CHALMERS

Development of Patient-Specific Electromagnetic

Model (PSEM) based on MR breast images

Master of Science Thesis

CHRISTIAN WEBER

Development of Patient-Specific Electromagnetic Model (PSEM) based on MR breast images

Christian Weber

Department of Informatics
Karlsruhe Institute of Technology

Supervisor: Hoi-Shun Lui

Department of Signals and Systems
Chalmers University of Technology

27 September 2010

Abstract

Patient-Specific Electromagnetic Breast Models are of use in several applications such as breast cancer detection and treatment. This thesis will be broken up into three parts. The first part will discuss and review the paper by astrow et al. [14] called "Development of Anatomically Realistic Numerical Breast Phantoms With Accurate Dielectric Properties for Modeling Microwave Interactions with the Human Breast". Secondly, a classification algorithm based on image neighbourhood statistics will be adapted to the specifics of breast tissue in order to improve the accuracy of the developed eletromagnetic breast model. Finally the results of both approaches will be compared and discussed. This comparison will be achieved with respect to consistency of the results; this includes a visual inspection of the numerical model. In addition the difficulty of segmenting fibro-connective breast tissue will be addressed. The results suggest the proposed approach is more stable than the original method.

Contents

Acknowledgements

I would like to thank some people that helped me while working in this thesis:

Firstly I'd like to thank my parents for their ongoing support and encouragement throughout my studies.

My supervisor Hoi-Shun for all his constructive input and patience.

Rachel for checking and correcting my writings,

and Fernando for all the fun time working together in the library.

1 Introduction

Among women breast cancer is the most occuring cancer, with 1.38 million newly diagnosed cancer cases in 2008, and the second most cancer overall[5]. Currently X-ray mammography and breast Magnetic Resonance Imaging (MRI) are the most common methods for breast cancer diagnostics. Since the X-ray is an ionizing radiation which is harmful to humans and MR scans are economically expensive, researchers are looking into alternative means of breast cancer detection. One such method is microwave tomography; it uses non-ionising radiation and is economically cheap. Patient-specific electromagnetic breast models are providing a platform for numerical studies for microwave imaging techniques of breast cancer detection. In relation to microwave tomograpghy, a priori knowledge of dielectric properties is shown to improve imaging results. It also enables the reconstruction of small objects using a lower frequency than the one that would be needed without a priori knowledge [4]. Another application is the enhancement of hyperthermia algorithms for cancer treatment [11]. The dielectric properties of breast tissue are determined by the tissue type. The two main types of breast tissues are the fatty and glandular. In some regions the glandular tissue assumes a very thin, threadlike shape, the so called fibro-connective tissue. In order to build a numerically accurate model of the dielectric properties of breasts it is necessary to treat the glandular tissue appearing compacted and the glandular tissue assuming a fibro-connective shape differently.

1.1 Challenges in segmentation task

Tissue segmentation in MR images is a non-trivial task as the image is degraded in several ways. The main influences are inhomogeneities in the magnetic field of an MR machine, noise and the partial volume effect. Depending on the location within the magentic field, inhomogeneities cause the intensity of one tissue type to vary. This leads to an overlap in intensity ranges of different tissue types. Even though very good methods for nonuniformity removal exist, it is usually impossible to completely remove all effects of inhomogeneity in the field. The current trend in MR imaging is to use higher magnetic field strength to increase the SNR; however this also increases the effects of field inhomogeneity. The partial volume effect describes the effect where due to limited spatial resolution, the signal is influenced by more than one tissue type. This occurs at all transitions between tissue types as well as at the air-breast boundary. Details of this process will be discussed in section 3.4.2. These factors make a simple segmenta-

tion method, such as thresholding using nothing but the intensity values, impractical in most cases.

1.2 Dielectric Properties

1.3 Thesis Layout

This thesis is structured as follows: A literature review of the method by Zastrow et al. [14] will be given and the mathematical method used in their work will be presented. The following section will introduce an additional segmentation step and its motivated choice. The method will then be explained in detail and the implementational issues will be discussed. The intermediate results from the first implementation will be further explored, as well as the problems that arose in the context of the breast tissue segmentation. Finally the results of both approaches will be evaluated and then concluded with a discussion of the original and proposed methods.

2 Literature Review

In this section a literature review of a proposed method for the "Development of Anatomically Realistic Numerical Breast Phantoms With Accurate Dielectric Properties for Modelling Microwave Interactions With the Human Breast" by Zastrow et al. [14] is given. A minor change in the smoothing process will also be introduced.

2.1 Preprocessing

In order to determine the parameters for the mapping from MRI intensity to the dielectric properties, some general preprocessing steps are applied in order to prepare the images for further processing. In the first step, the breast region is manually extracted from the MR images; a homomorphic filter is then applied to remove the effects of non-uniformity in the magnetic field during the imaging process. In the reimplementation of the algorithm in this thesis, the MINC toolbox[1] from the McConnell Brain Imaging Centre is used to remove inhomogeneities. After this the breast volume is segmented from the background and a smooth three dimensional breast border model is built.

2.1.1 3D Border Model

Firstly each slice is processed to determine a threshold value that discriminates between breast volume and background. For this task, Otsu's histogram based method, for automatic thresholding of two classes, gives consistently good results. Based on these thresholds a breast mask is built. This is created through examining all images and parsing them from all four borders towards the centre. All pixels that are below the threshold are set to zero until a pixel is found that exceeds the threshold. After this, all pixels that are left over in this row/column are set to one. This provides four masks, one for each direction, that can be combined by using an AND operation. Once applied on all slices, this finally allows a 3D mask to mark the breast volume with ones and the background with zeros. Additionally, each time the parsing hits the first non-background pixel this position is stored in an array. These vectors are used in the border smoothing process. Figure 1 shows an example of the four different masks and the resulting recognised border on a coronal[2] breast slice.

[1]http://www.bic.mni.mcgill.ca/ServicesSoftware/MINC
[2]see Appendix, Fig. 16
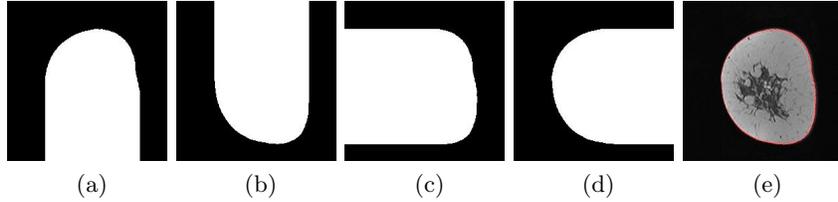
Figure 1: Creating the breast mask. Fig. 1a-d showing the four direction-masks, Fig. 1e shows the recognised border (in red) on the original slice.

2.1.2 Smoothing

a) Smoothing the border in the coronal plane
In order to remove artificially introduced roughness, the breast boundaries are smoothed. Zastrow et al. [14] smoothed the border by applying a best-fit ellipse on each slice and setting them as new breast border. Since this prevents the algorithm from processing images of compressed breasts, that do not comply with an elliptic shape, there is another approach that is less dependant on a specific shape of the breast.

In this approach the border points of each slice are converted from Cartesian into polar coordinates and then sorted by their angle. An energy function depending on the deviation of the radius of each element to its two direct neighbours is then defined. An energy function based on the radii also implies a circular structure, as the best-fit ellipse. However, such an energy function is more flexible to adapt to the natural variations in breast shapes and also in particular to compression induced deformation. Since the sole minimisation of the deviation of the radii alone would always lead to a perfect circle, a penalty value is added, which depends on the difference between the smoothed border and the initial border. The optimisation process is undertaken using an iterative process. This is where the border points are divided into two sets, even and odd numbers. Updating a border point directly accects the energy values that its neighbouring border points hold; to simplify the calculations in each iteration only one set of border points is updated, eighter odd numbers or even numbers.

The energy function for moving the border point inwards and outwards is calculated and compared with its current energy value. The point is then updated to the option that holds minimal energy. In every iteration, the set that is used is switched, in order to avoid oscillation. The iterations are stopped as soon as there are no more updates on the border points. This works well on most parts of the breast volume; it allows a smooth border

8

that is very close to the objective border is found. The the breast nipple often creates a highly asymmetrical shape in one or two slices and the algorithm tends to extend the border into the background region. Since this affects only a very small region in the model the effects are neglected here.

b) Interpolating the border in the sagittal plane
In order to create a numerical phantom with isotropic voxels it is often necessary to interpolate between slices. This is especially important when the slice thickness does not match the dimensions of spatial resolution within the plane. Border interpolation is required to approximate a continuous

(a)   (b)
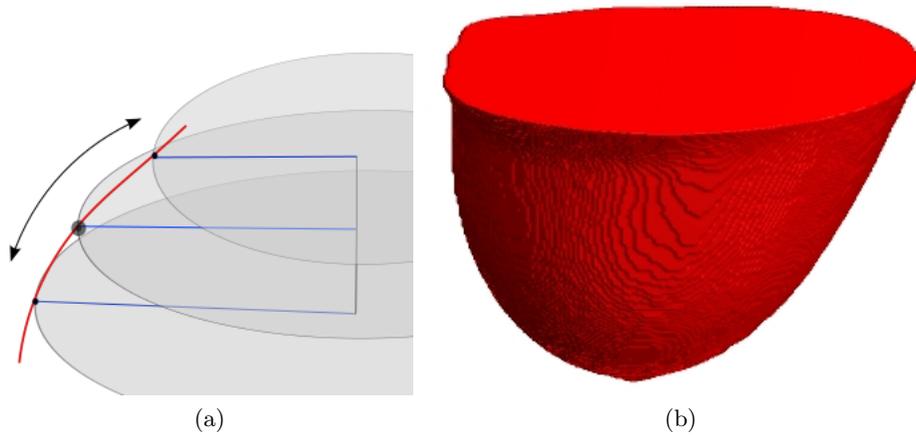
Figure 2: Fig. 2a Interpolating breast border between MRI slices.   Fig. 2b 3D Breast Volume after smoothing and border interpolation.

breast border between slices. For a given angle the radii of the breast border of three consecutive slices are taken and cubic spline interpolation is used between those points. This process is used for every second slice; the breast border is extended to the next and the previous slice respectively. Fig. 2a shows a schematic of the idea and Fig. 2b shows a 3D model of a fully segmented and smoothed breast volume.

2.2   Tissue Intensity Estimation

After breast volume has been separated from the background, the next step is to determine the intensity values for the different tissue types, i.e. glandular and fatty tissue. It is assumed that the distributions of the intensity

values of the two tissues follow a Gaussian distribution. Therefore the intensity histogram can be modelled as a mixture of Gaussians (GMM); the Gaussian with the higher mean value represents the fat tissue and the one with the lower mean value corresponds to the glandular tissue. The parameters $\mu_{fat}, \mu_{gland}, \sigma_{fat}$ and σ_{gland} of the GMM need to be estimated as only the intensity values observed; due to the fact that it is unknown which tissue type produced which intensity value. This is obtained using the Expectation Maximization (EM) algorithm. The EM algorithm can be used in various optimisation problems, it is presented here with respect to optimising the *goodness* of the GMM parameters.

2.2.1 Expectation Maximization Algorithm

Let $X = \{x_1, \ldots, x_N\}$ be a set of intensity values and f be a mixture of K Gaussians with $\Phi = \{\Phi_1, \ldots \Phi_K\} = \{(p_1, \mu_1, \sigma_1), \ldots, (p_K, \mu_K, \sigma_K)\}$ describing the parameters of the Gaussians and $p_k(x)$ the probability that a random sample $x \in X$ was generated by tissue class $1 \le k \le K$. Then f takes the following form:

$$f(x \mid \Phi) = \sum_{k=1}^{K} q(x, k) \tag{1}$$

with

$$q(x, k) = p_k(x) \frac{1}{\sqrt{2\pi}\sigma_k} e^{-\frac{1}{2}\left(\frac{\|x - \mu_k\|}{\sigma_k}\right)^2} \tag{2}$$

The Maximum Log-Likelihood Estimation (MLE) can be used to describe the probability of the parameters Φ given the observed data X. Mathematically the function $L(\Phi \mid X)$ can be given by

$$L(\Phi \mid X) = \log \prod_{n=1}^{N} f(x_n \mid \Phi) = \sum_{n=1}^{N} \log f(x_n \mid \Phi) \tag{3}$$

This allows us to formalise the problem of finding the unknown parameters Φ The objective is to find a Φ^* such that the probability of the MLE is maximised (4).

$$\Phi^* = \arg\max_{\Phi} \; L(\Phi \mid X) \tag{4}$$

Since it is difficult to maximise this equation directly, the EM algorithm is used for finding the maximum in an iterative manner. The EM algorithm

was published in a formalised form in 1977 in an article from Dempster, Laird and Rubin [3]. The main idea is to perform the optimisation iteratively and divide it into two steps. At first, instead of working on the target function directly, a function is constructed that gives a lower bound of the target function. This lower bound behaves similarly to the target function in the proximity of the currently estimated parameters (here Φ). This step is called the *Expectation Step*. In the second step this expected value is maximised, hence it is called the *Maximization Step*. For the problem at hand, the so-called membership probabilities $\omega_{k,n}^r$ are introduced. They indicate the degree to which a sample x_n belongs to the Gaussian numbered k in iteration round r. Given a Φ they can be easily calculated by the following equation:

$$\omega_{k,n}^r = \frac{p_k^r N(x_n, \mu_k^r, \sigma_k^r)}{\sum_{c=1}^{K} p_c^r N(x_n, \mu_c^r, \sigma_c^r)} \tag{5}$$

with

$$N(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{\|x-\mu\|}{\sigma}\right)^2}$$

A lower bound can then be constructed by introducing the estimated membership probabilities. Replacing the $q(x,k)$ with $\frac{\omega_{k,n}^r}{\omega_{k,n}^r} q(x,k)$ in equation 3 and applying Jensen's inequality provides such a lower bound b_r (right hand side of the inequation 6) of the log likelihood. This can be rewritten to (7).

$$L(\Phi \mid X) = \sum_{n=1}^{N} \log \sum_{k=1}^{K} \frac{\omega_{k,n}^r}{\omega_{k,n}^r} q(x,k) \geq \sum_{n=1}^{N} \sum_{k=1}^{K} \omega_{k,n}^r \log \frac{q(x,k)}{\omega_{k,n}^r} = b_r(\phi) \tag{6}$$

$$b_r(\Phi) = \sum_{n=1}^{N} \sum_{k=1}^{K} \omega_{k,n}^r \log q(x,k) - \underbrace{\sum_{n=1}^{N} \sum_{k=1}^{K} \omega_{k,n}^r \log \omega_{k,n}^r}_{\text{indepedent of } \Phi} \tag{7}$$

Equation 7 is the resulting lower bound of the *Expectation Step*. The second step (Maximisation) now aims to optimise the lower boundary through optimising Φ. Since the second part of equation (7) is independent of Φ it is irrelevant for the optimisation and it can therefore be ignored. The maximisation can be completed on a linear combination of logarithms rather than on the logarithm of a sum (as completed in (3)). The equations (8,9,10)

11

that maximise the given lower boundary b_r can be obtained through deriving the first part of equation (7) with respect to μ_k, σ_k, p_k and combining the results.

$$\mu_k^{r+1} = \frac{\sum_{n=1}^{N} \omega_{k,n}^r x_n}{\sum_{n=1}^{N} \omega_{k,n}^r} \tag{8}$$

$$\sigma_k^{r+1} = \sqrt{\frac{\sum_{n=1}^{N} \omega_{k,n}^r \parallel x_n - \mu_k^{r+1} \parallel^2}{\sum_{n=1}^{N} \omega_{k,n}^r}} \tag{9}$$

$$p_k^{r+1} = \frac{1}{N} \sum_{n=1}^{N} \omega_{k,n}^r \tag{10}$$

These updated values are used in the next iteration of the algorithm. The procedure is terminated when it converges to a local maximum of the lower bound, this also is a local maximum of the log likelihood function. For the first iteration the algorithm needs to be initialised with guessed or random values. [2], [10]
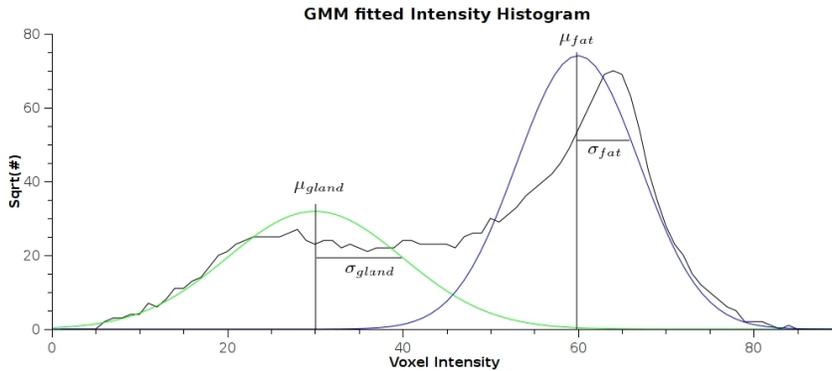
Figure 3: Example of a histogram fitted with two Gaussians.

Figure 3 shows a histogram that has been fitted with two Gaussians. Parameters have been estimated using the EM Algorithm. Due to the domination of fat tissue the square root of the occurrences of intensity values are used in the histogram.

2.3 Dielectric properties of normal breast tissue

The following mapping of intensity values to dielectric properties is based on the Wisconsin-Calgary study [6]. In this large-scale study the dielectric

properties of samples obtained by breast reduction surgeries are measured. The results show that the dielectric properties are mainly influenced by the fat content in the samples. The study also finds that there is a high variation of dielectric properties. This variation is to be taken into account when building numerical breast phantoms. The observed results are fitted to a Cole-Cole model and the resulting parameters are listed in table 8.

2.4 Intensity value mappings

Based on the findings of the Wisconsin-Calgary study the intensity range of the MRIs is split into seven ranges, representing the seven tissue types. The ranges are as follows: three for glandular tissue, three for fatty tissue and one for the transitional region between glandular and fatty tissue. The tissue types are assigned to the seven intensity ranges as shown in table 1. The intensity range is based on the parameters of the GMM. In Fig. 4 the regions are indicated by dashed lines. An example of a piecewise linear mapping from intensity values to dielectric constants is given in Fig. 5

tissue type	min	max
glandular/friboconnective-1	$min(I)$	$\mu_{gland} - \sigma_{gland}$
glandular/friboconnective-2	$\mu_{gland} - \sigma_{gland}$	μ_{gland}
glandular/friboconnective-3	μ_{gland}	$\mu_{gland} + \sigma_{gland}$
transitional	$\mu_{gland} + \sigma_{gland}$	$\mu_{fat} - \sigma_{fat}$
fatty-1	$\mu_{fat} - \sigma_{fat}$	μ_{fat}
fatty-2	μ_{fat}	$\mu_{fat} + \sigma_{fat}$
fatty-3	$\mu_{fat} + \sigma_{fat}$	$max(I)$

Table 1: List of tissue types according to voxel intensity ranges.
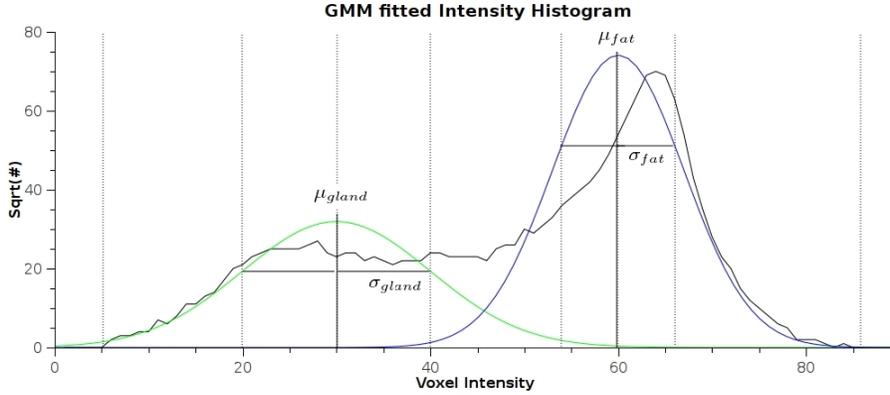
Figure 4: Voxel Intensity to Dielectric Property Mapping
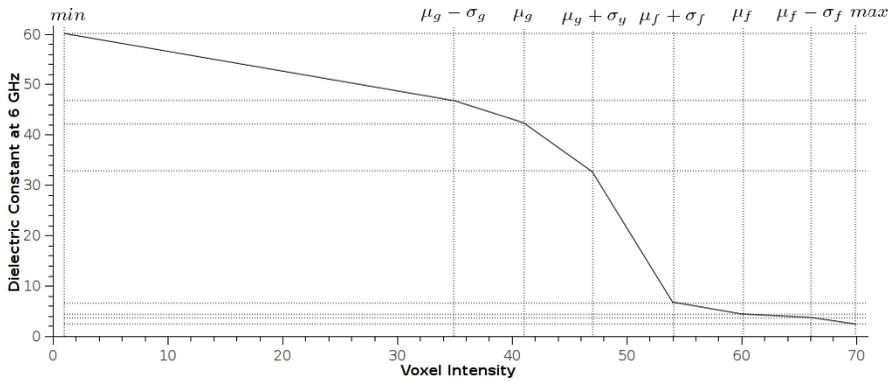
Figure 5: Example of Voxel Intensity to Dielectric Property Mapping

3 Proposed Modifications

3.1 Motivation

As shown in the approach above, an explicit assumption is made regarding the distribution of the intensity values; this is namely the normal distribution, that only partly reflects the true nature of that distribution. Due to noise, the partial volume effect and inhomogeneities that are left after pre-processsing the two Gaussians are never fully separable. This overlap of the Gaussians inherently introduces some misclassification. Through introducing an artificial tissue type that covers the transition from fatty tissue to glandular tissue by linear interpolation between the two dielectric prop-

14

erty values, the misclassifications can be addressed. This aims to reduce the error in the transitional range in the dielectric model. However, the introduction of this intensity range leads to a general overestimation of the dielectric properties for fatty tissue and an underestimation for glandular tissue.

The following method modification aims to counter the drawbacks by including additional information in the classification process. This will help deciding whether a voxel represents fatty or glandular tissue. For this task, a classification algorithm working on the statistics of pixel neighbourhoods has been chosen. The following sections will include an explanation of the basic idea behind the classification approach; a detailed presentation of the algorithm and then a discussion of the modifications of the methods in [9]. The aim of these modifications is to adapt the method of neighbourhood statistics to the specific characteristics of breast tissue classification.

3.2   Neighbourhood Statistics

Suyash and Whitaker [1] present a method for image denoising and reconstruction, based on higher order image statistics. Through working with so-called neighbourhoods they are able to denoise images and reconstruct fine structures of degraded fingerprint images. Tasdizen et al. [9] extended this method into a classification algorithm and applied it on brain MR images. There are two main ideas behind this approach; the first is that neighbourhoods, belonging to the same object, share the same or similar patterns of pixel intensities. This suggests that pixel-neighbourhoods of natural images are not truly random. This means means that through observing a neighbourhood, this provides some information about its centre pixel. The second idea is that the addition of two random variables (e.g. signal and additive noise) increases entropy [1]. The reduction of entropy can lead to a reduction of the influence of noise and thereby improve the segmentation result.

The rest of this section deals with the mathematical formulation of the above method. The notion of a digital image interpreted as a random field and the definition of an image neighbourhood is introduced. A definition of entropy on neighbourhoods will be presented which is then used to give a formalisation of an optimal segmentation. A discussion of the means to achieve this segmentation and their implementation will be presented.

15

3.2.1 Random Field

Let $\mathbb{X}(\Omega, T)$ be a family of random variables, where T is a set of indices and Ω is the sample space. If we now consider T to be a set of pixel indices in an image and Ω to be the set of all possible intensity values, then a digital image can be described as the realisation of a random field where $\mathbb{X}(\omega, t)$ gives the intensity of pixel $t \in T$ for fixed $\omega \in \Omega$ [1]. Furthermore, we define the neighbourhood $N_t \subset T$ of a pixel t as a set of pixels for which the following holds:

$$u \in N_t \iff t \in N_u$$

Figure 6 shows some commonly used neighbourhood shapes. Now we define another random vector $Z(t) = \{\mathbb{X}(\Omega, t)\}_{t \in N_t}$, we use the short term $z(t)$ for its realization. In other words, the vector $z(t)$ corresponds to the intensity values of the pixels t's neighbourhood.
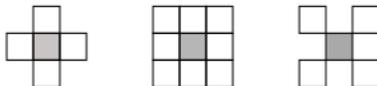

Figure 6: Example of commonly used neighbourhoods

3.2.2 Entropy

In information theory, entropy refers to the measure of information content of a random variable [13]. In other words, entropy describes how much information one gains by observing an event based on a random variable. For example if a fair coin is tossed, information gained by observing the event is maximal since nothing is known about the outcome beforehand. However, if the coin was unfair, for example the coin had heads on both sides, the average information gained by actually observing the event would be zero because the outcome is known without even observing the actual event.

For a discrete random variable X with values $x_i \in \{x_1, \ldots, x_n\}$, the entropy H is defined as:

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log p(x_i) \tag{11}$$

3.2.3 Neighbourhood Entropy

Let $k \in \{1 \ldots K\}$ denote the different tissue classes in an MRI, and the sets $\{T_k\}_{k=1}^{K}$ be mutually exclusive and a complete decomposition of T into

these tissue classes. Furthermore let $p_k(Z(t) = z(t))$ be the probability the specific neighbourhood $z(t)$ to be observed at pixel t assuming that this pixel belongs to the tissue class k [9].

The neighbourhood entropy of an image can then be computed as

$$H(T_1, \ldots, T_K) = -\sum_{k=1}^{K} \sum_{t \in T} \Big(p_k\big(Z(t) = z(t)\big) \cdot \log p_k\big(Z(t) = z(t)\big) \Big) \quad (12)$$

This measure can be used for evaluating a given segmentation of an image, by first calculating $p_k(x)$ for all $x \in \mathbb{X}$ (see 3.3) and then applying 12. The result indicates the information content of the image neighbourhoods. A high entropy characterises a more random classification with respect to the pixel neighbourhoods whereas a lower entropy indicates a high correlation between classification of the centre pixel, and its observed neighbourhood.

3.2.4 Classification by Entropy Optimisation

The actual pixel intensity is possibly degraded by influences such as noise or partial volume effect, thus reducing the information content carried in its neighbourhood. Therefore, in this approach a classification is considered optimal when it restores the information carried by the image neighbourhoods. This is the equivalent to reducing the neighbourhood entropy. This can be completed through assigning the classification label for each pixel to the tissue class that holds the highest probability; given the observed neighbourhood. Equation 13 describes mathematically how the new classification sets $\tilde{T}_k$ are constructed using the tissue probabilities.

$$\tilde{T}_k = \{ t \in T \mid p_k(z(t)) \geq p_i(z(t)), \forall i \neq k \} \quad (13)$$

Classification is achieved through an iterative process, each time the labels are updated and the probabilities p_k are re-estimated. Repeating this process will reduce the neighbourhood entropy as well as the influence of noise. In [9], it is proposed to stop this iterative process when the reduction of entropy drops below 0.1%.

3.3 Nonparametric Multivariate Density Estimation

Estimating the neighbourhood probabilities p_k correctly is an essential part in the segmentation algorithm, as described above. Studies on the statistics of 3×3 patches in natural images show that the distribution of the data is very sparse and exhibits nontrivial topologies [7]. This motivates the use of

17

a nonparametric multivariate density estimation to calculate the probability density function. This allows the algorithm to adapt to the specific topologies of the different tissue types, and to adapt automatically to different modalities and the use of contrast agent in MR imaging.

3.3.1 Parzen Window Method

The Parzen Window method is a statistical instrument with which it is possible to estimate the probability density function of a random variable given a set of events. This method works without any priori assumptions relating to the distributional nature of the random variable. Therefore, the Parzen Window method can be used to estimate the probability p_k of a pixel t with neighbourhood $z(t)$ belonging to tissue class k based on the observed neighbourhoods from other pixels that are initially classified as tissue class k. Mathematically this is given by:

$$p_k(z(t)) \approx \frac{1}{\mid A_k(t) \mid} \sum_{t_j \in A_k(t)} e^{-\frac{\|z(t) - z(t_j)\|}{2\sigma^2}} \tag{14}$$

where $A_k(t)$ is a subset of T_k consisting of pixels in the vicinity of t that are classified as tissue class k. If $\mathbb{X}$ was a truly stationary process it would be sufficient to have global A_k's but in practical applications it is often necessary to build A_k in dependence of t.

3.3.2 Estimating the window size σ

Choosing the right window size σ is critical for a good density estimation. The window size determines the distance around the value t. All points within this distance are relevant to the probability density estimation concercing this point. Therefore, a window size that is too small or large for the given data will lead to an under-smoothed or over-smoothed probability density function. To evaluate the performance of the density estimation for a given σ the neighbourhood entropy is used as a metric and the Newton-Raphson method is used to choose a σ that minimises the neighbourhood entropy. As can be seen in equation (14) a σ unrestrictedly approaching 0+ will lead to generally decreasing probabilities in all classes which in turn lead to a decrease in neighbourhood entropy. Therefore, a penalty value of τ is added when the sum of probabilities over all classes for a pixel falls under a threshold of ρ.

$$\tau = \begin{cases} 1 & \text{if } \sum_{k=1}^{K} p_k(z(t)) \leq \rho \\ 0 & \text{else} \end{cases}$$

By this, values for σ that give low probabilities for all classes are punished and those that give low probabilities for one class and high for another are favoured. This process leads to consistently good choices of σ that provide a good discrimination between the different tissue types.

3.4 Implementation Issues & Modifications

3.4.1 Initialisation

Before the first iteration the algorithm needs to be initialised, so that the neighbourhood probabilities can be estimated. Here, the results obtained by the EM Algorithm are used in the initialisation. In the first step of the initialisation the intensity values are divided into three tissue classes according to table 2.

tissue type	min	max
glandular	$min(I)$	$\mu_{gland} + (.85 \cdot \sigma_{gland})$
friboconnective	$\mu_{gland} + (.85 \cdot \sigma_{gland})$	$\mu_{fat} - \sigma_{fat}$
fat	$\mu_{fat} - \sigma_{fat}$	$max(I)$

Table 2: Initialisation of three tissue classes.

This way, the three tissue types of the former glandular tissue approach and the three tissue types for fatty tissue are each compacted into one single tissue class. What was formerly the transitional tissue class now roughly corresponds to the fibro-connective tissue class. The classification algorithm is quite sensitive to fat tissue that is falsely classified as fibro-connective tissue. Therefore a second step is introduced to evaluate the initialisation in which the intensity ranges for the tissue types are modified to ensure that each slice contains at least 70% fat tissue and no more than 10% amounts of fibro-connective tissue. This way the initial misclassification consists mainly of fibro-connective tissue that is falsely classified as fat tissue, but usually is reclassified to fibro-connective tissue after a few iterations. The reasons for the preferential use of fatty tissue in the initial classification will be explained in section 3.6. The discussion concerns the implications of the characteristics of fibro-connective tissue.

3.4.2 Tissue Probability Model

Due to the substantial influence of the partial volume effect in the breast border region, especially on the saggital slices, a very simple tissue probability model was introduced. The observed drop in intensity values in the

border region is visualised in figure 7. In the first image 7a the segmented breast volume is marked in blue and the result of applying a manually determined threshold for fat tissue on the original image is overlayed in red; this indicates the drop of intensity values in the border region. In figure 7b the segmented breast volume of the preceding slice is overlayed in yellow, showing that the drop of intensities coincides with the lack of neighboring breast tissue; thereby indicating the partial volume effect as the reason for the intensity drop in the breast border region.

(a)   (b)
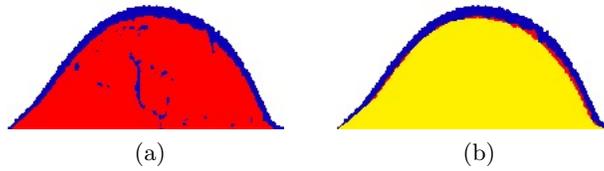
Figure 7: Threshold segmentation of two succeeding slices

To counter these effects a simple tissue probability model is introduced that calculates the probability $p_{tissue}(t)$ of a voxel t in dependence of the distance from the centre of the breast $d(t)$. For glandular- and fibro-connective tissues the probability of a voxel is given by:

$$p_{tissue}(t) = 1 - \left(\frac{\tan\left(\frac{d(t)}{r(t)}\right)}{c} \right)^{b} \tag{15}$$

where $r(t)$ is the distance from the centre to the border of the breast model in direction of t; b, c are a constants to control the probability values at the very border. This model is motivated by the anatomy of a breast where glandular tissue becomes rare in the breast border region. The probability of fat tissue is assumed to be independent of the voxels position. In this model the values for b and c are manually adjusted, a more complex model that calculates the distance from the breast centre, where the intensity drop is expected, based on the contours of the adjacent slices, can be used to automate this process, in addition to being sexpected to give better results as the effect varies over the different slices.

3.5   Discussion of Intermediate Results

The segmentation algorithm presented above is presented by Tasdizen et al. [9] with respect to tissue segmentation in brain MR images. The charac-

ter of how the tissue is clustered within the brain MR images is different in those images compared to breast MR images. Whereas in brain MR images the tissue is mostly clustered, the fibro-connective tissue in breast MR images runs through the breast in thread-like shapes. This significantly reduces the amount of *same-tissue-type* samples in the near proximity of fibro-connective tissue. Due to the fine structures of fibro-connective tissue a significantly higher percentage of this tissue type is affected by the partial volume effect which in turn leads to a less compact distribution of neighbourhoods compared to other tissue types. Another aspect is the possibility
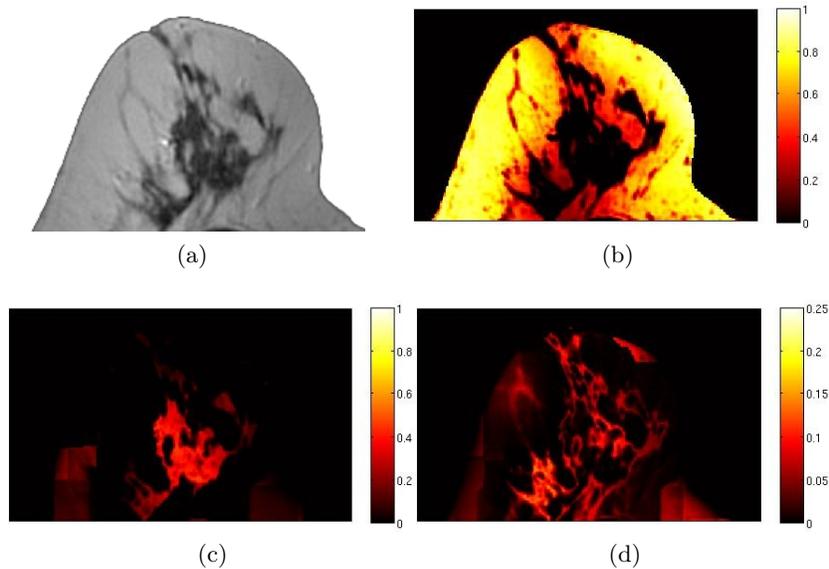

Figure 8: (a) Original MRI, Probability Matrix for (b) fatty tissue, (c) glandular tissue and (d) fibro-connective tissue.

that a fibro-connective tissue thread is initialised as fat tissue with no other fibro-connective tissue nearby. Therefore, it is not possible to compile a local tissue sample collection $A_{fib}(t)$ to estimate a probability density function that has the necessary sample size. When looking at the resulting probability matrices for the different tissue types shown in Fig. 8, it becomes quite evident that there is a high discrepancy of the certainty by which the different tissue types are estimated. Note that the scale of the colour map is adapted for fibro-connective tissue to improve visibility.

3.6 Implications of scattered nature of fibro-connective tissue

The special character of fibro-connective tissue influences the results of the neighbourhood statistics based approach. Its charateristic motivates the modifications to the standard implementation that are proposed later on. As already mentioned the fibro-connective tissue is by definition, threadlike and very thin. This leads to an extremely high percentage of its voxels being border voxels. This means that in at least one direction they are neighboring a different tissue type. To understand the magnitude of this effect, such a fibro-connective structure including some surrounding fat tissue is extracted from an image; as shown in Fig. 9, this region is then binarised into fat and fibro-connective tissue. On the binarized image the total number of voxels in each tissue type are set in relation to the ones that are border voxels.


(a) (b) (c)
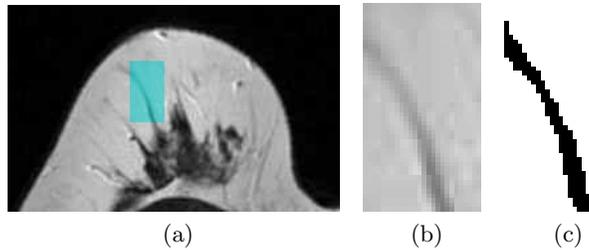
Figure 9: Exemplary calculation of border tissue ratio, for fat and fibro-connective tissue. 9a) showing original image , 9b) shows a magnification of the blue region and 9c) shows the blue region segmented for fat tissue.

$$r_{border} = \frac{\#\text{border voxels}}{\#\text{all voxels}}$$

For this region the percentage of border voxels r_{border} in fatty tissue is 11.48% and for fibro-connective tissue 76.69%, which supports the assumption that neighbourhoods of fibro-connective tissue are heavily influenced by the partial volume effect. This induces high variability between different neighbourhoods, that is observed when examining the samples that are used to to estimate the probability density functions. Fig. 10 shows the distribution of such samples when mapped in the one dimensional space, as it is done in the KDE.

It can be seen that the neighbourhoods of fatty tissue neatly cluster around a single maximum, whereas the neighbourhoods of the fibro-connective
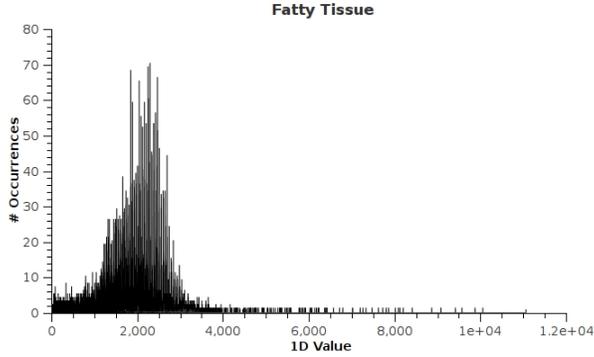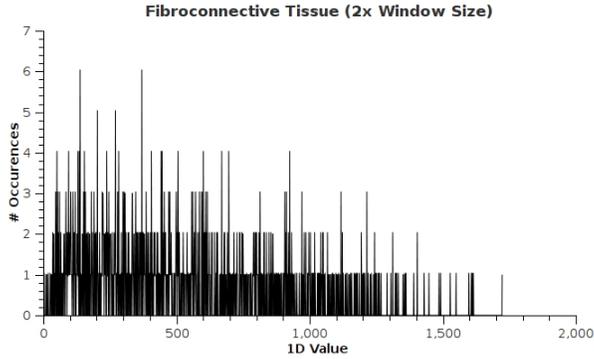
Figure 10: Distribution of neighbourhood Samples in 1D Space ($\parallel z(t) - z_j(t) \parallel$) for (a) fatty tissue and (b) fibro-connective tissue.

tissue are scattered to several maxima. Since a PDF integrates to one, a PDF with a single maximum and low standard deviation has a higher peak than PDFs with several nearly equally strong maxima. This provides an explanation for the low probability values obtained for fibro-connective tissue. Another problem that arises from high variability of neighbourhoods is that the KDE for this tissue is easily affected by misclassified tissue. In a neighbourhood collection A_k that induces a very flat probability function due to scattered and a small amount of samples (such as the fibro-connective tissue neighbourhoods) misclassifications of samples, that are clustered (such as fatty tissue neighbourhoods), can easily dominate the probability function. This is consistent with observations that the PDFs p_{fib} estimated from a neighbourhood collection of fibro-connective tissue converge to one that de-

scribes fatty tissue p_{fat}; as soon as a certain amount of misclassified fat tissue is introduced into fibro-connective neighbourhood collections. This causes the classifier for fibro-connective tissue to *learn* the characteristics of fatty tissue. As a result the standard implementation of neighbourhood statistic based segmentation becomes an unreliable approach for breast MR images segmentation. The modifications that are proposed in the following, are able to provide better stability in the process of segmentation.

3.7 Modifications

3.7.1 Normalisation of tissue probabilities

The generally lower probabilities of fibro-connective tissue make it difficult to compare the probabilities in the sense of *belonging to a certain tissue type*. The probability matrices make it clear that even though the KDE supplies low probabilities for fibro-connective tissue throughout the slice, the KDE is quite confident that some regions are not fat tissue. Here, the idea of tissue probability normalisation is proposed. This is to combine the complete knowledge gained from all three estimates and accordingly update the probability values. The updated probability $\hat{p}_k(t)$ values are calculated


(a)

(b)

(c)
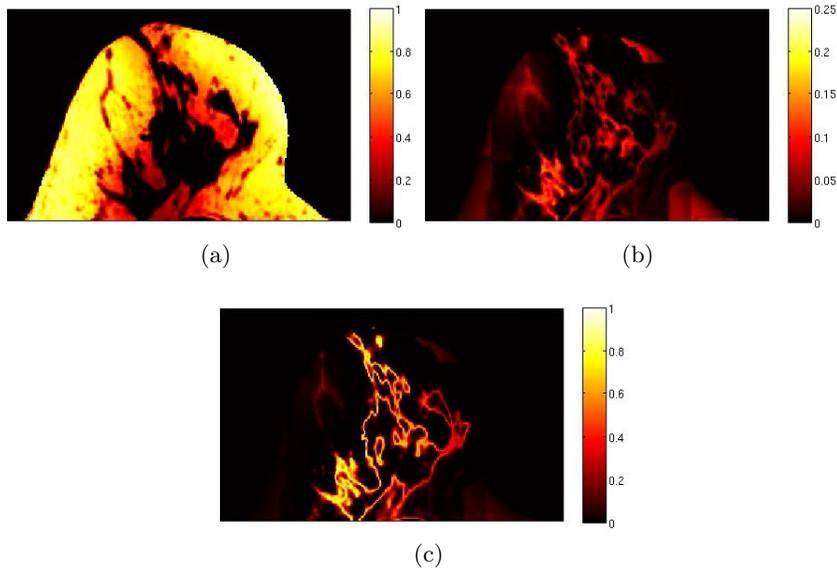
Figure 11: Shows the probability values of 11a) fatty tissue, 11b) fibro-connective tissue and 11c) fibro-connective tissue after normaliaation.

24

as the ratio of class probability to total probability given by:

$$\hat{p}_k(t) = \frac{p_k(t)}{\sum_{c=1}^{3} p_c(t)} \qquad (16)$$

As can be seen in Fig. 11c this increases the probability for fibro-connective tissue in regions where the other two tissue types have low probabilities. Also it decreases the probabilities in regions where one other tissue type has high certainty, for example in the bottom right of Fig. 11b. Here, artificially high values are introduced due to the lack of sufficient neighbourhood samples. Since this update rule does not affect the order of values, it has no effect on the classification when applied in the standard implementation of neighbourhood statistics based segmentation. However, in the following two proposed modifications of the standard implementation it generally improves the results.

3.7.2 Weighed neighbourhood Collections

During standard implementation the neighbourhood samples for a given tissue type k are compiled by collecting those neighbourhoods of which the center pixel has been classified as k in the previous iteration. This collection serves as a reference when estimating the probability of the current voxel being of the same tissue type. In this process all samples of the neighbourhood are considered to represent the tissue class equally well.

This method can be extended to incorporate the certainty by which each pixel has been classified, thus emphasising the influence of neighbourhoods that have a high certainty in the process of evaluation and decreasing the influence of those that have low probability. This is done by multiplying each addend with the previous probability p_k^* of its sample. The adapted equation to calculate the probability estimate is given in equation 17. In the first iteration the probabilities are unknown and are therefore achieved using the standard method.

$$p_k(z(t)) = \frac{1}{P} \sum_{t_j \in A_k(t)} p_k^*(t_j) e^{-\frac{\|z(t) - z(t_j)\|}{2\sigma^2}} \qquad (17)$$

with

$$P = \sum_{t_j \in A_k(t)} p_k^*(t_j)$$

25

3.7.3 Fuzzification of Neighbourhood Collections

In this variation the set of pixels is extended for each tissue type to a fuzzy set $(\mathbb{X}, d_{tissue})$. Here $\mathbb{X}$ is a set that corresponds to the indices of the MR images' pixels, and d_{tissue} is called the membership function that provides the grade of membership for each element $x \in \mathbb{X}$. Instead of classifying each pixel to exactly one tissue type it is now possible to assign a pixel to more than one tissue type with different degrees of membership. The role of the membership function is taken by the KDE for tissue probability; realizing the fuzzy set for a tissue class k as $(\mathbb{X}, p_k(x))$. The normalisation presented in section 3.7.1 ensures that the total degree of membership for each pixel summed up over all classes equals one. When estimating the new tissue probability all neighbourhoods in the local surrounding are taken into account weighted based on the tissue class in question and degree of membership.

3.8 Mapping of Dielectric Properties

3.8.1 Segmentation based mapping

After completion of the segmentation process, the glandular and fibro-connective tissue classes are united; as they have similar dielectric properties. One mapping table for each tissue type is then compiled. This is accomplished by first determining the highest and lowest intensity value that is contained in each tissue class. The dielectric properties for a given frequency are then calculated using the Cole-Cole parameters listed in table 8. The ranges are $maximum - glandular\text{-}low$ and $fat\text{-}high - fat\text{-}low$ for glandular/fibro-connective and fatty tissue respectively. Following on, the dielectric property ranges are linearly mapped between the according intensity ranges of the tissue types. Fig. 12 presents an example of such a mapping.

3.8.2 Probability based mapping

In this approach the neighbourhood probabilities are incorporated into the mapping procedure. The dielectric properties are again linearly interpolated between minimal and maximal glandular tissue intensity as well as minimal and maximal fatty tissue intensities. Instead of assigning the according dielectric property depending on the final classification; the assigned dielectric property is calculated as a mix of both dielectric values according to their probability. This is inspired by the probability maps for the different tissue types, shown in Fig. 8. The maps have areas of very high certainty
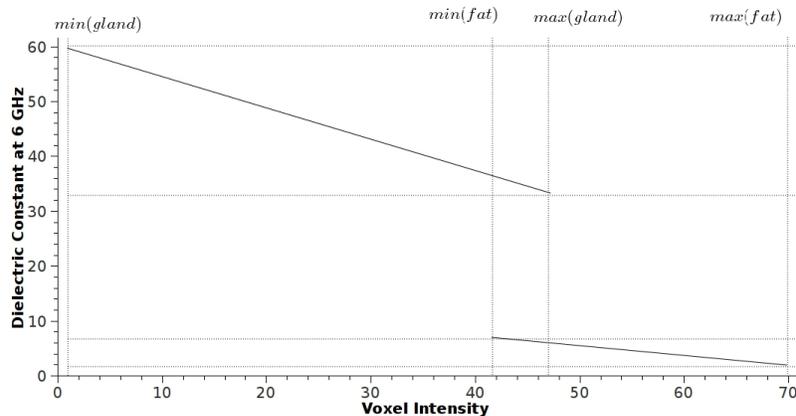
Figure 12: Illustration of Two Tissue Type Mapping

for one specific tissue type; as well as, some areas where connected structures can be recognised in the maps, however, only part of the structure can compete with the high probabilities of other tissue types. This is often the case for fibro-connective tissue competing with fatty tissue. For these cases the probability based mapping is introduced to create a qualitatively good representation of these recognised structures, with respect to the dielectric properties. For a voxel index $t \in T$ the dielectric property value $d(t)$ is given by

$$d(t) = ((p_{gland}(t) + p_{fib}(t)) \cdot map_{gland}(x(t))) + (p_{fat}(t) \cdot map_{fat}(x(t)))$$

4 Results

In order to evaluate and compare both algorithms, they are applied to the same set of three MR scans. These scans have been provided by the Sahlgrenska University Hospital MR centre in Göteborg. An overview of the data sets used for evaluation is given in table 3.

Set	voxel size	contrast agent	scan direction
Set I	.625 × .625 × 3 mm	none	sagittal
Set II	.625 × .625 × 3 mm	15ml	sagittal
Set III	.625 × .625 × 3 mm	15ml	sagittal

Table 3: List of evaluated MRI scans

4.1 Comparison

Due to the lack of ground truth it is not possible to directly judge the performance of the two different approaches. Therefore, other characteristics are examined to get an indication about the method's performance. The comparison is divided into two parts; a statistical comparison that evaluates the consistency of the results within each method and then compares them with each other. In the second part, the performance will be evaluated visually on sample MRI slices. The comparisons here are based on three MRI scans that were performed on the same patient and breast.

4.1.1 Statistical comparison

To compare the different results the statistical measures *misclassification rate (MCR)* and *exact agreement* are chosen. MCR is based on the *sensitivity* of a tissue type; weighted on the probability of a tissue type to occur. Sensitivity describes the percentage of tissue that is correctly recognised given a classification reference. It is calculated as a ratio of *true positives* and *false negatives*.

$$\text{sensitivity} = \frac{\#\text{true positives}}{\#\text{true positives} + \#\text{false negatives}} \quad (18)$$

The *true positives* describe voxels that are classified as a certain tissue type when they indeed are, and *false negatives* describe the voxels that are classified as not being a certain tissue when they actually are.

$$MCR_{t_1} = p_{tissue_1} \cdot (1 - \text{sensitivity}_1) + p_{tissue_O} \cdot (1 - \text{sensitivity}_O) \quad (19)$$

For the MCR *tissue*$_1$ represents the tissue under investigation and *tissue*$_O$ all other tissue.

The exact agreement simply gives the ratio of voxels that are classified the same to the total number of voxels. Here, all voxels of the breast volume are taken into account, and two voxels agree if they are assigned to the same tissue class. These two measures are used to compare the different results. Due to a slight displacement of the patient's breast between the scans, there is not a perfect one-to-one relation between the slices of the different sets. In order to minimise these effects six slices of each set have been selected. For each slice the corresponding slices in the other two sets have been manually selected based on high correspondence. The comparisons are calculated using only these slices. For the first comparison the images are segmented into two tissue types, glandular and fatty tissue. For the original method,

the images were segmented into three main-tissue types. For comparison purposes, we have to merge the transitional tissue region with the other two regions. This is done using a threshold determined by the intersection of the two Gaussians's given by the GMM. The transitional tissue region can then be split up in a way that minimises the classification error. For the proposed method fibro-connective and glandular tissues are united. The comparison is conducted with the division into these two tissue types, instead of into the seven used in the original method. This is motivated by the substantial difference in the dielectric properties between these two tissue types. The results are listed in the tables 4 and 5.

	I-II	I-III	II-III
glandular & fibro	9.8 (11.4)	10.2 (12.5)	3.3 (4)
fatty	13.2 (15.3)	14.3 (17.4)	5.3 (6.5)
exact agreement	88%	87 %	95%

Table 4: Comparison of segmented MRI scans, using original method.

	I-II	I-III	II-III
glandular & fibro	19.7 (14.5)	19.4 (15.7)	6.5 (8.1)
fatty	6.4 (4.5)	6.1 (4.8)	2.1 (2.6)
exact agreement	92%	92 %	96%

Table 5: Comparison of segmented MRI scans, using proposed method.

The first listed model in the comparison is taken as the reference, the values in parenthesis show the MCR with the second model as the reference model. The results show that all three sets have an acceptable *exact agreement* in both methods. It also shows that there is a higher agreement between the sets with contrast agent (II + III) than between a set with and without contrast agent. The contrast agent leads to a higher contrast between glandular and fatty tissues, which helps make the parameter estimation with the EM algorithm more confident. Therefore it accounts for a higher agreement between the sets with contrast agent. Another observation is that the proposed method seems to be less affected by the lack of contrast agent than the original method.

For a second comparison a three tissue segmentation is chosen isolating the *transitional tissue* and combining all glandular and fatty tissue types in one each for the original method and using the segmentation results

of the proposed method directly, thereby separating into glandular, fibro-connective and fatty tissue. On this data the same measures are applied. The results of this comparison are shown in tables 6, 7.

	I-II	I-III	II-III
glandular	45.8 (19.7)	43.6 (15.6)	10.7 (22.9)
transitional	70.1 (78.1)	55.8 (70.4)	66.2 (69.2)
fatty	19.4 (17.2)	13.2 (13.8)	8.6 (11.9)
exact agreement	68%	76 %	81%

Table 6: Comparison of segmented MRI scans, using original method.

	I-II	I-III	II-III
glandular	29.2 (21.9)	28.2 (26.3)	7.3 (24)
fibro	55.5 (51.7)	56 (51.2)	25.7 (13.8)
fatty	6.4 (4.5)	6.1 (4.8)	2.1 (2.6)
exact agreement	86%	86 %	94%

Table 7: Comparison of segmented MRI scans, using proposed method.

The introduction of the *transitional tissue type* into the evaluation leads to a significant drop in total agreement as well as a substantial increase of the MCR, indicating the transitional and fibro-connective tissue types, as expected, as the main source of misclassifications. For the proposed method the MCR is generally lower but still shows very high disagreement between the different data sets concerning the fibro-connective tissue. On the other hand the part of of fibro-connective tissue of the overall tissue is very low which explains that the *exact agreement* is barely affected for the data sets with contrast agent even though the fibro-connective tissue has an MCR of 25 (13) %. These results indicate a higher intra-algorithm agreement in the proposed method over the original one, but also indicate that both methods fail to have a confident classification of fibro-connective/transitional tissue, as some examples in the next section are showing.

4.1.2 Visual comparison

In the visual inspection, the comparison of the results from the original method and the segmentation based method are extended by a comparison with the probability based mapping of the dielectric properties. The discussion is based on general observations of the results which are demonstrated

here on some exemplary slices. When looking at the results of the original method one can see that fatty and glandular tissue is very accurately handled but that on this data set the fibro-connective tissue is underestimated (see Fig. 13, box A). Another observation that is made overall data sets
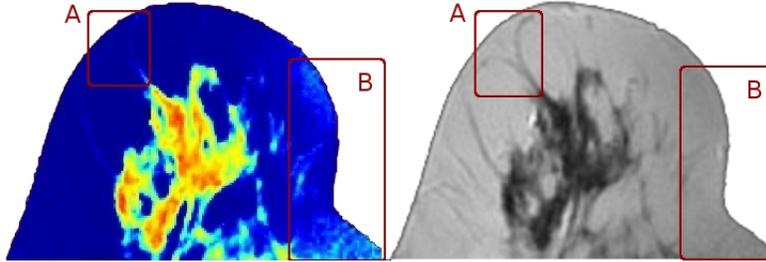

Figure 13: Comparing Original Method and Raw Image

are the influences of the inhomogeneities in the original image. Those are introduced into the dielectric property model. Here it causes some fatty tissue to be classified as transitional tissue (see Fig. 13,box B).

When looking at the results of the proposed method with segmentation based mapping one can see that the influences of light inhomogeneities are reduced (Fig. 14, box C). Also this approach is able to restore some fibro-connective structures, that are underestimated by the original approach, and therefore improve the dielectric model (see Fig. 14, ellipse B). Due to the problems of segmenting fibro-connective tissue, these structures are sometimes missed or only recognised partially (Fig. 14, ellipse A). A major drawback of the segmentation based mapping is that whenever this happens that structure is missed completely, and in the cases where these structures are partially correctly segmented unnatural roughness is introduced into the dielectric model. So while it is able to improve the accuracy in quantita-
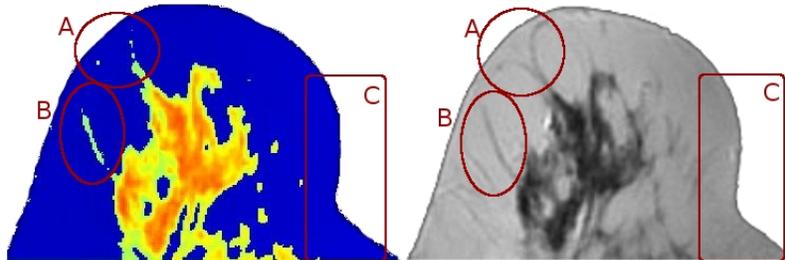

Figure 14: Comparing Segmetation Based Mapping and Raw Image

tively estimating the dielectric properties of some degraded structures, the introduction of such a roughness can destroy the qualitative order of the dielectric properties within a structure. This can be seen e.g. in Fig. 14, ellipse A.

The third approach, the proposed segmentation using probability based mapping, tries to compensate the artificially introduced roughness. The results of such a mapping can be seen in Fig. 15. This approach shows very promising results, as can be seen in Fig. 15. Circle A and box B model even fine structures of fibro-connective tissue quite well. Especially in *box B* it show structures that where too fine to be recognised by the two other approaches. Comparing the region in *box C* with the one from the original approach shows that the probability based mapping also is less affected by inhomogeneities. Only in some regions like the one marked by the *red circle D* the dielectric properties of fatty tissue is overestimated, which is again related to the problems of accurately estimating the fibro-connective tissue probabilities.
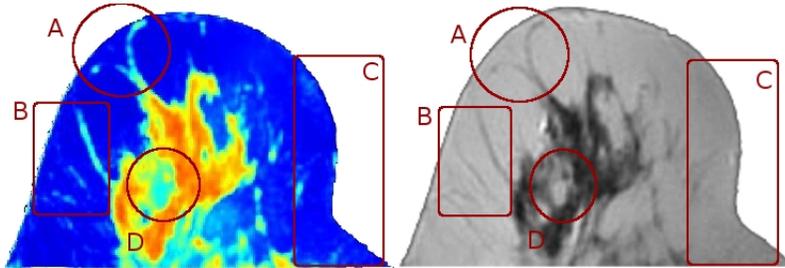

Figure 15: Comparing Probability Based Mapping and Raw Image

5 Discussion

5.1 Original Method

The original method is easy to implement and very robust as long as the histogram allows a good automatic separation of glandular and fatty tissue. But the assumption about the Gaussian nature of the tissue intensity distribution does not reflect the true distribution accurately. The content of glandular tissue in a slice plays a vital role in the accuracy of estimating the mean values of fatty and glandular tissue. The problem of estimating the parameters for the GMM in breast with a very high fatty tissue content is also reported by Zastrow et al. [14] in the paper presenting their work.

Comparing the results of the parameter estimation for the three MRI scans of the same breast shows variations in the parameters that had a substantial influence on the resulting model. This indicates that the quality of parameter estimation does not only depend on the fat content of the breast but also depends on other factors such as the use of contrast agent. These variations support the findings in the statistical comparison that show high variations between the different scans in the transitional tissue region. Another aspect is that this model does not account for the degrading influences of the partial volume effect which has a significant influence on fibro-connective tissue. Dielectric properties that are solely based on the intensity value lead, in the three data sets used for evaluation, either to an underestimation of fibro-connective tissue or an overemphasising of noise, with respect to the dielectric constant. The classification of compacted glandular and fatty tissue appears to be very accurate, but how the intensity values of fibro-connective tissue are classified depends on the characteristics of the histogram, and therefore on factors like fat tissue to glandular tissue ratio as this influences the results of the EM algorithm trying to fit the histogram with a GMM. Due to the piece-wise linear map from tissue intensity to dielectric property the quality of the numerical breast phantom also directly relates to the performance of non-uniformity removal and noise reduction. In this work not all nonuniformities could be removed, so when interpreting the results presented here about this method, it has to be taken into account that the method is intended to built a dielectric model from a perfectly uniform MR image and therefore does not account for any inhomogeneities within the mapping process.

5.2 Proposed method

In the statistical comparison the neighbourhood statistic based segmentation shows a higher intra-algorithm consistency between the three datasets than the original method. This suggests that this method is more reliable and stable process in building dielectric models from MR images. Even though part of this higher agreement relates to a more conservative handling of fibro-connective tissue, and therefore partly is achieved by a tendency to classify tissue as fatty tissue. It can be seen that the models of the same breast using different methods have a higher agreement using this method than using the original method. Comparing the probability estimation for the different tissue types with the original image gives optimistic results about segmenting the tissue types based on neighbourhood statistics. In the final results some fibro-connective structures are missed by the algorithm while

others are recognised quite good.

The main problem of this method is that when a fibro-connective structure is only partly recognised, the segmentation based mapping introduces high artificial roughness into the dielectric model because very similiar tissue is mapped to very different dielectric property values. Therefore the *solely segmentation based mapping* is not proposed to be used in building a numerical model, as long as the segmentation results are not improved. Another main problem in the classification process is the high degradation and small sample size of fibro-connective tissue neighbourhoods, which leads to this tissue type being *overshaddowed* by fatty tissue. The characteristic of fibro-connective tissue and its implication, as discussed in section 3.6, make it somewhat unfeasible to use the standard implementation of the neighbourhood statistics segmentation algorithm in the context of breast tissue. The proposed modifications, especially the *weighed neighbourhood collections* are able to stabilize the segmentation process, but they are still not able to completely overcome underestimation of fibro-connective tissue. In order to improve the segmentation results, it is suggested to look into the statistics of image neighbourhoods for the different tissues in more detail. For example we can determine if a higher dimensionality gives better discrimination between the different tissue types, as the fibro-connective neighbourhoods are not well clustered when mapped to a single dimension. Therefore it is suggested to combine the proposed method with the probability based mapping.

The probability based mapping that takes into account the degree of certainty to which a voxel is classified as a certain tissue is able to restore the smoothness that gets lost when mapping based solely on segmentation results. It is also able to recover fibro-connective structures qualitatively and presents a slight improvement over the original method, especially if the MR images are not completely free of inhomogeneities. Since it bases on the segmentation it gives more consistent results for the evaluated data sets, therefore the probability based mapping is suggested as a slight improvement, over the original method, that gives more stable results.

6 Appendix

	ϵ_∞	Δ_ϵ	τ (ps)	α	σ_s (S/m)
maximum	1.000	66.31	7.585	0.063	1.370
glandular-high	6.151	48.26	10.26	0.049	0.809
glandular-median	7.821	41.48	10.66	0.047	0.713
glandular-low	9.941	26.60	10.90	0.003	0.462
fat-high	4.031	3.654	14.12	0.055	0.083
fat-median	3.140	1.708	14.65	0.061	0.036
fat-low	2.908	1.2	16.88	0.069	0.02
minimum	2.293	0.141	16.40	0.251	0.002

Table 8: Single-Pole Cole-Cole Parameters for Eight Wideband Dielectric Properties Curves, [6]
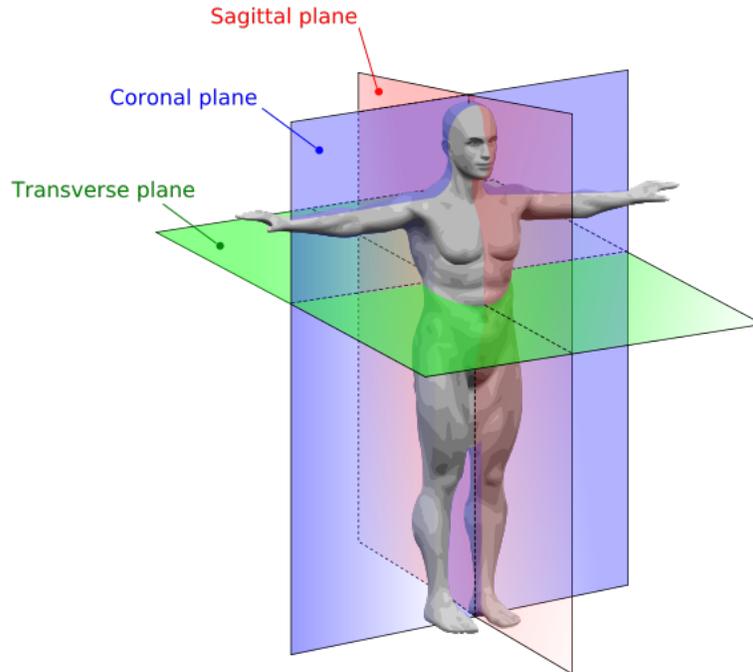

Figure 16: Planes of human anatomy, Source Wikipedia [12]

References

[1] Suyash P. Awate and Ross T. Whitaker. *Higher-Order Image Statistics for Unsupervised, Information-Theoretic, Adaptive, Image Filtering.*

[2] Jeff A. Bilmes. A gentle tutorial of the em algorithm and its application to parameter estimation for gaussian mixture and hidden markov models, 1998.

[3] 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)*, 39(1), 1977.

[4] Andreas Fhager and Mikael Persson. Using a priori data to improve the reconstruction of small objects in microwave tomography. *IEEE TRANSACTIONS ON MICROWAVE THEORY AND TECHNIQUES*, 55-11, 2007.

[5] GLOBOCAN. Cancer fact sheet, 2008. `http://globocan.iarc.fr`.

[6] Mariya Lazebnik, Leah McCartney, Dijana Popovic, Cynthia B. Watkins, Mary J. Lindstrom, Josephine Harter, Sarah Sewall, Anthony Magliocco, John H. Booske, Michal Okoniewski, and Susan C. Hagness. A large-scale study of the ultrawideband microwave dielectric properties of normal breast tissue obtained from reduction surgeries. *PHYSICS IN MEDICINE AND BIOLOGY*, 52, 2007.

[7] Ann B. Lee, Kim S. Pedersen, and David Mumford. The nonlinear statistics of high-contrast patches in natural images. *International Journal of Computer Vision*, 54, 2003.

[8] Stan Z. Li. Markov random field modeling in image analysis, 2009. p. 21-22, ISBN 978-1-8480-0279-1.

[9] Tolga Tasdizen, Suyash P. Awate, Ross T. Whitaker, and Norman L. Foster. *MRI Tissue Classification with Neighborhood Statistics: A Nonparametric, Entropy-Minimizing Approach.*

[10] Carlo Tomasi. Estimating gaussian mixture densities with em - a tutorial. `http://www.cs.duke.edu/courses/spring04/cps196.1/handouts/EM/tomasiEM.pdf`.

[11] Hana Dosicek Trefna, Jan Vrba, and Mikael Persson. Evaluation of a patch antenna applicator for time reversal hyperthermia. *International Journal of Hyperthermia*, 26-2:185–197, 2010.

[12] Wikipedia. `http://commons.wikimedia.org/w/index.php?title=File:Human_anatomy_planes.svg&oldid=43481201`.

[13] Wikipedia. Entropy (information theory) — wikipedia, the free encyclopedia, 2010. `http://en.wikipedia.org/w/index.php?title=Entropy_(information_theory)&oldid=378885832`.

[14] Earl Zastrow, Shakti K. Davis, Mariya Lazebnik, Frederick Kelcz, Barry D. Van Veen, and Susan C. Hagness. Development of anatomically realistic numerical breast phantoms with accurate dielectric properties for modeling microwave interactions with the human breast. *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, 55(12), 2008.