ISTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY

THE DERIVATION OF DIELECTRICAL BREAST MODELS FROM MRI IMAGES FOR BREAST CANCER DETECTION

M.Sc. Thesis by
Duygu SAP

Department : **Electronics & Communication Engineering**

Programme : **Telecommunications Engineering**

Thesis Supervisor: **Prof. Dr. İbrahim Akduman**

July 2011

ISTANBUL TECHNICAL UNIVERSITY ★ INSTITUTE OF SCIENCE AND TECHNOLOGY

THE DERIVATION OF DIELECTRICAL BREAST MODELS FROM MRI IMAGES FOR BREAST CANCER DETECTION

M.Sc. Thesis by
Duygu SAP
(504081306)

Date of submission : 06 May 2011
Date of defence examination : 15 June 2011

Supervisor (Chairman) : Prof. Dr. İbrahim Akduman(ITU)
Members of the Examining Committee : Assoc. Prof. Dr. Ali Yapar (ITU)
Assoc. Prof. Dr. Hülya Şahintürk (YTU)

July 2011

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

GÖĞÜS KANSERİ TETKİKİ İÇİN MRİ GÖRÜNTÜLERİNDEN DİELEKTRİK GÖĞÜS MODELLERİNİN ÇIKARIMI

YÜKSEK LİSANS TEZİ
Duygu SAP
(504081306)

Tezin Enstitüye Verildiği Tarih : 06 Mayıs 2011
Tezin Savunulduğu Tarih : 15 Haziran 2011

Tez Danışmanı : Prof. Dr. İbrahim Akduman (İTÜ)
Diğer Jüri Üyeleri : Doçent Dr. Ali Yapar (İTÜ)
Doçent Dr. Hülya Şahintürk (YTÜ)

Temmuz 2011

FOREWORD

I would like to thank my advisor, Prof Dr İbrahim Akduman, and declare my gratitude for all the support and knowledge he has provided me since the day I've started my Masters degree.

I would like to thank my colleague, Ahmet H. TUNÇAY, for his contributions to my thesis.

In particular, I would like to thank my family for having faith in my will and success, and being supportive and reassuring at all times.

May 2011 Duygu SAP

Mathematics Engineer

TABLE OF CONTENTS

Page

ABBREVIATIONS

MRI : Magnetic Resonance Imaging
GMM : Gaussian Mixture Model
DCE-MRI : Dynamic Contrast Enhanced Magnetic Resonance Imaging
ML : Maximum Likelihood
FDTD : Finite Difference Time Domain
CMI : Confocal Microwave Imaging
EM : Expectation Maximaziation

LIST OF FIGURES

Page

THE DERIVATION OF THE DIELECTRICAL BREAST MODELS FROM MRI IMAGES FOR BREAST CANCER DETECTION

SUMMARY

In this study, the objective is the investigation of the electrical properties' variations in the breast tissues in order to exhibit the heterogeneity of the breast tissues and the dispersive nature of the dielectric properties across the breast. In line with this purpose, MRI based breast models, which will presumably assist the breast cancer research carried out at microwave frequencies, were constructed. For precise results, several imaging processes such as the application of various smoothing and edge detection filters to the MRI images, which were obtained from Euromed Radiology Centre and Maltepe University, took place and MRI intensity values were transformed into uniform grids of dielectric values by using the MRI intensity histograms. Histograms were utilized in interpreting the tissue distribution of the relevant breast image and Gaussian curves were plotted to perceive the diversity between the two baseline breast tissue types, namely fibroglandular and fatty tissues. Then, the Gaussian mixture model and the basic statistical definitions were used in order to determine the intensity intervals. Afterwards, the intensity intervals computed were matched with the tissue–dependent dielectric intervals, subsequently, the intensity values were mapped to dielectric values, namely, dielectric constants and conductivity values, via piecewise linear interpolation and cubic spline interpolation. The dielectric distributions obtained from both of the approaches are compared and the resulting figures and graphs are illustrated here. Consequently, it is noticed that the piecewise linear interpolation results in a more scattered dielectric value distribution whereas cubic spline interpolation depicts a high contrast between the tissue regions. It is deduced that both of the methods give efficient results and realistic electrical breast models can be obtained from MRI images via these methods.

GÖĞÜS KANSERİ TETKİKİ İÇİN MRI GÖRÜNTÜLERİNDEN DİELEKTRİK GÖĞÜS MODELLERİNİN ÇIKARIMI

ÖZET

Bu çalışmada, amaç göğüs dokularının heterojen yapısını ve dielektrik özelliklerin göğüs içerisindeki değişken dağılımını gözlemleyebilmek için göğüs dokularının elektriksel özelliklerinin incelenmesidir. Bu amaç doğrultusunda, mikrodalga frekanslarda yürütülen göğüs kanseri çalışmalarında kullanılmak üzere MRI tabanlı göğüs modelleri oluşturulmuştur. Hassas sonuçlar elde edebilmek için, öncelikle Euromed Radyoloji Merkezi ve Maltepe Üniversitesi'nden alınan MRI görüntülerine pürüzsüzleştirme ve kenar belirleme filtreleri gibi çeşitli görüntü işleme teknikleri uygulanmış ve sonra MRI intensite değerleri MRI intensite histogramlarından yararlanılarak üniform bir dielektrik değer dağılımına dönüştürülmüştür. Bu kapsamda, histogramlar ilgili göğüs görüntüsünün doku dağılımı hakkında bilgi edinilmesinde ve Gauss eğrileri de iki temel doku türü olan yağ ve fibroglandular doku arasındaki ayrımın belirlenmesinde kullanılmıştır. Gauss karışım modeli ve temel istatiksel tanımlar aracılığıyla intensite aralıkları belirlenmiştir. Bundan sonra, hesaplanan intensite aralıkları doku-bağımlı dielektrik aralıklarla eşleştirilmiş ve ardından parçalı lineer interpolasyon ve kübik spline interpolasyonu kullanılarak intensite değerleri dielektrik değerlere, diğer bir deyişle, dielektrik sabit ve iletkenlik değerlerine, dönüştürülmüştür. Burada, her iki interpolasyon sonucunda elde edilen sonuçlar karşılaştırılmakta ve elde edilen grafik ve şekiller gösterilmektedir. Sonuçta, parçalı lineer interpolasyon daha dağınık bir dielektrik değer dağılımı verirken kübik spline interpolasyonunun dokular arasında daha yüksek bir kontrast gösterdiği fark edilmiştir. Böylece, uygulanan her iki yöntemin de etkin sonuçlar verdiği görülmüş ve bu yöntemler aracılığıyla MRI görüntülerinden yararlanılarak gerçeğe yakın elektriksel göğüs modelleri elde edilebileceği kanısına varılmıştır.

1. INTRODUCTION

Cancer is a kind of disease causing body cells to change and grow uncontrollably. Tumors are lumps or masses formed by cancer cells and named after the originating body parts. Breast tumors frequently encountered, known as benign, are not cancerous. Contrarily to the invasive or infiltrating cancerous tumors, they don't grow and spread in an uncontrollable manner. Breast cancer emerges in lobules of the breast, which consists of milk producing glands, or in the ducts connecting the lobules to the nipple and spreads by passing through the duct or glandular walls and penetrating the surrounding breast tissues. The severity of invasive breast cancer is determined by the stage of the disease; that is, the extent or spread of the cancer at the time of the first diagnosis. Therefore, early diagnosis is of great import [1].
American Cancer Society has declared that approximately 40,170 women had died from breast cancer in 2009. Breast cancer is the second most widespread cancer causing women deaths after the lung cancer. Moreover, according to the statistics of 2009, 1,910 cases of breast cancer, which accounts for 1% of all breast cancers, have occurred among men [1]. The incidence of the cancer has risen 0.5% annually resulting in 1.35 to 1.45 million new cases casted by 2010[2].

To date, there is no accurate method for precluding breast cancer, because of which regular mammograms are very popular. The decline in the risk of death from breast cancer as a consequence of mammography has been exhibited by population based screening assessments and many random trials. However, mammography does have limitations [1]. Approximately 70% of the tumors detected via mammography are benign. From 4% to 34% of the breast cancers are not detected by mammography [3].

Moreover, mammography may lead to further tracking examinations such as biopsies as a result of false positive test results [1]. Additionally, although mammography is a very successful technique at detecting cancers on fatty breasts just as ultrasounds, it does not provide sufficient sensitivity on dense breasts where fibroglandular tissues can conceal or be assumed as cancer [4]. Therefore, annual MRI screening supplementary to mammography is recommended to women at high risk by an expert panel assembled by the American Society in 2007. Thus, MRI tests suspicious cases encountered by mammography [1]**.** MRI can detect breast cancers hidden in mammography. Nevertheless, it is not replacing traditional mammography as a screening tool because of the relatively high running cost and high percentage of unsuccessful predictions as an imaging tool. However, it is used in addition to mammography in severe cases just as ultrasound, which is utilized in discovering whether a lesion noticed on a mammogram is a liquid cyst or a tumor [5], [1].

However, none of the previously mentioned techniques can provide much reliable results in the diagnosis of small cancerous tumors. A reliable detection necessitates the determination of a significant and steady contrast between malignant and normal breast tissues. It has been observed that the electrical contrast between the normal and cancerous breast tissues, especially the contrast between the malignant and normal tissues is more essential than the density contrast imaged by X-Rays. As a result, small tumors can be detected sensitively and suspicious areas can be specified as malignant or benign by benefiting from this contrast [3]. This study, which is based on the electrical properties of the breast tissues, is a part of the breast cancer detection research.

The varying nature of the electrical properties of the breast tissues at microwave frequencies has inspired the engineers to make researches in breast cancer detection during the last decades. Apart from being sensitive and the specific in detecting small tumors, microwave breast cancer detection tests are superior to the other commonly used techniques. For instance, as ionizing radiation and breast compression are avoided, the process takes place in safer and more comfortable conditions [3].

Currently, several methods operating at various microwave frequencies are investigated to distinguish small malignant tumors from the normal breast tissues accurately. The methods used in these researches are: passive, hybrid and active methods. Passive methods integrate radiometers to determine the differences in the temperature measurements in the breast in order to detect tumors due to the remarkable increase in their temperature. Hybrid methods use microwave energy to heat tumors in a fast and selective manner and ultrasound transducers to discover pressure waves generated by the heated tissues’ expansion. Active methods are based on illuminating the breast with microwaves and measuring the transmitted or reflected microwave signals afterwards. One of the active microwave methods, microwave tomography, aims to reconstruct the whole dielectric breast profile by means of a forward and inverse scattering model. Nevertheless, it is restricted by resolution, the amount of theoretic information required, and the imaging technique’s substantial computational requirements. Another active microwave method, Ultra-Wideband (UWB) Radar imaging, makes use of the reflected UWB signals in the determination of the location of microwave scatterers within the breast. Instead of reconstructing the whole dielectric profile of the breast as done in microwave tomography method, UWB Radar imaging follows the Confocal Microwave Imaging (CMI) approach in identifying the locations of scattering regions within the breast. Regions with high energy in the final images are predicted to be indicative of the cancerous tissues [6].

However, authoritative information about the dielectric properties of normal and diseased breast tissues at microwave frequencies has been limited. For modeling the propagation and scattering of microwave signals in the breast, accurate numerical breast models are needed. Commonly used breast models in modeling the propagation of electromagnetic signals in biological tissues are FDTD breast models. These breast models are expected to preserve the geometrical characteristics of the breast, the heterogeneity and the dispersive properties of the breast tissues [2].

During the enhancement of numerical breast models, MRI data based models have been noticed to provide more realistic distributions of the breast tissues compared to the models that assign the dielectric values randomly throughout the breast tissues. In the MRI data based models; dielectric values of the regions are assigned either directly or indirectly by using the intensity levels of the images. Accordingly, MRI data have become an efficient tool in establishing the dispersive nature of the breast tissues [2]. Several FDTD-based models have used MRI data in the construction of numerical phantoms. For instance, Time- Reversal (TR) based FDTD methods practiced in the breast cancer detection make use of realistic breast models derived from magnetic resonance imaging (MRI) data [7].

In this study, 2D anatomically realistic breast models illustrating the heterogeneity and the dispersive nature of breast tissues were derived from MRI images and the varying dielectric properties across the breast tissues were assigned via piecewise linear and non-linear functions. Tissue-dependent dielectric properties of breast tissues were obtained from the intensity values of the filtered MRI images.

During this process, tissue-dependent dielectric intervals were assumed to be the intervals published in [8]. Gaussian distributions were constructed to discern the distinction between the two main breast tissue types, namely, fibroglandular tissues and fatty tissues. The MRI images acquired from Euromed Radiology Center and Maltepe University are used as samples in the application of the derivation methods presented. The consequent breast models are expected to be utile in developing and testing the microwave techniques being enhanced for the breast cancer detection.

1.1 Purpose of the Thesis

In this study, it is aimed to investigate the electrical properties’ variations in the breast tissues to exhibit the heterogeneity of the breast tissues and the dispersive nature of the dielectric properties across the breast. In line with this purpose, MRI based breast models, which will assist the research in breast cancer detection carried out at microwave frequencies, were constructed. For precise results, MRI images were firstly enhanced through several imaging processes and then MRI intensity values were transformed into uniform grids of dielectric constants and conductivity values by means of the MRI intensity histograms. Gaussian curves were drawn to perceive the diversity between the two main breast tissues. In the end, the purpose is fulfilled by using the mappings from intensity intervals to tissue dependent dielectric intervals.

1.2 MRI Based Researches in Breast Cancer

MRI analysis is based on the creation of body images displayed as thin horizontal slices of the breast tissues that can be studied at several angles by using magnetic and radio waves where ionizing radiation is avoided. During the MRI analysis of each breast, a great number of images are acquired to be assessed by a radiologist. Similar to mammography, which uses X-ray machines equipped for specifically imaging breasts, MRI uses special equipment with higher qualities than the images obtained from the equipments used for MRI scanning of head or chest to create breast images.

In the researches related to breast cancer detection, several techniques based on MRI have been developed. A novel and auspicious technique introduced is dynamic contrast enhanced magnetic resonance imaging (DCE-MRI). It is expected to be particularly suitable for screening dense breasts. DCE-MRI processes involve a contrast agent, generally Gadolinium DTPA, which is useful in the improvement of tissue discrimination. The reason for using this paramagnetic compound is its existence at intravascular and extracellular fluid space and its ability to increase the luminance of these Gadolinium-enhanced tissues remarkably, because of which vascular tissues such as lesions can be detected with ease. Within a typical DCE-MRI process, the breast is imaged both before and after the injection of a contrast agent to observe the effects of the contrast agent on the tissues [9].

In a recent study, a computer-aided diagnosis (CAD) system, which has enhanced the sensitivity, precision and rapidity of MRI studies, has been developed for detecting tumors. Besides the images taken before the injection, at every minute after the injection MR images are taken for six minutes with the T1 mode, the FLASH 3D traverse setting of the MRI machine. The principal task carried out by the contrast agent during this time duration is changing the relaxation times of tissues as a result of which the radiologists can determine the existence and site of a tumor or a lesion. Sets of images acquired are analyzed considering the three probable cases: malignant, benign and suspicious. Within this analysis, a lesion is identified as:

i. benign if the increase in the signal intensity is monotonic over the period after the injection

ii. suspicious, if the peak signal intensity is observed before three minutes and preserved for the rest of test

iii. malignant if there’s an immediate decrease in the signal intensity right after the peak is reached [5].

Another research related to MRI in breast cancer detection introduced a temporal feature analysis. In accordance with the assertions mentioned previously, malignant lesions are distinguished with their rapid contrast enhancement followed by fast washout while benign lesions are identified by their monotonic contrast enhancement. Thus, it has been claimed that the automatic detection of the potential lesions on MRIs can be determined by the examination of the kinetic curves. In this method, firstly, suspicious areas are detected depending on the contrast enhancement properties of the tissues. Then an angiogenesis map is formed after analyzing the kinetic curves. After that, pixels are displayed in different colors in accordance with the categories they belong [10].

Besides these MRI based techniques used in the detection of the breast cancer, MRI data also serves as an efficient tool in the development of realistic numerical breast models by allowing the exhibition of the heterogeneous and dispersive nature of the breast tissues. The experiments made to determine the dielectric properties of the breast tissues have helped scientists to conceive the varying nature of dielectric values across the breast. Nevertheless, the complexity of the dispersion of the tissues throughout the breast still poses a difficult problem.

1.3 Dielectrical Properties of Breast Tissues

In the experiments made by Chaudhary et al, the dielectric properties of malignant and normal surgically removed tissues were measured at frequencies between 3 MHz and 3 GHz and the ratio of the relative permittivity and the conductivity of malignant tissues to normal tissues were observed to be 5:1 and 4.7:1 respectively[11].

In another research carried out by Joines et al, the dielectric properties of tissues excised from breasts and some other organs are measured over a frequency range extending from 50 to 900 MHz and the results were supportive of those asserted by [11]. The greatest dielectric contrast was observed between the normal and malignant tissues of the mammary gland, where the average ratio of the relative permittivity and the conductivity of malignant tissues to normal tissues were 6.4:1 and 3.8:1 respectively [12].

According to the measurements of the dielectric constants of fatty tissues and tumors between 1 and 10 GHz made by our research group in Istanbul Technical University, the ratio of the relative permittivity of malignant tissues to fatty tissues seemed to vary between 7:1 and 10:1. Some of the experiment results are given in the Appendix.

In addition to these, Surowiec at al measured the relative permittivity of diffusing carcinoma and the surrounding tissue by an automatic network analyzer and an end-of-line capacitor sensor at frequencies between 20 kHz and 100 MHz and matched the results to Cole-Cole dielectric relaxation models. It was noticed that the center of the tumor and the diffusing margins had significantly higher permittivity compared to the peripheral tissues of the tumor. The high permittivity associated with the diffusing margins of the tumor was attributed to the proliferation and it was predicted that it would cause a large microwave scattering as a result of which small tumors could be detected and identified using UWB radar [13].

Campbell and Land have examined four types of tissues: fatty, normal, benign and malignant via a resonant cavity technique. They've put forward that it might be impossible to define tumors as benign or malignant on account of their dielectric similarities and also pointed out that there'd been a greater variance within the normal tissues then had been claimed [14].

An extensive analysis of the dielectric properties of normal and cancerous tissues has aimed to describe the dielectric properties of a great number of newly excised breast reduction, biopsy, lumpectomy and mastectomy tissues across a broad frequency range of 0.5- 20 GHz by firstly correlating the measured dielectric properties with pathological analysis of the tissue samples, then applying a statistical analysis to check the data integrity and finally adjusting the data to Cole-Cole representations. Within this research, each sample was assigned to one of the three adipose tissue groups due to the percentage of adipose it contains. These adipose tissue groups were; Group 1(0-30% adipose tissue), Group 2, (31-84% adipose tissue) and Group 3 (85-100% adipose tissue). Median permittivity and conductivity curves were drawn by calculating the matched values for every sample in the group at 50 equispaced frequency points, and then the median values at specific frequencies were computed for every group. After that, Cole-Cole equations were used to match the median values calculated. When the results were compared to the previously published data for normal tissue, the median dielectric curves for Group 3 adipose tissue were noticed to be lower and the median dielectric curves for Group 1 adipose tissue were noticed to be higher than expected. These differences are attributed to the large heterogeneity in normal breast tissues [15], [16].

In the other studies, the normal tissue samples had been abducted from regions away from the area with the tumor. As the tumors usually appear in glandular tissues, these normal samples unsurprisingly have a higher adipose content compared to the glandular tissue around the tumor. Therefore, the heterogeneity in the dielectric properties of breast tissues had been underrated.

In the consecutive studies defining the dielectric properties of a breast tissue as a function of the relevant tissue's adipose content, the permittivity difference is estimated as 8% at 5 GHz. Thus, realistic breast models are anticipated to provide more accurate comprehension of the dispersion of the breast tissues due to the variances in the electrical properties [2].

1.4 Breast Models Derived from MRI

Commonly used models in the microwave investigation techniques are the finite-difference time-domain (FDTD) computational electromagnetic breast models, which serve as beneficial tools in displaying the structural complexities, tissue heterogeneity, and varying dielectric properties of the breast. FDTD models of other body parts, specifically the ones composed of tissue types delineated well and featuring spatially constant dielectric properties, are derived from MRIs handily and are efficiently used. Following this approach, in several recent researches, numerical breast phantoms have been derived from MRIs [2].

Lie et al constructed a FDTD model of the breast using MRI images of the breast. Within the construction process, a high resolution breast MRI was taken while the patient was in the prone position and a low resolution MRI was taken while the patient was in the supine position. The low resolution MRI was used in horizontally expanding and vertically compressing the high resolution MRI so that both scans' shapes matched and the reformed MRI scan displayed the breast's fibroglandular and adipose tissues coherently.

In the MRI images used by Lie at al, the dark regions depicted the adipose tissue and the ligh regions depicted the fibroglandular tissue. In the constructed models, these tissue distributions were maintained by the assignment of dielectric values in accordance with the corresponding MRI pixels' intensities [17]. During the assignment of the dielectric properties, the variance is taken as the upper bound of breast tissue variability, i.e. 10% as assumed in [12], [6] and [11].

The modeling methodology Lie at al reinforced has been in use for several years [2].

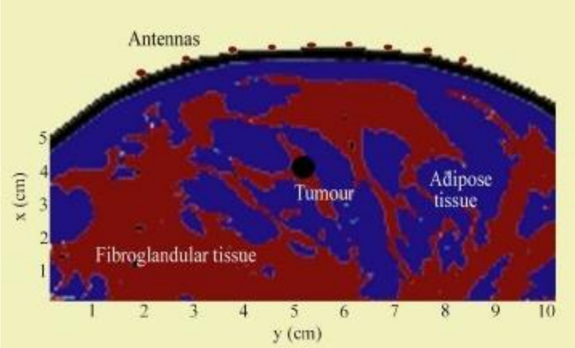


Figure 1.1: MRI based breast model where the tissues are apparent [2].

A recent research in the construction of the realistic breast models with electrical properties was carried out at the University of Wisconsin, Madison. Within this research, a collection of anatomically and dielectrically realistic 3-D numerical breast phantoms of different geometry and radiographic density are derived from 3-D MRIs of normal breast tissues whereas the frequency-dependent and tissue-dependent dielectric properties are deduced from the Wisconsin–Calgary study [15],[16]. Breast MRI voxel intensities were linearly mapped to dielectric properties of normal breast tissue accurately using a two-component Gaussian mixture model (GMM). As in Li et al.'s 2D FDTD model, the highly correlated nature of fibroglandular tissue distribution in the breast was conserved and the model was claimed to be more successful in representing the structural heterogeneity of normal breast tissue [2].

Nevertheless, accurate results are unlikely to be achieved in these derivations because of the complex distribution of glandular, adipose, and fibro-connective breast tissues and the remarkable heterogeneity of dielectric characteristics of normal breast tissues, but the presented numerical breast phantoms can serve as representative breast models to be used in developing new microwave techniques for breast cancer detection [2].

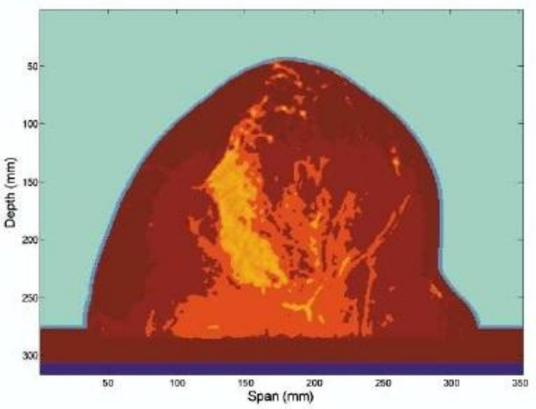


Figure 1.2: A cross section of a MRI-based breast model showing the tissue distribution [2].

2. IMAGE PROCESSING METHODS

2.1 Histograms

The histogram of a digital image illustrates the amount of pixels associated with each intensity level defined in the image. It is represented by the function $h(r_k) = n_k$

where r_k denotes the relevant intensity level and n_k denotes the number of pixels associated with r_k The domain of the intensity levels is $[0, G]$ where G is 255, 65535 and 1.0 for images of class uint8, uint16 and double respectively. Assuming that L intensity levels are defined in each of these intervals, r_1 would correspond to the intensity level 0 and r_L would correspond to the intensity level G. Additionally $G = L - 1$ holds for classses of uint8 and uint16.

Normalized histograms are obtained in a similar fashion, The normalized histogram given by $p(r_k) = \frac{h(r_k)}{n} = \frac{n_k}{n}$ is equivalent to the probability estimate of the intensity level r_k

Histograms are generated by the *imhist* function of Matlab as:

$h = imhist(f, b)$ where f denotes the input image and b denotes the number of bins, in other words, the subintervals of the intensity axis and normalized histograms can be attained as follows:

$p = \frac{imhist(f, b)}{numel(f)}$ where $numel(f)$ returns the number of elements i.e. pixels in f.

Furthermore, images can be translated into images with higher contrast and augmented dynamic range by means of histograms. This process is known as histogram equalization. The transformation function for a given image with the histogram $p_r(r_j)$ is: $s_k = T(r_k) = \sum_{j=1}^{k} p_r(r_j) = \frac{n_j}{n}$ which converts the intensity value r_j of the input image to the intensity value s_k of the output image, and corresponds to the cumulative density function of the intensity values.

As a result, the image is enhanced by expanding the intensity levels of the given image over a broader range. However, the result is not always successful.

At times, a processed image with a specified histogram can be practical to generate. The method used for this process is called histogram matching. In order to generate an image with a specified density, $p_z(z)$ a new variable, z denoting the intensity levels of the image under construction is introduced.

$$H(z) = \int_0^z p_z(w)dw = s \tag{2.1}$$

$$z = H^{-1}(z) = H^{-1}\left[T(r)\right] \tag{2.1a}$$

where $T(r)$ is the transformation function defined previously.

As H^{-1} exists for all valid $p_z(z)$ s and its components are non-zero, z intensity levels can be found handily [18].

histeq function of Matlab can be used for both histogram equalization and histogram matching processes.

2.2 Image Filtering

2.2.1 Homomorphic filter

During the production of the images, considerable shading effect occurs across the field of view due to the associations between the objects, the illumination and the imaging tools. For instance, the image’s brightness may be decreasing from the center of the image to the edges of the field of view or from one side of the image to its other side. Eliminating the effect of this shading, which might be caused by the nonuniformity of the illumination or imaging tools, is often necessary for consecutive image processing schemes.

The following model can illustrate the shading effect:

$b(x,y) = I_{ill}(x,y) \cdot a(x,y)$ where $a(x,y)$ stands for the object, $b(x,y)$ stands for the image and I_{ill} stands for illumination.

For low concentrations,

$$a(x,y) = gain(x,y) \cdot b(x,y) + offset(x,y) \qquad \textbf{(2.2)}$$

where $offset(x,y)$ and $gain(x,y)$ are the contributions of the imaging tool.

The resultant shading $= gain(x,y) \cdot I_{ill}(x,y) \cdot a(x,y) + offset(x,y)$ **(2.3)**

$I_{ill}(x,y)$ is ordinarily presumed to be changing moderately relative to $a(x,y)$

One of the two conditions where the intention is the estimation of shading terms $gain(x,y) \cdot I_{ill}(x,y)$ and $offset(x,y)$ needed to determine $a(x,y)$ from $c(x,y)$ is a posteriori estimation. Under these circumstances, the only data available is the recorded image and the aim is to remove the shading estimate from $c(x,y)$. Low-pass filtering, homomorphic filtering and morphological filtering are the methods used for this purpose.

In homomorphic filtering, considering that $offset(x,y) = 0$ and $gain(x,y) \cdot I_{ill}(x,y)$ varies slowly compared to $a(x,y)$, the logarithm of $c(x,y)$ is taken to produce low frequency and high frequency terms. Then, the shading is restrained by the high pass filter applied to the logarithm of $c(x,y)$. Finally, the image is retrieved by taking the exponent of the high-pass filtered results. Consequently, the brightness over an image is normalized, multiplicative noise is removed and the contrast is increased [19]. Here is the description of this filtering process developed by Oppenheim and Stockham [19]:

Step-1: $c(x,y) = gain(x,y) \cdot I_{ill}(x,y)$ **(2.4a)**

Step-2: $\ln[c(x,y)] = \ln[gain(x,y) \cdot I_{ill}(x,y)] + \ln a(x,y)]$ **(2.4b)**

Here the first term varies slowly whereas the second term varies fast.

Step-3: $HighPass\{\ln[c(x,y)]\} \approx \ln[a(x,y)]$ **(2.4c)**

Step-4: $a^{*}(x,y) = e^{HighPass\{\ln[c(x,y)]\}}$ **(2.4d)**

2.2.2 Mean filter

Mean filters smooth regional variations in an image and reduces the noise on account of blurring. There are five types of mean filters: arithmetic, geometric, harmonic and contraharmonic.

Arithmetic mean filter helps calculating the average values of the original image $g(x,y)$ in the region defined by S_{xy} which denotes the coordinate array of a rectangular subimage window of size $m \times n$. In the aftermath, the values of the retrieved image at (x,y) corresponds to the arithmetic means of the pixel values of the areas determined by S.

The mathematical representation for the filtered image is:

$$f(x,y) = \frac{1}{mn} \sum_{(s,t)\in S_{xy}} g(s,t) \tag{2.5}$$

Geometric mean filter computes the products of the pixels in the subimage windows, raised to the power $\frac{1}{mn}$ as the retrieved values for all the pixels. Geometric mean filters are not more successful than arithmetic filters in smoothing, however they preserve more details.

The mathematical representation for the filtered image is:

$$f(x,y) = \left[\prod_{(s,t)\in S_{xy}} g(s,t) \right]^{\frac{1}{mn}} \tag{2.6}$$

Similarly, the mathematical expressions for Harmonic and Contraharmonic mean filters, which are commonly used in salt-pepper noise elimination, are as shown respectively as follows:

$$f(x,y)=\frac{mn}{\sum_{(s,t)\in S_{xy}}\frac{1}{g(s,t)}} \tag{2.7}$$

$$f(x,y)=\frac{\sum_{(s,t)\in S_{xy}} g(s,t)^{Q+1}}{\sum_{(s,t)\in S_{xy}} g(s,t)^{Q}} \tag{2.8}$$

In the above equations, Q denotes the order of filter [20].

2.2.3 Laplace filter

Regional equalization of image contrast mentioned in section 2.1 yields increment in regional contrast at boundaries as a result of which the edges become more vivid and the image sharpens. However, there're other methods of edge enhancement which are more insensitive to overall brightness levels, noise, and the type or scale of detail present than the equalization.

The conventional $3x3$ Laplacian filter is:

$$\begin{pmatrix} -1 & -1 & -1 \\ -1 & +8 & -1 \\ -1 & -1 & -1 \end{pmatrix}$$

This filter nullifies the gray level in a uniformly bright region or a region with a uniform brightness gradient of an image by deducting the brightness values of the pixels in the neighborhood from eight times the pixel in the center. In the emergence of a discontinuity in the neighborhood such as a point or a line, the Laplacian gives either a positive or a negative value due to the central point's position relative to edge. While displaying the result, commonly an average gray value, which is 128 for an image where the gray values are depicted in the range [0, 255], in order to make the zero points appear middle gray, and the brighter and darker values generated by the Laplacian visible is accepted. Another approach is drawing the absolute value of the result; however, this creates double lines causing clutter in both the human vision and the consecutive processes along edges.

Additionally, as the name of the filter hints, the Laplacian filter is based on estimating the second derivative of brightness B , which does not vary due to rotation, and therefore not susceptible to the direction of discontinuity. Consequently, the points, lines, and edges are exposed, and the uniform and smoothly varying regions are restrained. By varying the weights of the kernel, the image can be sharpened. Here is a kernel that can be defined as a sharpening operator increasing the image contrast at edges.

$$\begin{pmatrix} -1 & -1 & -1 \\ -1 & +9 & -1 \\ -1 & -1 & -1 \end{pmatrix}$$

A clear image B can be retrieved by applying the Laplacian filter to the blurred image. Laplacian filter can be interpreted as a high pass filter, which authorizes the high frequencies a low pass filter would discharge or suppress in order to prevent the variations in brightness of the neighboring pixels to penetrate the filter and dismisses the low frequencies implying the gradual overall variation in brightness.

There are various kernels that can be used with the Laplacian operator. For instance, an elementary kernel is:

$$\begin{pmatrix} 0 & -1 & 0 \\ -1 & +4 & -1 \\ 0 & -1 & 0 \end{pmatrix}$$

The drawback of using simple kernels is the noise caused by the casual differences between adjacent pixels; even the homogeneous parts of an image is highlighted with this approach [21].

2.2.4 Edge detection filters

As well as smoothing, the capability of taking spatial derivatives of an image is a basic operation of image processing. However, digitized images are discrete functions $a(x, y)$ of the integer spatial coordinates instead of being continuous functions of the spatial variables; therefore, real spatial derivatives can only be approximated in the case of image processing.

In addition to this, derivation causes the high frequency noise to be highlighted in the final image. This issue is handled with the combination of the derivative operation with an operation covering up the high frequency noise, thus the smoothing operation is combined with the derivative operation required.

As the image values vary in multiple directions, the directional derivatives needs being considered. Thus, a random angle derivative matrix (filter) is identified as follows:

$$[h_\theta] = \cos\theta \cdot [h_x] + \sin\theta \cdot [h_y] \tag{2.9}$$

where h_x denotes a horizontal derivative matrix and h_y denotes a vertical derivative matrix.

Vector derivatives can also be written by using the gradient operator as:

$$\nabla a = \frac{\partial a}{\partial x} i_x^{\rightarrow} + \frac{\partial a}{\partial y} i_y^{\rightarrow} = (h_x \times a) i_x^{\rightarrow} + (h_y \times a) i_y^{\rightarrow} \tag{2.10}$$

where $i_x^{\rightarrow}$ and $i_y^{\rightarrow}$ are horizontal and vertical unit vectors respectively.

Furthermore, gradient direction and gradient magnitude are as follows:

$$\psi(\nabla a) = \arctan\left[\frac{(h_y \times a)}{(h_x \times a)}\right] \tag{2.11}$$

$$\|\nabla a\| = \sqrt{(h_x \times a)^2 + (h_y \times a)^2} \tag{2.12}$$

As it can be seen from above, the results of these computations depends on the h_y and h_x chosen. Gradient filters differ in the selection of these matrices. h_x and h_y matrices used by basic derivative filters are:

$$[h_x] = [h_y]^t = [1, -1] \tag{2.13}$$

$$[h_x] = \left[h_y\right]^t = [1, 0, -1] \tag{2.14}$$

Within **Sobel gradient filtering**, the filters are dissociable, each deriving in one direction by using (2.13) and smoothes in the orthogonal direction. The derivative filters used by Sobel gradient filter are:

$$[h_x] = \frac{1}{4}\begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} = \frac{1}{4}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \bullet \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}^T \tag{2.15}$$

$$\left[h_y\right] = \frac{1}{4}\begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} = \frac{1}{4}\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \bullet \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}^T \tag{2.16}$$

This filter is one of the most commonly used techniques, even though it requires a modest amount of computation to perform correctly.

Similar to Sobel gradient filters, **Prewitt gradient filters** are dissociable and each derives in one direction by utilizing (2.14). However, they use a uniform filter while smoothing in the orthogonal direction [18].

The derivative filters used by Prewitt gradient filter are:

$$[h_x] = \frac{1}{3}\begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{pmatrix} = \frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \bullet \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}^T \tag{2.17}$$

$$\left[h_y\right] = \frac{1}{3}\begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{pmatrix} = \frac{1}{3}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \bullet \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}^T \tag{2.18}$$

3. MATHEMATICAL METHODS

3.1 Gaussian Distribution

Gaussian distribution function can be used when the number of events or features are extremely high. It is a continuous function approximating the accurate binomial distribution of events or features.

The probability density functions of the Gaussian distribution with mean μ and standard deviation σ is:

$$p(x;\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}e^{(\frac{-(x-\mu)^2}{2\sigma^2})} \qquad \textbf{(3.1)}$$

The cumulative distribution function for the Gaussian distribution with mean μ and standard deviation σ is:

$$\phi(z;\mu,\sigma)=\int_{-\infty}^{z} p(x;\mu,\sigma)dx \qquad \textbf{(3.2)}$$

The probability density function for the standard Gaussian distribution where $\mu = 0$ and $\sigma = 1$ is:

$$p(x)=\frac{1}{\sqrt{2\pi}}e^{\left(\frac{-x^2}{2}\right)} \qquad \textbf{(3.3)}$$

The cumulative distribution function for the standard Gaussian distribution is:

$$\phi(z)=\int_{-\infty}^{z} p(x)dx \qquad \textbf{(3.4)}$$

In general, large numbers of independent random variables (approximately) has a normal distribution.

3.2 Gaussian Mixture Models

A Gaussian Mixture Model (GMM) is a probability density function represented by the weighted sum of N component Gaussian densities as follows:

$$p(x \mid \lambda) = \sum_{i=1}^{N} w_i g(x \mid \mu_i, \Sigma_i) \tag{3.5}$$

where x is a K dimensional data vector(standing for measurements or features), w_i, $i = 1, 2, \ldots N$, are the mixture weights and $g(x \mid \mu_i, \Sigma_i)$ for $i = 1, 2, \ldots N$ are the Gaussian densities of mixture components. For the i$^{\text{th}}$ component, the Gaussian density is of the form:

$$g(x \mid \mu_i, \Sigma_i) = \frac{1}{(2\pi)^{D/2} \mid \Sigma_i \mid^{1/2}} e^{\{-\frac{1}{2}(x-\mu_i)' \Sigma_i^{-1} (x-\mu_i)\}} \tag{3.6}$$

where μ_i is the mean vector and Σ_i is the covariance matrix. The mixture weights satisfy:

$$\sum_{i=1}^{N} w_i = 1 \tag{3.7}$$

The parameters of the Gaussian mixture model are denoted by: $\lambda = \{w_i, \mu_i, \Sigma_i\}$ for $i = 1, 2, \ldots N$

Several different configurations are thinkable for this model. For instance, the covariance matrices Σ_i can be either full rank or restrictedly diagonal, parameters can be mutual, or tied, among the Gaussian components. The amount of data and the usage of the GMM are the decisive factors in the determination of a certain model.

Even if the full covariance matrices are not used, the correlations between statistically dependent features can be illustrated as the linear combination of Gaussians with diagonal covariance basis, as using a set of M Gaussians with full covariance matrix is equivalent to using a larger set of Gaussians with diagonal covariance.

Besides the smooth overall distribution fit provided, the multi-modality of the density is distinctly given by the parameters of GMM. An example for this multi-modality attribution is shown in Figure 3.1

The most well-known and substantial method used in the estimation of GMM parameters is Maximum Likelihood (ML) estimation.

With training vectors $X = \{x_1, x_2, ..., x_T\}$ and a GMM configuration given, estimating the parameters of the GMM is achieved by maximizing the likelihood of the given GMM training data. Considering that the vectors are linearly independent, the GMM likelihood can be shown as:

$$p(X \mid \lambda) = \prod_{t=1}^{T} p(x_t, \lambda) \tag{3.8}$$

As the likelihood is not a linear function of the GMM parameters, direct maximization is impossible. Nevertheless, parameters can be estimated iteratively by utilizing a special case of the EM algorithm. At every iteration, the likelihood is ensured to increase monotonically and a new model is assigned to be the initial model until a convergence to a local maximum of the likelihood function is achieved. Here are the modeling parameters for each GMM component at each iterative step assuring a monotonic augmentation in the model's likelihood [23]:

$$w_i = \frac{1}{T} \sum_{t=1}^{T} \Pr(i \mid x_t, \lambda) \tag{3.9}$$

$$\mu_i = \frac{\sum_{t=1}^{T} \Pr(i \mid x_t, \lambda) \cdot x_t}{\sum_{t=1}^{T} \Pr(i \mid x_t, \lambda)} \tag{3.10}$$

$$\sigma_i^2 = \frac{\sum_{t=1}^{T} \Pr(i \mid x_t, \lambda) \cdot x_t^t}{\sum_{t=1}^{T} \Pr(i \mid x_t, \lambda)} - \mu_i^2 \tag{3.11}$$

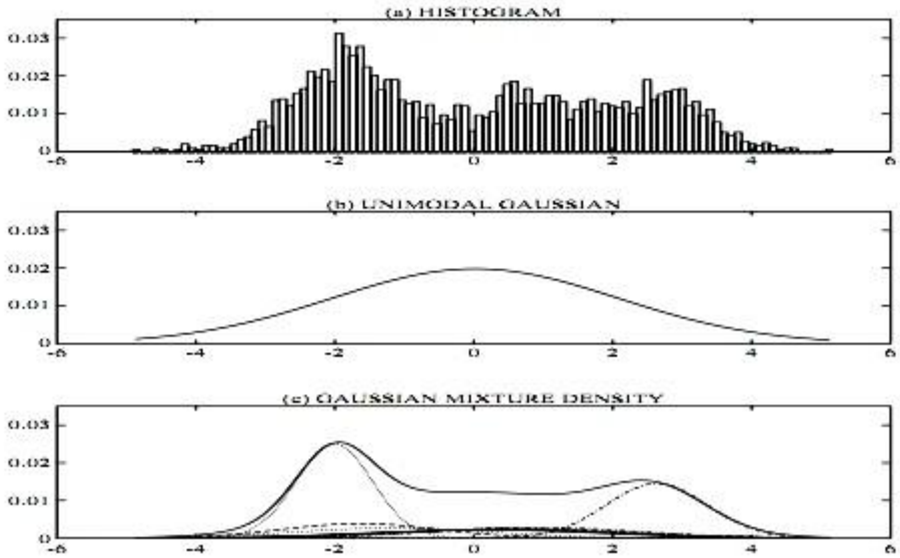


Figure 3.1: (a) Histogram (b)Unimodal Gaussian (c) Gaussian Mixture Density

3.3 Piecewise Cubic Spline Interpolation

Splines can be interpreted as thin flexible strips attached with pins, which stand for the knots of the spline in the mathematical definition, from mechanical point of view. As, in the mechanical model, the slope and the bending moment is continuous at the pins, the first and the second derivatives are continuous at the knots, namely, the data points. Similarly, as the bending moment vanishes at the pins at edges, the second derivatives also vanish at the two end data points. Because of these similarities between the spline curves and their corresponding mechanical models, these curves are called natural cubic splines [24].

While fitting a straight line to each interval between two successive points in a given data set, (x_i, f_i) and (x_{i+1}, f_{i+1}), only the function values at the data points can be preserved on each side of the interval. Additionally, higher order polynomials are not very effective in interpolation, too. The major drawback of polynomial interpolation is in the subintervals, especially between the first two and last two data points; the function varies extremely, overshooting the changes in the original data. Therefore, full-degree polynomial interpolation does not have a widespread use in fitting curves and data; instead, it is mainly used in the enhancement of other numerical techniques. Furthermore, polynomials of high degree are prone to oscillation dreadfully between the data points whereas smoother results are likely to be reached by using a cubic interpolant satisfying the continuity constraints in derivatives.

Each partition of a cubic spline curve is a cubic polynomial. Nevertheless, splines are less likely to oscillate between data points than the polynomials. Therefore, cubic splines are the most commonly used curves for this purpose [24], [25].

For a cubic spline spanning n knots, each of the $n-1$ cubics is in the following form:

$$f_i(x) = a_i + (x - x_i) b_i + (x - x_i)^2 c_i + (x - x_i)^3 d_i \tag{3.12}$$

Thus, $4n-4$ unknowns are needed to be found.

Using the derivative constraints and the nullity of curvature of natural splines at the two end points, the system becomes determined with $4(n-1)$ equations.

For $i = 1, 2, ... n-2$

$$f_i(x_i) = y_i \tag{3.13}$$

$$f_i(x_{i+1}) = y_{i+1} \quad \text{for } i = 1, 2, ... n-1 \tag{3.14}$$

$$f_i^{'}(x_{i+1}) = f_{i+1}^{'}(x_{i+1}) \tag{3.15}$$

$$f_i^{''}(x_{i+1}) = f_{i+1}^{''}(x_{i+1}) \tag{3.16}$$

$$f_1^{''}(x_1) = f_{n-1}^{''}(x_n) = 0 \tag{3.17}$$

On either side of a knot, although the second derivatives have different formulas, because of the continuity, they have the same value at the knots, i.e. $f_{i-1,i}^{''}(x_i) = f_{i,i+1}^{''}(x_i) = k_i$ where $k_1 = k_n = 0$ and k_i is unknown for $i = 1, 2, ... n-1$

In order to write the second derivative explicitly, Lagrange two-point interpolation can be used as follows:

$$f''_{i,i+1}(x) = k_i l_i(x) + k_{i+1} l_{i+1}(x) \text{ where } l_i(x) = \frac{x - x_{i+1}}{x_i - x_{i+1}} \text{ and } l_{i+1}(x) = \frac{x - x_i}{x_{i+1} - x_i} \qquad \textbf{(3.18)}$$

$$f''_{i,i+1}(x) = \frac{k_i(x - x_{i+1}) - k_{i+1}(x - x_i)}{x_i - x_{i+1}} \qquad \textbf{(3.19)}$$

By integrating with respect to x,

$$f_{i,i+1}(x) = \frac{k_i(x - x_{i+1})^3 - k_{i+1}(x - x_i)^3}{6(x_i - x_{i+1})} + A(x_i - x_{i+1}) - B(x - x_i) \qquad \textbf{(3.20)}$$

$$A = \frac{y_i}{x_i - x_{i+1}} - \frac{k_i}{6}(x_i - x_{i+1})$$

$$B = \frac{y_{i+1}}{x_i - x_{i+1}} - \frac{k_{i+1}}{6}(x_i - x_{i+1})$$

Considering the condition:

$$\begin{aligned} f_{i,i+1}(x) &= \frac{k_i}{6}\left[\frac{(x - x_{i+1})^3}{x_i - x_{i+1}} - (x - x_{i+1})(x_i - x_{i+1})\right] \\ &- \frac{k_{i+1}}{6}\left[\frac{(x - x_i)^3}{x_i - x_{i+1}} - (x - x_i)(x_i - x_{i+1})\right] \\ &+ \frac{y_i(x - x_{i+1}) - y_{i+1}(x - x_i)}{x_i - x_{i+1}} \end{aligned} \qquad \textbf{(3.21)}$$

In order to find the second derivative values at the interior knots, the first derivative constraint is used:

$$
\begin{aligned}
&f'_{i-1,i}(x_i) = f'_{i,i+1}(x_i) \\
&\frac{k_i}{6}\left[\frac{3(x_i - x_{i+1})^2}{x_i - x_{i+1}} - (x_i - x_{i+1})\right] - \frac{k_{i+1}}{6}\left[\frac{3(x_i - x_i)^2}{x_i - x_{i+1}} - (x_i - x_{i+1})\right] + \frac{y_i - y_{i+1}}{x_i - x_{i+1}} \\
&= \frac{k_{i-1}}{6}\left[\frac{3(x_i - x_i)^2}{x_{i-1} - x_i} - (x_{i-1} - x_i)\right] - \frac{k_i}{6}\left[\frac{3(x_i - x_{i-1})^2}{x_{i-1} - x_i} - (x_{i-1} - x_i)\right] + \frac{y_{i-1} - y_i}{x_{i-1} - x_i}
\end{aligned}
\tag{3.22}
$$

$$
\begin{aligned}
&\frac{k_i}{6}\left[3(x_i - x_{i+1}) - (x_i - x_{i+1})\right] + \frac{k_{i+1}}{6}(x_i - x_{i+1}) + \frac{y_i - y_{i+1}}{x_i - x_{i+1}} = \frac{k_{i-1}}{6}(x_i - x_{i-1}) \\
&+ \frac{k_i}{6}\left[3(x_{i-1} - x_i) - (x_{i-1} - x_i)\right] + \frac{y_{i-1} - y_i}{x_{i-1} - x_i}
\end{aligned}
\tag{3.22a}
$$

$$
\begin{aligned}
&k_i\left[2\left(x_i - x_{i+1}\right)\right] + k_{i+1}\left(x_i - x_{i+1}\right) + 6\frac{y_i - y_{i+1}}{x_i - x_{i+1}} = k_{i-1}\left(x_i - x_{i-1}\right) - k_i\left[2\left(x_{i-1} - x_i\right)\right] + 6\frac{y_{i-1} - y_i}{x_{i-1} - x_i} \\
&2k_i\left(x_{i-1} - x_{i+1}\right) + k_{i+1}\left(x_i - x_{i+1}\right) - k_{i-1}\left(x_i - x_{i-1}\right) = 6\left[\frac{y_{i-1} - y_i}{x_{i-1} - x_i} - \frac{y_i - y_{i+1}}{x_i - x_{i+1}}\right] \\
&k_{i-1}h + 2k_i 2h + k_{i+1}h = 6\left[\frac{y_{i-1} - y_i}{h} - \frac{y_i - y_{i+1}}{h}\right] \\
&k_{i-1} + 4k_i + 4k_{i+1} = \frac{6}{h^2}(y_{i-1} - y_i - y_{i+1})
\end{aligned}
\tag{3.22b}
$$

Here there are $n-2$ equations for n knowns. “Not-a-knot “strategy is commonlyused around the ends of the interval to make the system determined. This strategy yields the reduction of the number of unknowns by using single cubics on the first two and last two subintervals [26].

Splines can be fit to a given data set by using the *spline* function written by Carl De Boor in MATLAB.

4. MODELLING

4.1 Methodology

In the first step of this study, the MRI breast images provided from Euromed Radiology Centre and Maltepe University were processed morphologically in order to reduce imaging artifacts. Then, in the second step, the Gaussian distribution of the two main breast tissues, namely, fibroglandular and fatty tissues, were constructed to ascertain the distinctive intensity values for these tissues. After that, the piecewise mapping parameters: the minimum, maximum, intermediate and mean intensity values of the two main tissue types were computed. Additionally, the intensity intervals determined by the minimum, maximum, intermediate and mean intensity values were matched with seven distinct tissue types. Finally, the intensity values were mapped to the dielectric constants and conductivity values via piecewise linear mapping and cubic spline interpolation.

4.2 MRI Data Manipulation

At this stage, edge detection methods and smoothing operations are applied to the magnetic resonance images provided from the radiology center and the hospital stated previously.

Firstly, the details of images are observed by using the Sobel or Prewitt gradient filters and then via homomorphic filtering the low frequency spatial variations are removed in order to suppress the slowly varying intensity gradients, which occur especially in the tissues besides the coils and are caused by the nonuniform magnetic fields. Figure 4.4, Figure 4.5 and Figure 4.6 illustrates the filtered versions of the MRI images, “z”, ”e” and ”m” given in Figure 4.1., Figure 4.2 and Figure 4.3

Then, the image is masked using a binary mask in order to acquire more accuracy after the application of several morphological operations. The images shown in Figure 4.7, Figure 4.8 and Figure 4.9 are the resulting images of these imaging processes.

Apart from these, Figure 4.10-4.12 and Figure 4.13-4.15 respectively show the results of the Sobel and Prewitt edge detection filters applied to the MRI images and Figures from 4.16 to 4.18 show the results of the Laplace filter application.

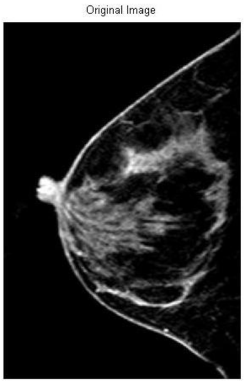


Figure 4.1: MRI image "z"

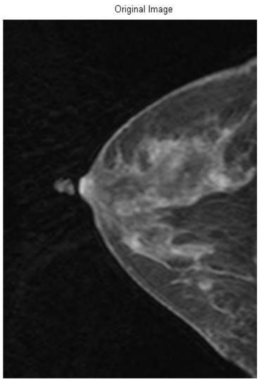


Figure 4.2: MRI image "e"

Figure 4.3:MRI image "m"

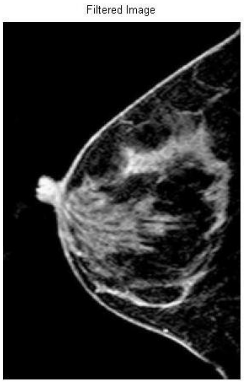


Figure 4.4: Fıltered Image "z"

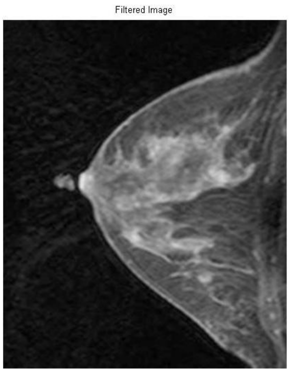


Figure 4.5: Filtered Image "e"

Figure 4.6: Filtered Image "m"

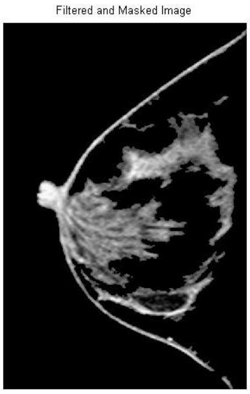


Figure 4.7: Filtered & Masked Image “z”

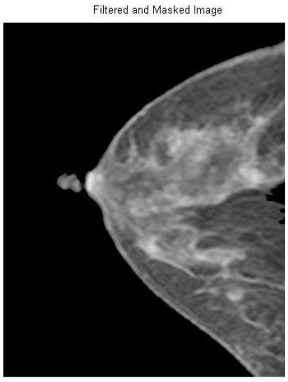


Figure 4.8: Filtered & Masked Image“e”

Figure 4.9: Filtered & Masked Image “m”

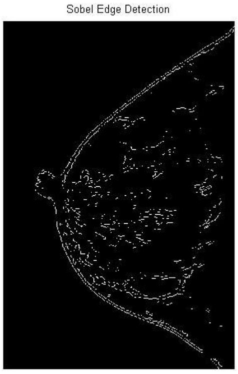


Figure 4.10: Sobel Edge Detection Filter applied to “z”

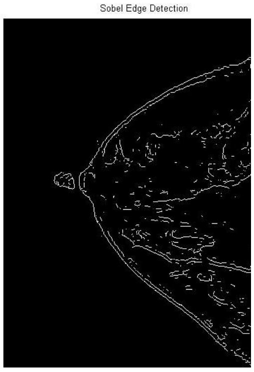


Figure 4.11:Sobel Edge Detection Filter applied to “e”

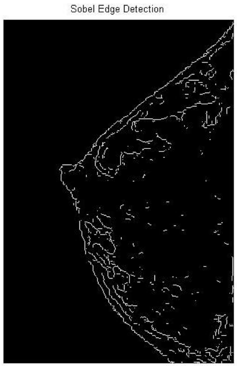


Figure 4.12: Sobel Edge Detection Filter applied to “m”

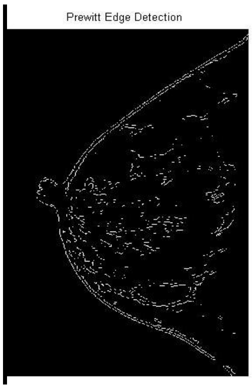


Figure 4.13: Prewitt Edge Detection Filter applied to “z”

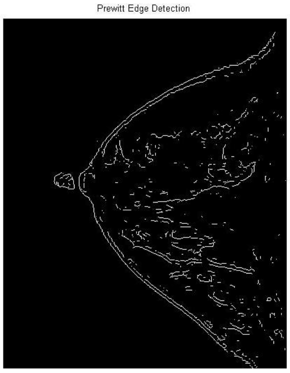


Figure 4.14: Prewitt Edge Detection Filter applied to” e”

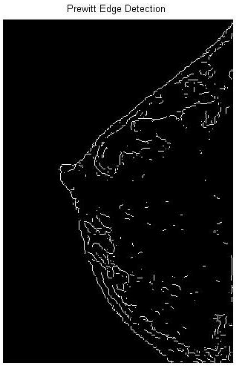


Figure 4.15: Prewitt Edge Detection Filter applied to “m”

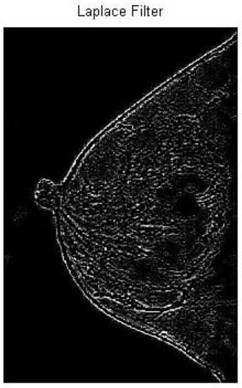


Figure 4.16: Laplace Filter applied to “z”

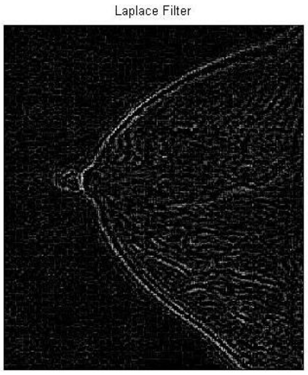


Figure 4.17: Laplace Filter applied to “e”

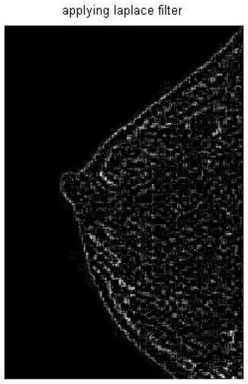


Figure 4.18: Laplace Filter applied to “m”

4.3 The Construction of Breast models

At this stage, the minimum, maximum, intermediate and mean intensity values of the two main tissues are computed by using the histograms of the images and the intensity intervals are mapped to the dielectric intervals via both piecewise linear and cubic spline interpolation.

Firstly, the tissues are dissociated depending on the expected diversity between their intensity levels by utilizing the threshold values determined in accordance with their Gaussian distributions.

Then, the minimum, maximum, intermediate and mean intensity values of the two distinct tissues, namely, fibroglandular tissue and fatty tissue, are derived from the Gaussian distribution model.

After that, the intensity intervals are matched with the appropriate tissue classes and the intensity values of each interval are mapped to the dielectric constant and conductivity values of the relevant tissue stated by [8] firstly by linear interpolation and then by cubic spline interpolation.

Furthermore, the dielectric intervals are also determined by using 10% variances within the same tissue groups and 15% variance between the distinct tissue groups (fibroglandular and fatty tissues) to illustrate an alternative approach for modeling. Within this approach, the dielectric and conductivity values stated by Converse et al are assumed valid [28] instead of the dielectric intervals determined by [8].

4.3.1 Piecewise linear interpolation

At this stage, the intensity values in the intervals determined depending on the Gaussian distributions of the fibroglandular and fatty tissues similarly to done in [8], are linearly mapped to the corresponding tissues' dielectric values.

Here are the eight mapping parameters taken as the edges of the seven intensity intervals, which correspond to the seven designated breast tissues: fibroglandular-1, fibroglandular-2, fibroglandular-3, transitional, fatty-1, fatty-2 and fatty-3, for MRI images where the fatty regions have low and glandular regions have high intensity values:

$$m_f = \inf(x \mid x \in B) \tag{4.1}$$

$$m_g = \mu_g - \sigma_g \tag{4.2}$$

$$m_{\sigma g} = \mu_g + \sigma_g \tag{4.3}$$

$$M_f = \mu_f + \sigma_f \tag{4.4}$$

$$m_{\sigma f} = \mu_f - \sigma_f \tag{4.5}$$

$$M_g = \sup(x \mid x \in B) \tag{4.6}$$

where B denotes the interior region of the breast.

Expected values and variances are firstly computed by means of the basic methods separately:

$$\mu_g = \sum\nolimits_{x \in B_g} x \cdot p_g(x) \tag{4.7}$$

$$\mu_f = \sum\nolimits_{x \in B_f} x \cdot p_f(x) \tag{4.8}$$

$$\sigma_g{}^2 = \sum\nolimits_{x \in B_g} (x - \mu_g)^2 \cdot p_g(x) \tag{4.9}$$

$$\sigma_f{}^2 = \sum\nolimits_{x \in B_f} (x - \mu_f)^2 \cdot p_f(x) \tag{4.10}$$

where $p_g(x)$ and $p_f(x)$ denote the probabilities of glandular and fatty tissues respectively.

Secondly, these parameters are obtained from the two-component GMM, which utilizes the EM algorithm to estimate the mixture parameters, by using the *gmdistribution* function of Matlab. The rest of the parameters are derived from the Gaussian distribution as before.

The histograms and Gaussian distributions of the glandular and fatty tissues of each breast sample are displayed in Figures 4.19, Figure 4.20 and Figure 4.21.

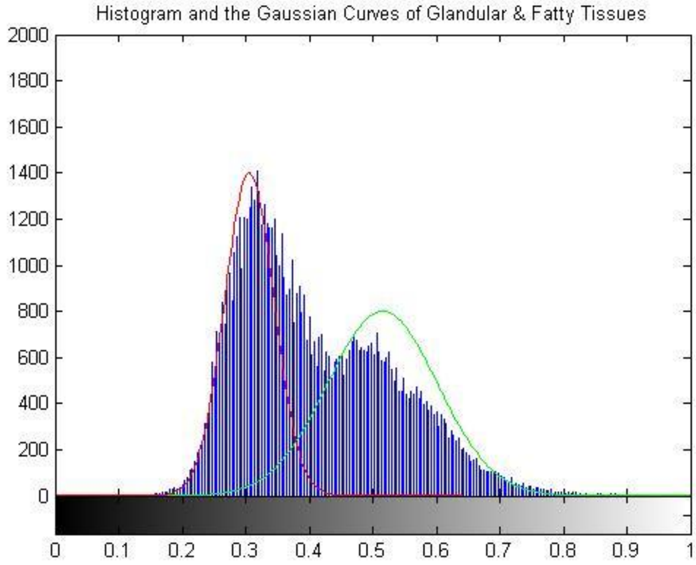


Figure 4.19: Histogram of MRI image “e” is shown by the blue curves and the glandular and fatty tissues’ Gaussian curves are denoted by green and red curves respectively.

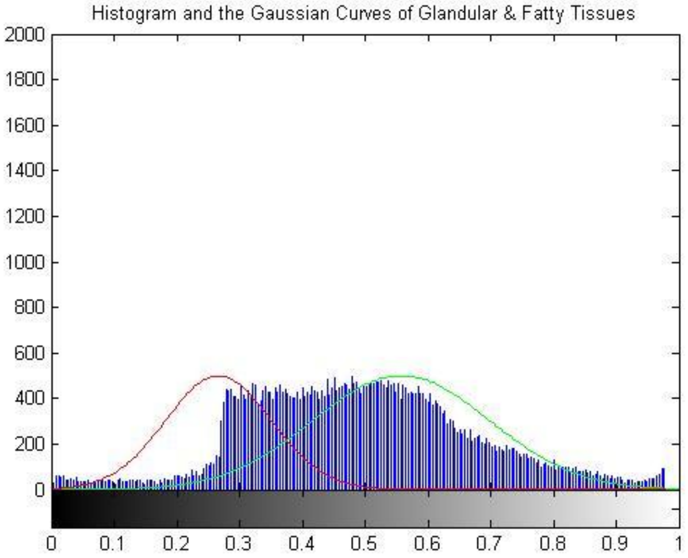


Figure 4.20: Histogram of MRI image “z” is shown by the blue curves and the glandular and fatty tissues’ Gaussian curves are denoted by green and red curves respectively.

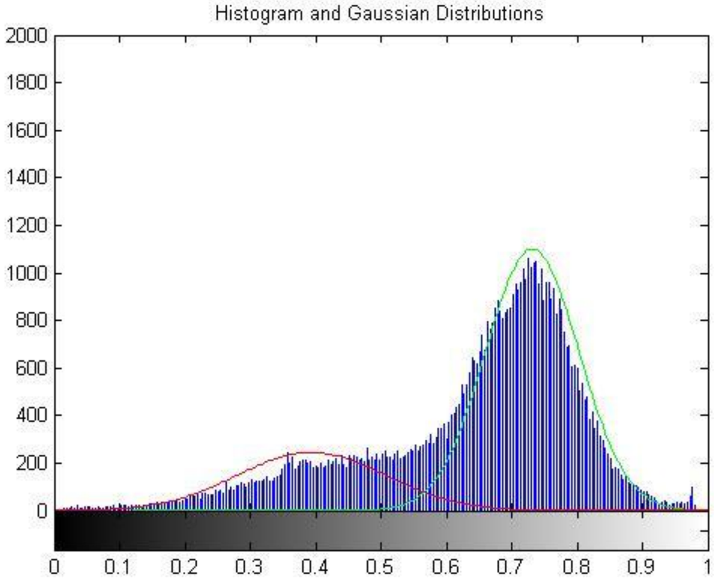


Figure 4.21: Histogram of MRI image” m” is shown by the blue curves and the glandular and fatty tissues’ Gaussian curves are denoted by green and red curves respectively.

The resulting models obtained by applying this mapping to the previously shown MRI images are illustrated in the figures below. Figures 4.22, Figure 4.23, Figure 4.26, Figure 4.27, Figures 4.30 and Figures 4.31 show the resulting models obtained by using the the dielectric intervals in accordance with [8]. Figure 4.24, Figures 4.25, Figures 4.28, Figures 4.29, Figures 4.32 and Figures 4.33 show the resulting models obtained by using the dielectric intervals computed by applying 10% and 15% variances to the values stated by [28].

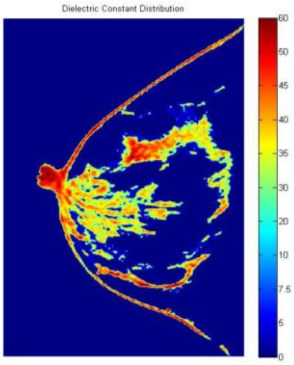


Figure 4.22: Dielectric distribution for MRI image “z"

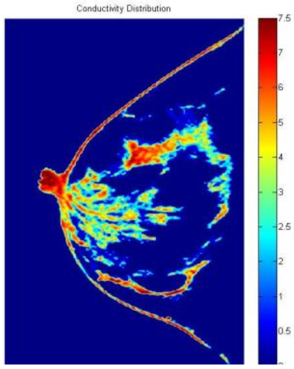


Figure 4.23: Conductivity distribution for MRI image “z"

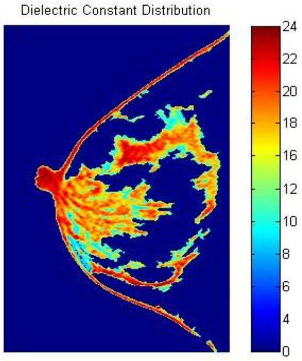


Figure 4.24: Dielectric distribution for MRI image “z"

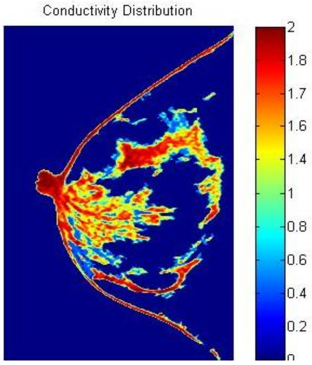


Figure 4.25: Conductivity distribution for MRI image “z"

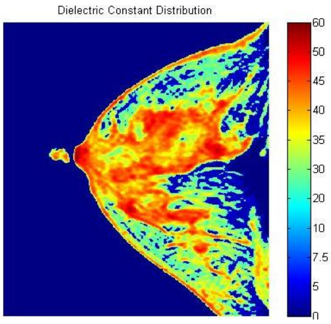


Figure 4.26: Dielectric distribution for image "e"

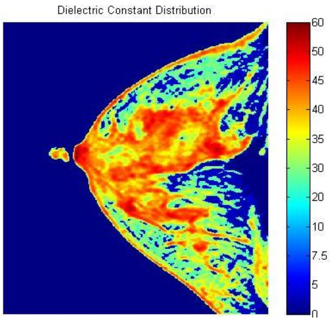


Figure 4.27: Conductivity distribution for MRI image “e”

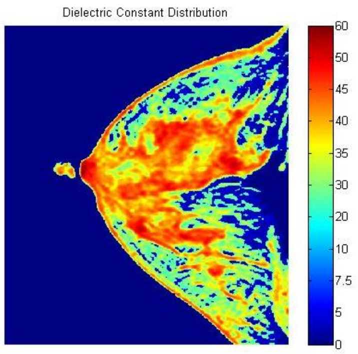


Figure 4.28: Dielectric distribution for MRI image “e"

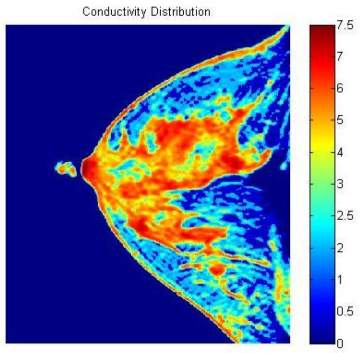


Figure 4.29: Conductivity distribution for MRI image “e"

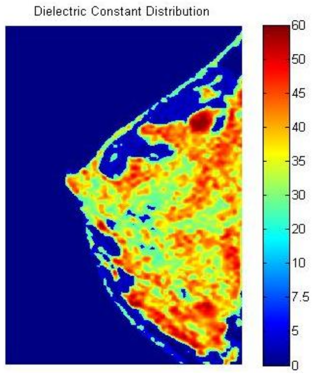


Figure 4.30: Dielectric distribution for MRI image “m"

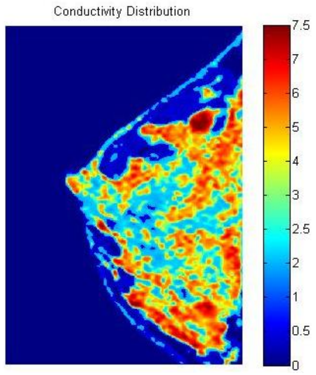


Figure 4.31: Conductivity distribution for MRI image “m"

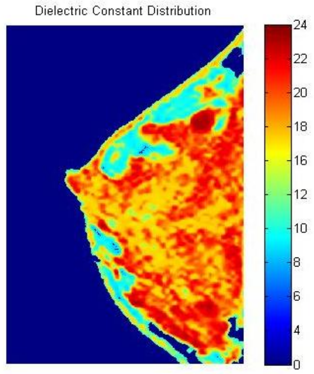


Figure 4.32: Dielectric distribution for MRI image “m"

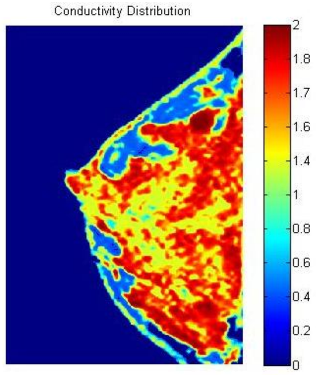


Figure 4.33: Conductivity distribution for MRI image “m"

4.3.2 Cubic spline interpolation

This stage differs from the previous stage in the way the intensity values are mapped to the corresponding tissues' dielectric values. Instead of using a piecewise linear function, piecewise cubic splines are occupied to present the relation between the MRI intensity values and dielectric constant and conductivity values.

The mapping parameters are firstly assigned by means of both customary methods and secondly by using the GMM as done in the piecewise linear interpolation application.

Moreover, as done in the previous technique two different dielectric intervals are used. Figure 4.34, Figure 4.35, Figure 4.38, Figure 4.39, Figure 4.42 and Figure 4.43 show the results obtained by using the dielectric intervals appropriate to the measurements made by [8] and Figure 4.36, Figure 4.37, Figure 4.40, Figure 4.41, Figure 4.44 and Figure 4.45 show the models obtained by applying 10% and 15% variances to the values stated by [28] conveniently.

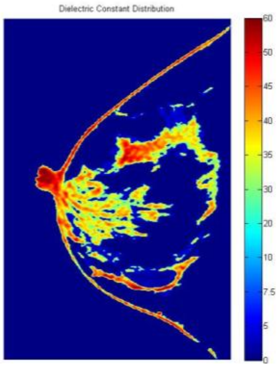


Figure 4.34: Dielectric distribution for MRI image "z"

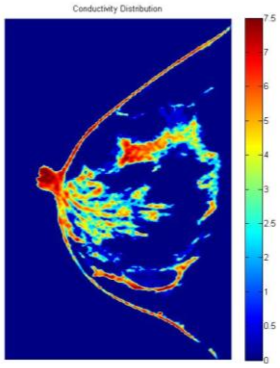


Figure 4.35: Conductivity distribution for MRI image "z"

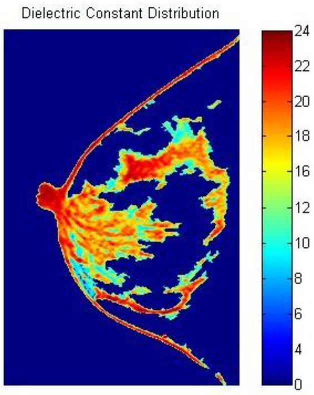


Figure 4.36: Dielectric distribution for MRI image “z”

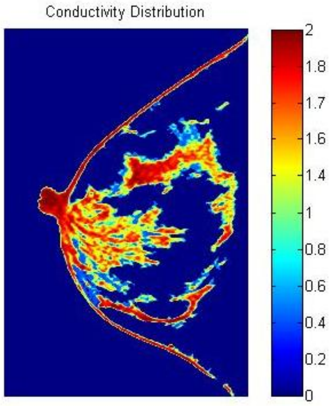


Figure 4.37: Conductivity distribution for MRI image “z”

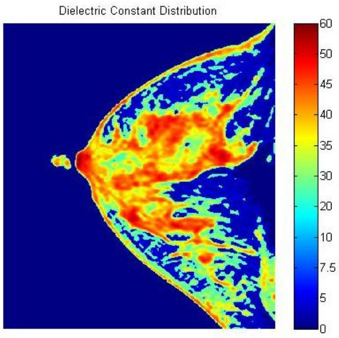


Figure 4.38: Dielectric distribution for MRI image “e”

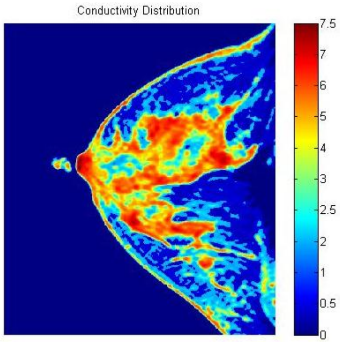


Figure 4.39: Conductivity distribution for MRI image “e”

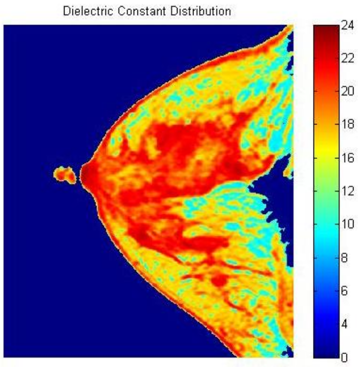


Figure 4.40: Dielectric distribution for MRI image “e”

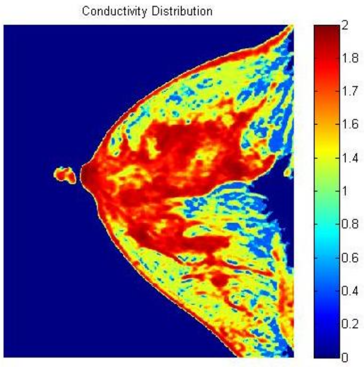


Figure 4.41: Conductivity distribution for MRI image “e”

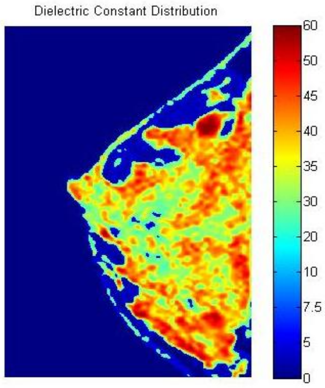


Figure 4.42: Dielectric distribution for MRI image “m”

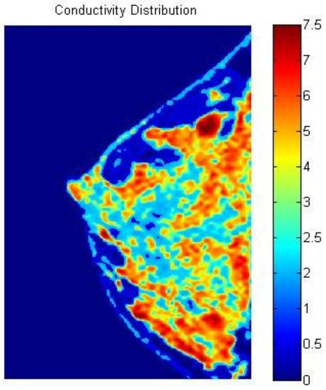


Figure 4.43: Conductivity distribution for MRI image “m”

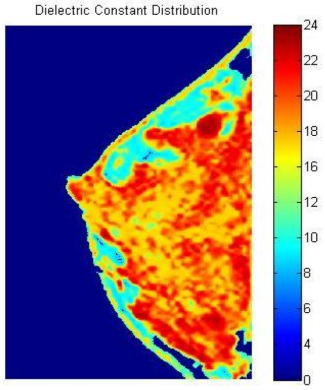


Figure 4.44: Dielectric distribution for MRI image “m”

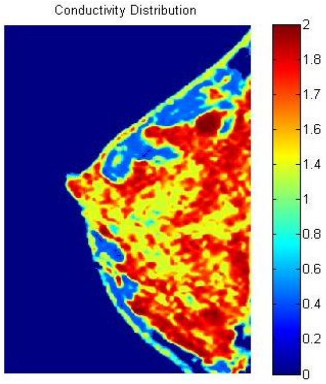


Figure 4.45: Dielectric distribution for MRI image “m”

4.4 Results and Discussion

As it can be seen from Figure 4.22 to Figure 4.45, piecewise linear interpolation and cubic spline interpolation give similar results. However, piecewise linear interpolation seem to be more powerful in exhibiting the dispersive nature of the breast tissues as the dielectric values differ more than they do in the cubic spline interpolation. On the other hand, cubic spline interpolation may be more useful in determining the location of the regions with high dielectric values, which also have higher probability of being suspicious or cancerous, as the contrast between the tissues are sharply illustrated.

Moreover, using the expected values and variances obtained from GMM did not give consistent results. It is assured that the estimated parameters of GMM via the EM algorithm vary depending on the initial parameters, thus, the consequent models differ. Therefore, distinct Gaussian distributions for both of the main tissues, which are illustrated in Figure 4.19, Figure 4.20 and Figure 4.21, are used in estimating the mapping parameters indirectly. Additionally, these graphs help in perceiving whether a given breast model is composed mostly of fatty or glandular tissues. For instance, it can be presumed that the MRI image denoted by “m” is mostly glandular whereas the MRI images denoted by “e” and “z” have more moderate distributions by having a look at the Gaussian curves in the histograms and these presumptions are compatible with the reports of the clinics.

Furthermore, two different dielectric intervals are considered. Firstly, the dielectric constant and conductivity values at the boundaries of the tissues computed at 6 GHz by [8] are used in the definition of the dielectric range and then these boundaries are computed by applying the commonly accepted variance values (10% and 15%) to the dielectric constant and conductivity values stated by [28] for 6 GHz. It’s been shown that there hasn’t been convenience between the dielectric values obtained by different researches.

According to the report on the MRI image denoted by “e”, there is a solid, localized, hypoechogenic lesion in two o’clock radial direction and scattered tubular hypoechogenic areas in four o’clock radial direction [26]. In all the models obtained for the breast image “e” (See Figure 4.26, Figure 4.27, Figure 4.28, Figure 4.29 and Figure 4.38, Figure 4.39, Figure 4.40, Figure 4.41), these areas correspond to the regions with high dielectric values, in other words, suspicious regions as well. Nevertheless, cubic spline approach provides more accuracy in the boundaries of these regions.

Additionally, in the report for the MRI image “z”, it’s stated that there is a solid lesion of size 11x10 mm with spicular contours, localized in 11 o’clock radial direction exhibiting malignant properties and there are two similar lesions which may possibly be malignant in 1 and 6 o’clock radial directions. Moreover, multiple small cysts of varying sizes are present [26]. For the slice processed as “z”, these results are presumable, but more slices should be assessed for a further interpretation. On the other hand, this also shows the expected difference between the two interpolation techniques. As piecewise linear interpolation is tending to disperse the dielectric values, the severity of the situation is hard to guess from Figure 4.22 - Figure 4.25

whereas cubic spline approach has given more vivid results in Figure 4.34 - Figure 4.37.

Furthermore, the report for MRI image “m” asserts that the breast contains multiple small circular lesions with radius smaller or equal to 2 cm and cysts smaller than 2 or 3 cm [27]. Figure 4.30 – Figure 4.33 and Figure 4.42 - Figure 4.45 are in line with these assertions as they all present highly heterogeneous dielectric distributions. Furthermore, the circular structure of the lesions and cysts are also visible in the images, especially in Figure 4.42 – Figure 4.45.

5. CONCLUSION

The major purpose of this study was to investigate the variations in the electrical properties of breast tissues to exhibit the heterogeneity of the breast and the dispersive nature of the dielectric properties across the breast tissues. For this reason, MRI based breast models which will assist the research in breast cancer detection were constructed. Several filters were applied to the MRI images to remove noise and obtain more clear images for further analysis. Then by plotting the histograms of each image, the distribution of the intensity values and accordingly the distribution of the dielectric values were demonstrated. After that, Gaussian distributions of each of the two baseline tissues were used to identify the tissue types due to the intensity values in the images.

Moreover, several methods were tested in mapping the intensity values to dielectric values. For instance, the mapping parameters were derived in two ways. Firstly, they were computed by means of the Gaussian distribution for both of the basic tissues and secondly, they were computed by constructing GMM models. When the results were compared, the first way for determining the coefficients seemed to give more stable, rapid and realistic results. On the other hand, two different interpolations, namely, piecewise linear interpolation and cubic spline interpolation, were used to transform the intensity values into dielectric values. It is deduced that piecewise linear interpolation is more advantageous than cubic spline interpolation in demonstrating the dispersive nature of the breast tissues. However, cubic spline interpolation may be superior to piecewise linear interpolation in determining the locations of the regions with high dielectric values, which also have higher probability of being suspicious or cancerous, as it results in greater contrast between the tissues and provides more accuracy in the transition boundaries.

REFERENCES

[1] **American Cancer Society,** *Cancer Fact & Figure 2009.*

[2] **O'Halloran, M., Conceiçao, R., Byrne, D., Glavin, M. and Jones, E., 2009.** FDTD Modeling of the Breast: A Review, *Progress In Electromagnetics Research B,* Vol. 18, 1-24.

[3] **Li, X., Davis, S. K., Hagness, S.C., van der Weide, D.W. and Van Veen, B.D., 2004.** Microwave Imaging via Space-Time Beamforming: Experimental Investigation of Tumor Detection in Multilayer Breast Phantoms, *IEEE Trans. On Microwave Theory and Tcehniques,* Vol.52, No.8.

[4] **Cardillo, F.A., Starita, A., Caramella, D. and Cilotti A.**, 2002. A Tool for Breast MRI Images Processing and Classification, *IEEE.*

[5] **Sim, K.S., Chia, F.K., Chong, S.S., Tso, C.P., Abbas, S.F. and Omar, S.,** 2010.Spatial Based MRI Images Lesion Detection and Colourisation, *IEEE EMBS Conference on Biomedical Engineering & Sciences, IECBES 2010,* Kuala Lumpur,Malaysia.

[6] **Fear, E.C., Hagness, S.C., Meaney, P.M., Okoniewski, M. And Stuchly, M.A.,2002.** Enhancing Breast Tumor Detection with Near-Field Imaging, *IEEE Microwave Magazine*, pp 48-55.

[7] **Kosmas, P. and Rappaport, C.M., 2005.** Time Reversal with the FDTD Method for Microwave Breast Cancer Detection, *IEEE Trans. On Microwave Theory and Techniques,* Vol.53, No.7, pp. 2317-2322

[8] **Zastrow, E.,Davis, S. K., Lazebnik, M., Kelcz, F., Van Veen, B.D. and Hagness, S. C., 2008.** Development of Anatomically Realistic Numerical Breast Phantoms with Accurate Dielectric Properties for Modeling Microwave Interactions with the Human Breast, *IEEE Trans.on Biomedical Engineering*, Vol.55, No.12, pp 2792-2799.

[9] **Varini, C., Nattkemper, T. W., Degenhard A. and Wismüller, A., 2004**. Breast MRI Data Analysis by LLE, *IEEE*

[10] **Wang, H., Huo, Z. and Zhang, J.,2005.** Computerized Classification Method for Differentiating Between Benign and Malignant Lesions on Breast MR Images, *IEEE Engineering in Medicine and Biology 27th Annual Conference,* Shangai, China.

[11] **Chaudhary, S. S., Mishra, R. K. , Swarup, A. and Thomas, J. M., 1994**. Dielectric properties of normal and malignant human breast tissue at radiowave and microwave frequencies, *Indian J. Biochem. Biophys.*, Vol. 21, pp 76-79.

[12] **Joines, W. T., Zhang, Y., Li, C. and Jirtle, R. L., 1993.** The measured electrical properties of normal and malignant human tissues from 50 to 900 MHz, *Med. Phys*., Vol. 21, pp 547-550.

[13] **Surowiec, A. J., Stuchly, S. S., Barr, J. R. and Swarup, A., 1988**. Dielectric properties of breast carcinoma and the surrounding tissues, *IEEE Trans. Biomed. Eng.*, Vol. 35, No. 4, pp 257-263.

[14] **Campbell, A. M. and Land, D. V. ,1992.** Dielectric properties of female human breast tissue measured in vitro at 3.2 GHz, *Phys. Med. Biol.*, Vol. 37, pp 193-210.

[15] **Lazebnik, M., McCartney, L., Popovic, D., Watkins, C. B., Lindstrom , M. J., Harter, J., Sewall, S., Magliocco, A., Booske , J. H., Okoniewski, M. and Hagness, S. C., 2007.** A large-scale study of the ultrawideband microwave dielectric properties of normal breast tissue obtained from reduction surgeries, *Phys. Med. Biol.*,Vol. 52, pp 2637-2656.

[16] **Lazebnik, M., Popovic, D., McCartney, L., Watkins, C. B., Lindstrom, M. J., Harter, J., Sewall, S., Ogilvie, T., Magliocco, A., Breslin, T. M. , Temple, W., Mew, D., Booske, J. H., Okoniewski, M. and. Hagness, S. C., 2007.** A large-scale study of the ultrawideband microwave dielectric properties of normal, benign and malignant breast tissues obtained from cancer surgeries, *Phys. Med. Biol.*,Vol. 52, pp 6093-6115.

[17] **Li, X. and Hagness, S. C. , 2001.** A confocal microwave imaging algorithm for breast cancer detection, *IEEE Microwave and Wireless Components Letters*, Vol. 11, 130-132.

[18] **Gonzalez, R.C. and Woods, R.E., 2002**. *Digital Image Processing*, 2nd ed.,pp 76-89, Upper Saddle River, NJ: Prentice Hall

[19]**Madisetti, V. and Williams, D.,1999.** *The Digital Signal Processing Handbook*, pp 1077-1079

[20] **Url-1 <** http://masters.donntu.edu.ua/2007/kita/gett/library/eng.htm> accessed at 15/04/2011….

[21] **Russ, J. C.,** *Image Processing Handbook*, 5th Ed., pp 281-287

[22]**Url-2 <** http://www.mathworks.com/help/toolbox/images/f19-12807.html> accessed at 20/04/2011

[23] **Reynolds, D.***Gaussian Mixture Models*

[24] **Kıusalaas, J.,2010.** *Numerical Methods in Engineering with Matlab*, 2nd Ed.,pp 116-118

[25] **Otto, S.R. and Denier, J.P.,2005.** *An Introduction to Programming and Numerical Methods in Matlab*, pp150-152

[26] **Euromed Radiology Center**

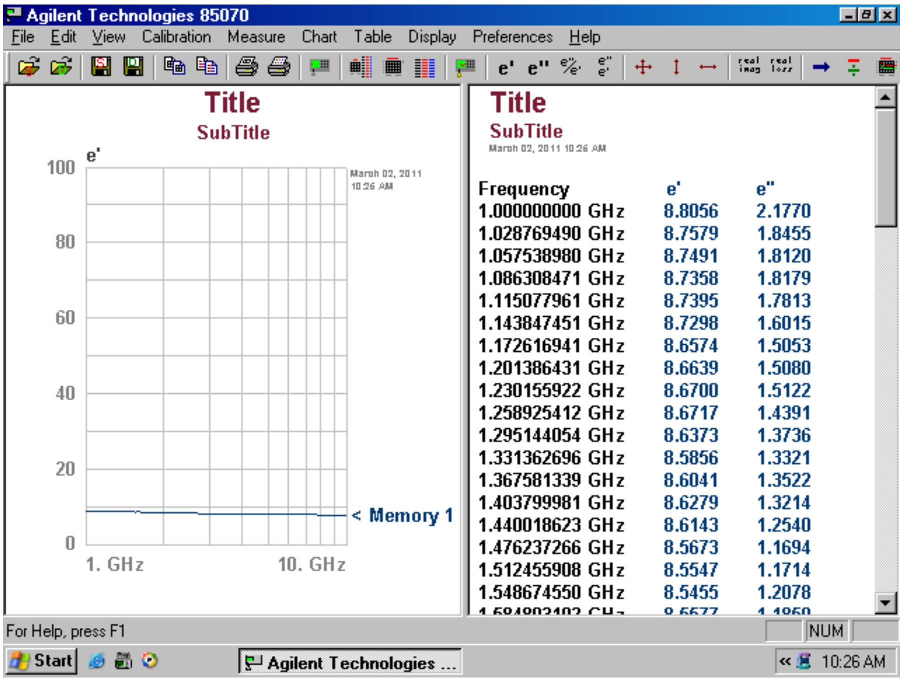
[27] **Maltepe University**

[28] **Converse, M., Bond, E.J., Van Veen B.D., Hagness, S.C., 2006.** A computational study of ultrawideband versus narrowband microwave hyperthermia for breast cancer treatment, *IEEE Trans.on Microwave theory &Techniques*, Vol.54, No.5, pp. 2169-2179.

APPENDICES

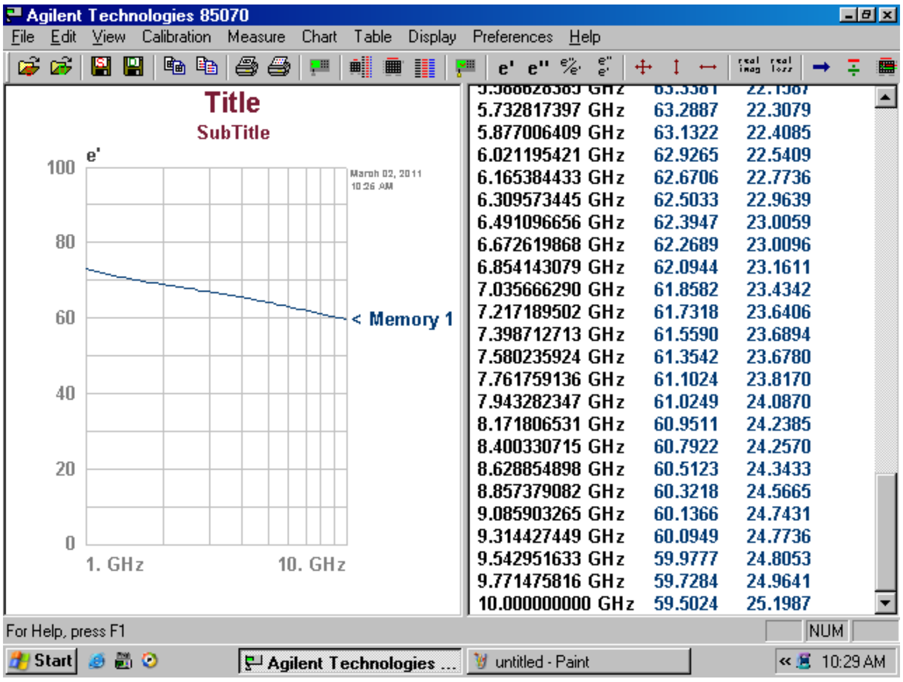
APPENDIX A.1 : Fatty & Cancerous Breast Tissue Measurements, Acquired from Electromagnetics Research Group, Istanbul Technical University

APPENDIX A.1



Frequency	e'	e''
1.000000000 GHz	8.8056	2.1770
1.028769490 GHz	8.7579	1.8455
1.057538980 GHz	8.7491	1.8120
1.086308471 GHz	8.7358	1.8179
1.115077961 GHz	8.7395	1.7813
1.143847451 GHz	8.7298	1.6015
1.172616941 GHz	8.6574	1.5053
1.201386431 GHz	8.6639	1.5080
1.230155922 GHz	8.6700	1.5122
1.258925412 GHz	8.6717	1.4391
1.295144054 GHz	8.6373	1.3736
1.331362696 GHz	8.5856	1.3321
1.367581339 GHz	8.6041	1.3522
1.403799981 GHz	8.6279	1.3214
1.440018623 GHz	8.6143	1.2540
1.476237266 GHz	8.5673	1.1694
1.512455908 GHz	8.5547	1.1714
1.548674550 GHz	8.5455	1.2078

For Help, press F1



Frequency	e'	e''
5.732817397 GHz	63.2887	22.3079
5.877006409 GHz	63.1322	22.4085
6.021195421 GHz	62.9265	22.5409
6.165384433 GHz	62.6706	22.7736
6.309573445 GHz	62.5033	22.9639
6.491096656 GHz	62.3947	23.0059
6.672619868 GHz	62.2689	23.0096
6.854143079 GHz	62.0944	23.1611
7.035666290 GHz	61.8582	23.4342
7.217189502 GHz	61.7318	23.6406
7.398712713 GHz	61.5590	23.6894
7.580235924 GHz	61.3542	23.6780
7.761759136 GHz	61.1024	23.8170
7.943282347 GHz	61.0249	24.0870
8.171806531 GHz	60.9511	24.2385
8.400330715 GHz	60.7922	24.2570
8.628854898 GHz	60.5123	24.3433
8.857379082 GHz	60.3218	24.5665
9.085903265 GHz	60.1366	24.7431
9.314427449 GHz	60.0949	24.7736
9.542951633 GHz	59.9777	24.8053
9.771475816 GHz	59.7284	24.9641
10.000000000 GHz	59.5024	25.1987

For Help, press F1

Figure A.1 : Fatty & Cancerous Breast Tissue Dielectric Constant Measurements