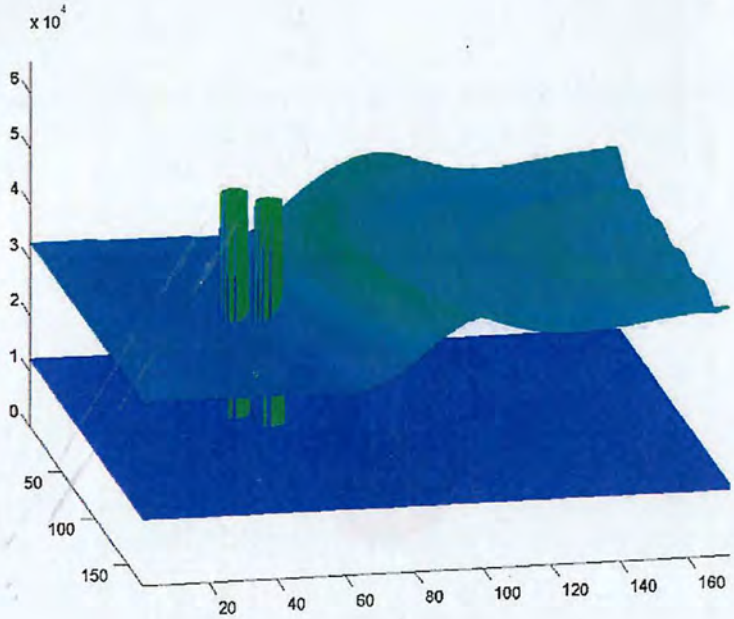
CHALMERS



Strategies for improving the image reconstruction in Microwave Tomography

Master of Science Thesis in Engineering Physics

DANIEL BERNHARDSSON

Biomedical Electromagnetics Group
Department of Signals and Systems
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden, 2009
Report No. EX024/2009

REPORT NO. EX024/2009

Strategies for improving the image reconstruction in Microwave Tomography

Diploma thesis by
Daniel Bernhardsson

Engineering Physics
Thesis project accomplished at
Department of Signals and Systems
Chalmers University of Technology
Göteborg, Sweden

August 2009

Supervisor and examiner: Assistant Prof. Andreas Fhager
Biomedical Electromagnetics Group
Biomedical Engineering
Department of Signals and Systems

Strategies for improving the image reconstruction in Microwave Tomography
DANIEL BERNHARDSSON

Technical report no EX024/2009
Department of Signals and Systems
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone + 46 (0)31-772 1000

Cover:
An illustration of an electromagnetic wave propagating in a dielectric profile that contains two scattering tumors.

Department of Signals and Systems
Göteborg, Sweden, 2009

Abstract

Microwave tomography is a technique for imaging dielectric properties. It has gained increasing attention during last decades, in particular by its potential application to detect breast cancer. The purpose of this thesis is to investigate different optimization methods in the image reconstruction procedure. The conjugate gradient (CJG), Gauss-Newton (GN), Quasi-Newton (QN) and One step secant (OSS) method have been tested and compared. The result is that the conjugate gradient method (CJG) is the best choice among the investigated methods. Furthermore a method is proposed for making use of a priori data based on knowledge about the dielectric properties of the investigated object. In numerical tests this method has proven to enhance the convergence rate. A method for choosing the step length in each iteration is also proposed that requires approximately 55% less calculation time for a typical computation scenario compared to a successive parabolic interpolation. Moreover an interrogating pulse with Gaussian frequency distribution is compared to a pulse with more uniformly distributed energy. Using the latter in a numerical simulation gives a better image reconstruction in the first few iterations. An analysis of experimental data is also made as an attempt of detecting a real tumor. It turns out to be impossible and possible sources of error are identified.

Keywords: microwave tomography, image reconstruction, optimization methods, conjugate gradient, Gauss-Newton, One step secant, Quasi-Newton, inverse problem, mammography, breast cancer detection

Strategies for improving the image reconstruction in Microwave Tomography
DANIEL BERNHARDSSON
Department of Signals and Systems
Chalmers University of Technology

Contents

Chapter 1

Introduction

1.1 Background

According to WHO breast cancer causes 519,000 deaths worldwide [1] which makes it one of the most common causes of death among women today. The number of cases worldwide has also increased significantly since the 1970s. Among the different types of cancer, breast cancer is one of the more curable if it is caught early. Imaging by X-ray mammography is the most available method in clinics nowadays to detect breast cancer. There are however some negative aspects, such as missing up to 15% of breast cancers [2]. There is also a small risk that the radiation exposure may induce cancer. Breast compression as a part of the examination is in some cases an uncomfortable experience for the patient. Moreover there could be some problems examining women with radiographically dense breasts. There is a need for an alternative method for screening and according to U.S. Institute of Medicine (IOM) report [2], an ideal breast screening tool

- has low health risk
- is sensitive to tumors and specific to malignancies
- is noninvasive and simple to perform
- involves minimal discomfort
- detects breast cancer at a curable stage
- is cost effective and widely available
- provides easy to interpret, objective, and consistent results

Although there already exist complementary imaging methods that avoid ionizing radiation such as ultrasound and magnetic resonance imaging (MRI) none of these are suitable for a mass screening programme. An alternative technique for screening that is believed to have the potential of fulfilling the above criteria is *Microwave Tomography*. A microwave systems avoids ionizing radiation and breast compression. It is also believed to be both sensitive and specific to small tumors. It would not be as expensive as magnetic resonance imaging and the examination is expected to be very fast.

1.2 Microwave Tomography

For a long time, microwave engineers have dreamed of imaging the human body with non-ionizing electromagnetic waves in order to detect cancers. The reason that microwaves look so promising for the detection of breast tumors is that the tumors have different dielectric properties, *i.e.* permittivity and conductivity, compared to healthy tissue at microwave frequencies. This fact has actually been known since 1926 when Fricke and Morse reported that the permittivity of cancerous breast tissue is higher than that of healthy breast tissue at 20 kHz [3]. Since then there have been several studies [4] [5] [6] [7] confirming the deviant properties of tumors. Many of them were studies of excised samples but more recently there have also been made in vivo measurements. A large scale study of different types of breast tissue in the microwave spectrum from 0.5 GHz to 20 GHz was made by M. Lazebnik *et al* [8]. The contrast in microwave-frequency dielectric properties shows to be 10:1 between malignant tissue and normal adipose-dominated tissues. For glandular/fibroconnective tissues however this contrast is no more than about 10%.

Doctors today meet a demanding task interpreting low contrast X-ray images. Apart from the observed high contrast in dielectric properties that is suitable for diagnosis, normal breast tissue is also more translucent to microwaves than many other tissues, such as muscle or brain. In addition, the breast is also more suitable for examination than internal organs that are less accessible. With all these promising benefits of using microwaves instead of x-rays one could ask why a microwave imaging system was not developed many years ago with more than eighty years since the first research results. And the answer lies in the fundamental difference between microwave and x-ray propagation as Fig. 1.1 shows.

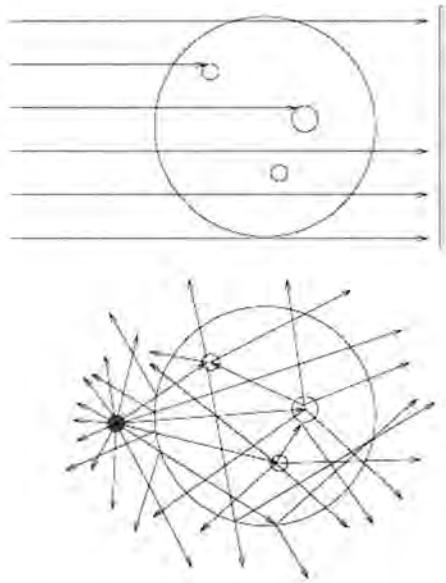


Figure 1.1: A comparison between x-ray and microwave propagation. The X-ray photons in the upper figure give an image that is a projection of the absorption through the object. The microwaves propagate in a more complex way from an omnidirectional antenna that is depicted by a blue circle in the figure.

The upper figure shows how x-rays propagate. They move straight forward and when they encounter an object like a tumor the intensity decreases due to absorption and Compton scattering. The photons created by Compton scattering are removed by a filter called a collimator. The remaining photons have gone linearly through the object giving a projection on the film and making different objects visible.

The bottom figure shows correspondingly the behavior of microwaves. Their wavelength is on the order of 1 cm compared to about 1 nm for x-rays. As the wavelength of the microwaves is comparable to the size of the investigated objects their propagation is more complex. An omnidirectional antenna, which is the blue circle in Fig. 1.1 transmits radiation in all directions which is then reflected and scattered by the objects it encounters. In this manner radiation can be considered to bounce several times between tumors or other objects before it is being measured. So after some radiation has been transmitted and received at different positions one will need an advanced algorithm for image reconstruction that takes this scattering into

account. The propagation of the microwaves gets even more complex because of the fact that the permittivity and conductivity of biological tissues have a strong dispersive behavior, *i.e.* frequency dependence. The frequency span can be divided into α, β and γ regions where different interactions cause the dielectric dispersion. In the γ region ($>$ 100 MHz) the dielectric properties are largely determined by the water content of the tissue and the presence of ions and small molecules. For a more thorough approach refer to the work by Fhager [9].

A typical microwave imaging system involves two phases: *acquiring data* and *solving the inverse scattering problem.* In the data acquisition phase, a number of antennas are being used to transmit electromagnetic waves in the microwave spectrum. The waves pass through the investigated object which will scatter the waves in different directions. The fields are then being measured at different positions by other antennas. The measured data needs further processing which includes solving the inverse scattering problem which in turn involves using a reconstruction algorithm. The aim is to get a complete image of the spatial distribution of the permittivity ϵ and/or the conductivity σ of the investigated object. In the work by Fhager [9], an image reconstruction algorithm has been implemented and a prototype for microwave tomography has been constructed as shown in Fig. 1.2.

Figure 1.2: The experimental prototype at Chalmers. A network analyzer connected to an antenna array.

Microwave tomography had a breakthrough in 1979 when Jacobi [10] developed a water-immersed antenna system that was able to image the canine kidney [11]. Back then the research was focused on linear reconstruction algorithms based on the Born and Rytov approximations. It was shown that these algorithms were limited to successful reconstruction of small, low contrast objects [12]. As breast tumors have high contrast to healthy tissue the main focus has shifted towards nonlinear iterative reconstruction. Such algorithms are more computationally intensive but possible to work with using modern computers. Microwave Tomography has gained much attention during the last years due to increasing computational power and improved numerical techniques.

1.3 Problem formulation

1.3.1 The direct problem

The direct problem refers to calculating the evolving of an electromagnetic field in a material with known dielectric properties. Examples of numerical methods for solving this problem are Finite Element Method (FEM) and Finite Difference Time Domain (FDTD) applied on Maxwell's equations. The FDTD method is used throughout this thesis. In this thesis a reconstruction algorithm is used that is nonlinear and iterative. It starts with an initial guess of the permittivity and conductivity image which is updated in each iteration. To be able to improve the image in each iteration one simulates how the radiation from an antenna would propagate given the current guess.

Moreover a computational region needs to have some kind of boundary conditions. To simulate a large space without having to set up an equally large computational region one possibility is to use an absorbing boundary such as Perfectly Matched Layers (PML). Waves within a large span of frequencies will gradually slow down and attenuate as they pass through these layers, without any reflections.

1.3.2 The inverse problem

Compared to the direct problem, the situation is now reversed. The transmitted signal is known and we have measured the electromagnetic field at certain positions. But instead it is the dielectric properties of the material at

each point that are unknown. The inverse problem refers to the calculation of these material properties. This is usually an ill-posed problem as a small change in the input data can give rise to a large change in the output data. In practice this means that even small measurement errors and noise can cause huge errors in the output data. One way of dealing with ill-posed problems is to apply a regularization to the problem.

The antennas in the array are used as both transmitters and receivers. Every antenna transmits a pulse once and each one of the other antennas measures a received pulse. Experimentally though, a Vector Network Analyzer (VNA) is being used. The VNA does not sample the received signal in short time intervals but it measures the steady-state response in magnitude and phase at a number of distinct frequencies for each pair of antennas. Each pulse can then be synthesized from these measurements. Without going deeper into the experimental procedure, for any pair of antennas, we claim to know how a transmitted pulse will be picked up at the receiving position. The aim is to create a dielectric image such that the pulses transform in the same manner in simulation as in the real measurement.

The pulse that is used in most of the cases in this thesis is a transient pulse of a sinusoidal carrier of a certain frequency, ω, with a Gaussian amplitude modulation:

$$E_m^{source}(t) = E_0 e^{\left(-\frac{1}{2}\frac{(t-t_1)^2}{t_0^2}\right)} \sin(\omega t) \tag{1.1}$$

Here t_0 is the pulse width and t_1 is set to $4t_0$. The pulse is plotted in Fig. 1.3. This pulse is sent from a transmitter, labeled m. The electric field on several receivers, $\mathbf{R_n}$, is correspondingly acquired and denoted by, $\mathbf{E_m^{meas}}(\mathbf{R_n}, t)$.

The basic idea behind solving of the inverse problem is that the difference between simulated and measured pulses should be as small as possible, see Fig. 1.4. The solution to the inverse problem is defined as exact correspondence between these pulses which will in practice never be obtained. Mathematically we define a cost functional, Eq 1.2, to be minimized.

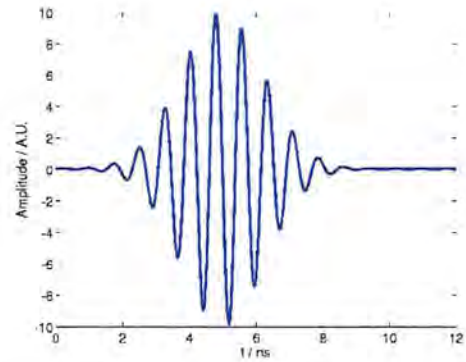


Figure 1.3: A plot of a sinusoidal pulse with a Gaussian amplitude modulation.

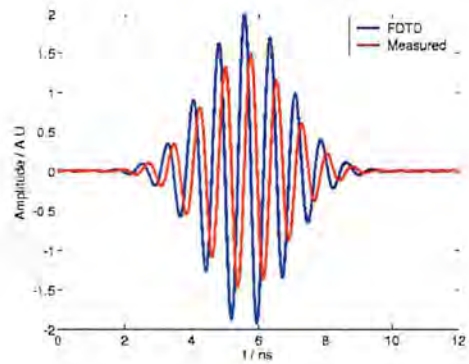


Figure 1.4: A correct reconstructed image should give a simulated pulse that is identical with the measured pulse at each antenna.

$$F(\epsilon(\mathbf{x}), \sigma(\mathbf{x})) = \int_0^T \sum_{m=1}^{M} \sum_{n=1}^{N} \left(|\mathbf{E}_m \left(\epsilon(\mathbf{x}), \sigma(\mathbf{x}), \mathbf{R}_n, t\right) - \mathbf{E}_m^{meas} \left(\mathbf{R}_n, t\right)|^2 \right) dt, \tag{1.2}$$

Where $\mathbf{E}_m \left(\epsilon(\mathbf{x}), \sigma(\mathbf{x}), \mathbf{R}_n, t\right)$ is the calculated field from the computational model of the setup and $\mathbf{E}_m^{meas} \left(\mathbf{R}_n, t\right)$ is the measured field. M is the number of transmitters and N is the number of receivers. In order to minimize this cost functional, gradients need to be computed. These gradients are calculated by considering a small increment in the dielectric profile and

then the corresponding change in the functional can be evaluated using a perturbation analysis. In the expression for the gradients the solution to the adjoint problem of Maxwell's equation is also used. The residuals between the measured and the simulated fields are used as the driving sources in the adjoint problem.

Using the Fréchet derivative of the functional like in [9] the gradients can be written as

$$\begin{aligned} G_{\epsilon}(\mathbf{x}) &= 2\sum_{m=1}^{M}\int_{0}^{T}\tilde{\mathbf{E}}_{\mathbf{m}}(\epsilon,\sigma,\mathbf{x},t)\cdot\partial t\mathbf{E}_{m}(\epsilon,\sigma,\mathbf{x},t)dt \\ G_{\sigma/\langle\sigma\rangle}(\mathbf{x}) &= 2\langle\sigma\rangle\sum_{m=1}^{M}\int_{0}^{T}\tilde{\mathbf{E}}_{\mathbf{m}}(\epsilon,\sigma,\mathbf{x},t)\cdot\mathbf{E}_{m}(\epsilon,\sigma,\mathbf{x},t)dt \end{aligned} \tag{1.3}$$

In the expressions above $\tilde{\mathbf{E}}_{\mathbf{m}}$ is the solution to the adjoint problem. Furthermore $\langle\sigma\rangle$ is a scaling parameter compensating for the different scaling of the permittivity and conductivity in the functional. The performance of the reconstruction algorithm depends on the choice of scaling parameter.

The gradients can be calculated quite efficiently and by using an optimization method such as the conjugate gradient method (CJG) in combination with a line search to find the step length, the reconstruction domain can be updated iteratively in search for the minimum of F.

Originally a conjugate gradient method has been used for the minimization problem but there exist a variety of optimization methods. The high dimensionality of the problem makes it impossible to use methods that search through entire volumes of the search space. Due to the problem of getting trapped in local minima [13] that do not correspond to the correct solution it has been proposed to use a global search approach. In this thesis however, the focus is on local optimization methods.

1.4 Purpose of this thesis

The aim in this report is at first to study different local optimization methods. The following methods are studied in Chapter 2:

- Conjugate Gradient (CJG);
- Quasi-Newton (QN);
- One-Step-Secant (OSS);
- Gauss-Newton (GN)

The chosen optimization method gives the search direction in each iteration. After the search direction has been found then an optimal step length has to be found. This was done by implementing a line search that successively calculated the cost functional at different positions along the search direction. This procedure is fairly ineffective because each evaluation of the functional needs the solving of a direct problem described in Chapter 1.3.1. This is solved by an FDTD simulation for each antenna. The line search needs about 5-10 evaluations of the functional. A method based on approximating the cost functional with a second order polynomial is presented in Chapter 4 using only one more calculation of the gradients which in terms of calculation time is comparable to only two evaluations of the functional. This is because calculating the gradients involves solving the direct problem for both $\mathbf{E}_m$ and $\tilde{\mathbf{E}}_\mathbf{m}$ in Eq 1.3.

In order to reconstruct the dielectric profile more efficiently and perhaps to avoid getting trapped in local minima, one wish to make use of *a* priori information, any information that we have about the solution before solving the inverse problem. Such information could be the position or shape of the investigated object or specific values of the dielectric parameters. It is still quite unexplored how to use this kind of information together with the optimization algorithm to improve the reconstruction. In Chapter 3 a simple way of taking *a* priori data in form of specific dielectric values into account is proposed that mainly enhances the connection between the permittivity and conductivity images.

In Chapter 5 the interrogating pulse is investigated and a different pulse is proposed. Test results are shown in Chapter 6 and the conclusions are found in Chapter 8. An attempt of reconstructing an image of an excised sample from an operation containing a tumor was made and is described in Chapter 7 together with an identification of the error sources.

Chapter 2

Optimization methods

Once the gradients have been found using Eq 1.3 an optimization method is needed to move towards the minimum of the cost functional and thereby improving the reconstructed image. The simplest optimization method used to find a minimum of a function is the steepest descent method which in every iteration goes in the negative direction of the gradient. There are however many methods that more wisely choose the search direction for each iteration by taking advantage of more information than the present gradients leading to faster convergence. In this thesis three different optimization methods have been implemented and tested together with the already implemented conjugate gradient method.

The optimization procedure can be summarized as follows:

1. Set $k := 0$ and select an initial guess $\mathbf{x_0}$.

2. Find a search direction $\mathbf{d_k}$ using one of the optimization methods. If $k = 0$ then $\mathbf{d}_k = -\mathbf{g}_k$. Here $\mathbf{g}_k$ is the current gradient.

3. Find the optimal step length $\lambda \geq 0$ that minimizes the cost functional F.

$$\begin{aligned} \lambda_k &= minF(\mathbf{x}_k + \lambda \mathbf{d}_k) \\ \mathbf{x}_{k+1} &= \mathbf{x}_k + \lambda_k \mathbf{d}_k \end{aligned} \tag{2.1}$$

In practice the optimal step length is never found but approximated by

a line search method which is further discussed in Chapter 4.

4. Compute

$$\begin{aligned} \Delta \mathbf{x}_k &= \lambda_k \mathbf{d_k} \\ \Delta \mathbf{g}_k &= \mathbf{g}_{k+1} - \mathbf{g}_k \end{aligned} \tag{2.2}$$

5. Set $k := k + 1$ and if the limited number of iterations has not been reached go to step 2.

2.1 Conjugate gradient, CJG

A version of the conjugate gradient optimization method was already implemented in the program for solving the inverse problem. The search direction is here determined by

$$\mathbf{d}_k = -\mathbf{g}_k + \beta_k \Delta \mathbf{x}_{k-1} \tag{2.3}$$

All the vectors contain information about both permittivity and conductivity so the calculated gradients for permittivity and conductivity are both included in the vector $\mathbf{g}_k$ for instance. $\Delta \mathbf{x}_{k-1}$ is the previous step taken by the algorithm.

For the already implemented algorithm by Polak and Ribiére [14] the constant β_k is computed by

$$\beta_k = \frac{\Delta \mathbf{g}_{k-1}^T \mathbf{g}_k}{\mathbf{g}_{k-1}^T \mathbf{g}_{k-1}} \tag{2.4}$$

Consequently the resulting change in permittivity and conductivity is a linear combination of the corresponding present gradient and previous change.

2.2 Quasi-Newton, QN

Quasi-Newton methods are based on Newton's method that assumes that the function can be approximated as a quadratic around a stationary point. In multi-dimensions a quadratic approximation is expressed with use of the Hessian matrix of second derivatives. The Quasi-Newton methods do not make use of the Hessian matrix explicitly. Instead an approximation of the Hessian matrix or its inverse is used that is updated in each iteration. There are various update methods but I chose to implement the Broyden-Fletcher-Goldfarb-Shanno (BFGS) method. This method is further explained in [14].

The Hessian matrix $\mathbf{B}$ is an $n \times n$ real matrix and is defined by

$$\mathbf{B}_{i,j} = \frac{\partial^2 F}{\partial x_i \partial x_j} \tag{2.5}$$

In the above expression F is the cost functional and x_i is the i:th element of the vector $\mathbf{x}$. The index i denotes a specific dielectric parameter and should not be confused with denoting a specific iteration, k. It is also important to understand that $\mathbf{B}$ as well as its inverse $\mathbf{H}$ are both symmetric and contain information about both permittivity and conductivity. The inverse Hessian matrix $\mathbf{H}$ can be written as a block matrix with four equally large submatrices in the following way.

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_{\epsilon,\epsilon} & \mathbf{H}_{\epsilon,\sigma} \\ \mathbf{H}_{\sigma,\epsilon} & \mathbf{H}_{\sigma,\sigma} \end{pmatrix}$$

The indices should be understood in the following manner. $\mathbf{H}_{\epsilon,\sigma}$ contains information about how the search direction for ϵ depends on the gradient for σ. Since the block matrix is symmetric, the submatrices $\mathbf{H}_{\epsilon,\epsilon}$ and $\mathbf{H}_{\sigma,\sigma}$ are symmetric as well. It also holds that

$$\mathbf{H}_{\epsilon,\sigma} = \mathbf{H}_{\sigma,\epsilon}^T \tag{2.6}$$

The search direction $\mathbf{d_k}$ for iteration k is given by

$$\mathbf{d}_k = -\mathbf{B}_k^{-1}\mathbf{g}_k \tag{2.7}$$

The above expression shows that it is useful to work with the inverse of the Hessian matrix so that the step can be calculated easier. The BFGS update method for the inverse Hessian matrix is

$$\mathbf{H}_{k+1}^{BFGS} = \mathbf{H}_k + \left(1 + \frac{\Delta\mathbf{g}_k^T \mathbf{H}_k \Delta\mathbf{g}_k}{\Delta\mathbf{g}_k^T \Delta\mathbf{x}_k}\right) \frac{\Delta\mathbf{x}_k \Delta\mathbf{x}_k^T}{\Delta\mathbf{g}_k^T \Delta\mathbf{x}_k} - \frac{\mathbf{H}_k \Delta\mathbf{g}_k \Delta\mathbf{x}_k^T + (\mathbf{H}_k \Delta\mathbf{g}_k \Delta\mathbf{x}_k^T)^T}{\Delta\mathbf{g}_k^T \Delta\mathbf{x}_k} \tag{2.8}$$

with the first approximation

$$\mathbf{H}_0 = \mathbf{I} \tag{2.9}$$

The idea is to successively get better and better approximations of the Hessian matrix and take advantage of the second order polynomial approximation. However for a general nonlinear objective function, convergence to a solution cannot be guaranteed. When the method fails to converge the approximated inverse Hessian has to be reset. In the tryouts I reset the algorithm every time the functional decreases with 1% or less of the functional value.

The Quasi-Newton method requires a lot of memory for matrix storage. The Hessian matrix and its inverse are both $(n \times n)$ matrices. The number of unknown parameters n, is equal to two times the number of pixels in the reconstruction region. Hence, the amount of memory required is growing fast with resolution and for implementation in high resolution and 3D it is probably necessary to use some memory saving strategy. I utilized a very basic one, noticing that the matrix to be stored is always symmetric. With this in mind it is possible to just allocate memory for the upper triangular matrix. An area of 100 × 100 pixels will for example give 2 × 100 × 100 = 20000 unknown variables. The Hessian matrix will be a (20000 × 20000) matrix. If one only stores the upper triangular matrix it will require about 20000 × 20000 / 2 = 200 millions of stored variables. If each of them uses 8 bytes the amount of memory needed for just this matrix will be 1.5 GB. A full 3D scan would of course demand a lot of memory.

2.3 One Step Secant, OSS

This method is an attempt to bridge the gap between the conjugate gradient and the Quasi-Newton algorithm. In the OSS method the same update formula as for Quasi-Newton is used but unlike this approach it does not store the complete inverse Hessian matrix but assumes that for each iteration the previous Hessian was the identity matrix. So by replacing the inverse Hessian matrix with the identity matrix in Eq 2.8 we get

$$\mathbf{H}_{k+1}^{OSS} = \mathbf{I} + \left(1 + \frac{\Delta\mathbf{g}_k^T \Delta\mathbf{g}_k}{\Delta\mathbf{g}_k^T \Delta\mathbf{x}_k}\right) \frac{\Delta\mathbf{x}_k \Delta\mathbf{x}_k^T}{\Delta\mathbf{g}_k^T \Delta\mathbf{x}_k} - \frac{\Delta\mathbf{g}_k \Delta\mathbf{x}_k^T + (\Delta\mathbf{g}_k \Delta\mathbf{x}_k^T)^T}{\Delta\mathbf{g}_k^T \Delta\mathbf{x}_k} \tag{2.10}$$

which can be simplified by defining

$$\delta_k = \Delta\mathbf{g}_k^T \Delta\mathbf{x}_k \tag{2.11}$$

and

$$\phi_k = 1 + \frac{\Delta\mathbf{g}_k^T \Delta\mathbf{g}_k}{\delta_k} \tag{2.12}$$

so the expression for the inverse Hessian matrix is

$$\mathbf{H}_{k+1}^{OSS} = \mathbf{I} + \phi_k \frac{\Delta\mathbf{x}_k \Delta\mathbf{x}_k^T}{\delta_k} - \frac{\Delta\mathbf{g}_k \Delta\mathbf{x}_k^T + \Delta\mathbf{x}_k \Delta\mathbf{g}_k^T}{\delta_k} \tag{2.13}$$

Using Eq 2.7 and the expression for the inverse Hessian matrix the search direction can be written

$$\begin{aligned} \mathbf{d}_{k+1} &= -\mathbf{B}_{k+1}^{-1} \mathbf{g}_{k+1} = -\mathbf{H}_{k+1}^{OSS} \mathbf{g}_{k+1} \\ &= -\mathbf{g}_{k+1} - \phi_k \Delta\mathbf{x}_k \frac{\Delta\mathbf{x}_k^T \mathbf{g}_{k+1}}{\delta_k} \\ &+ \Delta\mathbf{g}_k \frac{\Delta\mathbf{x}_k^T \mathbf{g}_{k+1}}{\delta_k} + \Delta\mathbf{x}_k \frac{\Delta\mathbf{g}_k^T \mathbf{g}_{k+1}}{\delta_k} \end{aligned} \tag{2.14}$$

In order to simplify this expression even more we define

$$\alpha_k = \frac{\Delta\mathbf{x}_k^T \mathbf{g}_{k+1}}{\delta_k} \tag{2.15}$$

and

$$\beta_k = -\phi_k \alpha_k + \frac{\Delta\mathbf{g}_k^T \mathbf{g}_{k+1}}{\delta_k} \tag{2.16}$$

which finally gives

$$\mathbf{d}_{k+1} = -\mathbf{g}_{k+1} + \alpha_k \Delta\mathbf{g}_k + \beta_k \Delta\mathbf{x}_k \tag{2.17}$$

The new search direction is accordingly dependent on the present gradient, the difference between the present and previous gradient but also the previous step.

2.4 Gauss-Newton, GN

The Gauss-Newton optimization method can be implemented in order to minimize a sum of squares. When looking back at the problem formulation we see that the cost functional, Eq 1.2, that we want to minimize is just that, a sum of squares. The integral sign is in practice replaced by a summation over discrete time intervals.

$$F\left(\epsilon(\mathbf{x}), \sigma(\mathbf{x})\right) = \int_0^T \sum_{m=1}^{M} \sum_{n=1}^{N} \left(|\mathbf{E}_m(\epsilon(\mathbf{x}), \sigma(\mathbf{x}), \mathbf{R}_n, t) - \mathbf{E}_m^{meas}(\mathbf{R}_n, t)|^2 \right) dt, \tag{2.18}$$

But the Gauss-Newton method needs the Jacobian matrix for the cost functional which is difficult to calculate. This matrix consists of the gradient for each value in the above sum *i.e.* each contribution to the cost functional for a specific transmitter, receiver and time step.

$$\mathbf{J}_F = \frac{\partial \left(|\mathbf{E}_m(\epsilon(\mathbf{x}), \sigma(\mathbf{x}), \mathbf{R}_n, t) - \mathbf{E}_m^{meas}(\mathbf{R}_n, t)| \right)}{\partial x_n} \tag{2.19}$$

Calculation of this matrix would surely take a lot of time. Instead we consider the transmitters separately. The contribution from transmitter m is denoted F_m. To be able to still write the cost functional as a sum of squares we define

$$\begin{aligned} f_m &= \sqrt{F_m} = \\ &= \sqrt{\int_0^T \sum_{n=1}^{N} \left(|\mathbf{E}_m(\epsilon(\mathbf{x}), \sigma(\mathbf{x}), \mathbf{R}_n, t) - \mathbf{E}_m^{meas}(\mathbf{R}_n, t)|^2 \right) dt} \\ \mathbf{f} &= (f_0, f_1, f_2, ..., f_M)^T \end{aligned} \tag{2.20}$$

With this definition the cost functional can be written

$$F = \sum_{m=1}^{M} F_m = \sum_{m=1}^{M} f_m^2 \tag{2.21}$$

And the resulting Jacobian matrix for f

$$\begin{aligned}\mathbf{J} = \frac{\partial f_m}{\partial x_n} &= \frac{\partial(\sqrt{F_m})}{\partial x_n} \\ &= \frac{1}{2\sqrt{F_m}} \cdot \frac{\partial F_m}{\partial x_n}\end{aligned} \tag{2.22}$$

As before x_n denotes the n:th argument of the vector $\mathbf{x}$. This vector contains all unknown values of permittivity and conductivity.
The Gauss-Newton step for an underdetermined system is, according to [15]

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \mathbf{J}^T \left(\mathbf{J}\mathbf{J}^T\right)^{-1} \mathbf{f} \tag{2.23}$$

In this expression $\mathbf{f}$ and $\mathbf{J}$ are calculated in the point $\mathbf{x} = \mathbf{x}_k$. In order to solve this equation and not calculate the inverse matrix explicitly, let us denote

$$\mathbf{w} = \left(\mathbf{J}\mathbf{J}^T\right)^{-1} \mathbf{f} \tag{2.24}$$

then the Gauss-Newton step can be written

$$\Delta\mathbf{x}_k = \mathbf{x}_{k+1} - \mathbf{x}_k = -\mathbf{J}^T\mathbf{w} \tag{2.25}$$

If the vector $\mathbf{w}$ can be found then the step can be calculated easily by multiplication with the Jacobian matrix $\mathbf{J}$. By defining

$$\mathbf{A} = \mathbf{J}\mathbf{J}^T \tag{2.26}$$

we get a linear equation system

$$\mathbf{A}\mathbf{w} = \mathbf{f} \tag{2.27}$$

The matrix $\mathbf{A}$ is both symmetric and positive definite so the system can be solved using Cholesky decomposition. Matrix $\mathbf{A}$ is decomposed into the product of two triangular matrices $\mathbf{L}$ and $\mathbf{L}^T$.

$$\mathbf{A} = \mathbf{L}\mathbf{L}^T \tag{2.28}$$

giving two triangular systems of linear equations, the first is

$$\mathbf{L}\mathbf{L}^T\mathbf{w} = \mathbf{f} \tag{2.29}$$

with a solution

$$\mathbf{L}^T\mathbf{w} = \mathbf{y} \tag{2.30}$$

which finally can be solved for $\mathbf{w}$. This results in the following step

$$\Delta \mathbf{x}_k = -\mathbf{J}^T \mathbf{w} \tag{2.31}$$

In this method the step length is actually found in the algorithm but I implemented the method together with the usual line search method. It should be mentioned that using the Jacobian matrix for each squared value in the cost functional would be a more correct implementation of the Gauss-Newton method. Interpreting Eq 2.31 one sees that $\mathbf{w}$ is a vector that applies a weighting to the gradients from different antennas. This implementation is thus a steepest descent method with weighted contributions from the antennas.

Chapter 3

Using a priori information

3.1 Directing the value of each pixel towards a priori values

In the reconstruction algorithm a gradient is calculated which is used by an optimization method to find a search direction. The search direction is given for both the permittivity and the conductivity. An appropriate scaling parameter is needed between the two. Sometimes there are two quite different images that come out for permittivity and conductivity even if the scattering object contains a strong coupling between the two.

One way to create a stronger correlation between permittivity and conductivity images is to consider each pixel of the image as a point in an $\epsilon - \sigma$-plane. After the search direction has been calculated, this point will move along a line in this plane and end up at new position in the next iteration. If one in advance knows the possible dielectric properties of the measured object, then these properties can be defined as points or regions in the $\epsilon - \sigma$-plane.

The connection between permittivity and conductivity is now made by finding the closest direction towards an a priori range. I have chosen to define closest direction in terms of smallest angle. After the direction has been found, the original search direction vector is projected onto this new direction. If the new direction has an angle of more than 90° to the original then the new direction vector is simply set to zero. As before a line search method is used to find the optimal step length. In order to ensure that the vector in the $\epsilon - \sigma$-plane does not pass the region it is moving towards a maximum step length can be found for each pixel. In Fig. 3.1 the vector $\mathbf{d}$ is

projected onto the direction towards the corner of a rectangular region in the $\epsilon - \sigma$-plane. The maximum step length in this case would be λ_{max} that makes the pixel reach the values of permittivity and conductivity corresponding to the corner of the rectangular region.

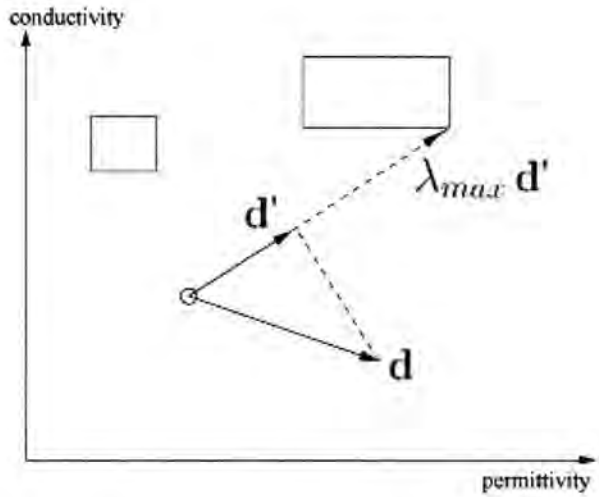


Figure 3.1: An example with three defined regions in the plane. The background medium is defined by a point and two other rectangular regions are defined. The direction **d** is being projected onto the closest allowed direction.

Chapter 4

Step length estimation

In order to find the optimal step length a search algorithm based on successive parabolic interpolation was implemented. With this technique the cost functional Eq 1.2 is evaluated at two points in the given search direction. The functional value in the present point is already known. By successively fitting a parabola to three points the extremum of this parabola replaces the oldest point and a new parabolic interpolation with three points can be done. This technique however has the disadvantage of being quite time consuming. In each iteration the program has to calculate about 5-10 functional values. Every calculation of this kind involves an FDTD-simulation for each transmitting antenna. Therefore it would be useful to estimate the optimal step length in a more efficient way. In this thesis I have implemented a simple mathematical expression for making an estimation of the optimal step length similar to an implementation in a paper by M.F. Moller [16].

The idea is to approximate the cost functional from Eq 1.2 with a second order polynomial in all unknown variables so that

$$F_{est}(\epsilon_1, ..., \epsilon_N, \sigma_1, ..., \sigma_N) = F_0 + \sum_{n=1}^{N} \left(A_n(\epsilon_n - \epsilon_n^*)^2 + B_n(\sigma_n - \sigma_n^*)^2 \right) \quad (4.1)$$

This approximation does not directly relate to the cost functional but is an assumed local behavior around a minimum that is unknown. The second order polynomial has the coefficients A_n and B_n. The variables $\epsilon_1, ..., \epsilon_N$ and $\sigma_1, ..., \sigma_N$ are the values of permittivity and conductivity respectively for each pixel $1, ..., N$. The minimum is found in the point defined by $\epsilon_n = \epsilon_n^*$ and $\sigma_n = \sigma_n^*$. In this point $F_{est} = F_0$.

A corresponding approximation to the gradient is found by taking the derivative for each variable

$$\begin{aligned}\mathbf{g_{est}} &= \nabla F_{est} = \\ &= (2A_1(\epsilon_1 - \epsilon_1^*), ..., 2A_N(\epsilon_N - \epsilon_N^*), 2B_1(\sigma_1 - \sigma_1^*), ..., 2B_N(\sigma_N - \sigma_N^*))\end{aligned} \tag{4.2}$$

The current approximation to the unknown dielectric parameters are defined by the point $\mathbf{x_0}$ and the search direction given by an optimization method is denoted $\mathbf{d}$. It is along this direction that we want to search for the minimum value of the functional. The gradient in Eq 4.2 is a linear function in all variables so following a line in the search space will produce a linear change of the gradient

$$\mathbf{g_{est}}(\mathbf{x}_0 + \lambda\mathbf{d}) = \mathbf{g_{est}}(\mathbf{x_0}) + \lambda\mathbf{s_d} \tag{4.3}$$

The gradient in the point $\mathbf{x_0}$ is known so

$$\mathbf{g_{est}}(\mathbf{x}_0 + \lambda\mathbf{d}) = \mathbf{g}(\mathbf{x_0}) + \lambda\mathbf{s_d} \tag{4.4}$$

The vector $\mathbf{s_d}$ is the rate of change of the gradient following the direction vector $\mathbf{d}$. This vector can be approximated by carrying out a movement in the search direction by just a tiny step and once more calculating the gradient. So setting $\lambda = \delta$ and solving Eq 4.4 for $\mathbf{s_d}$ gives

$$\mathbf{s_d} = \frac{\mathbf{g}(\mathbf{x_0} + \delta\mathbf{p}) - \mathbf{g}(\mathbf{x_0})}{\delta} \tag{4.5}$$

Now by moving in the direction of $\mathbf{d}$ we do not necessarily intersect the minimum of F_{est} but since we know that F_{est} is a second order polynomial we know that following any line will give a parabolic curve. The derivative along the curve is simply a scalar product of the gradient and the search direction

$$\frac{dF_{est}}{d\lambda}(\mathbf{x} = \mathbf{x}_0 + \lambda\mathbf{d}) = \mathbf{g_{est}} \cdot \mathbf{d} \tag{4.6}$$

So the minimum along the search line is found where the derivative is zero. This in combination with the expression for $\mathbf{g_{est}}$ in Eq 4.4 gives λ_{est} such that

$$\mathbf{d} \cdot (\mathbf{g}(\mathbf{x_0}) + \lambda_{est}\mathbf{s_d}) = 0 \tag{4.7}$$

It is now a matter of solving the above expression for λ_{est}

$$\lambda_{est} = -\frac{\mathbf{d} \cdot \mathbf{g}(\mathbf{x_0})}{\mathbf{d} \cdot \mathbf{s_d}} \tag{4.8}$$

If F would be a second degree polynomial then the estimation $F_{est} = F$ would be exact and λ_{est} would give the minimum along the search line. In the real case however, one can question if the second order polynomial approximation is applicable. The results from tryouts of simple cases though, show that this estimation procedure works very well. Especially considering that the successive parabolic interpolation needs the evaluation of the functional value in 5-10 points whereas this procedure only needs the calculation of the gradient in one more point. As previously seen calculation of the gradient is done by two FDTD-simulations for each transmitting antenna. To estimate how much time this method saves the number of FDTD-simulations for each iteration is shown in Table 4.1 assuming that the successive parabolic interpolation needs 7 calculations of the functional in average. It turns out that this method saves approximately 55% of the calculation time.

Table 4.1: Number of FDTD-simulations for each iteration

Successive Parabolic Interpolation		Step length estimation with polynomial	
Calculating gradient	$2 \times N$	Calculating gradient	2 × N
Calculating functional 7 times	$7 \times N$	Calculating new gradient	2 × N

Chapter 5

The Pulse

The chosen sinusoidal pulse with Gaussian amplitude modulation is not the only possible pulse one can use in the reconstruction algorithm. Basically any pulse could be used in a simulation by the FDTD algorithm. For practical use there are however some aspects to take into account.

- The length of the pulse should be kept short as this will make the simulated time interval shorter, hence also the calculation time.
- The spectral content of the pulse should cover a region that can be measured and simulated accurately.
- A wide spectrum of frequencies can be problematic because of high frequency-dependence in dielectric properties of tissue. An algorithm that deals with dispersion should however be able to use a wide spectrum.

The first two requirements limit the pulse in the time and frequency domain respectively and a straight-forward choice is the already tested pulse from Eq 1.1. It is a fact that the reconstructed image is limited by the spectral content of the pulse. To see a particular object with this reconstruction algorithm a pulse is needed with a wavelength comparable to the size of the object. How the reconstructed image changes when introducing higher frequencies is demonstrated in the tryouts in Chapter 6.

With the sinusoidal pulse with Gaussian amplitude modulation one picks a center-frequency that gives a wavelength that matches the size of the object to be reconstructed. For practical purpose the algorithm should be able

to reconstruct an image containing both small and big structures which involves high and low frequencies. For higher frequencies, the non-linear effects get stronger. For this reason a frequency-hopping procedure has been used. Starting with a low center-frequency that has a wavelength that matches a large structure and successively moving up in frequency and hence down in wavelength to find smaller objects and to sharpen edges. For the higher frequencies one uses the outcome of the previous frequency as initialization. An example of this is found in Chapter 6.

When using a frequency-hopping approach one must decide which center-frequencies and bandwidths to use but also how many iterations that should be run before changing to next frequency. It is crucial to have a good approximation before moving on to next frequency. Otherwise the algorithm is incapable of finding the minimum. This has also been shown in [9]. More seriously, in experimental tryouts it has occurred that when jumping up to a higher center-frequency the reconstruction algorithm starts to destroy the image and is incapable of improving it.

With these observed problems one asks if it is possible to use one single pulse for the whole reconstruction. It seems likely that such a pulse includes a wide span of frequencies to be able to reconstruct both small and big structures. The bandwidth of the ordinary sinusoidal pulse is however fairly limited due to the demand of keeping the spectral content within certain boundaries. In order to have energy equally distributed in a region in the frequency domain we study the function denoted $sinc(t) = \frac{sin(\pi x)}{\pi x}$. If we wish to keep the spectral energy from ν_{min} to ν_{max} the following function is suitable:

$$\mathbf{E}(t) = \cos\Big(\pi(\nu_{min} + \nu_{max})t\Big)\mathrm{sinc}\Big((\nu_{max} - \nu_{min})t\Big) \tag{5.1}$$

As seen in the Fig. 5.1 this pulse has the desired property concerning spectral energy distribution but viewed in the time domain one sees that it is a very long pulse. By multiplying the pulse with a Gaussian window function it can be shortened significantly. The corresponding change in the frequency domain will be a convolution that only makes the rectangle to smear out a bit at the edges. Care should be taken so that the frequency content still lies in the appropriate region. The shortened pulse is shown in Fig. 5.2. As seen in the frequency plot the amplitude is not zero for zero frequency. For experimental image reconstruction, as already mentioned, it is necessary to use a range of frequencies that can be accurately measured

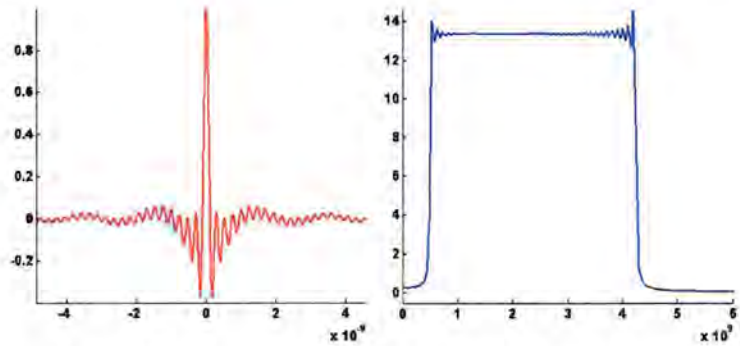


Figure 5.1: The pulse from Eq 5.1 in time domain to the left and frequency domain to the right.

and simulated.

$$\mathbf{E}(t) = e^{-\alpha t^2} \cos\Big(\pi(\nu_{min} + \nu_{max})t\Big)\mathrm{sinc}\Big((\nu_{max} - \nu_{min})t\Big) \tag{5.2}$$

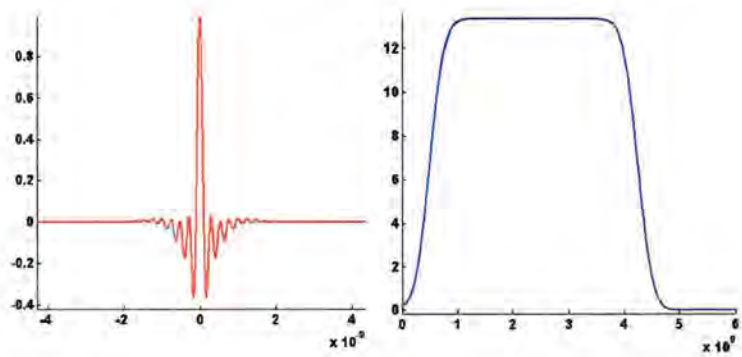


Figure 5.2: The pulse from Eq 5.2 with $\alpha = 8.92 \times 10^{17}$ in time domain to the left and frequency domain to the right.

The idea is that this pulse should be able to replace the frequency-hopping technique and simultaneously reconstruct both big and small structures. The waves with low frequencies can reconstruct large structures in the first iterations. Hopefully the waves with higher frequencies do not destroy the results from the lower frequency waves but bring in useful information when the large structures have been reconstructed. If the high frequency content makes reconstruction impossible due to the non-linearities then it would also

be possible to further modify the pulse by applying a low-pass filter for example. This could easily be done by integration.

$$\mathbf{E}(t) = \int_{-\infty}^{t} e^{-\alpha t_1^2} \cos\Big(\pi(\nu_{min} + \nu_{max})t_1\Big)\mathrm{sinc}\Big((\nu_{max} - \nu_{min})t_1\Big)dt_1 \qquad (5.3)$$

Chapter 6

Tryouts and results from synthesized data

In this chapter a number of tryouts are presented with the aim to test the different proposed methods from previous chapters. All tryouts in this chapter are based on numerical simulation and the experimental analysis is presented in Chapter 7.

If simulation and reconstruction are done on the same grid then the numerical errors for the two situations are equal. This could result in an unrealistically improved reconstruction. To avoid committing the *inverse crime*, a white Gaussian noise is added to the simulated measurement data $\mathbf{E}_m^{meas}(\mathbf{R}_n, t)$ with signal-to-noise ratio (SNR) defined as:

$$SNR = 10 \log_{10} \frac{\int_0^T E^2 dt}{\int_0^T N^2 dt},$$

where E is the electric field sampled at the receiver points and N is the added noise signal. Throughout these tryouts a value of SNR = 100 is being used.

6.1 Tryout 1: Reconstructing a triangle

Purpose:
In order to test the general performance of different optimization methods a simple test case with a triangular area was created. The triangle has sharp corners which should be challenging for the algorithm. The four optimization methods described in Chapter 2 were used in the reconstruction algorithm.

Settings:
To be able to use the QN-method and store the inverse Hessian matrix the reconstructed area was quite small, only 80 × 80 pixels. In spite of this relatively small area the number of unknown variables is $80 \times 80 \times 2 = 12800$ which makes the size of the Hessian matrix and its inverse (12800×12800). The total simulated area was 128 × 128 pixels. The total simulated area is shown as the outer rectangle and the reconstructed area is shown as the inner rectangle in Fig. 6.1. The simulated area needs to be bigger than the antenna array so that there is space for the absorbing boundary layers. The image to be reconstructed has a background with $\epsilon = 3.5\ \epsilon_0$ and $\sigma = 0.06$ S/m. The triangle has $\epsilon = 7\ \epsilon_0$ and $\sigma = 0.12$ S/m. The triangle is equilateral with the side 100 mm. Each pixel is 2 mm × 2 mm. The numerical simulation was made with 11 antennas using a pulse with 1 GHz center-frequency and 1 GHz bandwidth. In the rest of this thesis I will use the notation *center-frequency @ bandwidth* for all pulses.

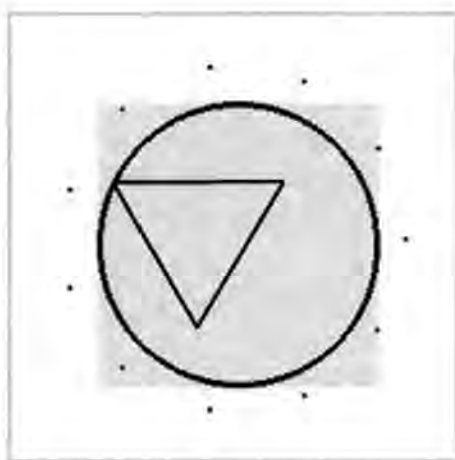


Figure 6.1: The antenna array inside the rectangle that depicts total simulated area. The circle shows the reconstructed area.

In these tryouts the relative error will be plotted on a logarithmic scale on the y-axis against number of iterations on the x-axis. The logarithmic relative error for permittivity is defined as follows and replacing ϵ with σ gives the corresponding error for conductivity.

$$\text{log rel err} = ln\left(\frac{\sum_{\text{all pixels}} |\epsilon_{reconstructed} - \epsilon_{true}|^2}{\sum_{\text{all pixels}} |\epsilon_{true} - \epsilon_{background}|^2}\right)$$

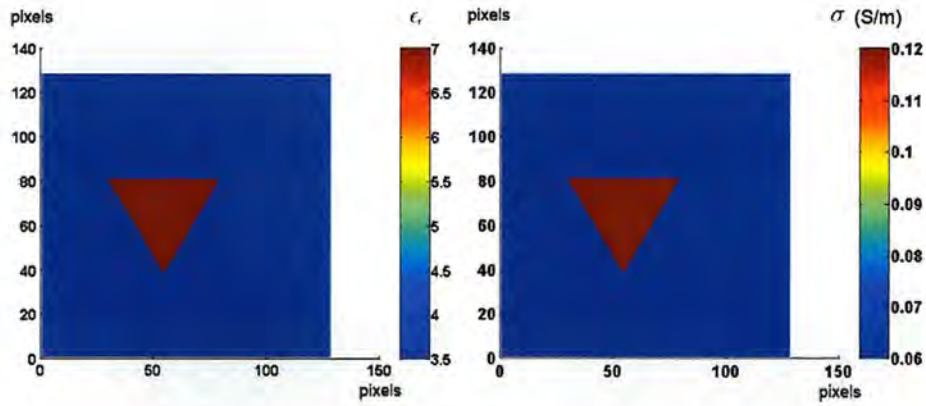


Figure 6.2: Original permittivity (left) and conductivity (right) images in triangle test case.

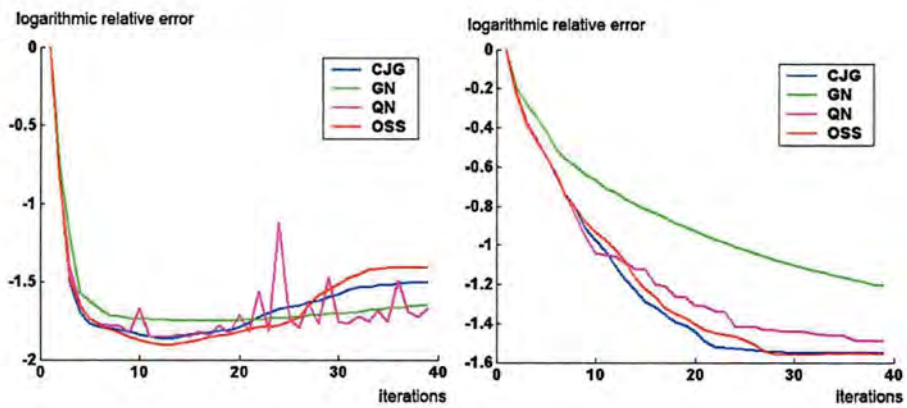


Figure 6.3: Logarithmic relative error for permittivity and conductivity from the four different optimization methods in the triangle test case.

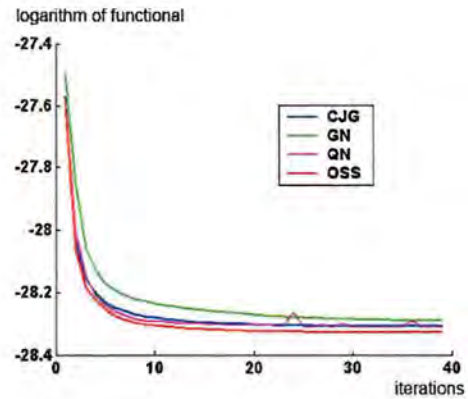


Figure 6.4: The logarithm of the cost functional in triangle test case.

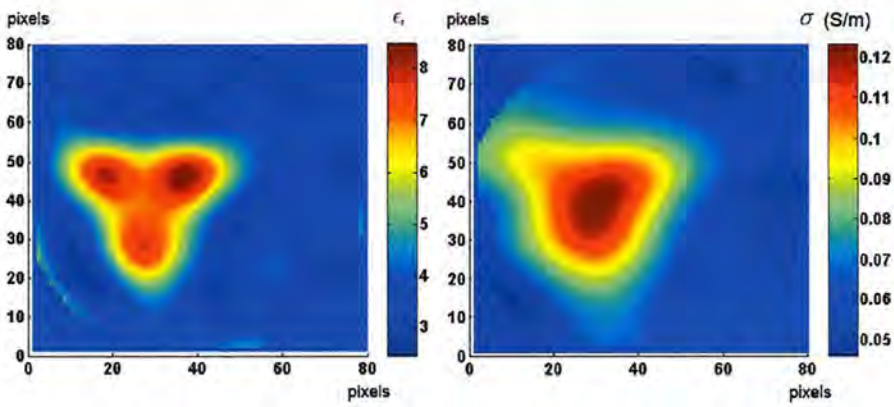


Figure 6.5: Reconstructed images for permittivity and conductivity after 15 iterations with CJG-method.

Results and Discussions:
In the Fig. 6.3 the error in permittivity reaches a minimum after about 15 iterations and then increases whereas the error in conductivity flattens out but continue to decrease. Since the shape of the triangle is exactly the same for permittivity and conductivity the difference in performance for the two must have to do with the contrast. With the chosen dielectric parameters for the triangle the algorithm can more easily see the contrast in permittivity. After about 15 iterations we have an approximation for the permittivity image that does not improve without more information from higher frequencies. But the conductivity image needs a few more iterations.

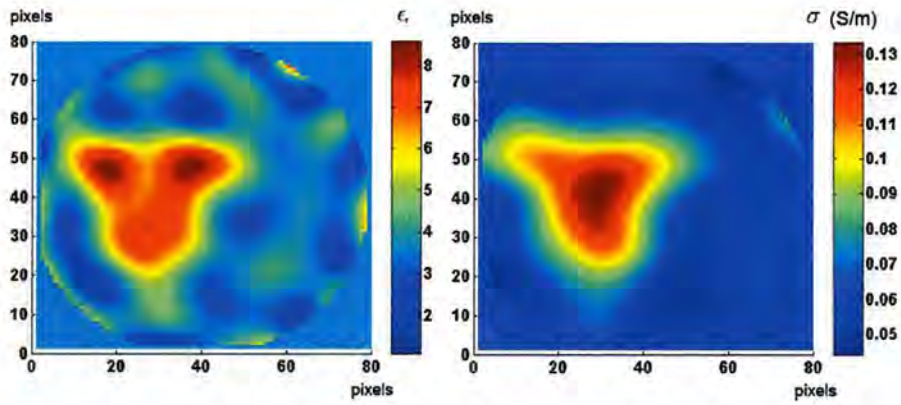


Figure 6.6: Reconstructed images for permittivity and conductivity after 39 iterations with CJG-method.

In a real reconstruction the error is unknown so there should be some other way to know how many iterations to do. The functional in Eq 1.2 is one possibility. The logarithm of the cost functional $ln(F)$ is plotted in Fig. 6.4. Here we can see that the functional is decreasing but at a very slow rate after about 10-15 iterations. Comparing the different optimization methods the CJG and OSS give the best result. After 15 iterations the reconstructed image with CJG looks as in Fig. 6.5. There is no big difference in the images among the different optimization methods. After 39 iterations the conductivity image has improved and the conductivity value 0.12 S/m is better approximated within the triangular area as shown in Fig. 6.6. The permittivity image has instead more artifacts around the triangle.

6.2 Tryout 2: Reconstructing circles

Purpose:
In this next tryout we investigate the different optimization methods ability to reconstruct eight different sized circles. See Fig. 6.8.

Settings:
In this tryout a bigger domain was chosen consisting of 220 $\times$ 220 pixels. The reconstruction was made on 172 $\times$ 172 pixels. Each pixel is now 1 mm $\times$ 1 mm and the number of antennas has been increased to 17. Fig. 6.7 is a corresponding figure to Fig. 6.1 for the new antenna array. The diameters of the circles are varying from 8 mm to 73 mm and they are seen in Fig. 6.8. The background is $\epsilon = 12\ \epsilon_0$ and $\sigma = 0$. The circles all have values of $\epsilon = 15\ \epsilon_0$ and $\sigma = 0.2$ S/m. Simulated measurements were done at 500 MHz @ 500 MHz, 1.25 GHz @ 1.25 GHz and 3 GHz @ 3 GHz. The reconstruction was made by starting at 500 MHz and doing 15 iterations. The result was used as initialization to 1.25 GHz with another 15 iterations. And finally 15 more at 3 GHz. Another reconstruction was also made using just the pulse 3 GHz @ 3 GHz.

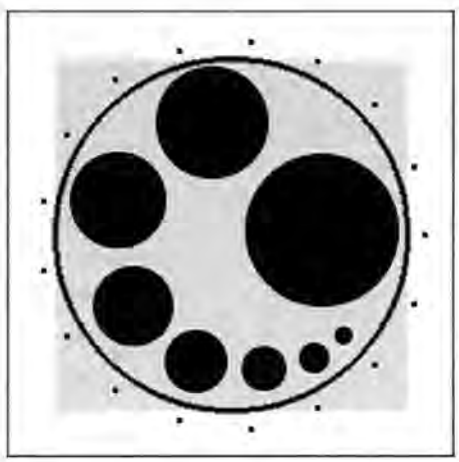


Figure 6.7: The antenna array inside the rectangle that depicts the simulated area. The big circle shows the reconstructed area.

Results and Discussions:
The error is plotted in the same way as in tryout 1. It is clear that moving up

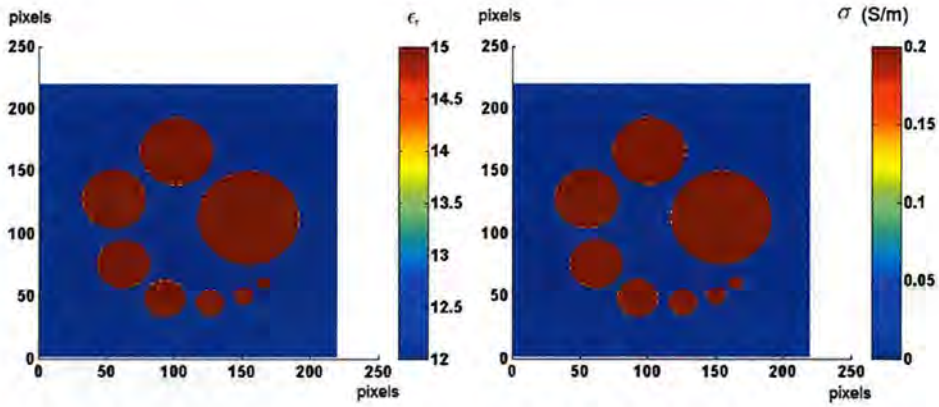


Figure 6.8: Original image in test case with circles.

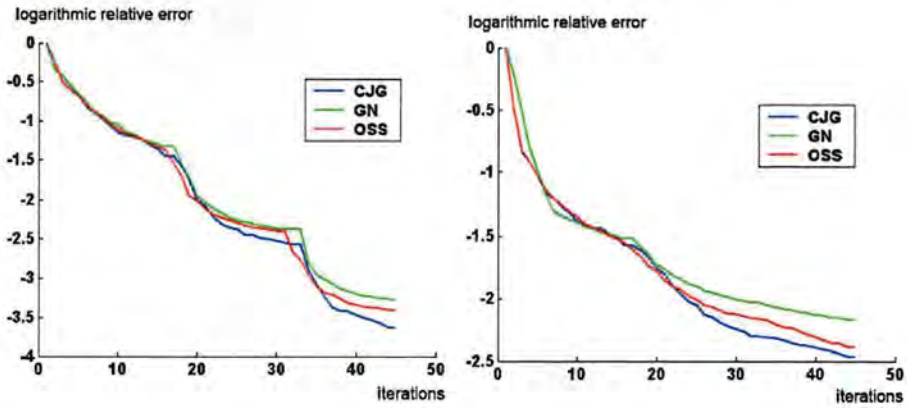


Figure 6.9: Results from three different optimization methods in the test case with circles.

to a higher frequency increases the downwards slope in Fig. 6.9. Each higher frequency brings in sharper edges and possibility to see smaller objects. But on the other hand a good approximation with low frequency is needed before using a pulse with higher frequency. Again the CJG method seems to be a good choice. The reconstructed images from different optimization methods are however very alike. The images from the CJG-method are shown in Fig. 6.10. From Fig. 6.11 we can see that starting with the high frequency pulse does not give a successful reconstruction. After 15 iterations the 3 GHz pulse has not been able to reconstruct one single circle.

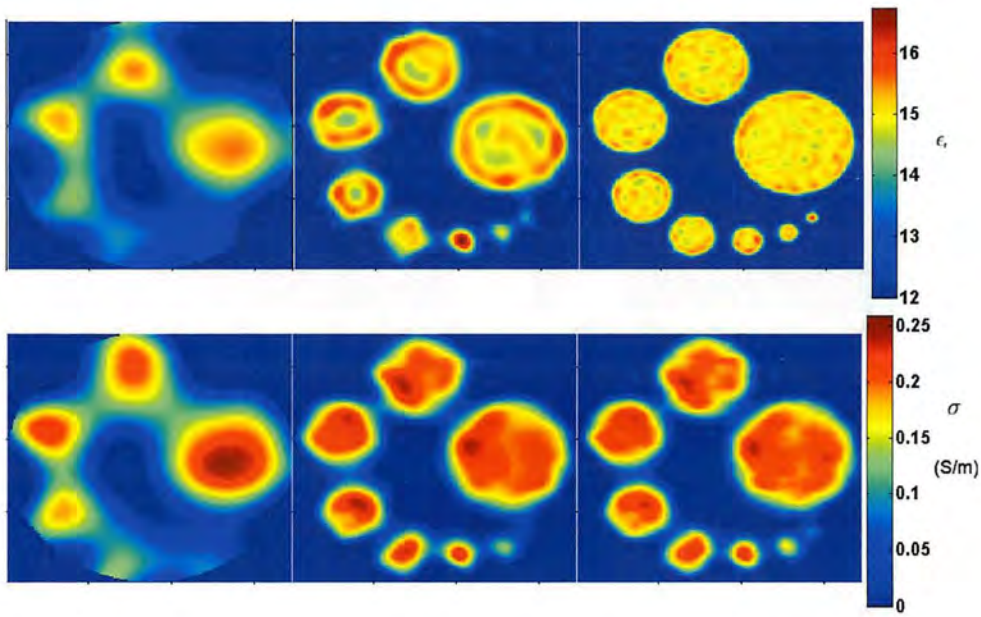


Figure 6.10: Reconstructed images after 15 iterations at each frequency with the CJG-method.

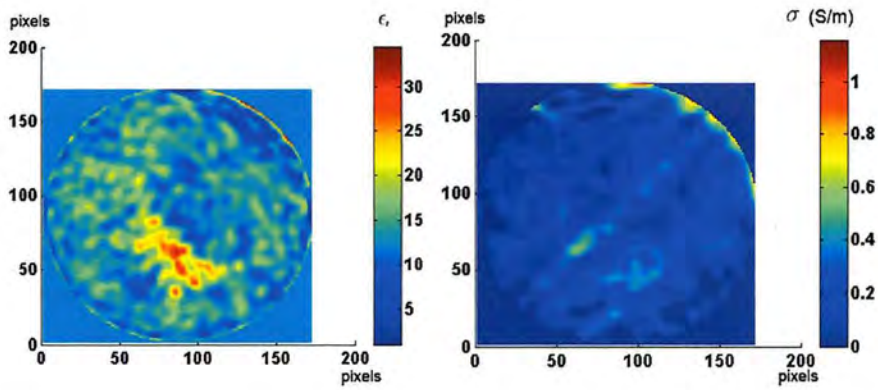


Figure 6.11: Reconstructed images after 15 iterations with the third high frequency pulse at 3 GHz.

6.3 Tryout 3: Using a priori information

Purpose:
With the aim of testing the a priori algorithm for redirecting the search direction described in Chapter 3 the test case with eight circles was run once again. This time using the a priori algorithm in each iteration.

Settings:
The same settings as tryout 2.

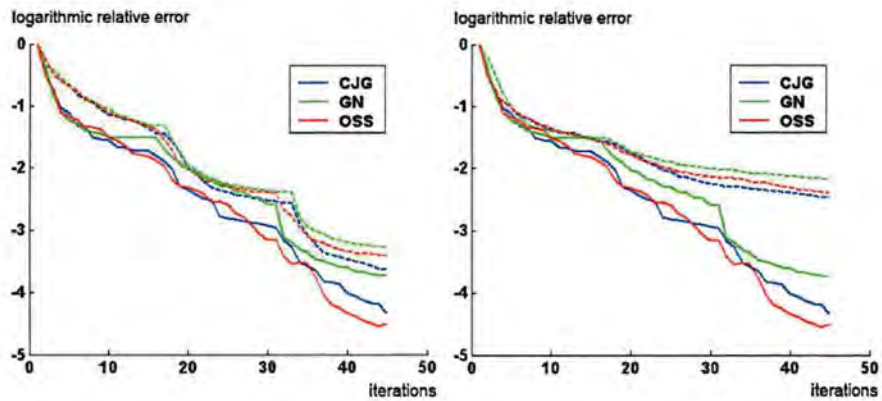


Figure 6.12: Results from three different optimization methods with the use of a priori information. The dashed lines represent the errors from tryout 2.

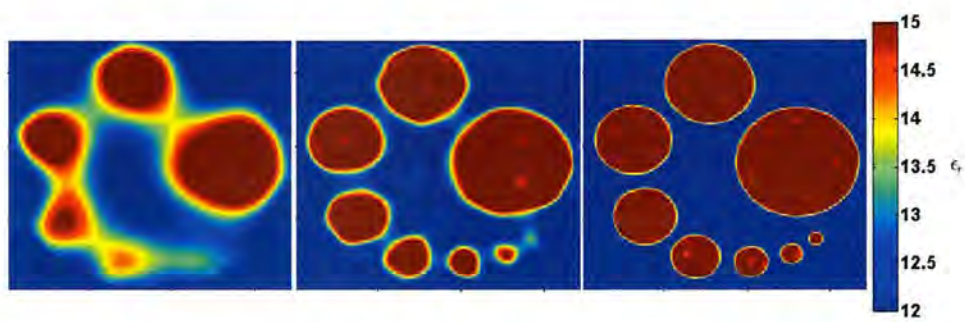


Figure 6.13: Reconstructed images after 15 iterations at each frequency with OSS-method and use of a priori information.

Results and Discussions:
Since the reconstructed images of permittivity and conductivity look the same only the permittivity is showed in the figure. The algorithm that redirects the gradients gives a better result in both permittivity and conductivity images but the big difference is in conductivity. This image had a slow convergence but is now being helped by the faster convergence in permittivity. The more accurate conductivity image can then improve the permittivity convergence as well. It should be pointed out that this is a very simple case with just two distinct dielectric values. In reality the dielectric properties can not be defined as points but as intervals. And with a larger number of possible different dielectric properties it would be more difficult for the algorithm to handle the connection between permittivity and conductivity. But in any case the results from this tryout shows that it is beneficial to use the correlation between permittivity and conductivity. And comparing the different optimization methods we can see that the OSS-method was more successful than the CJG-method in the final iterations. On the other hand there were a few iterations for which the CJG-method had the least error.

6.4 Tryout 4: Step length estimation

Purpose:
To test the step length estimation with polynomial approximation with the test case with a triangle was run again.

Settings:
The same settings as tryout 1.

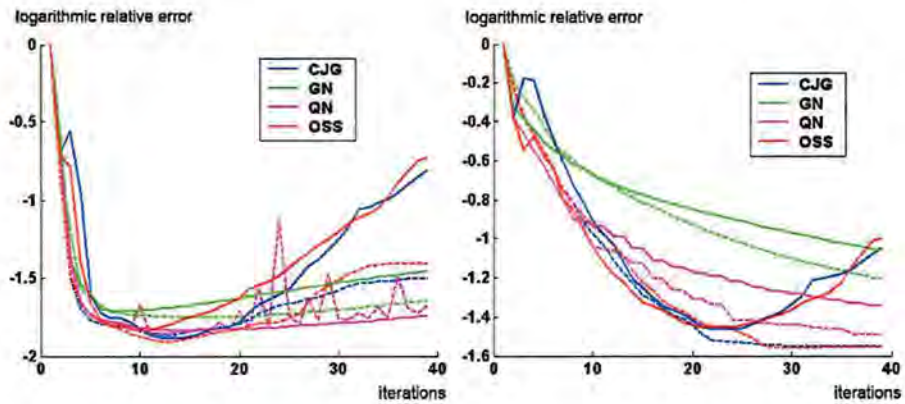


Figure 6.14: Results from the different optimization methods using the step length estimation with polynomial approximation. The dashed lines represent the errors from tryout 1.

Results and Discussions:
The described step length estimation goes much faster than the successive parabolic interpolation so even if the step length estimation would result in a slower convergence it could be worth it. During the same amount of time more iterations could be run with this approximation compared to when using the successive parabolic interpolation. However the results in Fig. 6.14 show that we do not loose much in convergence using the CJG and OSS-methods. When the errors start to grow however they grow faster by using the new approximation technique. It is interesting to see that the value of the cost functional remains fairly constant even when the error grows, see Fig. 6.17. The difference between this tryout and tryout 1 for iteration 20 with the CJG-method is shown in Fig. 6.16. It is the different artifacts that differ and no apparent connection to the triangular object is visible in these difference images.

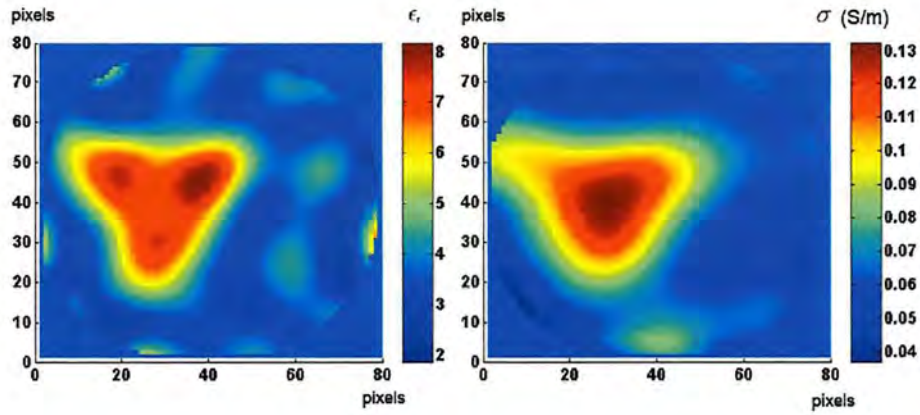


Figure 6.15: Reconstructed images after 20 iterations with the CJG-method and new step length estimation. No a priori information is being used.

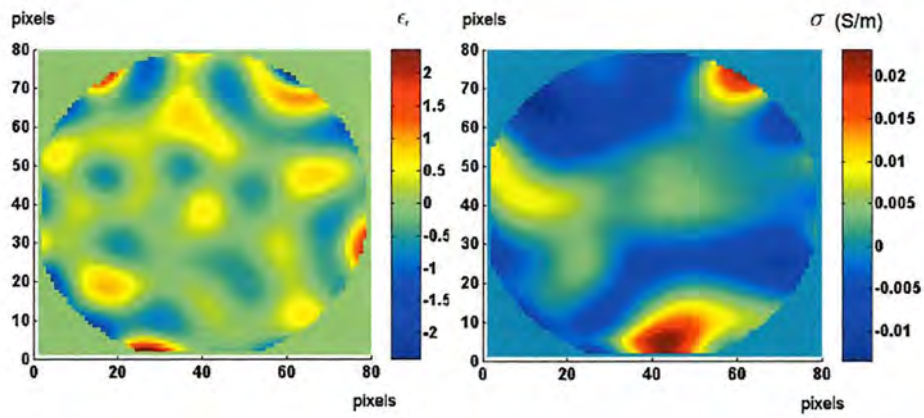


Figure 6.16: Difference images comparing images from Fig. 6.15 and results from tryout 1 after 20 iterations.

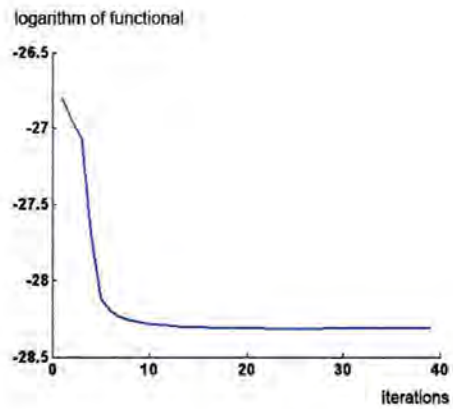


Figure 6.17: The logarithm of the cost functional plotted against iterations for the CJG-method with the new step length estimation.

6.5 Tryout 5: Step length estimation 2

Purpose:
Once more the step length estimation is tested but this time in the case with circles. Only the CJG-method is used this time with and without a priori algorithm.

Settings:
The same settings as tryout 2.

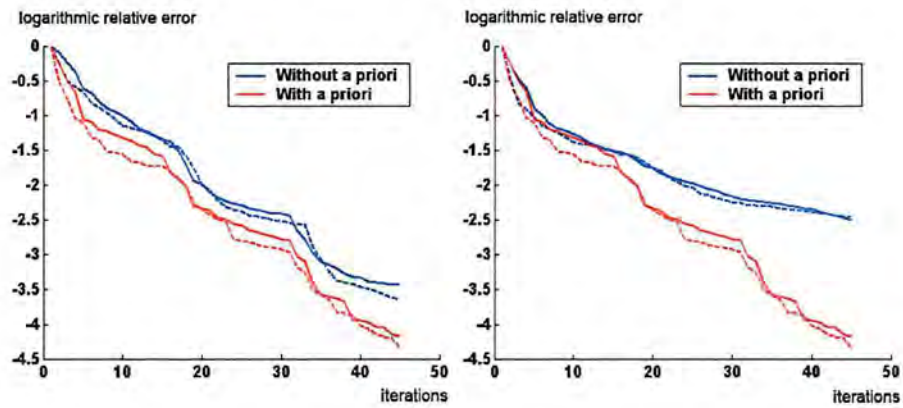


Figure 6.18: Results from the CJG-method with the new step length estimation with and without use of a priori information. The dashed lines represent the errors from tryout 2 and 3.

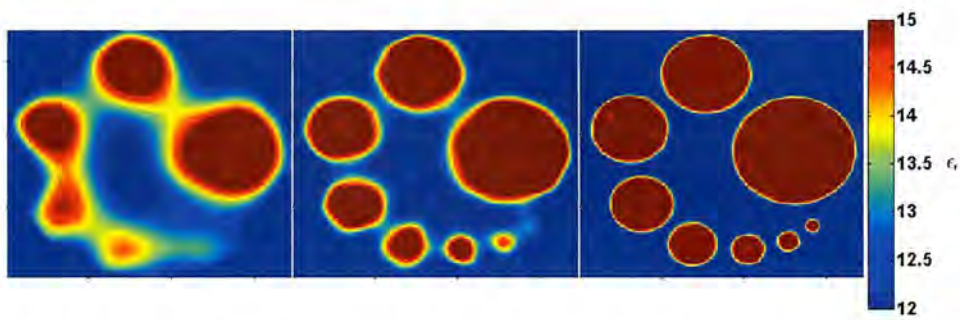


Figure 6.19: Reconstructed images after 15 iterations at each frequency with CJG-method and both a priori algorithm and step length estimation.

Results and Discussions:
As previously the permittivity and conductivity images are very similar so only the permittivity is shown in figure. No significant loss of accuracy can be seen using the new step length estimation. In both the case when using a priori information and not using it the error is slightly bigger when using the new approximation. But as the speed increases significantly this method seems promising. On average for this tryout the relative error only increased by 7.4% for permittivity and 6.6% for conductivity. As already discussed each iteration only takes about 45% of the calculation time needed using the successive parabolic interpolation.

6.6 Tryout 6: Two small objects

Purpose:
As both the a priori algorithm and the new step length estimation work fine another test case is made to test the performance in doing another kind of reconstruction. This test contains two small objects. The ability to see two small objects as two and not one big object is an important feature if the reconstruction should be used for detecting cancer tumors.

Settings:
Also in this tryout the domain was chosen to be 220 × 220 pixels. The reconstruction was made on 172 × 172 pixels. Each pixel is 1 mm × 1 mm and the number of antennas is 17. The diameters of the small circles are 8 mm. The background is $\epsilon = 12\ \epsilon_0$ and $\sigma = 0.32$ S/m. The both circles have values of $\epsilon = 55\ \epsilon_0$ and $\sigma = 1.0$ S/m. The original image is shown in Fig. 6.20. Reconstruction was found to be most successful on 1.25 kHz @ 1.25 kHz.

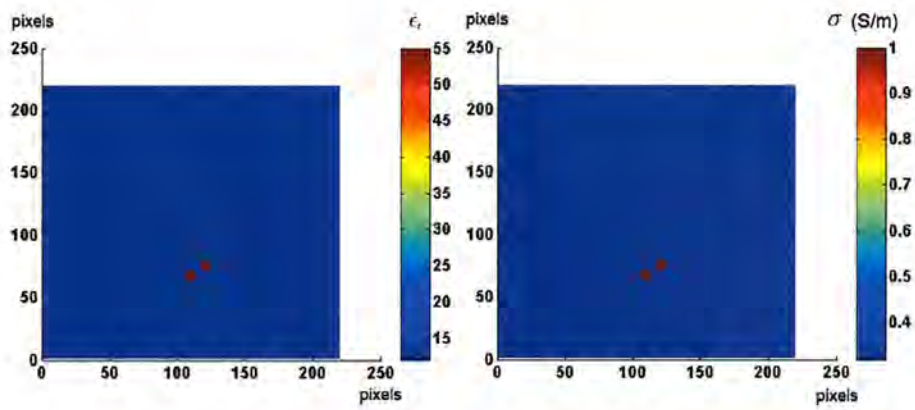


Figure 6.20: Original image in test case with two small objects.

Results and Discussions:
As previously using the a priori algorithm makes the images and errors for permittivity and conductivity similar so only the permittivity error and image are shown in the figures. The resulting error can be found in Fig. 6.21 and the reconstructed image with the CJG method after 60 iterations can be seen in Fig. 6.22. The corresponding image for the OSS-method is found in Fig. 6.23. The GN-method did not change the reconstructed image significantly after 12 iterations so this iteration is viewed in Fig. 6.24. It is interesting

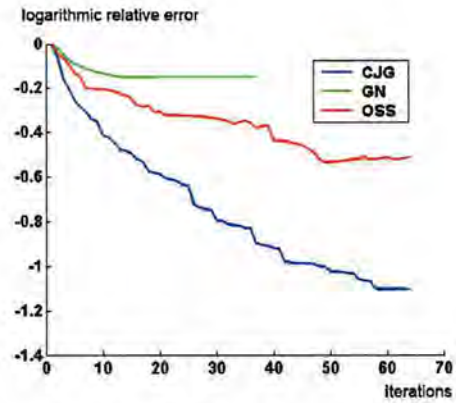


Figure 6.21: Results from three different optimization methods in test case with two small objects.

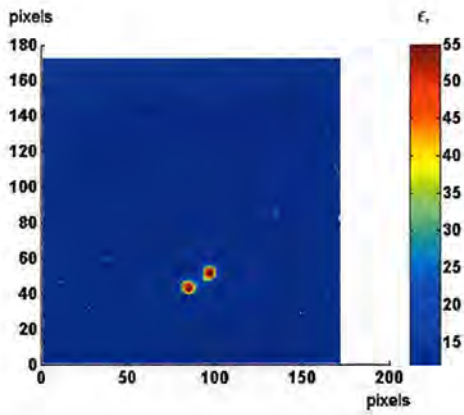


Figure 6.22: Reconstructed image after 64 iterations with CJG-method. A priori information and new step length estimation are being used.

to see the big difference comparing the three optimization methods. This time the CJG-method has a much better convergence than the OSS-method. Around 60 iterations are needed before the error plot flattens out in Fig. 6.21. The OSS-method does not reach the accurate dielectric values and the two circles spread out to form a single object. The CJG-method instead is capable of resolving the two objects.

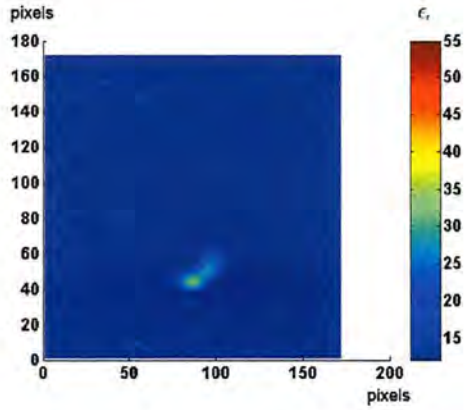


Figure 6.23: Reconstructed image after 64 iterations with OSS-method. A priori information and new step length estimation are being used.

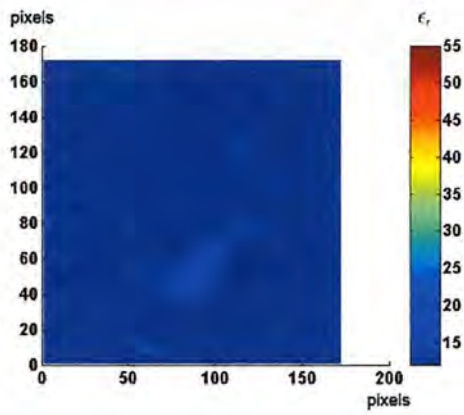


Figure 6.24: Reconstructed image after 12 iterations with GN-method. A priori information and new step length estimation are being used.

6.7 Tryout 7: The new pulse

Purpose:
To test the pulse from Eq 5.2 we run the test case with eight circles once again. We use the CJG-method with both a priori algorithm and new step length estimation.

Settings:
The settings are the same as in tryout 2 except that the pulse that is being used is the one given in Eq 5.2 and plotted in Fig. 5.2. It has its energy concentrated between 500 MHz and 4.25 GHz. This interval covers about the same interval as the three pulses from previous tryouts cover together. Here the same pulse is used in all iterations.

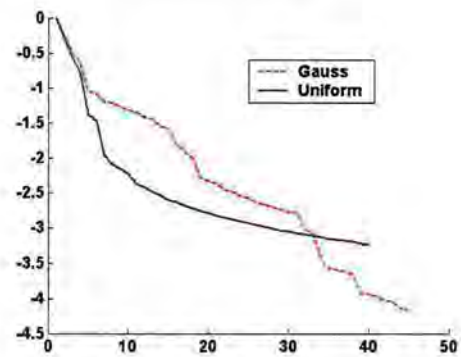


Figure 6.25: Remaining error using the pulse with Gaussian distribution from Eq 1.1 compared to using the pulse with uniform distribution in Eq 5.2.

Results and Discussions:
From this tryout several interesting facts can be noted. In Fig. 6.25 we note a faster convergence in the first iterations. The reconstructed images can be seen in Fig. 6.26 (uniform distribution) and 6.27 (Gaussian distribution). After just 5 iterations at least 7 of the 8 circles can be seen in the reconstructed image. After 15 iterations in the previous tryouts only 5 of the circles were visible. From iteration 5 to 20 the new pulse is actually around 10 iterations ahead of the previous reconstruction. It seems that this pulse can make use of higher frequencies earlier in the reconstruction and get a fast convergence early on. But it gets slower after a while which indicates that also this pulse needs to be tweaked to get a better performance. Ideally the rapid slope in the first iterations could be kept in the following iterations. Perhaps this could be done by modifying the spectral content of the pulse towards higher

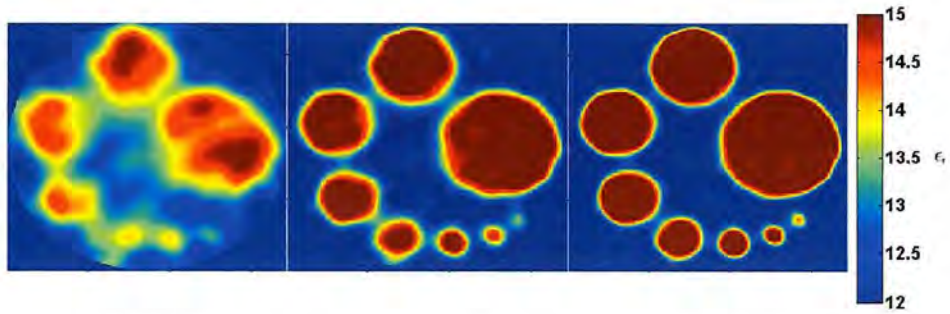


Figure 6.26: Reconstructed permittivity images after 5, 15 and 40 iterations with CJG-method using both a priori information and the step length estimation. This reconstruction uses the pulse with uniform distribution given in Eq 5.2.

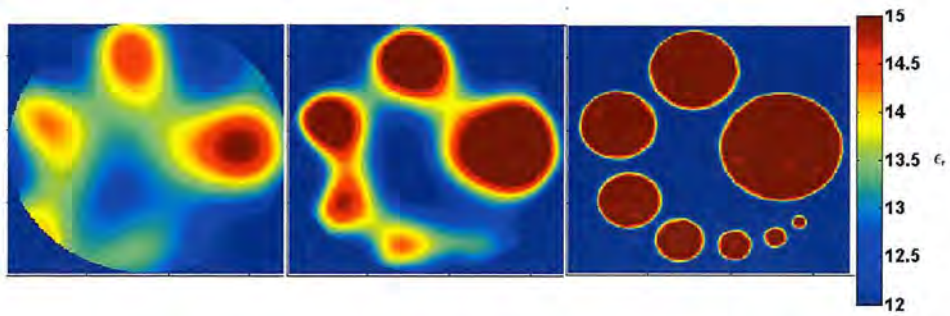


Figure 6.27: Reconstructed permittivity images after 5, 15 and 40 iterations with CJG-method using both a priori information and the step length estimation. This reconstruction uses the pulse with Gaussian distribution given in Eq 1.1.

frequencies for each iteration. The convergence rate is surely dependent on the pulse and this is a part of the reconstruction that should be possible to improve considerably with some optimization.

Chapter 7

Analysis of experimental data

7.1 Measurement

To test the applicability of the reconstruction algorithm for tumor detection a piece of breast that had been removed by surgery at Malmö University Hospital was placed in a plastic bag inside of the antenna array. The experimental setup shown in Fig. 1.2 was used. Measurements were made with battery water as coupling medium. The medium was measured to have $\epsilon = 79.27\ \epsilon_0$ and $\sigma = 0.001$ S/m at 300 MHz.

7.2 Analysis

Initially an attempt of reconstruction was made using a model of the antennas in battery water. The absorbing boundary layers discussed in section 1.3.1 were used to simulate an infinitely large box. The algorithm was incapable of finding the contour of the object as seen in Fig. 7.1. The calculation domain was 156 × 156 pixels and each pixel was 2 mm × 2 mm and the number of antennas was 20.

As described in section 1.3.2 the network analyzer does not sample the antenna signal but measures the steady-state response at distinct frequencies. Therefore the *measured* pulses that the reconstruction algorithm uses as input have to be synthesized from the real measurement data. These generated pulses were studied and something seemed to not work properly in the generation of pulses for antennas far from each other. In several cases the measured signal arrived to another antenna before it had been transmitted.

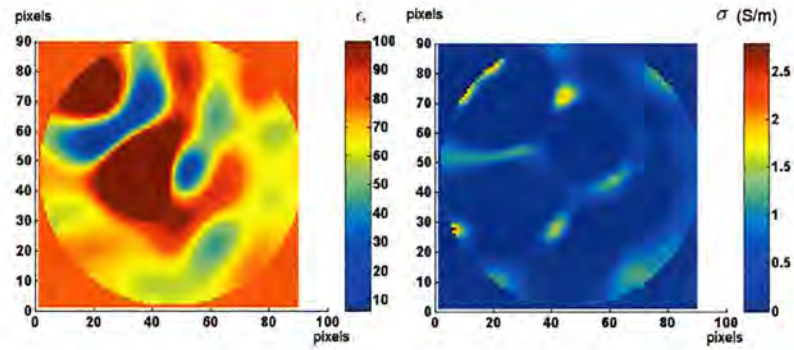


Figure 7.1: An example of failed reconstruction from experimental data. Reconstruction is attempted on 300 MHz @ 50 MHz. This is iteration 19.

In one reconstruction a pulse with 1.3 GHz @ 300 MHz was used. The signal from antenna 1 is plotted in Fig. 7.2.

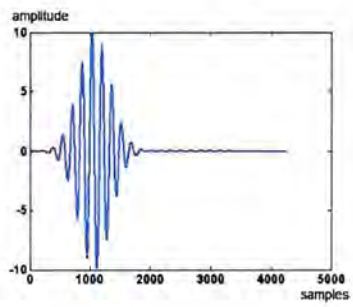


Figure 7.2: A pulse transmitted from antenna 1.

When the pulse from antenna 1 is received by a neighboring antenna it looks like in Fig. 7.3. We notice a delay in time because of the distance from the transmitter. We also notice that the amplitude has attenuated and that the signal has a longer duration. But the received signals by two other antennas further from the transmitter are shown in Fig. 7.4. In these cases there is a measured signal before the signal has been transmitted which is unrealistic. This indicates that the failure of the reconstruction algorithm may have to do with inadequate modelling of the system. In order to improve the model the plastic walls were used in the reconstruction program. The result however was not satisfactory. One example of reconstruction images is shown in Fig. 7.5. Reconstruction was made at 400 MHz @ 100 MHz.

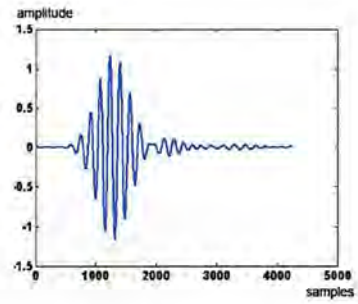


Figure 7.3: The pulse in Fig. 7.2 received by antenna 2.

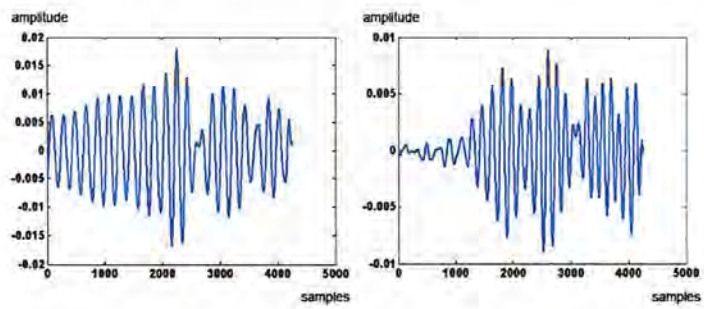


Figure 7.4: The pulse in Fig. 7.2 received by antenna 9 and 11.

7.3 Conclusion

In Fig. 7.6 we see the reflection coefficient from water to three other materials. The plastic used in the experimental setup is expected to have a reflection curve within this interval. The relative permittivity of water was set to 79 and the three curves represent corresponding values of 2, 2.5 and 3. We can see from the figure that at most 15% of the signal is being transmitted through the wall at zero angle. An angle of just two degrees gives total internal reflection.

If the signal from the receiving antennas would be sampled during a short time interval then the reflections from surrounding walls would not necessary

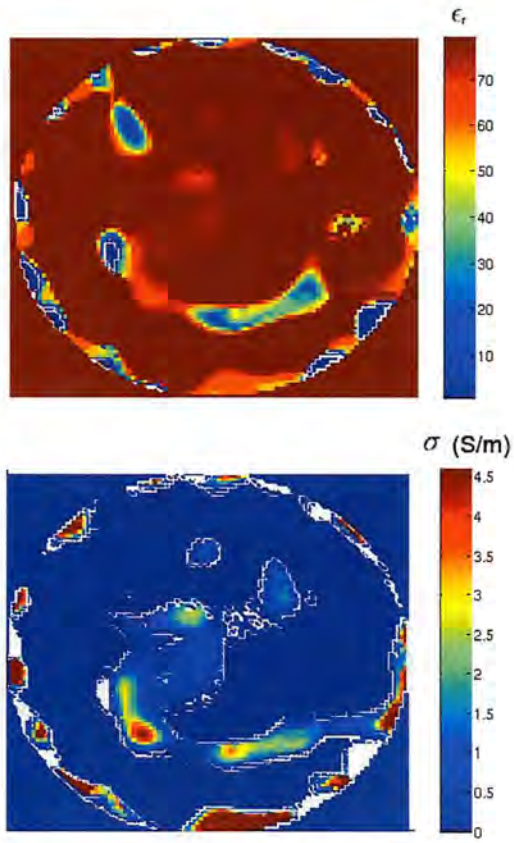


Figure 7.5: An attempt of reconstruction at 400 MHz @ 100 MHz with a model of the surrounding walls taken into account.

be a problem. In this case the box could be made big enough and the sampled time interval short enough so that the reflections would arrive to the antennas after the sampling had been done. Then the reflections would not influence the measurement at all. But as the network analyzer measures the response of each frequency when the signal has stabilized the measured signal will be influenced by all reflections arriving at any time. Using a medium like battery water in combination with reflecting walls should make a signal to stay in the system a very long time. In theory it would be possible to correctly model any surrounding material and reflections. But for two reasons it is beneficial to eliminate any reflections.

- Decreasing calculation time

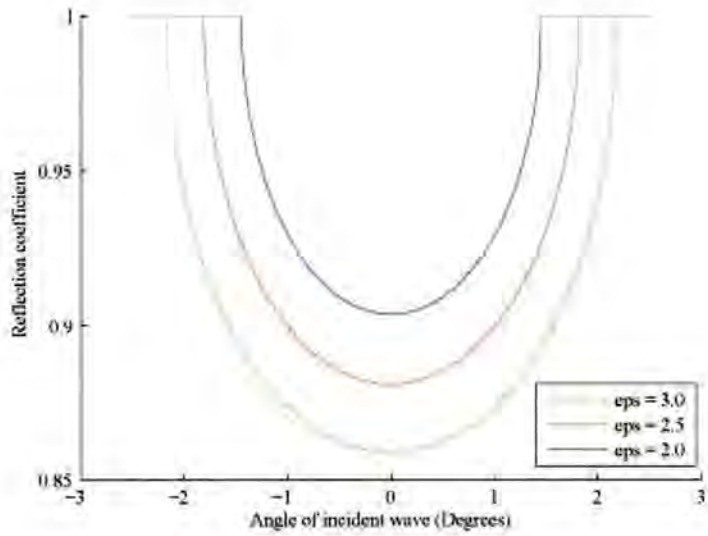


Figure 7.6: Reflection coefficient between water and three different materials plotted against incident angle.

- Reflection modelling is less precise

Regarding calculation time, the most time consuming part of the reconstruction is doing FDTD simulations. The time needed is proportional to the time interval that is being simulated. Hence, this interval should be kept as short as possible. When synthesizing the pulses from experimental data the direct problem has to be solved for the experimental setup including coupling medium and surrounding walls but without the investigated object. Since this involves FDTD simulations also this simulated time interval should be kept as short as possible. So even if reflections can be correctly modelled most of them influence the antennas with a large delay. These delay times would be much longer than the optimal time interval for the FDTD-method to simulate.

The second reason to eliminate reflections is that they will in all cases produce bigger errors to the reconstruction. If we assume that that we have a medium with $\sigma = 0$ so that there is no attenuation of the signal due to the medium and that the permittivity of the medium is measured with an error of just 0.5%. The wavelength is 2 cm then after traveling 20 cm, which is the diameter of the antenna array, the error in permittivity will cause an error of 9.0 degrees in the phase of the signal. As the phase error is proportional to

the distance travelled it turns out that the phase error is 90 degrees after 2 meters. The signal can be considered totally out of phase after just bouncing back and forth in the box two times.

Now the error in permittivity measurement of the coupling medium is just one of many. The estimated permittivity and position of walls as well as position of antennas will also produce increasing errors as the time interval is expanded. The 3D/2D-model errors are likely to also get bigger as the wave propagates. It seems that the most important improvement of the reconstructions at this time is to better handle reflections.

Two possibilities of solving the problems with reflections:

- Using a medium with higher conductivity. By doing so, signals will be attenuated so that less reflections can influence the measurement. This approach will have the disadvantage that less signal can reach the object.
- Using a surrounding environment that takes away all reflections.

The results from this investigation are not sufficient to conclude that two dimensional reconstruction is not applicable to resolve tumors. Further work needs to be done on the suggested sources of error. Unlike the idealized conditions of the simulated measurements. In order to successfully reconstruct an image with experimental data, several error sources have to be taken into account and minimized.

- Noise in measurements
- Response characteristics of the antennas
- Dispersion
- Reflections
- Model errors when simulating 3D waves in 2D

Chapter 8

Conclusions and Outlook

A number of tryouts have been made in order to test different optimization methods, the algorithm for taking use of a priori information, a polynomial approximation of the cost functional for step length estimation and finally a new pulse. To sum up the results are discussed below.

8.1 Optimization methods

From the tryouts 1-4 and 6 it seems that the CJG-method always works fine. Even though the OSS-method was more successful in tryout 3 there was no significant difference. The CJG and OSS method swapped places several times during the iterations in having least error. In tryout 6 the CJG-method had a big advantage over the OSS-method. The QN-method needs much more memory to save the approximated inverse Hessian matrix and does not seem to gain anything in convergence by doing so. The type of GN-method that was implemented in this thesis did not have any advantage compared to the other methods. So the conclusion from these tryouts is that the conjugated gradient method is best to use of the four tested methods.

8.2 Using a priori information

In tryout 3 we got a better convergence when using the a priori algorithm. Especially for the conductivity image that had a slower convergence. This result shows that it is beneficial to make use of the correlation between permittivity and conductivity. In this thesis a fairly simple method was used to do that. For practical purpose such an algorithm should be able to handle several combinations of allowed intervals in permittivity and conductivity. The tested algorithm is pixel based which means that two neighboring pixels

can take off in opposite directions introducing artifacts that makes the reconstruction more difficult. A low-pass filter could be used to not introduce faster variation in the reconstructed images than the pulse can handle. This was not a problem in tryout 3 because only two allowed regions were defined in the $\epsilon - \sigma$-plane.

8.3 Step length estimation

The polynomial approximation of the functional proved to be useful in estimating the optimal step length. Tryouts 4 and 5 clearly indicate that there is no significant loss in convergence rate. Instead the reconstruction time is much faster when the algorithm does not have to calculate the cost functional 5-10 times for each iteration in the line search. Approximately 55% of the calculation time can be saved.

8.4 The new pulse

The results from tryout 7 are very interesting. A pulse with its energy spread to a wider spectrum had a faster convergence in the first iterations. It also reconstructed smaller objects within the first five iterations without doing the reconstruction with low frequencies first. Perhaps the convergence rate in the first iterations could be kept by continuing to tweak the pulse in frequency domain for each iteration. The spectral content of the pulse could also be adapted to the spectral content of the residual signals. That is the difference between measured and simulated signal at each antenna. If a residual signal for example contains frequencies from 600 MHz then the structures corresponding to lower frequencies than 600 MHz should already be in place and the pulse should mainly lie over 600 MHz.

8.5 The future of Microwave Tomography

With constantly increasing computational power together with improved numerical techniques, Microwave Tomography definitely has potential to become reality and replace the X-ray examination. As mentioned in section 1.1 Microwave Tomography would have many advantages compared to the available breast screening tools today. There are still some challenges to overcome. Ideally one can see the results from the examination in real-time. To strive for this requires that every step in the reconstruction is optimized. I

am sure that many improvements can be done to make the image reconstruction more efficient. But it is not just about choosing a search method and finding ways to improve the convergence for the inverse problem but equally important is the choice of how to collect the data. Which antennas to use, how many, and where to place them? Optimal coupling medium? Should the surrounding walls be damping or reflecting? There are many choices concerning the experimental setup. So after some more years of clinical studies, improved algorithms, faster computers and more efficient measuring instruments I believe that the first Microwave systems will be available at the clinics.

Acknowledgements

I would like to thank my supervisor Andreas Fhager for helping me understanding the theory and giving me many good advices during the work. I would also like to thank Prof. Mikael Persson for giving me the opportunity to do this work. Patrik Dahlqvist at Medfield Diagnostics helped me extending the work towards experimental analysis. I also want to show my gratitude to my parents and friends that have supported me during the work.

Daniel Bernhardsson

Bibliography

[1] World Health Organization, http://www.who.int/mediacentre/ factsheets/fs297/en/index.html 2009

[2] S.J. Nass, I.C. Henderson, and J.C. Lashof, Eds: *Mammography and Beyond: Developing Technologies for the Early Detection of Breast Cancer*, National Cancer Policy Board, Institute of Medicine, and Commission on Life Studies, National Research Council, 2001.

[3] H. Fricke and S. Morse, "The electric capacity of tumors of the breast" *J. Cancer Res.*, vol. 10, pp. 340, 1940.

[4] S. S. Chaudhary, R. K. Misha, A. Swarup and J. M. Thomas, "Dielectric Properties of Normal & Malignant Human Breast Tissue at Radiowave and Microwave Frequencies," *Indian J. Biochem. Biophys.*, vol. 21, pp. 76-79, Feb. 1984.

[5] A. J. Surowiec, S. S. Stuchly, J. R. Barr, A. Swarup, "Dielectric Properties of Breast Carcinoma and the Surrounding Tissues," *IEEE Trans. Biomed. Eng.*, vol. 35, pp. 257-263, Apr. 1988.

[6] W. T. Joines, Y. Zhang, C. Li and R. L. Jirtle, "The Measured Electrical Properties of Normal and Malignant Human Tissues from 50 to 900 MHz," *Med. Phys.*, vol. 21, pp. 547-550, Apr. 1994.

[7] A. M. Campbell and D. V. Land, "Dielectric Properties of Female Human Breast Tissue Measured *in vitro* at 3.2 GHz," *Phys. Med. Biol.*, vol. 37, pp. 193-210, 1992.

[8] M. Lazebnik, D. Popovic, L. McCartney, C. B. Watkins, M. J. Lindstrom, J. Harter, S. Sewall, T. Ogilvic, A. Magliocco, T. M. Breslin, W. Temple, D. Mew, J. H. Booske, M. Okoniewski and S. C. Hagness, "A large-scale study of the ultrawideband microwave dielectric properties of normal, benign and malignant breast tissues obtained from cancer surgeries" *Physics in Medicine and Biology*, vol. 52, pp. 6093-6115, 2007.

[9] Andreas Fhager, *Microwave Tomography*, Ph.D. dissertation, Chalmers University of Technology, Göteborg, Sweden, 2006.

[10] J. H. Jacobi, L. E. Larsen and C. T. Hast, "Water-immersed microwave antennas and their application to microwave interrogation of biological targets," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-27, pp. 70-78, 1979.

[11] L. E. Larsen and J. H. Jacobi, "Microwave scattering parameter imagery of an isolated canine kidney," *Medical Physics*, vol. 6, pp. 394-403, 1979.

[12] M. Slaney, A. C. Kak, and L. E. Larsen, "Limitations of imaging with first order diffraction tomography," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-32, no. 8, pp. 860-873, Aug. 1984.

[13] E. Abenius, "Direct and inverse methods for waveguides and scattering problems in the time domain", Ph.D. dissertation, Dept. of Information Technology, Uppsala University, Uppsala, Sweden, 2005.

[14] Edwin K. P. Chong and Stanislaw H. Zak, "An introduction to optimization", Wiley-Interscience

[15] C. Wulff and A. Schebesch, "Numerical Continuation of Symmetric Periodic Orbits", SIAM J. Appl. Dyn. Syst. 5(3), pp. 435475, 2006.

[16] M.F. Moller, "A Scaled Conjugate Gradient Algorithm for Fast Supervised Learning", *Neural Networks*, vol. 6, pp. 525-533, 1993.