 0.79 $\pm$ 0.12. Results from different anatomies are not directly comparable but it provides a perspective on

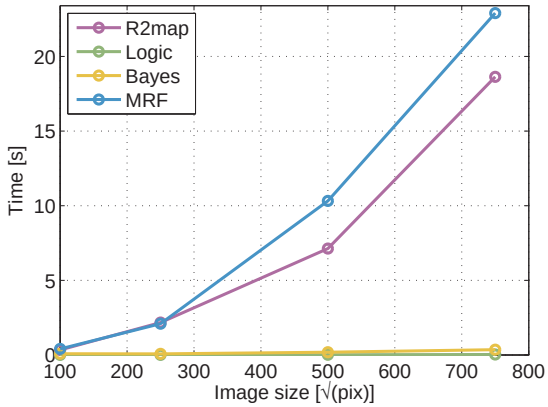


Figure 5.6: Comparison of computation times (excluding time spent on training and defining thresholds).

bulk segmentation strategies versus atlas-based approaches. Improvements to the current UTE bulk segmentation strategy are required.

A visual inspection of the classifications (Figure 5.7) show how *MRF* suffers from susceptibility artefacts, especially in the nasal cavity. The skull is generally segmented nicely, but there are areas of misclassification in the back of head, the neck and nose area. This can be contributed to signal loss due to a limited range of Flex-coils.

The generated DRRs (Figure 4.8) support the general picture. CT-estimates based on *MRF* are closer to the true CT. The advantage of *MRF* in comparison to *Bayes* is clearly seen in the DRRs. *MRF* is capable of reducing much of the noise that is seen outside of the skull. Further studies are required to investigate if the DRRs are accurate enough to be used for treatment set-up verification.

The dosimetric comparison was performed using the PTV (the whole brain) and two important OARs - the eyes and the hippocampus.
For the most part, the investigated DVHs showed a dose overestimation (Figure 4.7). This can be explained by the consequent low MR bone volume on segmented MRI (Table 4.3). Too little bone results in less attenuation and allows more radiation to reach the brain and hippocampus.
• **PTV:** Given that the overall segmentation is more accurate for *MRF* and *Bayes*, it is seen that the DHV deviates less from the CT-plan than with the other methods. The *R2map* plan is observed to behave much like the MRI_{Unit} plan. This is easily explained by the fact that *R2map* has a BoneMiss fraction of 0.78. All the missed bone is instead assigned an electron density equal to

water. The noted DVH points (Table 4.4) show small (<1.5%) dose differences from CT.

• **Eyes:** The *R2map* plan deviates markedly from the rest of the plans. The cause is inspected in Figure 5.7. It is observed that a part of the eyes is treated as air. It suggests that the threshold for defining air with *R2map* was set too high in this case. It would have been easy to fix, but it illustrates a weakness of the manual thresholding strategies. Tweaking of threshold can be difficult, especially when tissue groups tend to overlap on the intensity maps shown previously. Once a threshold has been chosen it is only valid for a specific set of parameters, coil, anatomy etc.
Dose difference on the noted DVH points were small (<2%) except for the *R2map*.

•**Hippocampus:** The $D_{40\%}$ dose to the hippocampus should not exceed 7.3 EQD_2 Gy_2[31]. The final treatment plan was within this dose constraint, but the pre-clinical CT plan used in the this report was not. The dose difference from CT at this DVH point was generally high (7.5-10%). There is no immediate explanation for this behaviour. No gross misclassifications could be observed in the area around the hippocampus. It is something that should be considered and investigated in future studies.

Summary For the most part, the dose distributions from UTE bulk segmented MRI are similar to the CT-based plan. The observations fit the expectations gained from similar studies using bulk segmented MRI [18, 19]. The dose differences are in the same order of magnitude (a few percent) and it is confirmed that the *Bulk* plans tend to be more accurate than *Unit* plans. This study is only based on a single patient and further studies are required to explore the reproducibility of the results. It is reasonable to conclude that this UTE bulk segmented images can potentially provide clinical acceptable dose plans.

The *MRF* classifier gives the most accurate CT-estimates based on the images acquired on the 1-T open MRI-scanner. Currently a maximum Dice coefficient of 0.57 was obtained for head anatomy. The approach is closely followed by *Bayes* as expected. The more advanced way of defining tissue classes have some advantages. The flexibility of the training is very useful for imaging other anatomies than the head. *MRF* suffers from very long computational times, and the automatic training could be further improved. These are technical issues that can be addressed in future studies.

The performance of *R2map* and *Logic* classifiers are possibly not the optimal. Tweaking of the manual thresholds and filter settings might improve the results.

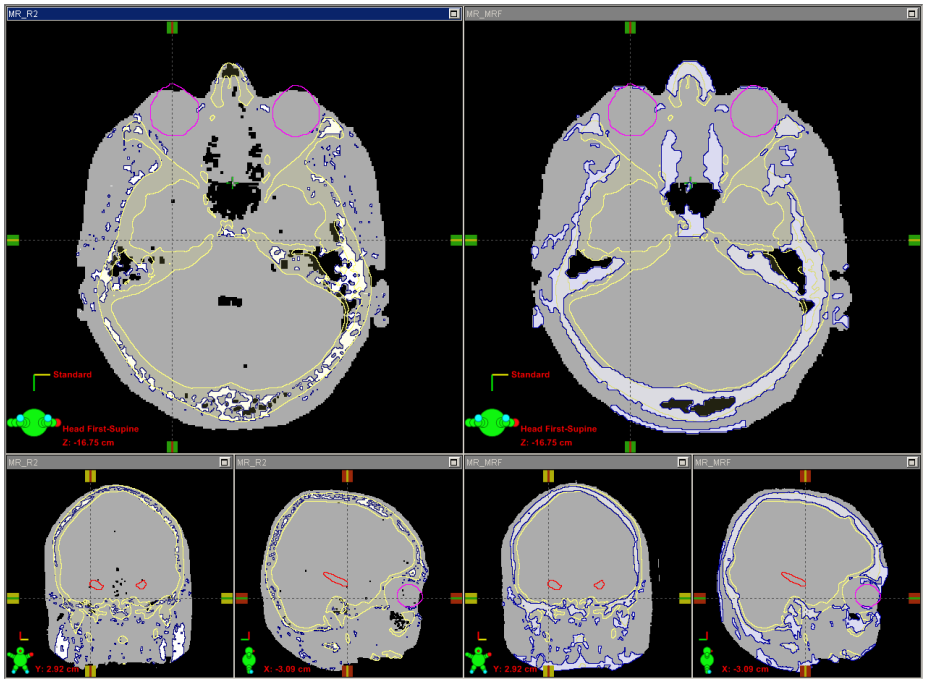


Figure 5.7: Comparison of CT, *R2map*(Left) and *MRF*(Right) segmentations of the head. The structures of true CT bone (yellow), the MR Bone (blue), the Eyes (pink) and Hippocampus (red) are shown.

It is, however, unlikely that they would suddenly perform markedly better. The linear decision boundaries that these method have, are too simple given the UTE images that can currently be acquired.

The performance of the classifiers are heavily influenced by the quality of the acquired images. In RT applications fixation devices in some cases prevents the use of the most optimal coil. Using the Large-coil instead of the Flex-coil for the head patient would for instance solve some of signal issues observed at the nose and the neck, but would result in a general higher noise level in the image. If UTE images in the future can be acquired in such a way that tissue groups are better separated on the intensity map, then all the tested classifiers could show increased performance. It might be more valuable to do research into better image acquisition rather than implementing more advanced classifiers. One way of improving the performance could be to include additional MR images, and thus expanding the intensity maps into higher dimensions. A promising example of this could be the Dixon sequence[14].

Chapter 6

Conclusion

The MRI UTE sequence shows a strong potential for making accurate CT-estimates of bone structures and the sequences will therefore play an important role in the realisation in radiation therapy planning based solely on MRI.

During the course of this project, the UTE sequence was tested on a 1 Tesla open MR-scanner dedicated for radiation therapy. The important trajectory delay values were found for a range of clinical useful field of views, but further studies could potentially be considered.

A bovine phantom was scanned with varying acquisition parameters and by using a contrast-to-noise ratio as a quantitative metric a set optimal parameters were found. The first echo time should be minimized, the second echo time is optimal close to 4 ms, and the ideal flip angle is 25°.

The performance of four different classification strategies was tested on data from a bovine phantom and a head patient. A classifier using Markov Random Fields gave the most accurate segmentations in all tested cases. For the patient a Dice coefficient of 0.57 was obtained for segmented bone volume. It would be preferable to use the Markov Random Field classifier, although it is computationally expensive.

It was shown that acquisitions with the optimal parameters provided the most

accurate bulk segmented MR images concerning both a geometric and a dosimetric comparison. It was further established that plans based on the bulk segmented MRI images were more accurate than uniform density assigned MRI plans.

While there is room for improvements on both classification and on the image acquisition side, the results are close enough to CT, that it is not unreasonable to consider further studies into MRI-only therapy.

Appendix A

Abstract for ASTRO 2012, accepted for poster presentation

Optimized Acquisition Parameters for MRI Only RT Using Ultrashort Echo Times

H.M. Kjer, R.H. Hansen, R.R. Paulsen, J.M. Edmund

Purpose/Objective

RT based on MRI only shows a promising potential if the information of electron-density from CT can be replaced with bulk density regions of tissue, bone and air in MRI. For this to be feasible, automatic bone segmentation is needed. This is troublesome on standard MRI sequences since no bone signal is available. The ultrashort echo time (UTE) sequence solves this issue and gives contrast between bone and tissue. We present a method for optimizing the UTE acquisition parameters using the contrast-to-noise ratio (CNR) and correlating the influence of this to bone segmentation and dosimetric agreement.

Materials/Methods

A cut-off bovine knee-joint phantom was CT scanned and scanned in a 1T open MR with a head coil.

With the UTE sequence, two images at echo times TE1 and TE2 were acquired. The contrast between bone and soft tissue recorded in a subtraction image (ΔTE = TE1-TE2), and an isotropic FOV/voxel size (200/1.6 mm) and flip angle 10° was used. The CNR dependence on TE1 and TE2 was explored. First, TE2 was varied from 1.6 to 6 ms with TE1/TR fixed at 0.11/6.5 ms. Secondly, TE1 was varied from 0.11 ms to 0.2 ms with TE2/TR fixed to 3.2/9.3 ms.

The image usability was quantified using a robust measure of CNR defined as $|M_B - M_T|/\sqrt{\mathrm{IQR}_B^2 + \mathrm{IQR}_T^2}$, where M_b, M_T and $\mathrm{IQR}_B, \mathrm{IQR}_T$ were the medians and interquartile ranges of the bone and tissue signals defined in regions of interest (ROI) on the ΔTE slices.

Segmentation was done using a two-step intensity classifier. First, the TE1 and TE2 images were used for a coarse logical segmentation of ΔTE. Voxels above a low threshold on TE1 can be considered non-air, and voxels below a threshold on TE2 to be air or bone. Secondly, ΔTE voxels were classified as bone if the intensity fell within a range defined as the mean $\pm$ 2 standard deviations of the bone ROI. Classified soft tissue was set to 0 HU and bone to 501 HU (CT average).

The bulk MR images were registered with those of CT in Eclipse which allowed comparisons of MR and true CT bone volume. A CT-based treatment plan was created giving 2 Gy at the isocenter using two opposing APPA fields. The plan was copied to the MR images and recalculated with fixed MUs.

Results

For the CNR, an almost inverse linear relationship with increasing TE1 was seen. When varying TE2, the CNR displays an almost parabolic behavior having a maximum CNR with TE2 between 3.5-4 ms.

Comparisons of low and high CNR segmentations showed the following: A significant increase of 13% (from 65 to 78%) more accurate bone classification, and a slightly lower dosimetric deviation of 0.3% and 1.2% from the CT based treatment plan at the isocenter and D98% (bone DVH).

Conclusion

UTE acquisition parameters that maximize the CNR was shown to improve the bone segmentation and minimize the dosimetric differences between CT and MR based treatment plans. A maximum of the defined CNR is obtained with the minimum possible TE1 and a TE2 around 3.5 ms.

Appendix B

Visionday 2012 poster

MRI mimicry of CT

Author: Kjer, H.M.[a]
Supervisors: Edmund, J.M.[b] and Paulsen, R.R.[a]
[a] DTU IMM, [b]Copenhagen University Hospital, Herlev

Abstract

The MRI UTE sequence gives the ability to visualize a signal from bone tissue. The sequence was recently installed on a Phillips Panorama 1T open MR scanner. Through a phantom-based study we have found a set of acquisition parameters that should optimize the contrast between bone and muscle tissue.

Using a Bayesian classifier with Markov Random Fields we segmented the UTE images from a patient into bone and soft tissue regions in order to produce a bulk image that mimics that of a CT.

Introduction

Radiation therapy (RT) is one of the common ways to treat a variety of cancers. Today, all RT planning is based on CT images (Figure 1), because it is geometrically accurate and the fact that CT can be used to calculate the radiation energy deposition in the body.

MRI-based RT would be advantageous in many aspects. The superior soft-tissue contrast over CT allows for better delineation of tumour and organs at risk[1]. However, conventional MRI sequences (Figure 2) are not able to visualize bone, which makes it difficult to translate an MRI into a map of electron densities used for dose calculations.

Figure 1: *CT image, transaxial slice of human head*

Figure 2: *MR T2 weighted image, transaxial slice of human head*

Ultrashort Echo Time (UTE) sequences can capture the signal from bone. The sequence was recently installed on the Philips Panorama 1T open MR system situated at the RT centre at Herlev Hospital. Several issues regarding this new sequence needed to be explored:

Which acquisition parameters should be used to optimize bone contrast?

Can bone be automatically segmented from UTE images ?

References

[1] R.C.Krempien et al., Open low-field magnetic resonance imaging in radiation therapy treatment planning, Int. J. Radiation Oncology Biol. Phys., Vol. 53, No. 5, pp. 1350–1360, 2002.

[2] M.D.Robson et al., Magnetic Resonance: An Introduction to Ultrashort TE (UTE) Imaging, J. Comput. Assist. Tomogr., 2003;27:825–846.

UTE Imaging

The bone signal is lost with conventional MRI sequences because the signal is decayed during the time spent to make the excitation, refocusing, build-up spatial gradients etc.

If the acquisition is placed *ultrashortly* after the excitation, a bone signal is recorded along with the signal from all other tissues (Figure 3, left)[2]. In order to make a distinction between bone and other tissues a popular approach is to acquired a second image shortly after the first (Figure 3, middle). Because bone decays much faster than other tissues it will appear bright on the subtraction image (Figure 3, right).

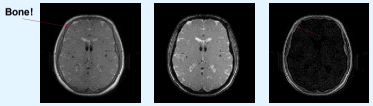


Figure 3: *MR UTE images, transaxial slice of human head.* ***Left:*** *Echo 1, 0.9 ms after excitation.* ***Middle:*** *Echo 2, 3.6 ms after the first.* ***Right:*** *Subtraction image.*

UTE imaging parameters: Being a new and relatively untested sequence, a study investigated which UTE acquisition parameter settings would be optimal for bone segmentation.

Method & Results: A cut-off bovine knee-joint phantom was imaged several times with varying parameters. A Contrast to Noise Ratio (CNR) between two manual annotated regions of muscle and bone tissue was calculated for each acquisition mode. The results are shown in Figure 4. Parameters that maximize the CNR are preferred.

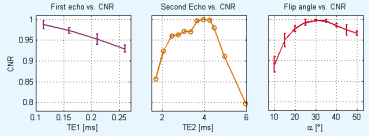


Figure 4: *Normalized CNR results.* ***Left:*** *The first echo time varied, average over 4 repetitions.* ***Middle:*** *The second echo time varied.* ***Right:*** *Varying flip angles, average over 3 repetitions.*

Conclusion on UTE imaging parameters

UTE images should be acquired with the **minimal first echo time**. The lower limit is due to hardware (coils). The **second echo time** should be ***~3.6 ms***, and the best choice for **flip angle** is approximately ***25°***.

Bone Classification

In order to mimic a CT image, the MRI needs to be classified into different tissues representing large differences in Hounsfield Units (HU). Each tissue is assigned a bulk value of either -1000 HU (air), 0 HU (water) or ~600 HU (or something similar high, bone).

Method: Segmentation is done with a voxel-driven approach based on the Bayes classifier and Markov Random Fields. Currently the classification is done using the voxel-intensity from both echoes, but the method can potentially be expanded to include other MRI sequences (for example a T2 weighted sequence).

The classifier training is currently done by manual labelling of four tissue types: Air, brain, bone & other soft-tissue (skin layer).

Results: The true CT and MRI mimic is shown in Figure 5.

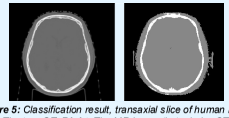
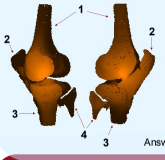


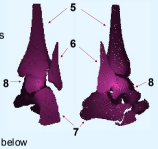
Figure 5: *Classification result, transaxial slice of human head.* ***Left:*** *The true CT.* ***Right:*** *The MR image that mimics CT.*

3D point clouds of the segmented bones from other UTE scans are shown below in order to inspect the classification result visually in the entire volume instead of just a slice-wise result.

Is the bone anatomy recognisable ? Which bones can be seen ?



a. Femur
b. Humerus
c. Calcaneus
d. Fibula
e. Ulna
f. Talus
g. Radius
h. Tibia
i. Patella



Answers are given below

Conclusion on bone classification

The UTE sequence demonstrated that bone segmentation is possible. Several issues remain to be addressed with the current approach. Especially how the method can be automated.

To further improve the results additional image information is needed. This could be extra MRI sequences or the inclusion of CT atlases.

Answers: 1a, 2i, 3h, 4d, 5f, 6d, 7c, 8f

Appendix C

Abstract for ESTRO 2012, accepted for oral presentation

Auto-segmentation of bone in MRI-only based radiotherapy using ultra short echo time.

J.M. Edmund[1], H.M. Kjer[2], R.H. Hansen[3]

1) Purpose/Objective

Treatment planning based on MRI-only has shown a promising potential if bulk density assignment for tissue, air and bone are taken into consideration. A major issue is the need to auto-segment these bulk tissue structures in the MRI image to make the approach feasible. This is, however, complicated by an extremely short T2 relaxation time (~ 1 ms) of the bone resulting in no signal using conventional MRI sequences. Here, we present an approach adapted from PET/MRI attenuation maps to automatically segment the bone using MRI sequences based on ultra short echo times (UTE).

(2) Material/methods

A cutaway from the front leg of a calf including the knee-joint was used as a phantom. The MR images were acquired on a 3.0-T MRI scanner (Philips Achieva) using a cardiac coil to cover the entire phantom. The UTE sequence applies two different echo times, TE1 and TE2, which were 0.2 and 1.9 ms, respectively, a flip angle of 10^{o}, and a TR of 4.0 ms. An isotropic voxel dimension of 1,8 mm was obtained with a FOV of 240mm A reference CT scan (Philips Big Bore CT) was also acquired for comparison.

Processing of the TE1 and TE2 MR images was done in MatLab using the DICOM toolbox. First, a TE1-2 image is created by subtracting TE2 from TE1 (figure, row 1). This image is then masked with a binary image of TE2 creating a TE1-2* image with a more well defined outer contour and discrimination between air and tissue (figure, row 2). Two different filters based on the most insensitive pixel (MIP) were applied to auto-segment TE1-2* into tissue and bone. Method 1 (M1): tissue (outer contour) > MIP/2.4 > bone > MIP/5 > tissue > 0. Method 2 (M2): bone > MIP/5 > tissue > 0 (figure, row 3).

The TE1-2* and segmented M1 and M2 scans were registered with the reference CT scan in Eclipse v.10 (Varian Medical Systems). A structure set containing (auto-segmented) bone and tissue (auto-segmented body - bone) were created for the CT-, M1- and M2-scan (figure, row 3). A CT based treatment plan containing two opposing APPA fields giving 2 Gy to the iso-center (middle of phantom) was created and re-calculated on M1 and M2 using tissue = 0 HU and bone = 362 HU (CT average).

(3) Results

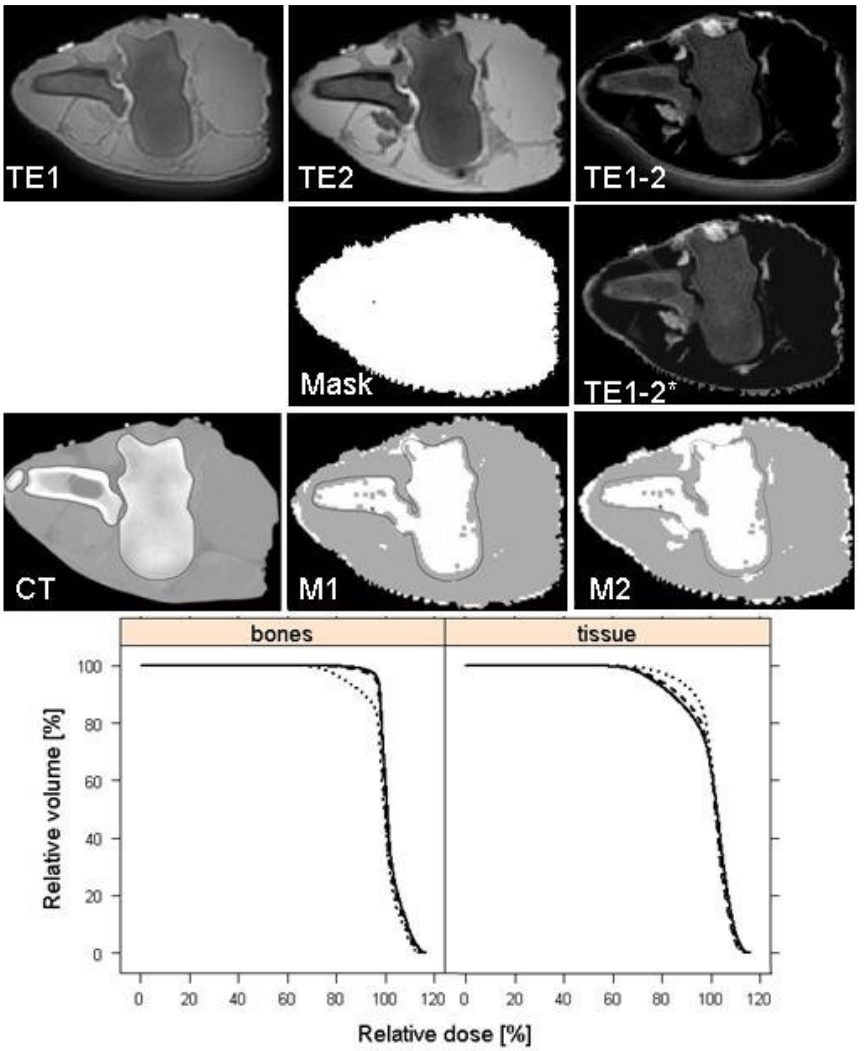
87 % and 72 % of the M1 and M2 bone agreed with the CT reference (intersection of volumes). Non-bone was 13% and 28 % for M1 and M2 bone (M1/M2 volume-intersection). Bone-miss was 28% and 15% for M1 and M2 (CT volume-intersection).

The dosimetric differences were less than 1.5 % in the iso-center. The DVHs of the bone and tissue (figure, row 4) show good agreement between M1 and CT, $\Delta D_{98\%}$ = 2/-4 % (tissue/bone) while M2 shows a distinct deviation from CT, $\Delta D_{98\%}$ = 12/-18 % (tissue/bone).

(4) Conclusion.

Although there's able room for improvement, the use of UTE sequences to create contrast for bone in MRI images was demonstrated. From a geometric and dosimetric point of view, it seems important to correct for the outer contour when combining the two UTE images. This can, however, result in a larger bone-miss and a compromise has to be found with the presented MIP approach.

Figure.



Row 1:TE1=0.2ms, TE2=1.9ms, TE1-2=Image difference TE1-TE2.
Row 2: Mask=Binary map for air/tissue, TE1-2*=Mask·TE1-2.
Row 3: CT=Reference, M1=Method 1, M2=Method 2.
Row 4: DVH of bones and tissue. Lines: Solid=CT, Dashed=M1, Dotted=M2.

Bibliography

[1] R. C. Krempien, K. Schubert, D. Zierhut, M. C. Steckner, M. Treiber, W. Harms, U. Mende, D. Latz, M. Wannenmacher, and F. Wenz, "Open low-field magnetic resonance imaging in radiation therapy treatment planning," *International Journal of Radiation Oncology*Biology*Physics*, vol. 53, no. 5, pp. 1350–1360, 2002.

[2] V. S. Khoo and D. L. Joon, "New developments in MRI for target volume delineation in radiotherapy," *British Journal of Radiology*, vol. 79, no. Special Issue 1, pp. S2–S15, 2006.

[3] T. Nyholm, M. Nyberg, M. G. Karlsson, and M. Karlsson, "Systematisation of spatial uncertainties for comparison between a MR and a CT-based radiotherapy workflow for prostate treatments," *Radiation Oncology*, vol. 4, no. 1, p. 54, 2009.

[4] M. Hofmann, F. Steinke, V. Scheel, G. Charpiat, J. Farquhar, P. Aschoff, M. Brady, B. Schölkopf, and B. J. Pichler, "MRI-based attenuation correction for PET/MRI: A novel approach combining pattern recognition and atlas registration," *Journal of Nuclear Medicine*, vol. 49, no. 11, pp. 1875–1883, 2008.

[5] J. A. Dowling, J. Lambert, J. Parker, O. Salvado, J. Fripp, A. Capp, C. Wratten, J. W. Denham, and P. B. Greer, "An atlas-based electron density mapping method for magnetic resonance imaging (MRI)-alone treatment planning and adaptive MRI-based prostate radiation therapy," *International Journal of Radiation Oncology*Biology*Physics*, vol. 83, no. 1, pp. e5 – e11, 2012.

[6] M. Hofmann, B. Pichler, B. Schölkopf, and T. Beyer, "Towards quantitative PET/MRI: a review of MR-based attenuation correction techniques," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 36, pp. 93–104, 2009.

[7] B. H. Kristensen, F. J. Laursen, V. Løgager, P. F. Geertsen, and A. Krarup-Hansen, "Dosimetric and geometric evaluation of an open low-field magnetic resonance simulator for radiotherapy treatment planning of brain tumours," *Radiotherapy and Oncology*, vol. 87, no. 1, pp. 100 – 109, 2008.

[8] J. Lambert, P. B. Greer, F. Menk, J. Patterson, J. Parker, K. Dahl, S. Gupta, A. Capp, C. Wratten, C. Tang, M. Kumar, J. Dowling, S. Hauville, C. Hughes, K. Fisher, P. Lau, J. W. Denham, and O. Salvado, "MRI-guided prostate radiation therapy planning: Investigation of dosimetric accuracy of MRI-based dose planning," *Radiotherapy and Oncology*, vol. 98, no. 3, pp. 330–334, 2011.

[9] T. Boettger, T. Nyholm, M. Karlsson, C. Nunna, and J. C. Celi, "Radiation therapy planning and simulation with magnetic resonance images," *Proceedings of SPIE*, vol. 6918.

[10] S. Thörnqvist, "Initial step toward MRI-based treatment planning for external radiotherapy," Master's thesis, Lund University, 2009.

[11] V. Keereman, Y. Fierens, T. Broux, Y. De Deene, M. Lonneux, and S. Vandenberghe, "MRI-based attenuation correction for PET/MRI using ultrashort echo time sequences," *Journal of Nuclear Medicine*, vol. 51, no. 5, pp. 812–818, 2010.

[12] C. Catana, A. van der Kouwe, T. Benner, C. J. Michel, M. Hamm, M. Fenchel, B. Fischl, B. Rosen, M. Schmand, and G. A. Sorensen, "Toward implementing an MRI-based PET attenuation-correction method for neurologic studies on the MR-PET brain prototype," *Journal of Nuclear Medicine*, vol. 51, no. 9, pp. 1431–1438, 2010.

[13] A. Johansson, M. Karlsson, and T. Nyholm, "CT substitute derived from MRI sequences with ultrashort echo time," *Medical Physics*, vol. 38, no. 5, pp. 2708–2714, 2011.

[14] Y. Berker, J. Franke, A. Salomon, M. Palmowski, H. C. Donker, Y. Temur, F. M. Mottaghy, C. Kuhl, D. Izquierdo-Garcia, Z. A. Fayad, F. Kiessling, and V. Schulz, "MRI-based attenuation correction for hybrid PET/MRI systems: A 4-class tissue segmentation technique using a combined ultrashort-echo-time/dixon MRI sequence," *Journal of Nuclear Medicine*, vol. 53, no. 5, pp. 796–804, 2012.

[15] P. D. Gatehouse, R. W. Thomas, M. D. Robson, G. Hamilton, A. H. Herlihy, and G. M. Bydder, "Magnetic resonance imaging of the knee with ultrashort TE pulse sequences," *Magnetic Resonance Imaging*, vol. 22, no. 8, pp. 1061–1067, 2004.

[16] M. A. Hall-Craggs, J. Porter, P. D. Gatehouse, and G. M. Bydder, "Ultrashort echo time (UTE) MRI of the spine in thalassaemia," *British Journal of Radiology*, vol. 77, no. 914, pp. 104–110, 2004.

[17] M. D. Robson, P. D. Gatehouse, M. Bydder, and G. M. Bydder, "Magnetic resonance: An introduction to ultrashort TE (UTE) imaging," *Journal of Computer Assisted Tomography*, vol. 27, no. 6, pp. 825–846, 2003.

[18] J. Jonsson, M. Karlsson, M. Karlsson, and T. Nyholm, "Treatment planning using MRI data: an analysis of the dose calculation accuracy for different treatment regions," *Radiation Oncology*, vol. 5, no. 1, p. 62, 2010.

[19] M. E. Korsholm, L. W. Waring, R. R. Paulsen, and J. M. Edmund, "Statistical analysis of MRI-only based dose planning," *Radiotherapy & Oncology*, vol. 103, no. Supplement 1, p. S134, 2012.

[20] D. G. Nishimura, *Principle of Magnetic Resonance Imaging*. lulu.com, 2010.

[21] L. G. Hanson, *Introduction to Magnetic Resonance Imaging Techniques*. 2009.

[22] J. Rahmer, P. Börnert, J. Groen, and C. Bos, "Three-dimensional radial ultrashort echo-time imaging with T2 adapted sampling," *Magnetic Resonance in Medicine*, vol. 55, no. 5, pp. 1075–1082, 2006.

[23] D. J. Tyler, M. D. Robson, R. M. Henkelman, I. R. Young, and G. M. Bydder, "Magnetic resonance imaging with ultrashort TE (UTE) pulse sequences: Technical considerations," *Journal of Magnetic Resonance Imaging*, vol. 25, no. 2, pp. 279–289, 2007.

[24] S. B. McGrayne, *The Theory That Would Not Die: How Bayes' Rule Cracked The Enigma Code, Hunted Down Russian Submarines, & Emerged Triumphant from Two Centuries of Controversy*. New Haven: Yale University Press, 2011.

[25] R. R. Paulsen and T. B. Moeslund, *Introduction to Medical Image Analysis*. DTU Informatics, 2011.

[26] K. V. Leemput and R. Larsen, "02505 course note - medical image analysis." 2011.

[27] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. pp. 1–38, 1977.

[28] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. pp. 297–302, 1945.

[29] R. Drzymala, R. Mohan, L. Brewster, J. Chu, M. Goitein, W. Harms, and M. Urie, "Dose-volume histograms," *International Journal of Radiation Oncology*Biology*Physics*, vol. 21, no. 1, pp. 71 – 78, 1991.

[30] G. G. Steel, *Basic Clinical Radiobiology, 3rd Edition.* Hodder Arnold Publishers, 2002.

[31] V. Gondi, B. P. Hermann, M. P. Mehta, and W. A. Tomé, "Hippocampal dosimetry predicts neurocognitive function impairment after fractionated stereotactic radiotherapy for benign or low-grade adult brain tumors," *International Journal of Radiation Oncology*Biology*Physics*, vol. 83, no. 4, pp. e487 – e493, 2012.

[32] *ICRU report nr. 46, Photon, electron, proton and neutron interaction data for body tissues.* International Commission on Radiation Units and Measurements., 1992.